# Electron Microscopic Studies of Sequence Arrangement: Poly(A) Mapping, RNA Tumor Viruses, and Slime Mold Actin Genes

In partial performance
of the rites of passage
Caltech

1978

Welcome Bender

to the most recent and perfect reincarnation

of the eternal feminine

Acknowledgements

For six enjoyable years in Pasadena I am most indebted to the Caltech biology community, to faculty who treat first year students as colleagues and staff who accepted my problems as their own. I learned most from watching the style of Norm Davidson. He cares a lot about science and more about people. What I know of biology I learned primarily from an endless string of weekly seminars with Dick Russell, Ed Lewis, and Eric Davidson. What I know of the witchcraft of electron microscopy, I learned from Sylvia Hu. Carl and Carole Johnson gave me a home, and Ed Low and friends added music. Financial support came from Geraldine and Eleonora, who got it, I think, from the National Science Foundation and the National Institutes of Health.

## Justifications and Disclaimers

The goal of my graduate research has been to develop methods of labeling poly(A) sequences for visualization in the electron microscope, and to apply the techniques to biologically interesting problems. This thesis includes six reprints or manuscripts. The first five deal with observations of the structures of the RNA genomes of RNA tumor viruses. The sixth is about sequence arrangement in a cloned slime mold actin gene. In all cases, the contributions have multiple authors, and what follows is an attempt to give credit where due. All manuscripts were extensively rewritten (and thus considerably improved) by Norman Davidson. In all the tumor virus papers, the non-Caltech authors supplied short-harvest virus or purified viral 70S RNA.

The first paper presents the poly(A) mapping technique and its use in labeling the 3' ends of RNA; this allowed us to orient the monomer subunits in the 70S RNA dimers from RNA tumor viruses. The basic dimer structure of the viral RNA from the cat virus, RD114, had already been well characterized by Hsing-Jien Kung and Jim Bailey. Hsing-Jien and Sylvia Hu provided the 70S RNA on which the poly(A) mapping was done.

The second paper, which was published back to back with the first, presents detailed secondary structure information on the RNAs from two tumor viruses, BKD (from baboons) and WoMV (from woolly monkeys). Both viral RNAs turned out to

look quite similar to the RD114 dimer. The work here is almost entirely that of Sylvia and Hsing-Jien, who studied BKD and WoMV, respectively. My contribution was in Poly(A) mapping of RNA dimers, from which molecules the measurements of figure 4 and table 1 were derived, and in writing the manuscript.

In the third paper, we looked at the RNA from a murine tumor virus to see if it also had the same structure. The Friend virus that we examined proved to be unexpectedly complicated, in that it was a mixture of several different viral components. But we found the dimer structure in all components and our data suggested that the dimers only formed between two identical monomers. All of the spreadings from urea plus formamide reported here and the micrographs of figure 3 were done by Hsing-Jien. The gel result of figure 1 was duplicated many times by Hsing-Jien and Shyam Dube.

The fourth manuscript describes a careful search for viral RNA heterodimers (dimers between two different sorts of monomers). We looked at a mixture of Moloney murine leukemia and sarcoma viruses because we knew that the leukemia and sarcoma RNAs were related or identical in the region where the dimers are held together, and because Jan Maisel had preliminary evidence for heterodimers in her gel results. Microscopy showed rather conclusively that heterodimers are very rare or nonexistent, and the paper also presents a sampling of the many inconclusive experiments to reconcile that finding with the initial gel observations. The gels of

figure 2 were done by Jan and were repeated many times by Sylvia and myself. The microscopy of the leukemia viral RNA and of the clone 3 RNA was primarily done by Sylvia; figure 10, table 1, and half of figures 3 and 4 are her work. Jan did the gel analysis of clone 124 RNA (figure 6) and Peter Duesberg did the gels of 124-5R RNA (figure 7) and the urea gels of clone 3 RNA. All of the authors wrote initial drafts; Norm and I worked out the present compromise. The paper was submitted to the Journal of Virology and has been accepted, pending the addition of hybridization evidence that clone 3 and clone 124-5R RNAs are what we think they are. That evidence has recently been gathered, and the revised manuscript has been returned to the Journal of Virology.

The fifth manuscript, which we call the scorecard paper, is a collection of all the RNA structures we saw in other viral RNAs which were yet to be published. Most important is the RNA dimer structure of REV, an avian virus, which suggests that the dimer linkage is not peculiar to mammalian viruses. We took the discussion here as an opportunity for uninhibited speculation on the nature and function of the various RNA structures seen. All the microscopy on both wild mouse viruses was done by Yueh-hsiu Chien. The manuscript presently awaits corrections and comments from outside authors before submission to the Journal of Virology.

The last manuscript presents initial mapping data for a recombinant plasmid which contains an actin gene from the

the cellular slime mold, Dictyostelium discoidium. The project began as an attempt to improve the poly(A) mapping technology to permit labeling of very short poly(A) sequences. The Dictyostelium actin gene was chosen to test the method and to see if the short poly(A) tracts have anything to do with messenger coding regions. The method works well, but poly(A) sequences do not have any obvious relation to this actin gene. In the process we had to locate and orient the actin coding region within the recombinant plasmid. All the plasmid constructions and preparations of plasmid DNA were done by Rick Firtel and members of his group. Most of the restriction mapping data is the work of Karen Kindle; in the final version of this paper the description of the restriction digestions will be expanded and a figure will be added. The recloning of the actin gene into $\lambda_{gt}$'s and the analysis of their orientation was done by Mike Silverman; again, the description will be expanded and a figure added in the final version. Bill Taylor collaborated in the heteroduplex studies of the H6, B1, and A1 plasmids. This manuscript is planned as the second of a group of three to be submitted together to Cell. The preceeding paper (by Kindle and Firtel) will describe the isolation of the actin plasmid and the evidence that the plasmid hybridizes to actin messenger RNA. The following paper (by McKeown et al.-I expect to be one of the minor authors) will contain hybridization data on the number and diversity of slime mold actin genes and a detailed mapping of a second actin plasmid.

Cell, Vol. 7, 595–607, April 1976, Copyright © 1976 by MIT

1

# Mapping of Poly(A) Sequences in the Electron Microscope Reveals Unusual Structure of Type C Oncornavirus RNA Molecules

**Welcome Bender* and Norman Davidson†**
Division of Biology* and Department of Chemistry†
California Institute of Technology
Pasadena, California 91125

## Summary

**We have synthesized a convenient electron microscope label for mapping poly(A) sequences. Short lengths of poly(dT) are polymerized onto nicked circular SV40 DNA with the enzyme terminal deoxynucleotidyl transferase. An RNA or DNA molecule of interest is treated with glyoxal, hybridized briefly with the poly(dT) circles, and spread for microscopy; poly(A) stretches are clearly marked because they are attached to the poly(dT) on the easily recognized SV40 duplex circles.**

**The RNAs of several type C oncornaviruses were examined by this method. The endogenous feline virus (RD-114), the endogenous baboon virus (BKD), and the wooly monkey sarcoma virus (WoMV) all contain a dimer of RNA subunits held together in a central secondary structure feature we call the dimer linkage structure. Both ends distal to the dimer linkage structure hybridize to the SV40-poly(dT). Assuming both poly(A)s are on the 3′ ends of the subunits and that both subunits are identical, the two identical subunits are held together by interactions between sequences close to the 5′ ends.**

## Introduction

We have developed a reliable, efficient method for identifying and mapping poly(A) stretches in nucleic acids by electron microscopy. We have used this technique to study the structure of the RNA of RD-114 and other tumor viruses.

Electron microscope studies in this laboratory (Kung et al., 1975; Kung et al., 1976) show that the high molecular weight RNA molecules extracted from several different RNA tumor viruses are dimers, with reproducible secondary structures. Each of the two monomer subunits of any one dimer has a molecular length of approximately 10 kb (kb = kilobase, 1000 nucleotides).

In the dimer, an end of one monomer is noncovalently joined to an end of the other monomer in a central secondary structure feature. This feature in RD-114 RNA was previously called the "rabbit ears" (Kung et al., 1975), but because its shape is different in different tumor virus RNAs, we now prefer the more general term, dimer linkage structure. Each dimer thus has two monomer ends joined in the dimer linkage structure and two free ends. We have determined the positions of poly(A) sequences in the dimers and in dissociated monomers.

The obvious method for electron microscope mapping of poly(A) stretches in a nucleic acid strand is to hybridize the poly(A) to a poly(dT) or poly(U) strand attached to a suitable electron microscope label. The initial problem encountered is that A·T and A·U base pairs are relatively unstable, so that the moderately denaturing conditions needed to extend single-strand nucleic acids in electron microscope spreads cause dissociation of the poly(A)·poly(dT) or poly(A)·poly(U) duplexes. This difficulty was overcome by Hsu, Kung, and Davidson (1973) by treating the nucleic acid with glyoxal. Under suitable reaction conditions, glyoxal is bound only to G residues and blocks their base pairing capability (Broude and Budowsky, 1971). This reaction disrupts the secondary structure of most nucleic acids sufficiently so that they can be extended for electron microscopy by weakly denaturing solvents (such as 30% formamide, 0.1 M Tris) in which poly(A)·poly(dT) duplexes are stable. Hsu and his co-workers used long poly(dT) strands to label the poly(A) sequences at the end of Sindbis virus RNA. This method was moderately effective, but discrimination between the poly(dT) strand and the Sindbis RNA strand was not always certain. Lizardi, Williamson, and Brown (1975) hybridized poly(dT) to glyoxal-treated silk fibroin mRNA, then added excess poly(A), and observed the resulting duplex regions at the ends of some of the mRNA molecules. In this method, the duplex region cannot always be clearly recognized. For whatever reasons, the overall efficiency of labeling by either method was rather low.

Carbon, Shenk, and Berg (1975) constructed a better label for poly(A) mapping by cleaving SV40 DNA circles with RI endonuclease, resecting the 5′ ends with λ-exonuclease, and polymerizing short poly(dT) on the exposed 3′ ends with terminal deoxynucleotidyl transferase. We used some of Carbon's extended SV40 DNA and found it very efficient at labeling poly(A) on Sindbis RNA. The poly(A) stretch was marked where a linear SV40 duplex DNA joined a linear Sindbis single-strand RNA, but the contrast between single strands and double strands in the electron microscope was seldom good enough to pinpoint the SV40 DNA to Sindbis RNA junction. We subsequently noticed that terminal transferase can polymerize poly(dT) at nicks in nicked circular SV40 DNA (as previously reported by Jackson and Berg in a personal communication to Lobban and Kaiser, 1973). We have found that the resulting SV40 circles with poly(dT) tails are excellent markers for poly(A) mapping.

## Results

### SV40-Poly(dT)

When SV40 DNA was incubated with terminal transferase and thymidine triphosphate under the conditions described in Experimental Procedures, the enzyme polymerized long poly(dT) tails on the circle, presumably at single-strand nicks (Figure 1). In trial reactions using an SV40 DNA concentration of 20 μg/ml, the average tail length was 650 bases after 90 min and 2000 bases after 270 min. After the 90 min reaction, there were, on average, two tails per circle. 75% of the treated SV40 DNA bound to poly(A)-sepharose and was eluted in 0.1 M NaCl at about 65°C, the expected melting temperature for poly(A)·poly(dT) duplexes (Riley, Maling, and Chamberlin, 1966). 24 hr incubations did not produce tails longer than about 2000 bases on average, but there were more tails per circle and many unattached pieces of poly(dT). About half the SV40 DNA was in the supercoiled form initially, but almost all the molecules contained poly(dT) after the incubation, so that some endonucleolytic nicking must occur during the terminal transferase reaction.

To prepare SV40 with very short tails, we used a 1 hr incubation with 500 μg/ml SV40 DNA. This produced an average of two visible tails per circle, each about 175 bases long.

### Sindbis RNA

Sindbis RNA was used as a first test system for poly(A) mapping techniques, since the RNA is long and had been studied previously (Hsu et al., 1973). Over 90% of Sindbis molecules contain poly(A), but the size of the poly(A) sequence is somewhat heterogeneous; the average length is about 60 nucleotides, with about 60% of the poly(A) molecules between 50 and 70 bases long (T. Frey and J. Strauss, manuscript in preparation). Sindbis RNA treated with glyoxal was first hybridized with a preparation of SV40 with long (about 2 kb) dT tails. Many clear hybrids were found (Figure 2), but the long tails of poly(dT) were a disadvantage, since the Sindbis RNA often hybridized to the end of the poly(dT). The resolution of these cytochrome spreads is not sufficient to distinguish the poly(A)·poly(dT) duplex region from the single strands, so that the poly(A) end of the Sindbis RNA could be anywhere within a few thousand bases of the SV40. In subsequent experiments, we used SV40 with much shorter tails —about 175 bases long. For the long RNA molecules of Sindbis or RNA tumor viruses, the length of the short poly(dT) tails was negligible, and the poly(A) end of the RNA was assumed to start at the SV40 circle.

In measuring the efficiency of poly(A) hybridization to the SV40-poly(dT), it was not practical to score only full-length (14 kb) Sindbis RNA molecules since very few molecules were unbroken. Of the RNA molecules greater than half size (7 kb), the fraction in clear hybrids with SV40-poly(dT) was only about 40%, even if the RNA was preselected for poly(A) sequences by binding to oligo(dT)-cellulose. It is probable that some of the shorter Sindbis poly(A) sequences would dissociate from the SV40-poly(dT) under our spreading conditions. Breakage of the RNA during the annealing procedure might also lower the apparent hybridization efficiency. Despite the incomplete labeling of Sindbis poly(A), it was clear from these initial studies that the SV40-poly(dT) circles would be effective for mapping poly(A) stretches in tumor virus RNA.

### RD-114

The RNA of the endogenous feline type C virus, RD-114, has been extensively studied in our laboratory (Kung et al., 1975). The 52S RNA from the virion appears to be a dimer (20 kb) whose two 26S subunits (10 kb) are held together in a Y shaped or Y shaped structure that we had previously named the rabbit ears, although we now prefer to describe it as the dimer linkage structure. An additional secondary structure feature observed in this RNA is a loop with a circumference of 3.8 kb located about 2.4 kb from the dimer linkage. The conditions normally used for electron microscope spreading are sufficiently denaturing so that a loop is opened up about 50% of the time.

When the 52S dimer RNA of RD-114 is treated with glyoxal and allowed to hybridize with SV40-poly(dT), the majority of the molecules have structures similar to those shown in Figure 3. The RNA forms a circular structure with the two free ends joined to a single SV40-poly(dT) circle at different (dT) tails and with the dimer linkage structure in the center of the RNA. Thus there is a poly(A) sequence at each of the two free ends of the 52S RD-114 RNA. In the molecule shown in Figure 3A, both loops have opened up; both loops are present in the molecule shown in Figure 3B.

It is rather rare (1 of 67 cases) to see an RD-114 dimer with each end joined to a different SV40-poly(dT) circle (such as is shown in Figure 5 for the RNA from BKD, a baboon virus). This is expected according to the theory of ring closure of linear chains (Jacobson and Stockmayer, 1950; Wang and Davidson, 1966), since the equivalent concentration of the unhybridized poly(A) end in the neighborhood of an end already bound to SV40 is about $1.4 \times 10^{-8}$ M [assuming a Kuhn statistical segment

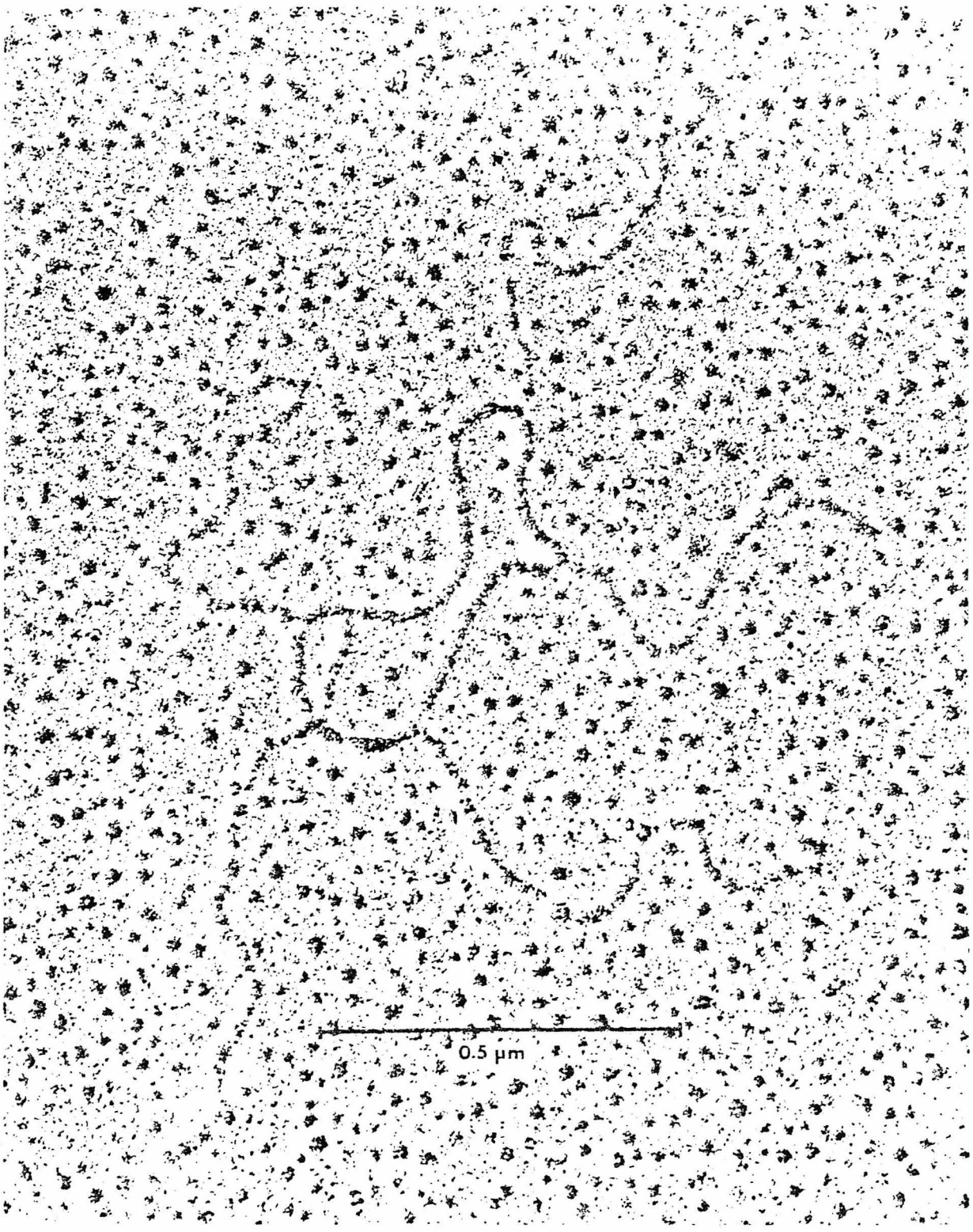



Figure 1. Electron Micrograph of a Molecule of Circular Double-Strand SV40 DNA (Contour Length 5.5 kb) with 5 Attached Poly(dT) Tails

The tails range in length from 2.2 to 3.0 kb. The molecule was produced by a 4.5 hr terminal transferase reaction with 20 µg/ml of SV40 DNA.


Figure 2. A Molecule of Sindbis RNA Hybridized to an SV40 Circle with Long Poly(dT) Tails
The arrow indicates the fork that we identify as the junction of the RNA with the poly(dT).

length (b) of $2 \times 10^{-6}$ cm for RNA (see appendix to Schmid, Manning, and Davidson, 1975)], whereas the concentration of free SV40 molecules is only about $5 \times 10^{-10}$ M.

If the glyoxal-treated RD-114 is heated to 90°C for 10 sec in 10 mM Tris, 1 mM EDTA before hybridization and spreading, the dimers dissociate at the dimer linkage structure, and the 3.8 kb loops also melt out, leaving 10 kb linear RNA subunits which hybridize to the SV40-poly(dT) at one end only (data not shown). Thus there are no poly(A) sequences buried within the dimer linkage structure.

We wish to count directly the fraction of RD-114 molecules with poly(A) stretches. RD-114 RNA was hybridized with SV40-poly(dT) and spread. A large random sampling of RD-114 molecules was photographed whether or not they were attached to SV40. The length of the RD-114 subunit was determined as the average distance from ends clearly hybridized on SV40 to a clearly defined dimer linkage structure ([L] = 9.8 kb; standard deviation, $\sigma$, = 1 kb). All molecules were measured from end to dimer linkage, or from end to end if there was no dimer linkage structure, and those whose lengths were within 1 standard deviation of [L] were scored (Figure 4). Of these, 76% (88/115) were hybridized to the SV40 circles. This is a minimum estimate of the percentage of monomers with poly(A), because monomer molecules with a few hundred bases at the 3′ end broken off would be scored as full length, and because the hybridization efficiency may be <100%. The 76% also gives a minimum estimate for the efficiency of hybridization and scoring with the SV40-poly(dT) circles.

### Other Tumor Viruses

We have examined the RNAs of several other type C viruses to see if they have similar secondary structure and poly(A) position. BKD, an endogenous baboon virus, contains RNA which after glyoxal treatment shows much the same structure as found with RD-114 (Kung et al., 1976). Again, we see a dimer linkage structure and 3.7 kb loops, 1.9 kb from the dimer linkage. The dimer linkage structure itself is slightly different, usually looking like ⅄ or + or sometimes ⅄. Hybrids with SV40-poly(dT) again show poly(A) regions at both free ends (Figure 5).

WoMV is a simian sarcoma virus isolated from a woolly monkey fibrosarcoma. Some investigators refer to this virus preparation as SSV-1. It should be noted that WoMV preparations contain a large excess of nontransforming helper virus (Wolfe, Smith, and Deinhardt, 1972), and so it is really the RNA of this associated virus that we are studying here. The usual glyoxal treatment dissociated most dimers into subunits, indicating that the cohesion of the WoMV monomers within the dimer linkage structure is less stable than for RD-114 or BKD. Nevertheless, many dimers were observed and measured. There is a dimer linkage structure, usually ⅄ or + shaped, and loops of about 3.9 kb, 1.4 kb from the dimer linkage (Kung et al., 1976). Both free ends hybridize to SV40-poly(dT) (Figure 6).

For Rous sarcoma virus, the usual 1 hr incubation with glyoxal causes dissociation of all of the 60–70S RNA to 9.5 kb monomers, indicating that the 60–70S complex dissociates more readily than does the WoMV dimer. It is difficult to find spreading conditions under which the RNA is sufficiently extended to be traced while still retaining structures larger than 9.5 kb monomers. The best results so far have been obtained by treating 60–70S RNA with glyoxal for only 10–15 min, dialyzing against 10 mM Tris, 1 mM EDTA as usual, heating that low salt solution to 37°C for 15 sec, and then hybridizing and spreading. This procedure, like the gene 32 spreading method of Mangel, Delius, and Duesberg (1974), gave some molecules with a total contour length of about 19 kb. Figure 7 is a plot of the length distribution of all molecules in a particular sample which were attached to an SV40-poly(dT) label. There is a clear peak at the monomer RNA length of about 9.5 kb. Of the molecules that are of dimer length (about 19 kb), about half are attached to SV40-poly(dT) circles by two poly(A) ends (solid bars). In these molecules, the structures are often not easily interpretable, and the non-poly(A) ends are usually not closely associated. Such a molecule is shown in Figure 8a. We do, however, find rare examples in which the two non-poly(A) ends of the RSV monomers are joined in an apparent dimer linkage structure (Figure 8b).

### Discussion

The method described here should be generally useful for mapping sufficiently long poly(A) sequences at the ends of single-strand RNA molecules. The efficiency of labeling is high for the poly(A) ends on Sindbis RNA and for the several tumor virus RNAs studied, although there is only a slight excess of SV40-poly(dT) molecules over viral RNA molecules in the hybridization mixture. The double-strand SV40 circles stand out clearly in the electron microscope; this permits rapid scanning for hybrids at low magnification. Since the single strand of the hybridized RNA stops right at the double-strand circle, there are very few accidental overlaps that would be scored as hybrids, and the positioning of the end is limited only by the size of the poly(dT) tails. In addition, the SV40 circle provides an automatic internal length standard. This marker should be very useful for map-

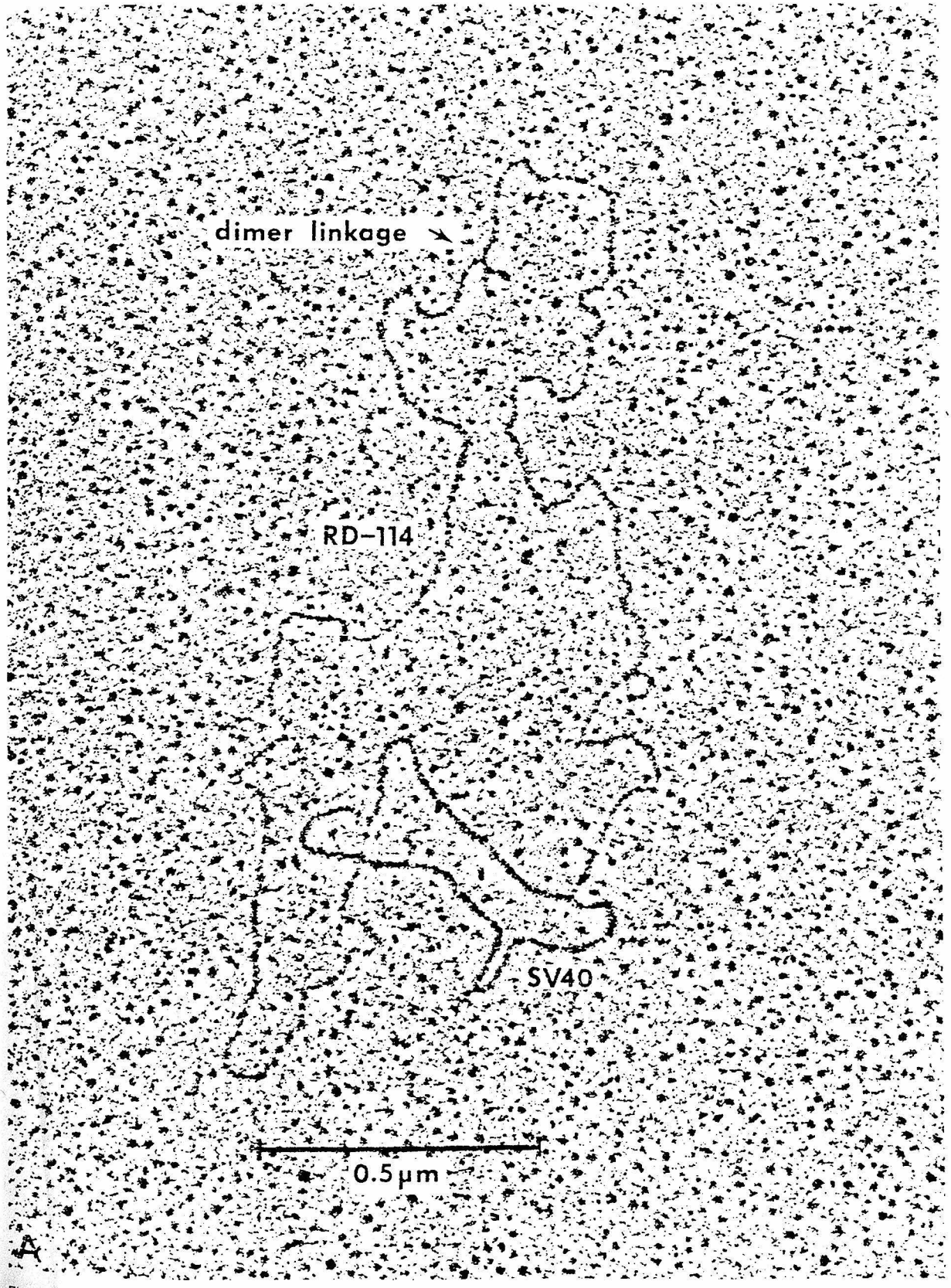
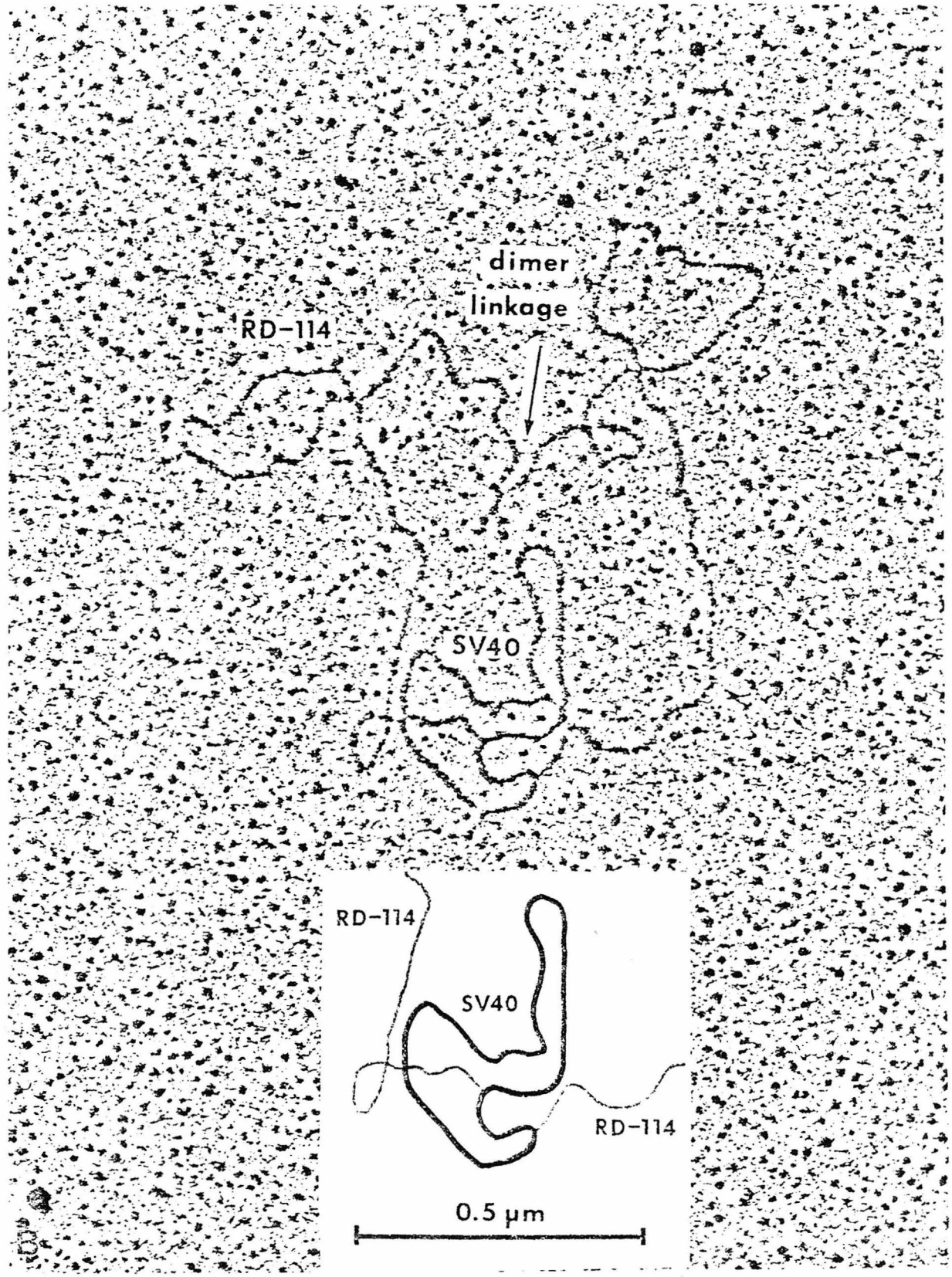



Figure 3. An RD-114 RNA Dimer without Loops (A) and with the Symmetrically Disposed Loops (B), Each Hybridized at Both Ends to an SV40 DNA Circle with Short Poly(dT) Tails

The dimer linkage structure is indicated. A tracing of the connections of the RNA with the SV40 circle is shown for (B).


dimer
linkage
RD-114
SV40
RD-114
SV40
RD-114
0.5 μm

ping secondary structure features or heteroduplex regions with respect to the poly(A) ends of viral RNAs, mRNAs, and heterogeneous nuclear RNAs. However, the technique relies on glyoxal treatment and low formamide spreading; the hybrids are not stable in 50% formamide.

We do not know how the efficiency of labeling varies as a function of the poly(A) length, but we suspect that poly(A)·poly(dT) hybrids shorter than about 50 base pairs would dissociate in 30% formamide. The present method will probably be poor for mapping internal poly(A) sequences within long single-strand molecules, since hybrid SV40 circles would often lie on top of the single-strand molecule. There would then be two or more crossover points, and it would not be clear which intersection is the poly(A) site.

Many oncornavirus RNAs have been shown to contain poly(A) (Lai and Duesberg, 1972; Green and Cartas, 1972; Ihle, Lee, and Kennedy, 1974; Wang and Duesberg, 1974; Quade, Smith, and Nicols, 1974), and in some cases (including RSV), the end of the poly(A) stretch has been identified as a 3′ end (Wang and Duesberg, 1974; Quade et al., 1974; Rho and Green, 1974). In RSV, it appears that there are only two subunits per 60–70S complex by biophysical (King, 1976) and electron microscopic measurements (Mangel et al., 1974; these results), and all 35S subunits are probably identical (Duesberg et al., 1974; Weissmann et al., 1974). The studies reported by Kung et al. (1975, 1976) show that RD-114, BKD, and WoMV viral RNAs all have a dimer structure with two monomer ends joined in the dimer linkage structure. The present study shows that each monomer has a poly(A) stretch on the free end. Assuming that RD-114, BKD, and WoMV have identical subunits with poly(A)s at the 3′ ends, then the dimer linkage structure must be a complex of two identical (not homologous) 5′ ends. Some speculations as to the nature of the binding of the monomers in the dimer linkage structure are given in the accompanying paper (Kung et al., 1976).

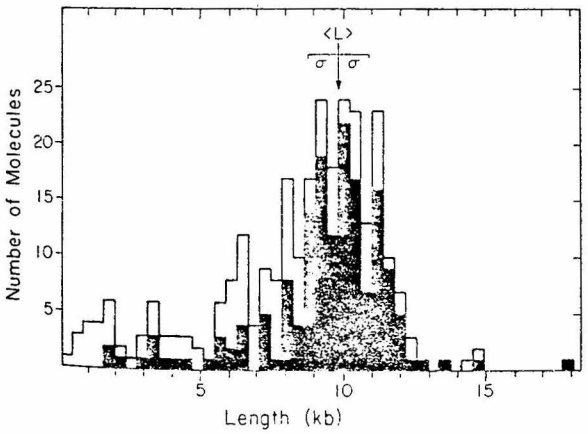



Figure 4. A Measurement of the Fraction of Full-Length Monomer Units Having Poly(A) Ends Attached to SV40-Poly(dT)

All long molecules in fields chosen at random were photographed. Lengths of monomer subunits were measured. The glyoxal treatment caused dissociation of about 25% of the RD-114 RNA into monomer-length molecules with no recognizable dimer linkage structure. These were measured directly. In molecules with a clear dimer linkage structure, the length from this feature to each end was measured. Open bars represent all molecules; solid bars represent molecules hybridized to SV40-poly(dT). The indicated average monomer length ([L]) and standard deviation ($\sigma$) were calculated from molecules which had a dimer linkage structure and which were attached to SV40-poly(dT) and thus must have been unbroken monomer subunits. Of molecules within 1 standard deviation of this average full length, 76% are attached to SV40-poly(dT).

Only a few dimer molecules with overlapping 5′ ends were observed in Rous sarcoma virus RNA, and there were no reproducible large loops. It may be that the 5′ to 5′ dimer linkage structure and the loops occur in native RSV 60–70S RNA, but that they are much more readily dissociated than for the other RNAs studied here. Alternatively, the few structures we have seen may be due to accidental overlaps.

There have been reports that not all viral subunits have poly(A) ends; only 65% of RSV subunits (Wang and Duesberg, 1974) and 40% of RD-114 subunits (Ihle et al., 1974) bound to oligo(dT)-cellulose or poly(U)-sepharose. In these several studies, however, the full-size subunits were purified by sucrose gradient sedimentation, and this procedure depends critically upon the resolving power of the gradient and on the quality of the RNA (that is, what fraction of the molecules are almost, but not quite, full-length). King and Wells (1976) have recently shown that when RSV is collected with very short (10 min) harvests and the RNA subunits are resolved on polyacrylamide gels, >90% of the subunits contain poly(A). We have shown here by direct counting of hybrids that at least three fourths of the intact RD-114 subunits have poly(A). Our ability to resolve full-length molecules from almost full-length is limited by the 10% variation in length of glyoxal-treated RD-114 RNA. Furthermore, as shown by King and Wells (1976) and many other investigators, the quality of the RNA can probably be improved by harvesting the virus at shorter intervals. Our results, therefore, tend to support the view that all intact subunits of RD-114 have poly(A) ends.

#### Experimental Procedures

SV40 DNA was a gift from Dr. Harumi Kasamatsu. It was extracted from SV40-infected TC-7 African Green Monkey cells. Sindbis virus, grown on chick embryo fibroblasts, was a gift from Dr. Jim Strauss, and the RNA was isolated as described previously (Hsu et al., 1973). RD-114, BKD, and WoMV viruses were prepared as described in the accompanying paper (Kung et al., 1976). RSV

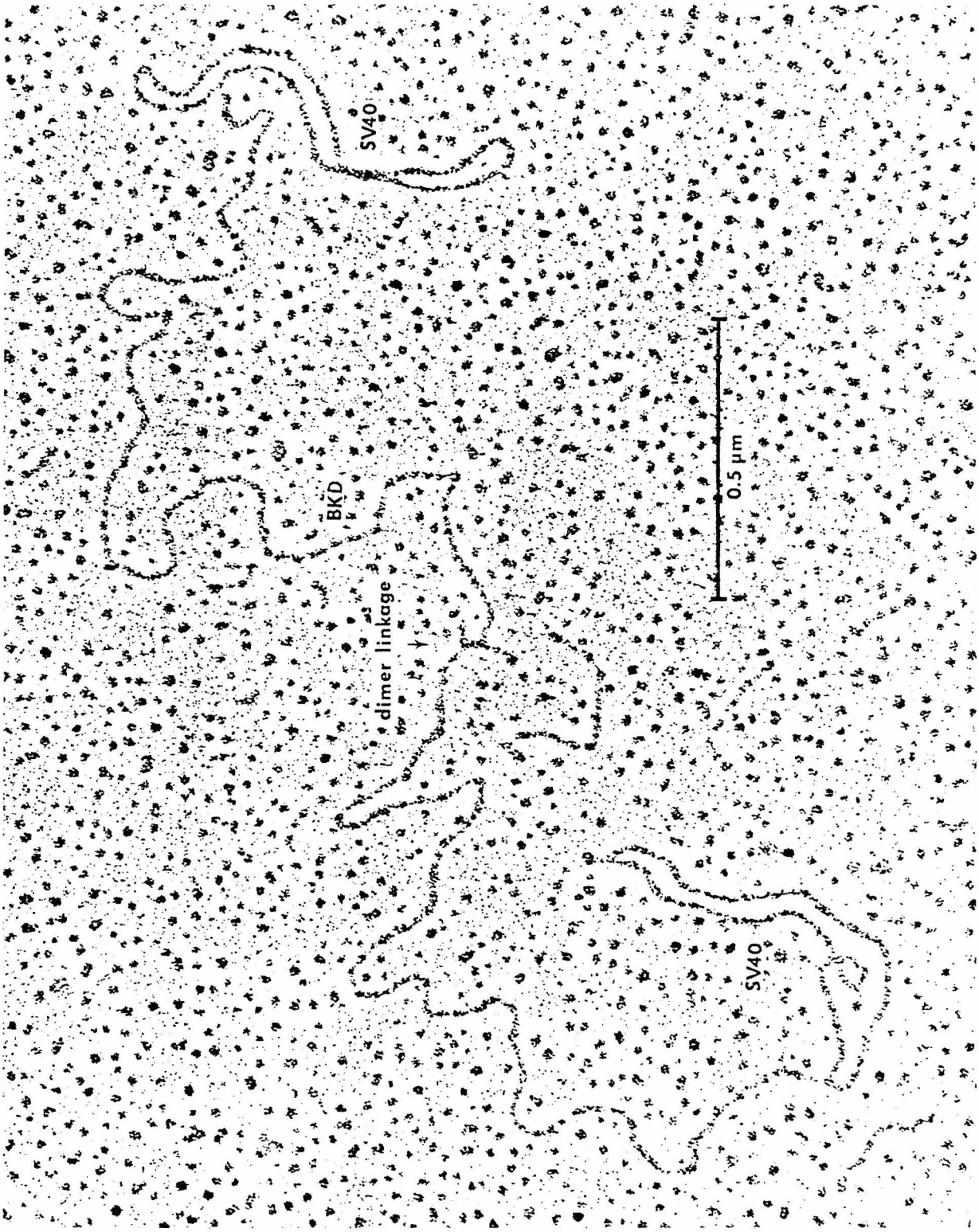



Figure 5. A BKD RNA Dimer Hybridized to SV40 Circles with Short Poly(dT) Tails

An arrow indicates the dimer linkage structure. This molecule is one of the unusual cases in which each poly(A) end is hybridized to a different SV40-poly(dT) circle.

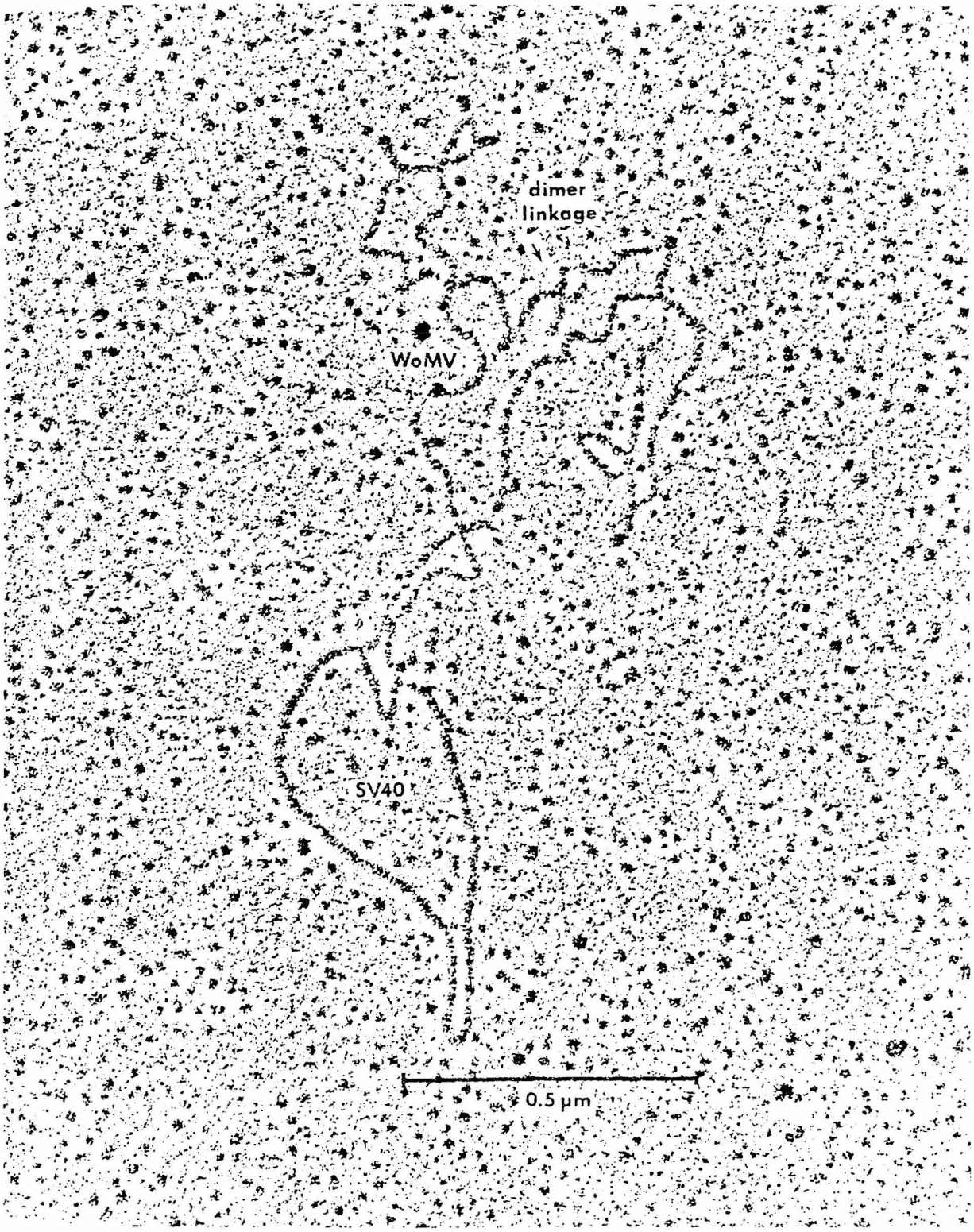



Figure 6. An WoMV RNA Dimer Hybridized to an SV40 Circle with Short Poly(dT) Tails

The RNA shows the symmetrically placed large loops. The dimer linkage structure is indicated.

(Rous sarcoma virus, Prague C. strain), grown in chick embryo fibroblasts, was obtained from Dr. Peter Vogt. All type C viruses were collected at 3 h intervals, and the cell culture medium containing the virus was frozen and stored at -70°C. The viruses were banded in sucrose, the RNA was phenol-extracted and ethanol-precipitated, and the 60-70S RNA was isolated by sedimentation as described previously (Kung et al., 1975). Terminal deoxynucleotidyl transferase, purified from calf thymus, was purchased from P & L Biochemicals, Inc. Oligo(dT)-cellulose, type T-2, was purchased from Collaborative Research, Inc. To minimize ribonuclease contamination, glassware was baked overnight at 150°C, and water was double-distilled.

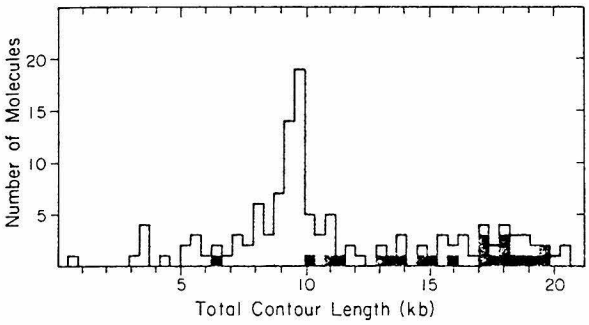



Figure 7. Histogram of Total Contour Lengths of Randomly Selected RSV RNA Molecules Which Are Attached to SV40-Poly(dT)

Open bars represent all molecules; solid bars indicate molecules attached to SV40 by two separate ends. There were no molecules found attached to SV40 by three or more ends.

**Preparation of SV40-Poly(dT)**

SV40 DNA (a mixture of supercoiled and nicked circular DNAs) was extended with terminal transferase in a reaction mix of 10-50 $\mu$l containing 0.1 M $KH_2PO_4$ and 0.05 M cacodylic acid adjusted to pH 7.0 with KOH, 1 mM $\beta$-mercaptoethanol, 1 mM $CoCl_2$, 3 mM thymidine triphosphate (TTP), 500 units per ml terminal transferase, and up to 0.5 mg/ml SV40 DNA. The solution was incubated at 37°C for 1-24 hr, depending upon the tail length desired. The reaction mixture was then diluted about 5 fold into 10 mM Tris, 1 mM EDTA (pH 8.5), and layered onto a gradient of 5-20% sucrose in the same buffer. Sedimentation (SW50.1 rotor, 44K rpm, 6 hr, 4°C) separated the SV40-poly(dT) from unincorporated TTP and unattached poly(dT). The SV40-poly(dT) was collected and dialyzed against 10 mM Tris, 1 mM EDTA (pH 8.5), and stored at -70°C. Sometimes the reaction was run using $^3$H-labeled TTP (10 Ci/mole), in which case incorporation could be measured by acid precipitation, or using $^3$H-labeled SV40 DNA (3.6 × $10^5$ cpm/$\mu$g), so that the fraction of circles extended could be measured by fractionation on poly(A)-sepharose. However, SV40-poly(dT) used for electron microscope labeling was not fractionated on poly(A)-sepharose.

**Poly(A)-Sepharose**

Poly(A)-sepharose was prepared by the CNBr coupling method of Berridge and Aronson (1973) using 200 mg CNBr/ml packed se-

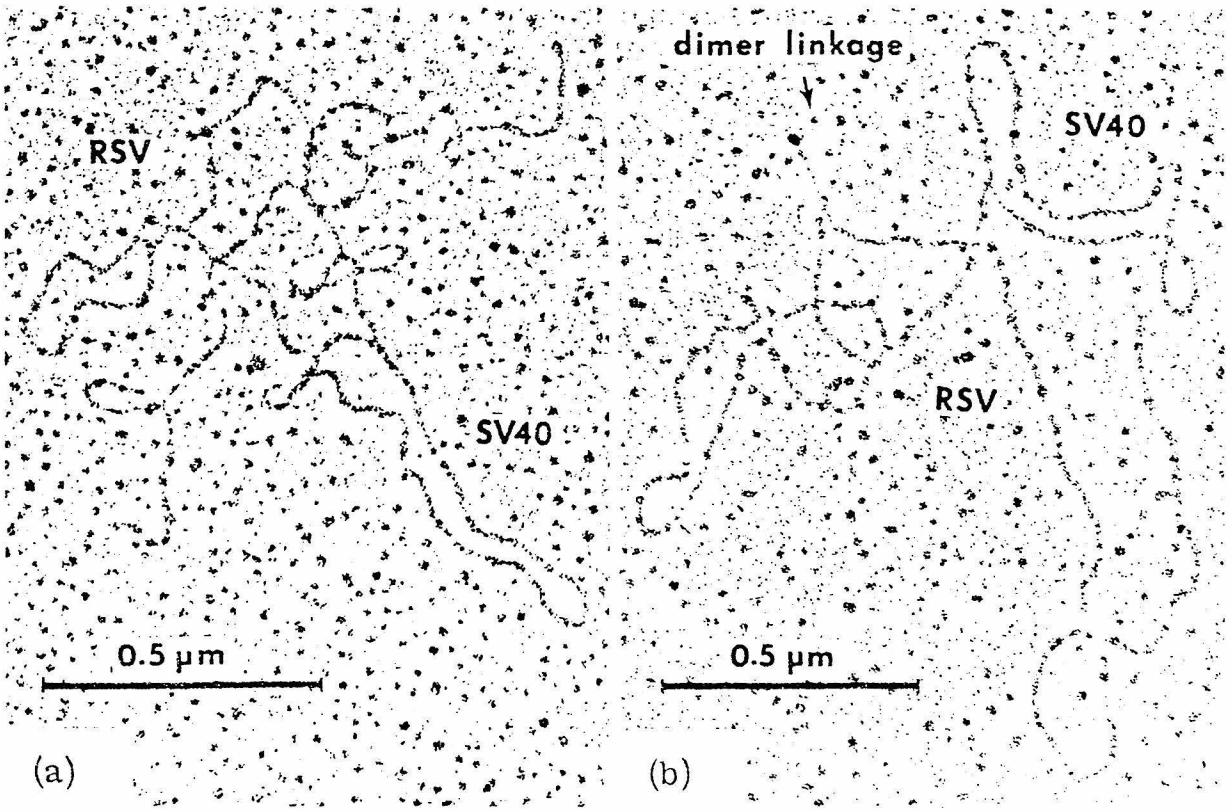



Figure 8. RSV RNA Molecules Hybridized to SV40 Circles with Short Poly(dT) Tails

The arrow in (b) indicates a presumed dimer linkage structure.

pharose. The sample to be absorbed to the poly(A)-sepharose was mixed with 1 ml of 0.1 M NaCl, 10 mM Tris, 1 mM EDTA (pH 7.5), and put over a column of 1 ml bed volume at a rate of about 0.2 ml/min at 25°C. The column was eluted with the same buffer at increasing temperature.

### Oligo(dT)-Cellulose

The oligo(dT)-cellulose column (1 ml bed volume) was loaded and washed as the poly(A)-sepharose, except that everything was done at 4°C. The column was eluted at 37°C with 5 ml of 10 mM Tris, 1 mM EDTA (pH 7.5) at a flow rate of 1 ml/min.

### Glyoxal Treatment and Hybridization

RNA was incubated with 1 M glyoxal, 10 mM sodium phosphate (pH 6.7) at 37°C for up to 1 hr. The mixture was then dialyzed against 10 mM Tris, 1 mM EDTA (pH 8.5) for 24 hr at 4°C to insure removal of glyoxal reversibly bound to A and C residues (Broude and Budowsky, 1971). About 0.01 $\mu$g of RNA was then mixed with about 0.02 $\mu$g SV40-poly(dT) in 5–20 $\mu$l of 0.4 M Tris, 0.04 M EDTA (pH 8.5). The mixture was allowed to stand at room temperature for 5–15 min before spreading. The expected $Cot_{\frac{1}{2}}$ for $poly(A)_{200} \cdot poly(dT)_{200}$ reassociation at a cation concentration of 0.25 M at 38°C is about $2 \times 10^{-6}$ M sec (Lee and Wetmur, 1972). Our incubations [about 100 ng/ml in poly(A) and poly(dT), about 0.24 M in cations, 10', 25°C] reach a Cot of about $1.8 \times 10^{-4}$ M sec, that is, a 90 fold excess over $Cot_{\frac{1}{2}}$ [neglecting corrections for temperature and for the unknown effects of the lengths of strands to which the poly(A) and poly(dT) are attached].

### Spreading

Hybrids were spread by the formamide, cytochrome C method (Davis, Simon, and Davidson, 1971) using 30 or 40% formamide in the hyperphase with 2 or 12% formamide in the hypophase, respectively. Grids were rotary-shadowed with platinum palladium and examined using a Phillips 300 electron microscope. Molecules of interest were photographed and the contour lengths were measured using a Hewlett-Packard digitizer. A diffraction grating was used to calibrate the length of SV40 circles (1.76 ± 0.06 $\mu$m), and SV40 was available as an internal length standard on all spreadings. Kilobase lengths of glyoxal-treated single-strand RNA were computed assuming 3.91 kb/$\mu$m as measured for glyoxal-treated E. coli 23S ribosomal RNA (Hsu et al., 1973). The lengths of RD-114 and RSV monomers given here are slightly larger than those previously reported (Kung et al., 1974, 1975). This correction resulted from a more accurate calibration of the microscope magnification.

### Acknowledgments

We are grateful to John Carbon for his advice and suggestions about the general method of using poly(dT) tails on duplex DNA as a label. We have profited from the counsel of our co-workers, Sylvia Hu, Hsing-Jien Kung, James Bailey, Margery Nicolson, Robert McAllister, and Peter Vogt.

This research has been supported by a research grant from the National Institute of General Medical Sciences and by a contract with the Virus-Cancer Program of the National Cancer Institute. W. B. has been supported by a NSF fellowship and a training grant from the NIH.

Received December 8, 1975; revised January 1, 1976

### References

Berridge, M. V., and Aronson, A. I. (1973). An Assay for the Endonucleolytic Cleavage of RNA to Large Oligonucleotides. Anal. Biochem. *53*, 603–612.

Broude, N. E., and Budowsky, E. I. (1971). The Reaction of Glyoxal with Nucleic Acid Components III. Kinetics of the Reaction with Monomers. Biochem. Biophys. Acta *254*, 380–388.

Carbon, J., Shenk, T. E., and Berg, P. (1975). Construction *In Vitro* of Mutants of Simian Virus 40: Insertion of a Poly(dA·dT) Segment at the *Hemophilus Parainfluenza* II Restriction Endonuclease Cleavage Site. J. Mol. Biol. *98*, 1–15.

Davis, R. W., Simon, M., and Davidson, N. (1971). Electron Microscope Heteroduplex Methods for Mapping Regions of Base Sequence Homology in Nucleic Acids. Methods Enzymol. *21*, 413–428.

Duesberg, P., Vogt, P. K., Beemon, K. B., and Lai, M. (1974). Avian RNA Tumor Viruses: Mechanism of Recombination and Complexity of the Genome. Cold Spring Harbor Symp. Quant. Biol. *39*, 847–857.

Green, M., and Cartas, M. (1972). The Genome of RNA Tumor Viruses Contains Polyadenylic Acid Sequences. Proc. Nat. Acad. Sci. USA *69*, 791–794.

Hsu, M. T., Kung, H. J., and Davidson, N. (1973). An Electron Microscope Study of Sindbis Virus RNA. Cold Spring Harbor Symp. Quant. Biol. *38*, 943–950.

Ihle, J. N., Lee, K., and Kennedy, F. T. (1974). Fractionation of 34S Ribonucleic Acid Subunits from Oncornavirus on Polyuridylate-Sepharose Columns. J. Biol. Chem. *249*, 38–42.

Jacobson, H., and Stockmayer, W. H. (1950). Intramolecular Reaction in Poly-condensations. I. The Theory of Linear Systems. J. Chem. Phys. *18*, 1600–1606.

King, A. M. Q. (1976). High Molecular Weight RNA's from Rous Sarcoma Virus and Moloney Murine Leukemia Virus Contain Two Subunits. J. Biol. Chem. *251*, 141–149.

King, A. M. Q., and Wells, R. D. (1976). All Intact Subunit RNA's from Rous Sarcoma Virus Contain Poly(A). J. Biol. Chem. *251*, 150–152.

Kung, H. J., Bailey, J. M., Davidson, N., Vogt, P. K., Nicolson, M. O., and McAllister, R. M. (1974). Electron Microscope Studies of Tumor Virus RNA. Cold Spring Harbor Symp. Quant. Biol. *39*, 827–834.

Kung, H. J., Bailey, J. M., Davidson, N., Nicolson, M. O., and McAllister, R. M. (1975). Structure, Subunit Composition, and Molecular Weight of RD-114 RNA. J. Virol. *16*, 397–411.

Kung, H. J., Hu, S., Bender, W., Bailey, J. M., Davidson, N., Nicolson, M. O., and McAllister, R. M. (1976). RD-114, Baboon, and Woolly Monkey Viral RNAs Compared in Size and Structure. Cell *7*, 609–620.

Lai, M. M. C., and Duesberg, P. H. (1972). Adenylic Acid-rich Sequence in RNAs of Rous Sarcoma Virus and Rauscher Mouse Leukemia Virus. Nature *235*, 383–386.

Lee, C. H., and Wetnur, J. G. (1972). On the Kinetics of Helix Formation Between Complementary Ribohomopolymers and Deoxyribohomopolymers. Biopolymers *11*, 1485–1497.

Lobban, P. E., and Kaiser, A. D. (1973). Enzymatic End-to-end Joining of DNA Molecules. J. Mol. Biol. *78*, 453–471.

Lizardi, P. M., Williamson, R., and Brown, D. D. (1975). The Size of Fibroin Messenger RNA and Its Polyadenylic Acid Content. Cell *4*, 199–205.

Mangel, W. F., Delius, H., and Duesberg, P. H. (1974). Structure and Molecular Weight of the 60–70S RNA and the 30–40S RNA of the Rous Sarcoma Virus. Proc. Nat. Acad. Sci. USA *71*, 4541–4545.

Quade, K., Smith, R. E., and Nicols, J. L. (1974). Poly(Riboadenylic Acid) and Adjacent Nucleotides in Rous Sarcoma Virus RNA. Virology *62*, 60–70.

Rho, H. M., and Green, M. (1974). The Homopolyadenylate and Adjacent Nucleotides at the 3′ Terminus of 30–40S RNA Subunits in the Genome of Murine Sarcoma-Leukemia Virus. Proc. Nat. Acad. Sci. USA *71*, 2386–2390.

Riley, M., Maling, B., and Chamberlin, M. J. (1966). Physical and Chemical Characterization of Two- and Three-Stranded Adenine-

Thymine and Adenine-Uracil Homopolymer Complexes. J. Mol. Biol. *20,* 359–389.

Schmid, C. W., Manning, J. E., and Davidson, N. (1975). Inverted Repeat Sequences in the Drosophila Genome. Cell *5,* 159–172.

Wang, J. C., and Davidson, N. (1966). On the Probability of Ring Closure of Lambda DNA. J. Mol. Biol. *19,* 469–482.

Wang, L., and Duesberg, P. (1974). Properties and Location of Poly(A) in Rous Sarcoma Virus RNA. J. Virol. *14,* 1515–1529.

Weissmann, C., Parson, J. T., Coffin, J. W., Rymo, L., Billeter, M. A., and Hofstetter, H. (1974). Studies on the Structure and Synthesis of Rous Sarcoma Virus RNA. Cold Spring Harb. Symp. Quant. Biol. *39,* 1043–1056.

Wolfe, L. G., Smith, R. K., and Deinhardt, F. (1972). Simian Sarcoma Virus, Type I (Lagothrix): Focus Assay and Demonstration of Nontransforming Associated Virus. J. Nat. Cancer Inst. *48,* 1905–1907.

Cell, Vol. 7, 609–620, April 1976, Copyright © 1976 by MIT


# RD-114, Baboon, and Woolly Monkey Viral RNAs Compared in Size and Structure

**Hsing-Jien Kung, Sylvia Hu, Welcome Bender, James M. Bailey, and Norman Davidson**
Department of Chemistry
California Institute of Technology
Pasadena, California 91125
**Margery O. Nicolson and Robert M. McAllister**
Department of Pediatrics
University of Southern California
School of Medicine
Childrens' Hospital of Los Angeles
Los Angeles, California 90054

## Summary

**The molecular weights, subunit compositions, and secondary structure patterns of the RNAs from an endogenous baboon virus and from a woolly monkey sarcoma virus were examined and compared to the properties of the RNA of RD-114, an endogenous feline virus. The high molecular weight RNA extracted from each of these three viruses has a sedimentation coefficient of 52S, and a molecular length, measured by electron microscopy, of 16–20 kb (kb = kilobase, 1000 nucleotides).**

**Each such RNA is a dimer, containing two monomer subunits of 8–10 kb in length (molecular weight 3 × $10^6$ daltons). The two monomer subunits are joined at their non-poly(A) ends in a structure called the dimer linkage structure. The appearance of this structure is somewhat different for the different viruses. The dimer linkage dissociates at a temperature estimated to be 87°C in aqueous 0.1M $Na^+$ for RD-114 and baboon viral RNAs, but at the lower temperature of 66°C for woolly monkey RNA. All three viral RNAs have two large loops of similar size and position symmetrically placed on either side of the dimer linkage structure.**

**Since the baboon virus is partially related to RD-114, and the woolly monkey virus is unrelated to either of the other two, the dimer linkage and symmetrical loops are surprisingly similar and may well be common features of type C virus RNAs.**

## Introduction

The RNA of the endogenous feline type C virus, RD-114, has been extensively studied in our laboratory and has been shown to have a novel structure (Kung et al., 1975). The RNA from the virion is a 20 kb dimer molecule with an end of one 10 kb monomer noncovalently joined to an end of the other in a secondary structure feature at the middle of the dimer molecule. The name "rabbit ears" was used to describe this central junction feature in RD-114, but as will be shown here, the detailed shape of this feature differs with different spreading conditions or with RNAs of different viruses under identical spreading conditions. We have therefore now adopted the more functional name "dimer linkage structure." RD-114 dimer RNA also has a large loop in each monomer half, and these loops are symmetrically placed with respect to the dimer linkage structure. Poly(A) mapping has located poly(A) sequences on both free ends of the dimer (Bender and Davidson, 1976). If the poly(A)s are on the 3′ ends of the monomer RNAs, then the two 5′ ends must be joined in the dimer linkage structure.

Initial electron microscopic studies had shown no RNA secondary structure features similar to those of RD-114 in either Rous sarcoma virus or Gardner-Arstein Feline leukemia virus (Kung et al., 1974). We chose therefore to examine the endogenous baboon virus, BKD (Todaro et al., 1974a), which is partially related to RD-114, to see if the RD-114 features would be conserved. We found that the structure and physical properties of BKD RNA were quite similar to those of RD-114 RNA. We therefore studied the RNA of another primate-derived virus, denoted WoMV, which is the simian sarcoma virus isolated from a woolly monkey (Theilen et al., 1971). This virus is not endogenous in primates and is not related to either RD-114 or BKD viruses by serological or nucleic acid hybridization tests. Nevertheless, it too shows a dimer structure strikingly similar to that of RD-114 and BKD.

## Results

### Baboon Endogenous Virus

#### *Sedimentation Analysis*

Total RNA extracted from the baboon endogenous virus, BKD, was sedimented in a sucrose gradient containing 0.1 M NaCl, and RD-114 RNA was sedimented in a parallel tube. RD-114 RNA was shown to have an S value of 52 under these conditions (Kung et al., 1975). Both RNAs sediment with the same velocity (Figure 1a), and so BKD RNA has a sedimentation coefficient of 52S in 0.1 M NaCl. BKD RNA also sediments identically to RD-114 RNA in a low salt (0.6 mM cations) sucrose gradient (data not shown) or on a glyoxal-sucrose gradient (Figure 1c). Because of the low electrolyte concentration and the effects of the glyoxal, the latter two sedimentation media are moderately denaturing and cause dissociation of some (but not all) secondary structure. These observations therefore suggest that the RNAs of RD-114 and BKD have the same molecular weight.

#### *Gel Electrophoresis*

It has been shown that the high molecular weight RNA of RD-114 is actually a dimer that can be dissociated into monomers only after treatment with

strongly denaturing conditions (65% of our standard urea plus formamide solvent, 0.06 M cations, 60°C, Kung et al., 1975). As will be shown in the results of electron microscope observations, the BKD high molecular weight RNA is also a dimer. We wished to determine the melting temperature of the dimer linkage as compared to that of RD-114. Samples of BKD RNA in 67% urea plus formamide, 0.024 M cations, were heated to various temperatures, cooled, and applied to a nondenaturing agarose gel for electrophoresis. As shown in Figure 2a, samples heated to 40°C or less gave a single high molecular weight peak; samples heated to 50°C were partially dissociated into a lower molecular weight peak, and samples heated to 60°C were completely dissociated to the lower molecular weight. It will be shown below that the high molecular weight peak is an RNA dimer and the lower molecular weight peak is monomer RNA. Making slight corrections for the different solvent conditions used in the melting experiments for RD-114 and BKD RNAs, we conclude that the BKD dimer linkage has about the same stability as the RD-114 dimer linkage.

We have found that methylmercury hydroxide can be used to denature completely RNA molecules, so that agarose gel electrophoresis of RNA in the presence of methylmercury hydroxide can be used to determine the molecular weight (Bailey and Davidson, 1976). As shown in Figure 3a, BKD RNA monomers on a methylmercury hydroxide agarose gel co-migrated with RD-114 RNA at a mobility corresponding to $3 \times 10^6$ daltons (or 9 kb).

***Electron Microscopy***

*Glyoxal Treatment* is an easy and effective way to disrupt most RNA secondary structure (Hsu, Kung, and Davidson, 1973), and we consider length measurements in the electron microscope of glyoxal-treated RNA to be our most reproducible method for determining molecular weights. BKD RNA was taken from the high molecular weight peak of a nondenaturing sucrose gradient, incubated with 1M glyoxal, 0.01 M potassium phosphate (pH 7.0) for 1 hr at 37°C, and spread for microscopy from 55%

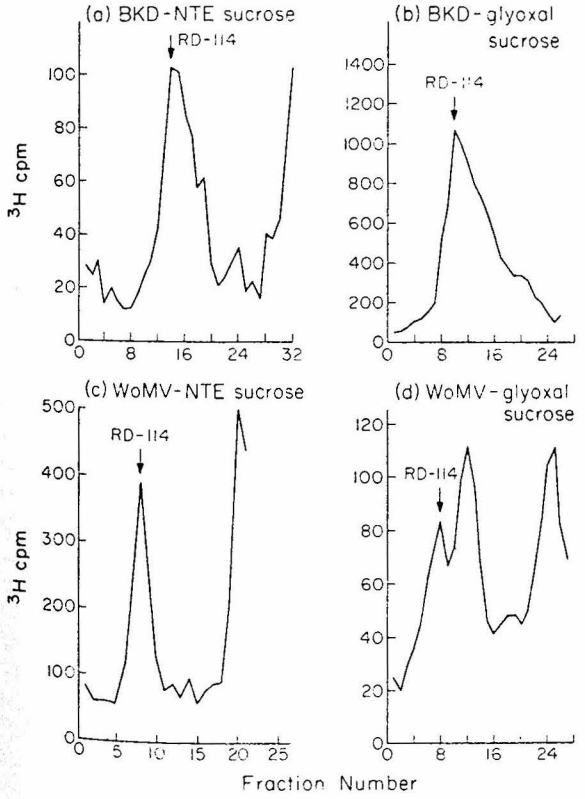



Figure 1. Sedimentation Profiles of BKD and WoMV RNA as Extracted from the Virion

Centrifugation was carried out in an SW50.1 rotor at 4°C. Arrows indicate RD-114 RNA marker. (a) BKD RNA in NTE-sucrose gradient, 44,000 rpm, 1.25 hr; (b) BKD RNA in glyoxal-sucrose gradient, 45,000 rpm, 5 hr; (c) WoMV RNA in NTE-sucrose gradient, 45,000 rpm, 1.75 hr; (d) WoMV RNA in glyoxal-sucrose gradient, 45,000 rpm, 5.5 hr. All sedimentations are from right to left.

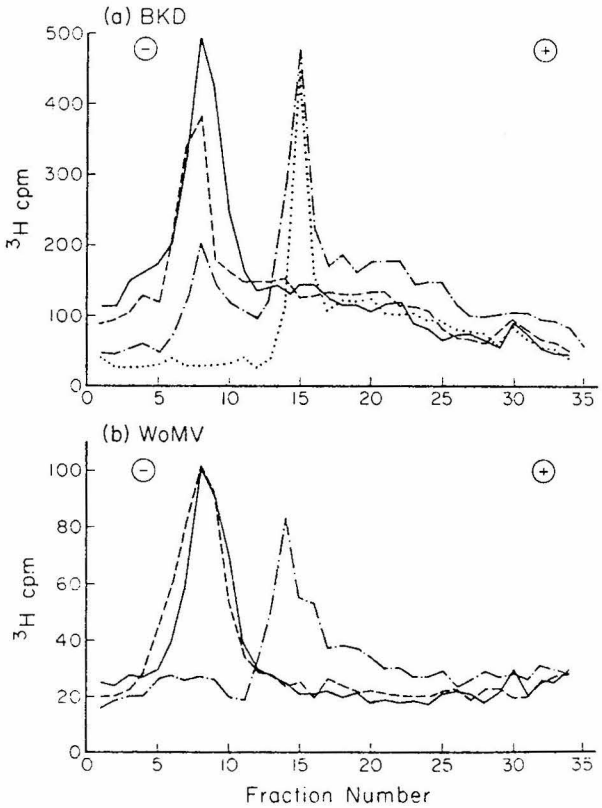



Figure 2. Agarose Gel Electrophoresis of Viral RNAs Showing Dissociation of the High Molecular Weight Complex

(a) BKD RNA in 67% (U + F), 0.024 M cations. (——) unheated; (--) heated to 40°C; (-·-) heated to 50°C; and (····) heated to 60°C.

(b) WoMV RNA at room temperature with 0.015 M cations and (——) 0% (U + F), (--) 65% (U + F), and (-·-) 75% (U + F).

formamide. The histogram in Figure 4a shows that the length distribution is rather broad, although there is a maximum at about 20 kb. We presume that there are many nicked molecules held together intramolecularly in the high molecular weight RNA complex under nondenaturing conditions and that these fragments of all sizes appear after glyoxal treatment. To estimate the full-size molecular weight of any partially degraded RNA or DNA, one must arbitrarily decide on a boundary above which molecules are considered full-size, and below which they are considered fragments. This is also a problem for size determinations by electrophoresis or sedimentation, although the difficulty is partially disguised by the usual weight average method of plotting results.

Fortunately, these RNA tumor viruses contain RNA dimers, and, as shown in the accompanying paper (Bender and Davidson, 1976), they have poly(A) sequences at both outside ends. Any continuous dimer molecule with both poly(A) ends intact must be full-length. Similarly, since monomers are joined in the dimer linkage very close to their non-poly(A) ends, the distance from a poly(A) end

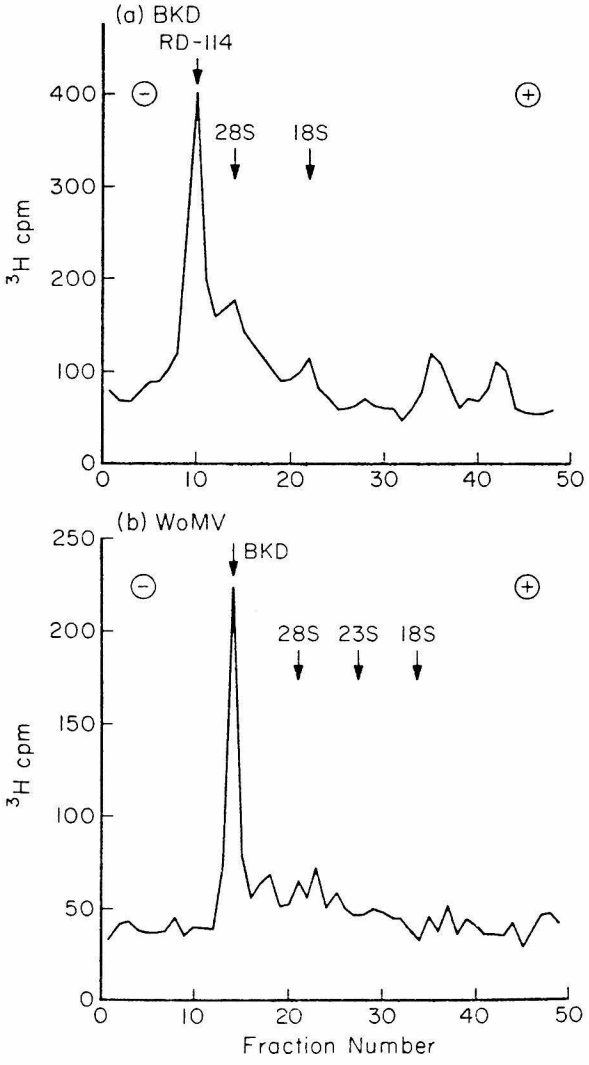



Figure 3. Electrophoresis of Viral RNA on Denaturing Methylmercury Hydroxide Gels

(a) BKD viral RNA, run at 5 mA per gel for 3 hr. The positions of HeLa 18S and 28S rRNA markers and of RD-114 viral RNA (run in a separate tube) are shown.

(b) WoMV viral RNA, run at 5 mA per gel for 5 hr. The positions of HeLa 18S and 28S and E. coli 23S rRNA markers and of BKD RNA (in a separate tube) are shown.

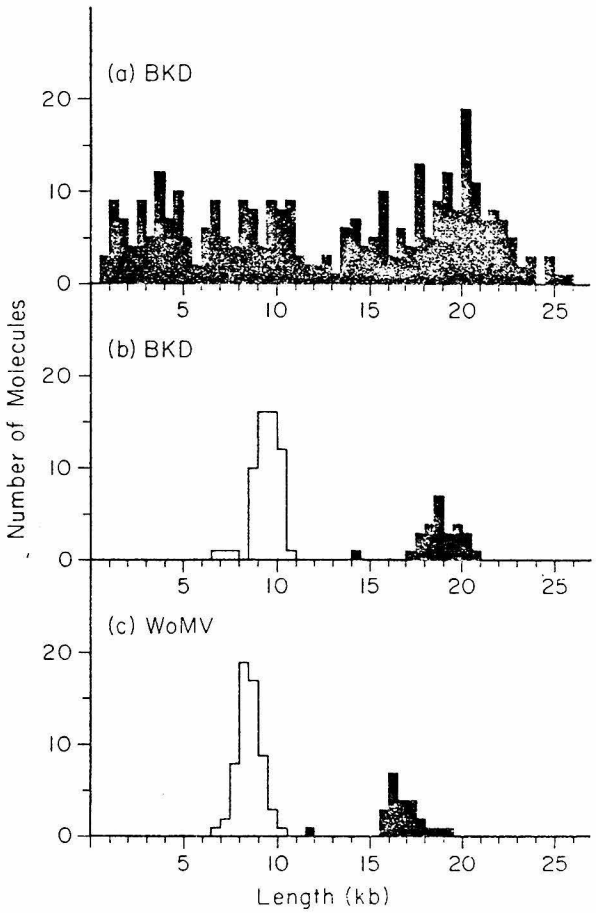



Figure 4. Histogram of Total Lengths of Glyoxal-Treated Viral RNAs

(a) BKD 52S RNA treated with 1M glyoxal for 1 hr and spread from 55% formamide. All molecules were measured.

(b) BKD 52S RNA treated with 1 M glyoxal for 1 hr at 37°C, dialyzed against 10 mM Tris, 1 mM EDTA for 24 hr, hybridized with SV40 circles with short poly(dT) tails to mark the poly(A) ends, and spread from 40% formamide. Dimers (solid bars) were scored as unbroken molecules with two poly(A) ends marked by the SV40-poly(dT). The lengths are measured from one SV40 attachment point to the other (excluding the dimer linkage structure). Monomers (open bars) were scored as molecules continuous from poly(A) end to dimer linkage; contour lengths were measured from the SV40 attachment site to the base of the dimer linkage.

(c) WoMV 52S RNA treated with glyoxal, dialyzed 24 hr, hybridized with SV40-poly(dT), and spread, as above. Dimers (solid bars) and monomers (open bars) were scored and measured as above

to the base of the dimer linkage structure can be taken as the full monomer length. Thus we can avoid scoring any broken molecules without biasing our selection of full-length molecules.

Glyoxal-treated BKD RNA was hybridized to SV40 circles with short poly(dT) tails (average length of 175 bases) to mark the poly(A) ends, and the mix was spread for microscopy (Bender and Davidson, 1976) from 40% formamide. Full-length dimers and monomers, as defined above, were measured and are plotted in Figure 4b. The dimer length is 18.8 ± 1.3 kb, and the monomer size is 9.4 ± 0.7 kb. RD-114 has dimers of 20.0 ± 1.7 kb and monomers of 9.8 ± 1 kb (data not shown).

We are not certain of the number of nucleotides within the dimer linkage structure; assuming that both the arms of the dimer linkage are single-stranded, we estimate that there are about 400 bases included in the BKD dimer linkage (about 600 for RD-114). Including the RNA in the dimer linkage, the BKD dimer length is thus 19.2 kb, and the monomer is 9.6 kb (20.6 and 10.1 kb for RD-114 dimers and monomers, respectively). These values for RD-114 are slightly higher than those reported previously (Kung et al., 1975); the difference is due to more accurate calibration of the microscope magnification.

There are a few reproducible secondary structure features which remain after glyoxal treatment. Representative BKD RNA molecules are shown in Figures 5a and 5b. Molecules of about 20 kb length always had a structural feature near the center which looked like Y, +, or X. This feature is similar in position to the dimer linkage structure observed in RD-114 RNA, although the RD-114 dimer linkage is usually shaped like Y (Kung et al., 1975). This central feature divided the RNA molecule into two equal subunits. There is frequently a loop present on one or both subunits; the loop occurred on 80% of the subunits in one particular spread. The size and location of the loops are very similar to what we observed previously on RD-114. On 70% of the subunits, within the region of the loops, there is a small hairpin-like structure with a length of about 150 base pairs. This structure was not present on RD-114 RNA. For about 30% of the monomer subunits of the dimers, there is an additional reproducible loop due to a crossover between a point 1 kb from the dimer linkage structure and a point 0.5 kb from the poly(A) end (Figure 5b). This attachment was not seen in glyoxal-treated RD-114 RNA. An overall diagram of the structure of BKD RNA is shown in Figure 7, and the contour lengths of the features of the glyoxal treated molecules are given in Table 1.

*Urea plus Formamide* (U + F) was used at different concentrations to give solvents of controlled denaturation strength to study the relative stabilities of the various secondary structure features of BKD RNA. The exact composition of these solvents is described in Experimental Procedures. When BKD RNA is spread at room temperature from 50% or 70% (U + F) (0.24 M cations), the molecules are still tangled and hardly traceable. Spread from 80% (U + F), the molecules are well extended, but the secondary structure features seen in glyoxal-treated RNA (the dimer linkage structure, the loops, and the hairpins within the loops) are still present on most molecules, and about 15% of the molecules have the outside ends looped around and attached near the dimer linkage. When BKD RNA in 67% (U + F) is heated to 40°C briefly, then cooled and spread at room temperature, the loops disappear from most molecules, although the dimer linkage and the hairpins remain intact. Heating to 55°C in 67% (U + F) causes most of the dimers to dissociate into monomers; this result is direct confirmation of the dimer melting temperature determined by the agarose gel electrophoresis method. The small hairpin formerly within the loop is still present after brief heating to 55°C in 67% (U + F), and it is near the center of the monomer RNA (Figures 5c and 5d); either it has a very high melting temperature, or it melts on heating but then renatures during the subsequent cooling.

After brief heating of BKD RNA to 55°C in 67% (U + F), a new secondary structure feature appears in most of the dissociated molecules. This feature, which is not seen in glyoxal-treated RNA, looks like a hairpin, usually with a loop at the end of the stem (see Figure 5c). It measures about 1.3 kb in contour length and is about 1 kb from the non-poly(A) end. A similar feature was seen in RD-114 RNA after heating and cooling (Kung et al., 1975). Rarely, we find an intact dimer after the above heating procedure with two of these new large hairpins symmetrically placed about the dimer linkage (Figure 5d). These molecules show that the new hairpin is formed in a region close to the dimer linkage end of the monomer and near the region that is normally the loop. We have also seen that in spreads from 80% (U + F) at room temperature, where almost all molecules are still dimers, about 20% of the monomer subunits have these new large hairpins instead of the usual large loop. Under these conditions, there are a few cases (about 1%) in which both the new hairpin and the large loop appear to be present on the same RNA subunit; in these molecules, the base of the hairpin is immediately adjacent to the base of the loop (Figure 5e).

**Woolly Monkey Virus**

The WoMV virus preparation (called SSV-1 by some investigators) contains mostly nontransforming

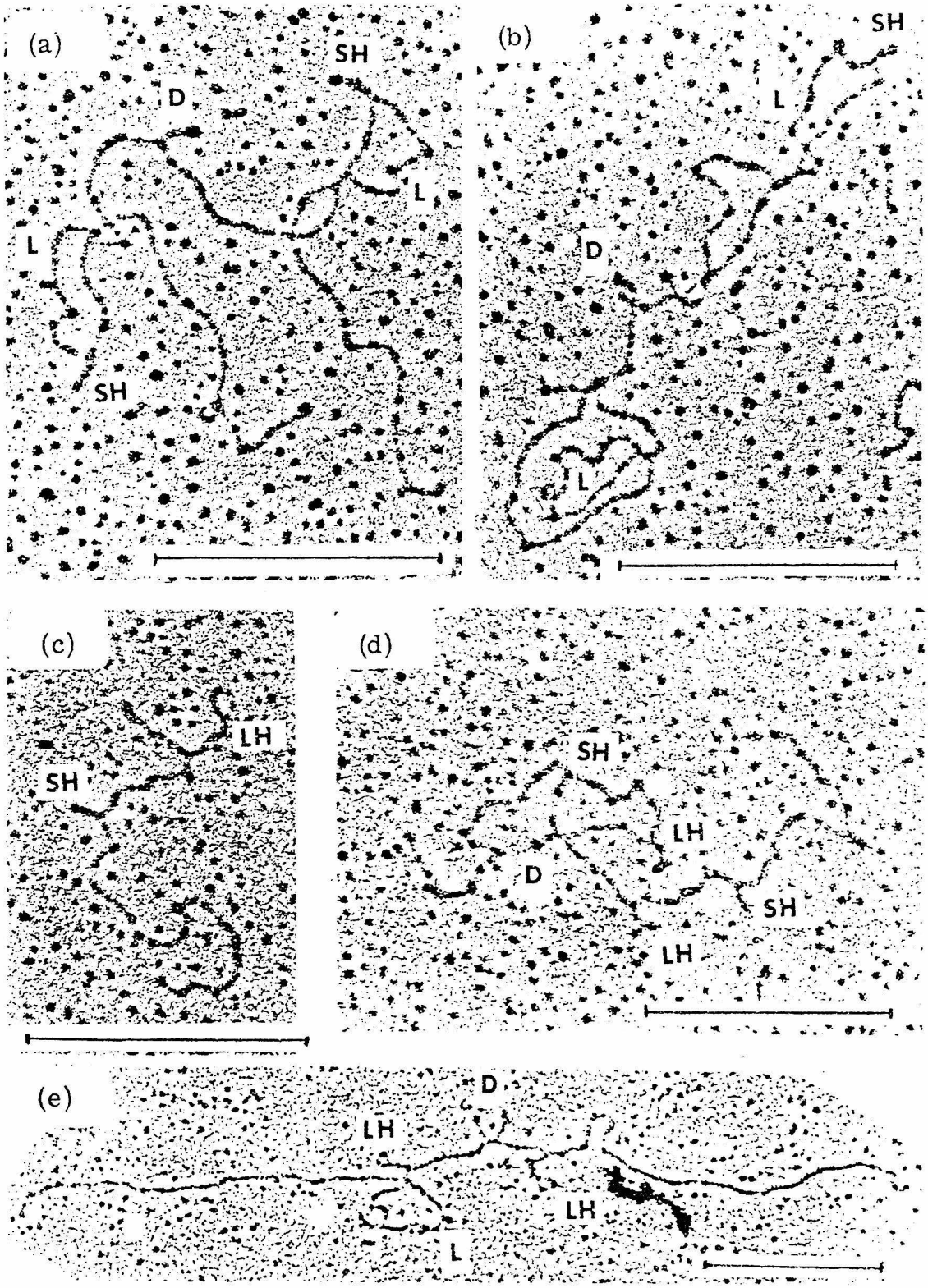



Figure 5. Electron Micrographs of BKD RNA.

(a) and (b) glyoxal-treated RNA spread from 55% formamide; (c) and (d) RNA in 67% (U + F) heated to 55°C, cooled, and spread at room temperature. (e) RNA in 80% (U + F) spread without prior heating. (D) indicates the dimer linkage structures; (L) indicates the loops; (SH) indicates the small hairpins within the loop region; and (LH) indicates the new large hairpin that appears under denaturing conditions. The bar lengths are 0.5 $\mu$m.

helper virus (Wolfe et al., 1972), so that these studies deal with the RNA of this nontransforming virus.

***Sedimentation Analysis***

The sedimentation of high molecular weight RNA extracted from WoMV was compared with RD-114 RNA on a sucrose gradient in 0.1 M NaCl. As shown in Figure 1b, the WoMV RNA co-sediments with the 52S RD-114 RNA. WoMV RNA also co-sediments with RD-114 RNA in gradients containing 0.01 M NaCl (data not shown). When WoMV RNA is first treated with glyoxal under our standard conditions and then run on a glyoxal sucrose gradient, two peaks appear (Figure 1d). In addition to the peak co-sedimenting with glyoxal-treated RD-114 RNA, there is a more slowly sedimenting peak. The electron microscope analysis, reported below, showed that the fast peak contained RNA dimers and the slow peak contained monomers. Since glyoxal-treated RD-114 RNA or BKD RNA gave only the dimer peak, the WoMV RNA dimer linkage is less stable to glyoxal treatment than that of RD-114 or BKD.

***Gel electrophoresis***

We used electrophoresis on non-denaturing agarose gels to follow the dissociation of the WoMV dimer. High molecular weight RNA was dissolved at room temperature in a solvent containing 0.015 M cations and varying concentrations of urea plus formamide, and was then applied to the gel. Figure 2b shows that RNA treated with 65% (U + F) at room temperature migrated with the untreated sample, and molecules in 75% (U + F) moved significantly faster. Heating the RNA to 40°C in 88% (U + F) did not produce any further dissociation. As will be shown, the slow peak contains dimer-length RNA molecules and the fast peak contains RNA monomers. BKD RNA dimers melted apart in 67% (U + F) at about 50°C; the dissociation of WoMV at 25°C in 75% (U + F) demonstrates that the dimer linkage is considerably less stable in WoMV RNA than in BKD or RD-114 RNAs (see Discussion).

To get a size determination of the WoMV RNA monomer, we ran the RNA on an agarose gel in the presence of methylmercury hydroxide. The WoMV RNA migrates as a single peak corresponding to a molecular weight of about $3 \times 10^6$ daltons or 9 kb (Figure 3b).

***Electron Microscope Studies***

*Glyoxal*-treated WoMV RNA sedimented as two separate peaks on a glyoxal sucrose gradient (Figure 1d), and these peaks were each isolated and spread from 30% formamide. The slower sedimenting peak gave molecules of about 9 kb contour length (Figure 6b), and the faster peak contained molecules which had a dimer linkage structure similar to those of RD-114 and BKD and which were about 17 kb in length (Figure 6a). We conclude that the slow and fast peaks of the glyoxal sucrose gradient contain monomers and dimers of WoMV RNA, respectively.

Again we attempted to get an accurate estimate of the RNA size by using the poly(A) mapping technique. Full-length dimers are defined as continuous molecules with poly(A) sequences at both outside ends; full-length monomers are molecules continuous from a poly(A) end to a clear dimer linkage structure. A histogram of the monomer and dimer lengths is plotted in Figure 4c; the dimer length is 16.7 ± 1.4 kb and the monomer length is 8.5 ± 0.7 kb. If the dimer linkage sequences (about 300 bases) are included, the dimer and monomer lengths are adjusted to 17 and 8.7 kb, respectively.

The data in Table 1 show that BKD RNA is nearly the same size as RD-114 RNA but that WoMV RNA is significantly smaller. We presume the length difference reflects an actual difference in molecular weights, but it is possible that WoMV RNA could have small-scale secondary structure which is resistant to glyoxal treatment and which shortens the total contour length of the RNA. The methylmercury hydroxide gels indicated the same molecular weights for the three different viral RNAs, but the accuracy of these measurements was insufficient to detect a 10% difference in molecular weights.

The secondary structure pattern seen in glyoxal-treated WoMV RNA is almost identical to that of BKD RNA (Figure 6a). The dimer linkage structure was shaped like ⊻, ⊹, or ⊥, and was always in the approximate center of the dimers. There were loops on many (50%) of the monomer halves and small hairpins present in most (70%) of the loops.

Table 1. Contour Lengths of RNA Structural Features

| | a | b | c | d | e | f | g | h | i |
|---|---|---|---|---|---|---|---|---|---|
| RD-114 | 20.0 ± 1.7 | 9.8 ± 1.0 | 3.9 ± 0.7 | 3.8 ± 0.4 | 2.4 ± 0.4 | | | | 0.3 ± 0.07 |
| BKD | 18.8 ± 1.3 | 9.4 ± 0.7 | 3.9 ± 0.4 | 3.7 ± 0.4 | 1.9 ± 0.1 | 1.2 ± 0.1 | 0.5 ± 0.1 | 1.9 ± 0.3 | 0.2 ± 0.04 |
| WoMV | 16.7 ± 1.4 | 8.5 ± 0.7 | 3.5 ± 0.5 | 3.9 ± 0.4 | 1.4 ± 0.3 | 1.0 ± 0.3 | 0.8 ± 0.4 | 2.1 ± 0.4 | 0.14 ± 0.04 |

The lengths and standard deviations of the features marked by the letters (a–i) in the diagram of Figure 7. The small hairpin feature (SH) is not present in RD-114, and its length (g) for BKD and WoMV is given in kilobases of single-strand RNA (twice the apparent double-strand length). As discussed in the text, our estimates of the full lengths, including the sequences in the dimer linkage structure, are a + 2i for dimers and b + i for monomers. We assume 1 kb equals 0.256 μm for glyoxal-treated RNA.

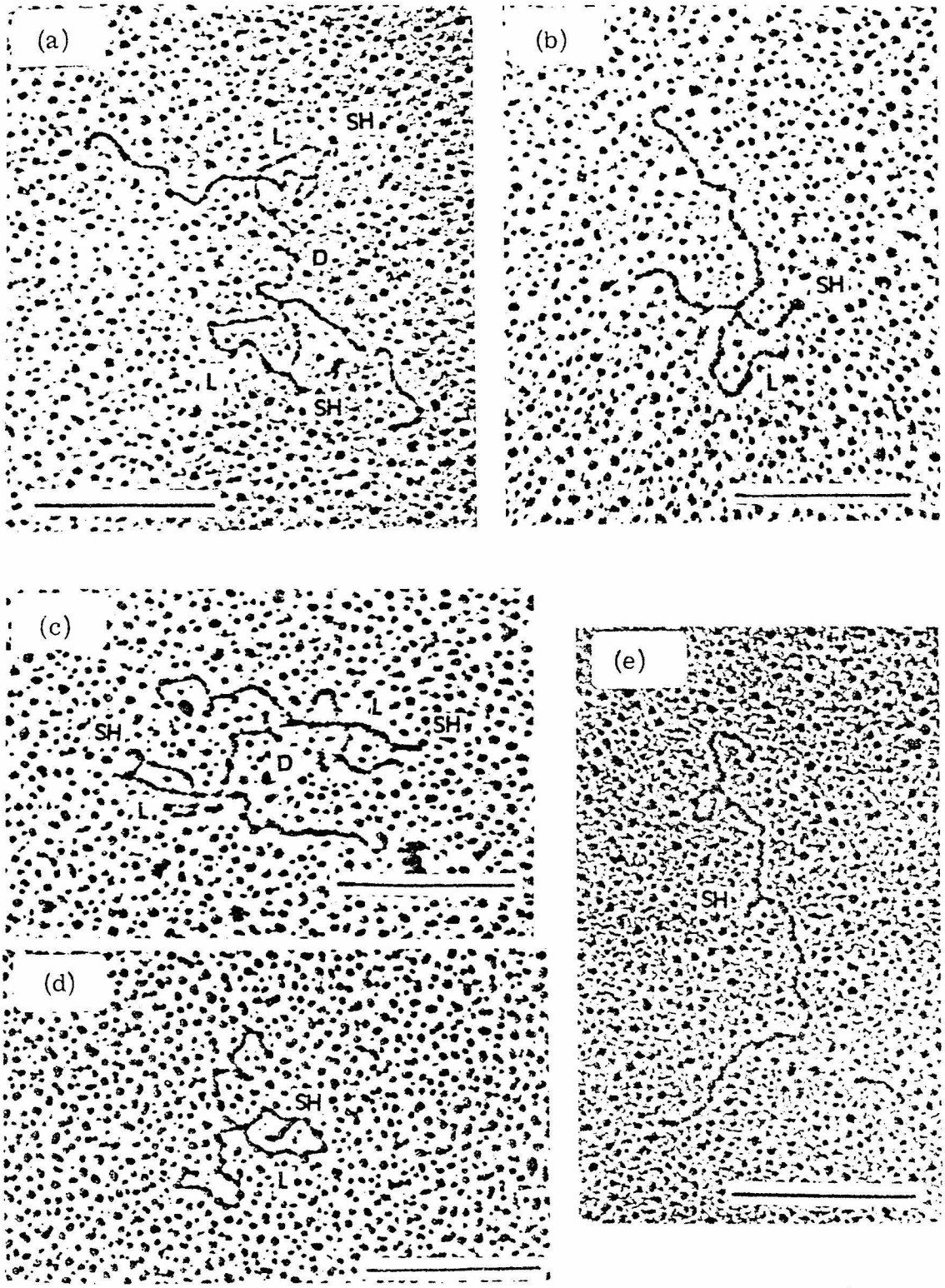



Figure 6. Electron Micrographs of WoMV RNA

(a) and (b) show RNA taken from a glyoxal-sucrose gradient (Figure 1d) and spread from 30% formamide. The molecule from (a) came from the fast sedimenting peak and in (b) from the lower sedimenting peak. (c–d) show molecules spread without prior heating from a solution with 0.03 M cations and 70% (U + F) [in (c) and (d)] or 88% (U + F) [in (e)]. (D) indicates the dimer linkage structures; (L) indicates the loops; and (SH) indicates the small hairpin. The bar lengths are 0.5 $\mu$m.

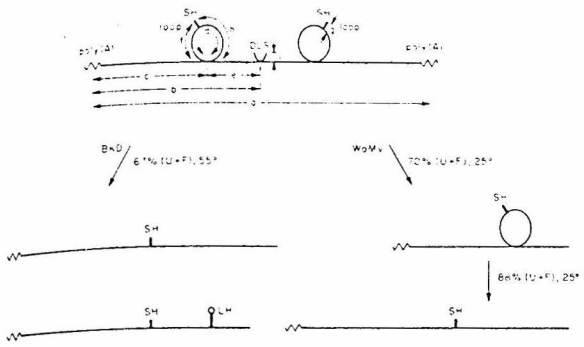


Figure 7. General Diagram of Secondary Structure Features in Viral RNA Dimers of RD-114, BKD, and WoMV

Contour lengths of various features (a–i) are given in Table 1. The dimer linkage structure (DLS) is not always shaped as drawn here, and the small hairpin (SH) with the loop is found in BKD and WoMV, but not in RD-114. Structures seen under denaturing conditions are shown separately for BKD and WoMV RNAs. (LH) indicates the new large hairpin seen in BKD under denaturing conditions.

The outside ends were never observed to fold back and attach near the dimer linkage as sometimes happened with BKD RNA. The diagram of Figure 7 also applies to the secondary structure features found in WoMV RNA, and the contour lengths of these features are given in Table 1. Note that the lengths from the poly(A) end to the loop and the loop circumference are approximately the same in WoMV as in BKD, but the distance from the loop to the dimer linkage structure in WoMV RNA is significantly shorter than in BKD RNA.

*Urea plus Formamide.* WoMV RNA was also spread with different concentrations of (U + F) in the hyperphase to study the stability of the various secondary structure features under varying denaturing conditions. In a 50% (U + F) spread (0.04 M cations), most WoMV RNA molecules were condensed and not easy to interpret. With 70% (U + F), the molecules were often still tangled but were usually traceable. About half of the molecules were of dimer length (about 17 kb) and showed the secondary structure features seen in glyoxal-treated RNA, that is, the dimer linkage structure, the loops, and the hairpins within the loops (Figure 6c). About 80% of the monomer halves had a loop and about 80% had the small hairpin within the loop region. Dissociated monomers also had these loops and hairpins (Figure 6d).

The partial dissociation of dimers observed in 70% (U + F) parallels the dissociation between 65% and 75% (U + F) seen by agarose gel electrophoresis. Spreading from 88% (U + F) caused most (85%) of the dimers to dissociate, and very few (8%) of the monomers still contained loops, but the small hairpin structure was observed near the center of 75% of the monomer RNA molecules (Figure 6d).

We also looked to see whether the WoMV RNA could form the same large hairpin structure as was observed in BKD and RD-114 molecules. The WoMV RNA in 80% (U + F) was heated to 60°C, cooled, and spread at room temperature. We could still identify the small hairpin located near the center in over 70% of the monomers, but no large hairpins were seen.

## Discussion

Our most striking result is that the high molecular weight RNAs from RD-114, BKD, and WoMV are all dimers and have very similar structures. There are differences in the observed morphologies of the dimer linkage structure and small differences in other secondary structure features. WoMV RNA also differs from that of RD-114 and BKD in that the dimer dissociates more readily.

We have used a urea–formamide solvent for studies of the dissociation of the dimers because it is also a good solvent for electron microscopy. We can estimate the dissociation temperatures in aqueous 0.1 M NaCl by using the equation $\Delta T_m = 0.4 \times \text{percent}(U + F) + 16.6°C \log[(\text{cation}) \times 10]$. The predicted effect of cation concentration in this equation is taken from Schildkraut and Lifson (1965) and is probably fairly accurate. The estimated effect of (U + F) concentration is based on very limited data. Friedrich and Feix (1972) show that the $T_m$ of the double-stranded replicative form of MS2 RNA is depressed by 0.3 to 0.5°C per percent formamide. The melting point depression of DNA by (U + F) and by pure formamide are approximately the same per volume percent solvent. If we assume the same is true for RNA duplexes, we arrive at the above equation.

This equation predicts $T_m$ for BKD, RD-114, and WoMV dimers in aqueous 0.1 M $Na^+$ of 87°C, 87°C, and 66°C. We regard these values as only approximate, but the relative order of dissociation temperatures is reliable.

RD-114 virus was isolated from a human rhabdomyosarcoma after an in vivo passage in a fetal cat brain (McAllister et al., 1972). It is thought to be an endogenous feline virus because it shows close sequence homology to the infectious type C virus that can be induced from a cat cell line in culture (Livingston and Todaro, 1973; Fischinger et al., 1973; Sarma et al., 1973), and because it shows sequence homology with cat cellular DNA (Baluda and Roy-Burman, 1973; Nieman, 1973; Ruprecht, Goodman, and Spiegelman, 1973).

BKD virus was isolated by co-cultivation of normal baboon kidney with a dog thymus cell line, and it appears identical to several other isolates of endogenous type C baboon viruses (Melnick et al.,

1973; Benveniste et al., 1974; Todaro et al., 1974a). BKD shows partial antigenic relationship to RD-114 with regard to viral reverse transcriptase and the viral core protein, p30 (Hellman et al., 1974; Sherr et al., 1974). The sequence homology between RD-114 and BKD, determined by hybridizing a cDNA probe made from one virus to an excess of RNA from the other, is 10–20% (Benveniste et al., 1974; Todaro et al., 1974b). Thus by several criteria, BKD and RD-114 are likely to be evolutionarily related. Todaro et al. (1974b) have suggested that the CCC/RD-114 family of endogenous cat viruses was introduced into the cat germ line by horizontal transmission of a primate virus, related to the baboon virus.

WoMV is the simian sarcoma virus isolated from a woolly monkey fibrosarcoma (Theilen et al., 1971). There is no detectable sequence homology between WoMV RNA and the RNAs of either RD-114 or BKD as measured by hybridization of labeled reverse transcriptase products to excess viral RNA (Benveniste and Todaro, 1973; East et al., 1975). Similarly, there is no antigenic cross-reaction between WoMV and either RD-114 and BKD for either the viral polymerase or the gs-1 determinants of the p30 protein (Sherr et al., 1974, 1975). WoMV does appear by both antigenic and sequence homology tests to be closely related to the gibbon-ape leukemia virus, and more distantly related to several murine type C viruses and an endogenous pig virus (Sherr et al., 1974, 1975; Benveniste and Todaro, 1973).

That such similar structures appear in the RNAs of RD-114, BKD, and WoMV suggests that all type C RNA viruses have RNA dimers with loops. These features, however, were very rarely observed in Rous sarcoma virus (Bender and Davidson, 1976). We have observed that the 60–70S RNA of Rous sarcoma virus dissociates to 30–35S RNA more readily than do WoMV dimers. (Bender and Davidson, 1976). Thus it is probable that if dimer structures exist in Rous sarcoma virus, it would be difficult to observe them, since denaturing conditions are required to extend the RNA for electron microscopy. If loops exist, they must also be easily disrupted.

The RNA that we observe has been through phenol extraction, ethanol precipitation, and spreading in denaturing solvents. We cannot be certain that what we see represents the "native" structure, that is, that which exists in the virion. It is improbable that virions contain single RNA monomers which associate into dimers after extraction, since we have never observed reassociation of dimers after they have been dissociated by heating. Comparisons of the relative sedimentation velocities of RD-114 RNA and several marker RNAs in nondenaturing and glyoxal sucrose gradients support the view that the 52S RNA complex is a single-dimer molecule (Kung et al., 1975). By analogy, the same is true for BKD and WoMV 52S RNAs. Our data say nothing about the number of 52S dimers per virion.

It seems clear that the loops and dimer linkage structures are formed by processes more complicated than simple association of complementary sequences. When the loops are dissociated by heating viral RNA in a low concentration of (U + F), the loops never reappear when the solution is cooled. If there were two complementary sequences at the base of the loop, they should reanneal very quickly, since they are always in close proximity. Similarly, the dimer linkage structure does not reform after heating and cooling. We have heat-denatured 52S RD-114 RNA and incubated it at a high RNA concentration to allow reassociation of dimer linkages, but reassociation has never been observed. The large hairpin structure of RD-114 and BKD is also peculiar in that it never appears until the virus is subjected to very denaturing conditions, and then it is present only on some of the molecules.

Several alternative models may be suggested to explain how the monomers are joined in the dimer linkage structure; similar models could account for the structure at the junctions of the loops. There may be a protein binding the two subunits together. Such a protein would have to be resistant to the phenol-SDS extraction used here to prepare all viral RNAs, and also resistant, in the case of RD-114, to pronase-SDS extraction (Kung et al., 1975). If dissociation of the dimers occurs upon heat denaturation of a protein linker, the monomers would probably not reassociate.

In proposing models to explain how the dimer linkage structure could be held together by base pairing, we accept two constraints: first, there is only standard Watson-Crick antiparallel base pairing; second, the dimer linkage involves an interaction between sequences close to the 5′ ends of identical monomers. The latter assumption is based on the following observations. The poly(A) sequences of mRNAs and some tumor viral RNAs are at the 3′ end (Wang and Duesberg, 1974; Quade, Smith, and Nicols, 1974; Rho and Green, 1974), and our poly(A) mapping shows that the dimer linkage joins the two non-poly(A) ends. Furthermore, oligonucleotide fingerprint evidence suggests that all subunits of Rous sarcoma virus RNA are identical in sequence (Beemon, Duesberg, and Vogt, 1974; Weissmann et al., 1974). We assume this is also the case for other tumor virus RNAs.

Figure 8a shows a structure with two identical monomers base-paired to each other. All such structures require the presence of inverted repeat sequences on each strand. Figure 8b shows how

a small RNA molecule might serve as a linker. The RNA linker could be the same tRNA that serves as a primer for the in vitro reverse transcriptase reaction (Dahlberg et al., 1974; Taylor and Illmensee, 1975). In both these models, the monomers would not be expected to reassociate after thermal dissociation. In Figure 8a, the inverted repeats would form base pairs intramolecularly, and the linker of Figure 8b would anneal completely to just one monomer strand. Both models therefore require special in vivo mechanisms to cause intermolecular base pairing for dimer formation. Figure 8c gives another possible structure with a small RNA linker which contains a tandem direct repeat so that it can base pair with the two identical monomers. This structure could reform after heat dissociation. Models for the dimer linkage similar to these were independently conceived and communicated to us by Dr. Andrew King of the University of Wisconsin.

We can only guess at the functions of the dimer linkage structure and the loop structures. They could have some role in packaging the RNA in the virus. The dimer linkage, for instance, may insure that every virion is diploid. It has been reported, however, that the 60–70S complex in avian sarcoma virus is unstable when the virus is first assembled and released (Canaani, Helm, and Duesberg, 1973; Stoltzfus and Snyder, 1975). The structures could also function in the synthesis of DNA from the viral RNA. It is less probable that the dimer linkage structures affect RNA synthesis or protein synthesis because integrated RD-114 proviral DNA is infectious (Nicolson et al., 1976), and as has been discussed, the structures mentioned in the above models would not be expected to form in the integrated DNA or in the messenger transcribed from it.

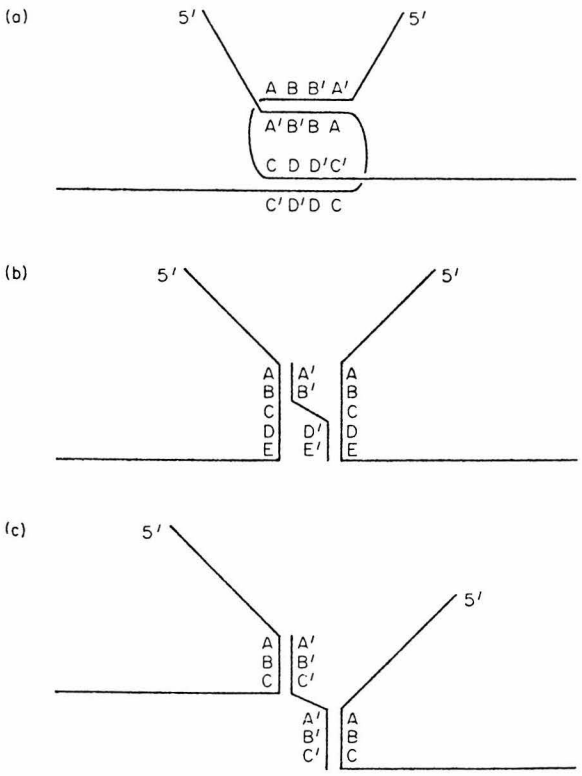



Figure 8. Diagrams of Possible Structures for the Dimer Linkage in Which the Binding Is through Normal Complementary Base Pairing

(a) shows a model in which identical monomers bind to each other at inverted repeat sequences. In (b), a small RNA linker joins two identical monomers; a second linker of the same sequence could be added to the diagram to give double bonding. (c) is a model in which the linker contains a tandem repeat. In all cases, the letters (A–E) indicate base sequences and the primed letters (A′–E′) indicate their complementary sequences.

## Experimental Procedures

### Virus and RNA Preparation

BKD virus, obtained from Dr. Raymond Gilden, is the Bab 8-K strain (Todaro et al., 1974a). It was grown in dog thymus cells and then by us in human HT-1080 cells (Rasheed et al., 1974). WoMV, obtained from Dr. Gilden, is derived from the simian sarcoma virus isolated by Theilen et al. (1971). It was grown in woolly monkey muscle cells and then by us in human RD cells. Nontransforming helper virions are usually in more than 10 fold excess over sarcoma virions (Wolfe, Smith, and Deinhardt, 1972; Scolnick et al., 1972; Aaronson, 1973). Dr. Gilden confirmed the identity of the BKD virus preparation by hybridizing cDNA transcripts from RD-114 or BKD to cellular RNA from our producer cells. The identity of our WoMV preparation was confirmed by Gilden in CF tests using antisera to WoMV p30 antigens.

The viruses were banded in sucrose, and the RNA was prepared by phenol-SDS extraction and ethanol precipitation as described previously (Kung et al., 1975).

### Sedimentation Analysis

NTE–sucrose gradient: 10–30% sucrose gradients containing NTE buffer [0.1 M NaCl, 0.01 M Tris (pH 7.0), 0.001 M EDTA] were prepared. $^3$H-labeled RNA samples (in NTE), about 50 $\mu$l, were layered onto the gradients and centrifuged in an SW50.1 rotor at 45,000 rpm and 4°C.

Glyoxal–sucrose gradient: $^3$H-labeled RNA samples were treated with glyoxal by dialyzing against 1 M glyoxal in 0.01 M phosphate buffer (pH 7.0) for 1 hr at 37°C, then by dialyzing against 0.1 M glyoxal in the same buffer for approximately 30 min at 4°C. A sample thus treated was sedimented through a 10–30% sucrose gradient in the presence of 0.1 M glyoxal, 0.01 M phosphate buffer (pH 7.2), at 4°C, 45,000 rpm in an SW50.1 rotor. RD-114 52S RNA was similarly treated and run in parallel as an external marker.

For preparative purposes, low salt [1 mM Tris (pH 7.2), 0.1 mM EDTA] sucrose gradients were used to fractionate high molecular weight RNA species. Phenol-extracted RNA samples were dissolved in the low salt buffer and sedimented through these gradients at 4°C. Under these conditions, BKD and WoMV RNAs both sediment as single components with the same sedimentation velocity as RD-114 RNA. Samples taken from the peak positions were used for electron microscope studies.

### Electron Microscopy

The preparation of RNA samples for electron microscopy and the details of the different spreading techniques have been previously described (Kung et al., 1975). The urea plus formamide (U + F) solvent was prepared by adding 480 g (8 mol) of urea to1l of recrystallized formamide; this gives a solution about 74% by volume formamide and about 5.9 M in urea. The volume percentage of this

solvent in a solution of aqueous buffer is quoted as percent (U + F). The techniques for mapping the poly(A) sequences on the viral RNA are presented in the accompanying paper (Bender and Davidson, 1976). For length measurements, SV40 DNA double-stranded circles were compared to a diffraction grating, and SV40 circles were used as internal length standards in the poly(A) mapping experiments to calibrate the lengths of viral RNA monomers and dimers (Figure 4). Lengths were converted to kilobases assuming 0.256 $\mu$m/kb, as measured for glyoxal-treated E. coli rRNA (Hsu et al., 1973).

### Agarose Gel Electrophoresis

BKD 52S RNA samples were dissolved in 67% (U + F) containing 0.033 M Tris (pH 7.3) (0.024 M cations), heated at desired temperature for 1 min, and then quickly cooled in ice. After adding 10% glycerol, the sample solutions were loaded onto 1% agarose gels containing E buffer [0.05 M boric acid, 0.005 M sodium borate ($Na_2B_4O_7 \cdot 10\ H_2O$), 0.01 M sodium sulfate, and 0.001 M EDTA (pH 8.2)], all as described (Kung et al., 1975). WoMV 52S RNA dissociates more readily than BKD 52S RNA, and so to determine the dissociation conditions, WoMV RNA samples were not heated but were dissolved in solutions with varying concentrations of (U + F) and with 0.025 M Tris, 0.0025 M EDTA (pH 8.5). These solutions were then applied to the gels.

For methylmercury hydroxide–agarose gel experiments, all RNA samples were applied onto the gel (1% agarose) in E buffer containing 5 mM $CH_3HgOH$ and 10% glycerol. 28S and 18S unlabeled HeLa rRNA and 23S E. coli rRNA were included as internal markers. To visualize rRNA bands after electrophoresis, gels were stained with 1 $\mu$g/ml ethidium bromide, 0.5 M $NH_4Ac$ for 30 min and examined by illumination with short wavelength ultraviolet light. The distances traveled by rRNA markers were recorded. The gels were then sliced into 2 mm fractions, and the radioactivity of $^3$H-labeled BKD or WoMV RNAs was determined.

### Acknowledgments

We are grateful to Dr. Raymond Gilden for his several contributions as noted in the text. This research has been supported by contracts to N. D. and R. M. from the Virus Cancer Program of the National Cancer Institute; J. B. has been supported by a fellowship from the Helen Hay Whitney Foundation, and W. B. by an NSF fellowship and a training grant from the NIH.

Received December 8, 1975; revised January 20, 1976

### References

Aaronson, S. (1973). Biologic Characterization of Mammalian Cells Transformed by a Primate Sarcoma Virus. Virology *52*, 562–567.

Bailey, J. M., and Davidson, N. (1976). Methylmercury as a Reversible Denaturing Agent for Agarose Gel Electrophoresis. Anal. Biochem., in press.

Baluda, M. A., and Roy-Burman, P. (1973). Partial Characterization of RD-114 Virus by DNA-RNA Hybridization Studies. Nature New Biol. *244*, 59–62.

Beemon, K., Duesberg, P., and Vogt, P. (1974). Evidence for Crossing-over Between Avian Tumor Viruses Based on Analysis of Viral RNAs. Proc. Nat. Acad. Sci. USA *71*, 4254–4258.

Bender, W., and Davidson, N. (1976). Mapping of Poly(A) Sequences in the Electron Microscope Reveals Unusual Structure of Type C Oncornavirus RNA Molecules. Cell *7*, 595–607.

Benveniste, R. E., and Todaro, G. J. (1973). Homology Between Type-C Viruses of Various Species as Determined by Molecular Hybridization. Proc. Nat. Acad. Sci. USA *70*, 3316–3320.

Benveniste, R. E., Lieber, M. M., Livingston, D. M., Sherr, C. J., Todaro, G. J., and Kalter, S. S. (1974). Infectious C-type Virus Isolated From a Baboon Placenta. Nature *248*, 17–20.

Canaani, E., Helm, K. V. D., and Duesberg, P. (1973). Evidence for 30-40S RNA as Precursor of the 60-70S RNA of Rous Sarcoma Virus. Proc. Nat. Acad. Sci. USA *70*, 401–405.

Dahlberg, J. E., Sawyer, R. C., Taylor, J. M., Faras, A. J., Levinson, W. E., Goodman, H. M., and Bishop, J. M. (1974). Transcription of DNA from the 70S RNA of Rous Sarcoma Virus. J. Virol. *13*, 1126–1133.

East, J. L., Knesek, J. E., Chan, J. C., and Dmochowski, L. (1975). Quantitative Nucleotide Sequence Relationships of Mammalian RNA Tumor Viruses. J. Virol. *15*, 1396–1408.

Fischinger, P. J., Peebles, P. T., Nomura, S., and Haapala, D. K. (1973). Isolation of an RD-114-Like Oncornavirus from a Cat Cell Line. J. Virol. *11*, 978–985.

Friedrich, R., and Felix, G. (1972). RNA-RNA Hybridization in Aqueous Solutions Containing Formamide. Anal. Biochem. *50*, 467–476.

Hellman, A., Peebles, P. T., Strickland, J. E., Fowler, A. K., Kalter, S. S., Droszlan, S., and Gilden, R. V. (1974). Baboon Virus Isolate M-7 with Properties Similar to Feline Virus RD-114. J. Virol. *14*, 133–138.

Hsu, M. T., Kung, H. J., and Davidson, N. (1973). An Electron Microscope Study of Sindbis Virus RNA. Cold Spring Harb. Symp. Quant. Biol. *38*, 943–950.

Kung, H. J., Bailey, J. M., Davidson, N., Vogt, P. K., Nicolson, M. O., and McAllister, R. M. (1974). Electron Microscope Studies of Tumor Virus RNA. Cold Spring Harb. Symp. Quant. Biol. *39*, 827–834.

Kung, H. J., Bailey, J. M., Davidson, N., Nicolson, M. O., and McAllister, R. M. (1975). Structure, Subunit Composition, and Molecular Weight of RD-114 RNA. J. Virol. *16*, 397–411.

Livingston, D. M., and Todaro, G. J. (1973). Endogenous Type C Virus from a Cat Cell Clone with Properties Distinct from Previously Described Feline Type C Virus. Virology *53*, 142–151.

McAllister, R. M., Nicolson, M., Gardner, M. B., Rongey, R. W., Rasheed, S., Sarma, P. S., Huebner, R. J., Hatanaka, M., Oroszlan, S., Gilden, R. V., Kabigting, A., and Vernon, L. (1972). C-type Virus Released from Cultured Human Rhabdomyosarcoma Cells. Nature New Biol. *235*, 3–6.

Melnick, J. L., Altenberg, B., Arnstein, P., Mivkovic, R., and Tevethia, S. S. (1973). Transformation of Baboon Cells with Feline Sarcoma Virus. Intervirology *1*, 386–398.

Nicolson, M. O., Hariri, F., Krempin, M., McAllister, R., and Gilden, R. V. (1976). Infectious Proviral DNA in Human Cells Infected with Non-transforming Type-C Viruses. Virology, in press.

Nieman, P. E. (1973). Measurement of RD-114 Virus Nucleotide Sequences in Feline Cellular DNA. Nature New Biol. *244*, 62–64.

Quade, K., Smith, R. E., and Nicols, J. L. (1974). Poly (Riboadenylic Acid) and Adjacent Nucleotides in Rous Sarcoma Virus RNA. Virology *62*, 60–70.

Rasheed, S., Nelson-Rees, W. A., Toth, E. M., Arnstein, P., and Gardner, M. B. (1974). Characterization of a Newly Derived Human Sarcoma Cell Line (HT-1080). Cancer *33*, 1027–1033.

Rho, H. M., and Green, M. (1974). The Homopolyadenylate and Adjacent Nucleotides at the 3'-Terminus of 30-40S RNA Subunits in the Genome of Murine Sarcoma-Leukemia Virus. Proc. Nat. Acad. Sci. USA *71*, 2386–2390.

Ruprecht, R. M., Goodman, N. C., and Spiegelman, S. (1973). Determination of Natural Host Taxonomy of RNA Tumor Viruses by Molecular Hybridization: Application to RD-114, a Candidate Human Virus. Proc. Nat. Acad. Sci. USA *70*, 1437–1441.

Sarma, P. S., Tseng, J., Lee, Y. K., and Gilden, R. V. (1973). Virus Similar to RD-114 Virus in Cat Cells. Nature New Biol. *244*, 56–58.

Schildkraut, C., and Lifson, S. (1965). Dependence of the Melting Temperature of DNA on Salt Concentration. Biopolymers *3*, 195–208.

Scolnick, E. M., Parks, E. P., Todaro, G. J., and Aaronson, S. A. (1972). Immunological Characterization of Primate C-Type Virus Reverse Transcriptase. Nature New Biol. *235*, 35–40.

Sherr, C. J., Lieber, M. L., Benveniste, R. E., and Todaro, G. J. (1974). Endogenous Baboon Type C Virus (M7): Biochemical and Immunologic Characterization. Virology *58*, 492–503.

Sherr, C. J., Fedele, L. A., Benveniste, R. E., and Todaro, G. J. (1975). Interspecies Antigenic Determinates of the Reverse Transcriptases and p30 Proteins of Mammalian Type C Viruses. J. Virol. *15*, 1440–1448.

Stoltzfus, C. M., and Snyder, P. N. (1975). Structure of B77 Sarcoma Virus RNA: Stabilization of RNA After Packaging. J. Virol. *16*, 1161–1170.

Taylor, J. M., and Illmensee, R. (1975). Site on the RNA of an Avian Sarcoma Virus at Which Primer is Bound. J. Virol. *16*, 553–558.

Theilen, G. J., Gould, D., Fowler, M., and Dungworth, D. (1971). C-type Virus in Tumor Tissue of a Woolly Monkey (Lagothrix spp.) with Fibrosarcoma. J. Nat. Cancer Inst. *47*, 881–889.

Todaro, G. J., Sherr, C. J., Benveniste, R. E., Lieber, M. M., and Melnick, J. L. (1974a). Type C Viruses of Baboons: Isolation from Normal Cell Cultures. Cell *2*, 55–61.

Todaro, G., Beneviste, R., Callahan, R., Lieber, M., and Scherr, C. (1974). Endogenous Primate and Feline Type C Viruses. Cold Spring Harbor Symp. Quant. Biol. *39*, 1159–1168.

Wang, L., and Duesberg, P. (1974). Properties and Location of Poly(A) in Rous Sarcoma Virus RNA. J. Virol. *14*, 1515–1529.

Weissmann, C., Parson, J. T., Coffin, J. W., Rymo, L., Billeter, M. A., and Hofstetter, H. (1974). Studies on the Structure and Synthesis of Rous Sarcoma Virus RNA. Cold Spring Harbor Symp. Quant. Biol. *39*, 1043–1056.

Wolfe, L. G., Smith, R. K., and Deinhardt, F. (1972). Simian Sarcoma Virus, Type 1 (Lagothrix): Focus Assay and Demonstration of Nontransforming Associated Virus. J. Nat. Cancer Inst. *48*, 1905–1907.

JOURNAL OF VIROLOGY, Oct. 1976, p. 264-272 Vol. 20, No. 1
 Printed in U.S.A.

# Size, Subunit Composition, and Secondary Structure of the Friend Virus Genome[1]

SHYAM DUBE, HSING-JIEN KUNG, WELCOME BENDER, NORMAN DAVIDSON,* AND WOLFRAM OSTERTAG

*Department of Chemistry* and Division of Biology, California Institute of Technology, Pasadena, California 91125, and Max Planck Institut für Experimentelle Medizin, Göttingen, Germany*

Received for publication 5 May 1976

Electron microscope and gel electrophoresis studies show that the high-molecular-weight (50 to 70S) RNA extracted from Friend virus (FV) is a dimer with the same basic structure previously observed for the RNAs from RD-114 virus, baboon virus, and woolly monkey virus. This observation greatly strengthens the inference that the dimer structure is a general characteristic of the RNAs of all mammalian type C viruses. The FV dimer is slightly less stable than the RNA dimer of woolly monkey virus, which is, in turn, much less stable than those of RD-114 and baboon virus. There are three FV monomer components, small (S), medium (M), and large (L), with molecular lengths of 6.7 ± 0.6, 7.7 ± 0.6, and 9.5 ± 0.6 kilobases, respectively. There are approximately equal amounts of the S and M components and much less of the L component. Most of the dimers are homodimers (SS, MM, and LL). The frequency of heterodimers (SM, SL, ML) is much less than expected for a random assortment model.

Friend virus (FV) (7) preparations are known to contain two viral components with different biological activities (1, 10, 16), the spleen focus-forming virus, and the lymphatic leukemia virus. The spleen focus-forming virus is oncogenic; it transforms erythroid precursor cells and induces spleen focus formation, thus resulting in erythroleukemia in mice. The spleen focus-forming virus is replication defective and needs a helper virus. The helper function is provided by the lymphatic leukemia virus component of the FV complex (16). The helper function can be provided by many other murine leukemia viruses (for example, Moloney leukemia virus), which do not induce erythroid disease (16). The endogenous virus released during dimethyl sulfoxide-induced differentiation of Friend cells (5) and upon exposure of these cells to bromodeoxyuridine also has helper activity (5, 14).

Earlier work on the FV genomic RNA has shown that: (i) the FV RNA contains two subunits whose sedimentation coefficients are approximately 35 and 32S; (ii) the amount of the larger subunit is ≤20%, whereas the smaller subunit constitutes 80% or more of the FV RNA; (iii) the genomic complexity is approximately $2.5 \times 10^6$ daltons; and (iv) the viral RNA contains polyadenylic acid [poly(A)] and also internal oligo(A) tracks (12). In addition, the 35 and 32S subunits have been tentatively assigned to lymphatic leukemia virus and spleen focus-forming virus RNA, respectively (11).

Electron microscope studies from this laboratory have shown that the high-molecular-weight (50 to 70S typically) RNA components extracted from several different type C viruses all have a common structure (3, 8, 9). They are dimers consisting of two monomers with molecular weights of about $3 \times 10^6$, joined together at their 5′ ends by a secondary feature called the dimer linkage structure (DLS). There is a loop close to the middle of each monomer, and there is a poly(A) stretch at each 3′ end. This structure has been observed for RD-114, an endogenous xenotropic cat virus; for BKD, an endogenous xenotropic baboon virus; and for WoMV, a virus horizontally transmitted in woolly monkeys. It was of interest, therefore, to ask whether a murine type C virus has the same structure. The Friend system is of additional interest because it contains both transforming and leukemogenic activities and because Friend cells are a system in which to study the relation between expression of viral and cellular genes.

## MATERIALS AND METHODS

**Stock solutions.** NTE buffer consisted of 0.1 M NaCl, 0.01 M Tris-OH, and 0.001 M EDTA, adjusted to pH 7.5 with HCl.

**(U+F) solvent.** A 480-g quantity of urea was dissolved in 1 liter of formamide, giving 1.35 liters of

[1] Contribution no. 5302 from the Department of Chemistry, California Institute of Technology, Pasadena, CA 91125.

solution. The volume percentage of this solution when mixed with aqueous electrolyte is reported as percent (U+F). (See reference 8 for explanation.)

**E buffer.** E buffer consisted of 0.05 M boric acid, 0.005 M sodium borate ($Na_2B_4O_6 \cdot 10H_2O$), 0.01 M $Na_2SO_4$, and 0.001 M EDTA (pH 8.2).

**TE buffer.** A y molar concentration of TE buffer contains (per liter) y moles of Tris OH and 0.1 y mol of Na EDTA, all adjusted to pH 8.5 with HCl. The cation concentration is 0.3 y in $Na^+$ ions, approximately 0.3 y in Tris $H^+$, and therefore contains 0.6 y total.

**Cells, virus, and RNA.** The cell line used, FSD-3, has been described previously (4). Briefly it originated as follows. FV containing cell-free supernatant from a cell culture of an FV-transformed cell line, FSD-1 (13), clone $F_4$, was injected intraperitoneally into BALB/c mice. Spleen cells from these mice were used to start the FSD-3 cell line. FV was isolated from these cells as described (5). FV twice purified by isopycnic banding in a 24 to 48% sucrose gradient was used for RNA isolation. The viral pellet was suspended in 1% sodium dodecyl sulfate, and to this an equal volume of a mixture of phenol, chloroform, and isoamyl alcohol (50:2:48) was added. Viral RNA was precipitated from the aqueous phase by ethanol and purified by sedimentation on a 10 to 30% NTE-sucrose gradient for 75 min at 4°C in a Beckman SW50.1 rotor. Fractions were collected and monitored for absorbance at 260 nm. The peak fractions in the 50 to 60S region were pooled, precipitated with ethanol, and used for the present studies.

**Agarose gel electrophoresis.** All gel electrophoresis experiments were conducted on 1% agarose gels in E buffer. The preparation of the non-denaturing gels and of the strongly denaturing $CH_3HgOH$ gels has been described (2, 8). In the experiments for studying the dissociation temperature of the 50 to 60S RNA, RNA samples were mixed with (U+F) solvent and NTE buffer to the desired concentrations. Samples (ca. 20 $\mu$l) were sealed in a capillary tube, incubated for 5 min at the desired temperature, mixed with glycerol and bromophenol blue, and loaded onto the gels. Electrophoresis was performed at 5 mA/tube for 3 h at room temperature. The gels were then stained with 1 $\mu$g of ethidium bromide per ml in 0.5 M $NH_4Ac$ for 30 min and examined by illumination with short-wavelength UV light.

For agarose gel electrophoresis experiments using $CH_3HgOH$ as a denaturing agent, RNA samples were mixed with $CH_3HgOH$ in E buffer to a final concentration of 5 mM $CH_3HgOH$ and loaded onto a 1% agarose gel containing 5 mM $CH_3HgOH$. Electrophoresis was performed for 4 h at 25°C at 5 mA/tube.

**Electron microscopy.** The urea-formamide and the glyoxal-formamide spreading techniques have been described previously (3, 8, 9). In the urea-formamide spreading, RNA samples were diluted into the spreading solution which contained about 30 $\mu$g of cytochrome *c* per ml and the desired concentration of (U+F) and electrolyte. The hypophase was double-distilled water. To study the structure of undissociated 50 to 60S RNA, the spreading was performed in 80% (U+F) solution containing 0.15 M TE ($\approx$0.09 M cations in the final solution). These conditions cause only partial melting of the 50 to 60S RNA. In the experiments to determine the molecular weights of dissociated subunits, RNA samples were first treated with 85 to 90% (U+F) in $\leq$0.05 M TE, and were then either spread directly from 85% (U+F) or diluted and spread from 60% (U+F). Essentially identical results have been obtained from these spreadings.

Our standard glyoxal-formamide technique involves treatment of the RNA with 1 M glyoxal and 0.01 M phosphate buffer, pH 6.7, at 37°C for 1 h. A less-denaturing treatment is with 1 M glyoxal, 0.01 M phosphate buffer, and 0.05 M NaCl at 37°C for 1 h. The glyoxalated RNA is then either diluted 10-fold or more into spreading solution of 30% formamide and 0.1 M TE or dialyzed into this solution and spread onto 5% formamide and 0.01 M TE.

The technique for mapping of poly(A) segments on RNA molecules using simian virus 40 (SV40) circles with polydeoxythymidylate [poly(dT)] tails has been described previously (3). Conditions of electron microscopy are as previously described. Magnifications were measured with a diffraction grating and/or using circular SV40 as an internal standard.

## RESULTS

**RNA components of FV complex from FSD-3 cells.** In non-denaturing (i.e., high salt, NTE) sucrose gradients, FV RNA sedimented as a single peak with a sedimentation coefficient of 50 to 60S. When this RNA fraction was collected and subjected to electrophoresis in non-denaturing agarose gels, a single diffuse band (Fig. 1b) with an apparent molecular weight of $5 \times 10^6$ to $6 \times 10^6$ was reproducibly observed. If, however, the RNA samples were first treated with a denaturing solvent [i.e., dissolved in 66% (U+F), 0.03 M cations] at 25°C and then applied onto the agarose gel, about 50% of the FV RNA dissociated into subunits. This was evidenced by the appearance of three faster-moving bands in addition to the original diffuse band. Upon heating to 40°C in the same solvent [66% (U+F), 0.03 M cations] the diffuse band dissociated completely; there were only the three faster bands plus a broad distribution of lower-molecular-weight fragments. (This gel pattern is not illustrated, but a similar pattern obtained in methyl mercury gels is described below.) Treatment of RNA with more denaturing conditions [90% (U+F), 0.03 M cations] at 40°C did not cause further dissociation.

The three-band pattern has been observed for the heat-dissociated FV RNA extracted from several different virus preparations. The two bands with higher mobility, referred to as S (for small) and M (for medium), had comparable fluorescent intensities when stained with ethidium bromide; L (large), the slowest moving band, had about one-third the intensity of the S or M bands.

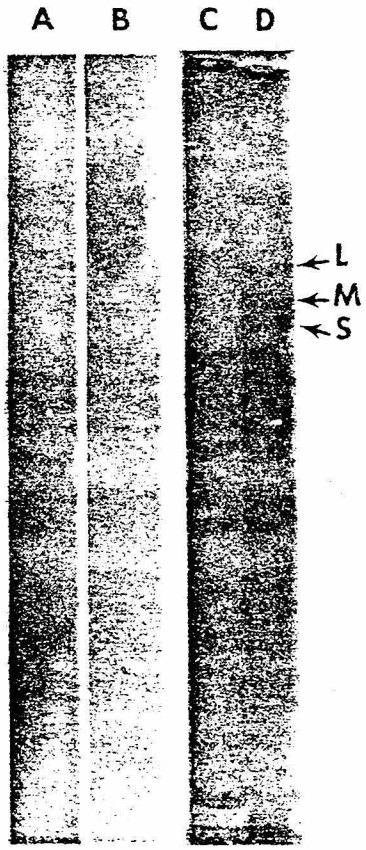



FIG. 1. *Electrophoretic gel patterns in 1% agarose. (A) and (B) are nondenaturing gels run in parallel; (C) and (D) are denaturing gels containing 5 mM $CH_3HgOH$ run in parallel. (A) HeLa 28 and 18S ribosomal RNA. (B) 50 to 60S FV RNA. (C) HeLa 28 and 18S rRNA and Escherichia coli 23S ribosomal RNA. (D) same as (C) plus 50 to 60S FV RNA. L, M, and S denote the large, medium, and small subunits of FV RNA.*

**Molecular weight estimates by gel electrophoresis.** $CH_3HgOH$ (5 mM) is a strong denaturant for nucleic acids and agarose gel electrophoresis. $CH_3HgOH$ is useful for measuring molecular weights of RNA polynucleotide strands with minimal effects due to secondary structure (2). An experiment on FV RNA is shown in Fig. 1D; again three discrete bands are observed, with the slow (L) component being the faintest.

A plot of log (molecular weight) versus mobility for the molecular weight standards in Fig. 1C and 1D and for the three FV components, S, M, and L, is shown in Fig. 2. A rough visual estimate of the band widths of the three components is also shown in the figure. This plot leads to molecular weight estimates of $2.33 \times 10^6 \pm 0.2 \times 10^6$, $2.65 \times 10^6 \pm 0.2 \times 10^6$, and $3.27 \times 10^6 \pm 0.2 \times 10^6$, corresponding to kilobase (kb) lengths of 6.7 ± 0.6, 7.7 ± 0.6, and 9.5 ± 0.6, for the S, M, and L components, respectively.

**Electron microscope studies.** To extend the molecules into filaments with a recognizable topology and reproducible length for electron microscopy, RNA must be spread from partially denaturing solvents. The dimer structures of RD-114 and BKD RNAs are quite stable and easy to preserve under suitable spreading conditions (8, 9). WoMV RNA dimers dissociate more readily, and it is necessary to delicately adjust the spreading conditions to observe the dimer structure (9). We observe that FV RNA has a dimer structure which is slightly less stable towards dissociation than is WoMV RNA.

The two partial denaturation techniques that we have found to be particularly useful for electron microscopy involve: (i) treatment with and spreading from a (U+F) solvent and (ii) treatment with glyoxal followed by spreading from a rather low-percent formamide solvent, usually 40%. The urea-formamide method permits quantitative control of denaturing conditions. The glyoxal method is used when it is desired to identify the poly(A) ends of the RNA molecules by hybridization to circular duplex SV40 DNA with attached poly(dT) tails (3). As explained below, a sample of molecules with labeled poly(A) ends is particularly useful for reliable length measurements on molecules known to be unbroken.

When FV RNA is treated with 90% (U+F),

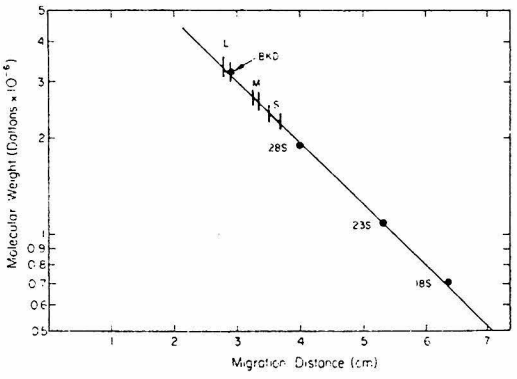



FIG. 2. *A plot of RNA electrophoretic mobility versus molecular weight for the $CH_3HgOH$ gel shown in Fig. 1D. Symbols: (●) Indicates the position of a molecular weight standard; (⊢⊣) indicates the band width of FV RNA. The baboon viral RNA (BKD) marker was labeled with $^3H$, and its position was determined by slicing the gel into 1-mm disks and counting.*

0.03 M cations, and then spread from 60% (U+F), 0.03 M cations, polynucleotide strands with a broad distribution of molecular lengths ranging from 0 to about 2.0 $\mu$m and with no discernible secondary structure are seen. Some typical examples are shown in Fig. 3a. A histogram of the length distribution shows a peak in the range of 1.7 to 1.9 $\mu$m, corresponding to molecular lengths in the range of 6.7 to 7.5 kb. As shown below these are monomer units of FV RNA.

When FV RNA is spread from 80% (U+F), 0.09 M cations, about 70% of the molecules are sufficiently extended to be traceable. Many of these molecules, such as the one shown in Fig. 3c, have about twice the length of the filaments shown in Fig. 3a and are, therefore, dimers. They have the same basic structure seen previously for the dimers of RD-114, BKD, and WoMV RNAs. There is a central DLS, which in this case is rather small, and sometimes, as in Fig. 3c, there is a secondary structure loop in the middle of a monomer. In the case of FV RNA, there are often several crossover points

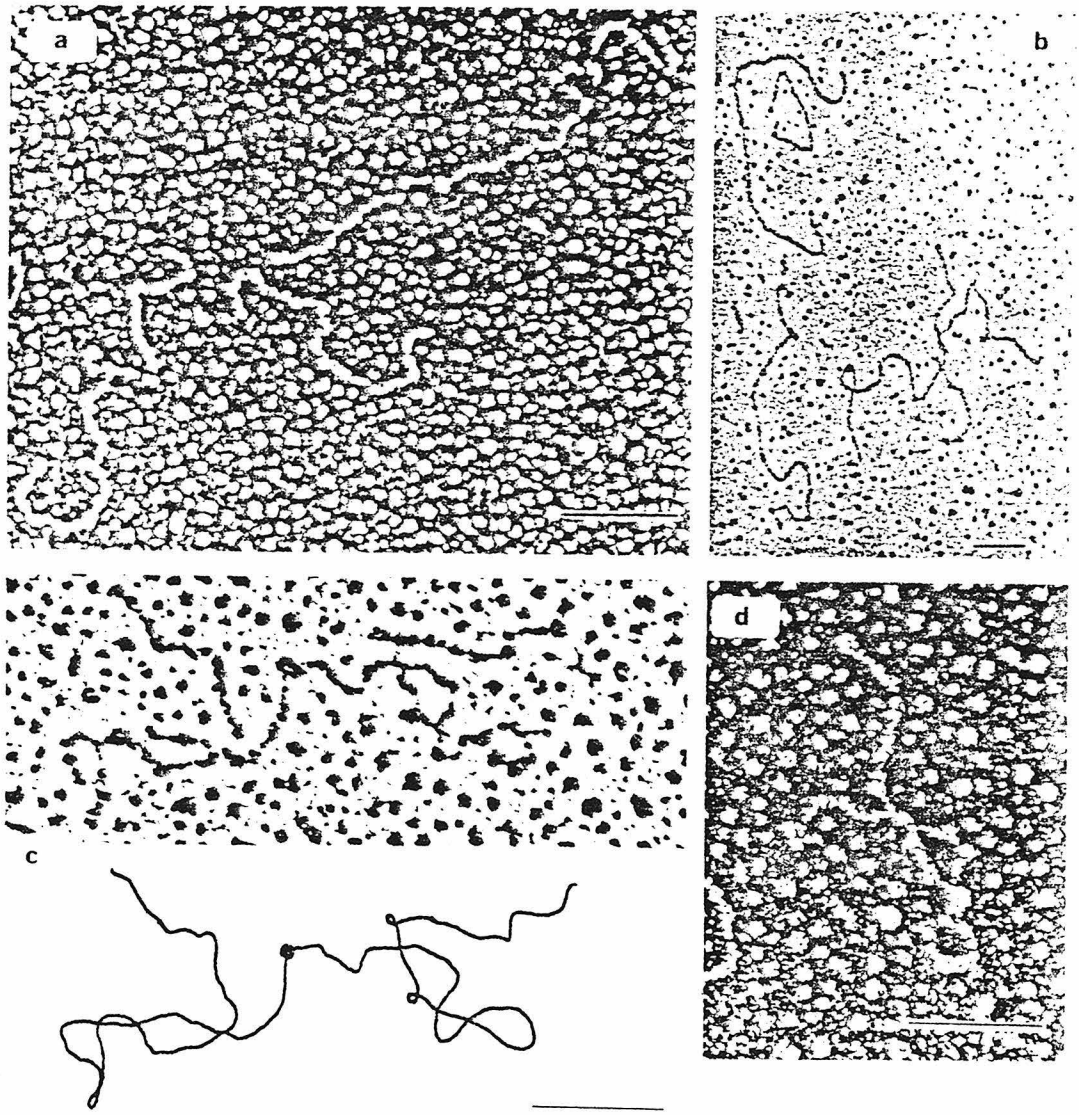



FIG. 3. *Electron micrographs of FV RNAs. (a) A 90% concentration of (U + F) solvent (0.03 M cations) treated and spread from 60% (U + F). (b) Glyoxal treated and spread from 30% formamide (0.06 M cations). (c) and (d) A 80% (U + F) solvent (0.09 M cations) treated and spread in the same medium. An interpretive tracing of the molecule in (c) is shown. All spreading procedures are described in Materials and Methods. The length marker is 0.2 $\mu$m.*

within the loops. In the same spreading some monomers are seen, and some of these still have the secondary structure loop (Fig. 3d). Under the spreading conditions used, about 50% of the traceable molecules are of monomer length or shorter.

If FV RNA is incubated with 1 M glyoxal in 0.01 M phosphate buffer at 37°C for 1 h, and the resulting RNA is spread from 30 or 40% formamide, 0.06 M cations, there are a few dimers, but most of the RNA molecules are smooth structureless filaments with molecular lengths of up to 1.9 $\mu$m. Thus, this treatment causes considerable dissociation of the FV RNA into monomers. However, if the RNA is treated with glyoxal at a higher salt concentration and spread from 40% formamide, 0.06 M cations, many more dimer molecules are seen. Poly(A) tails on RNA molecules treated with glyoxal can be mapped in the electron microscope by hybridizing to circular SV40 containing poly(dT) tails (3). Examples of such molecules are shown in Fig. 4. They contain a central DLS, loops close to the middle of each monomer component, and poly(A) ends attached to SV40-(dT). There are some incomplete dimers consisting of a complete monomer joined to a broken monomer. These have a recognizable DLS,

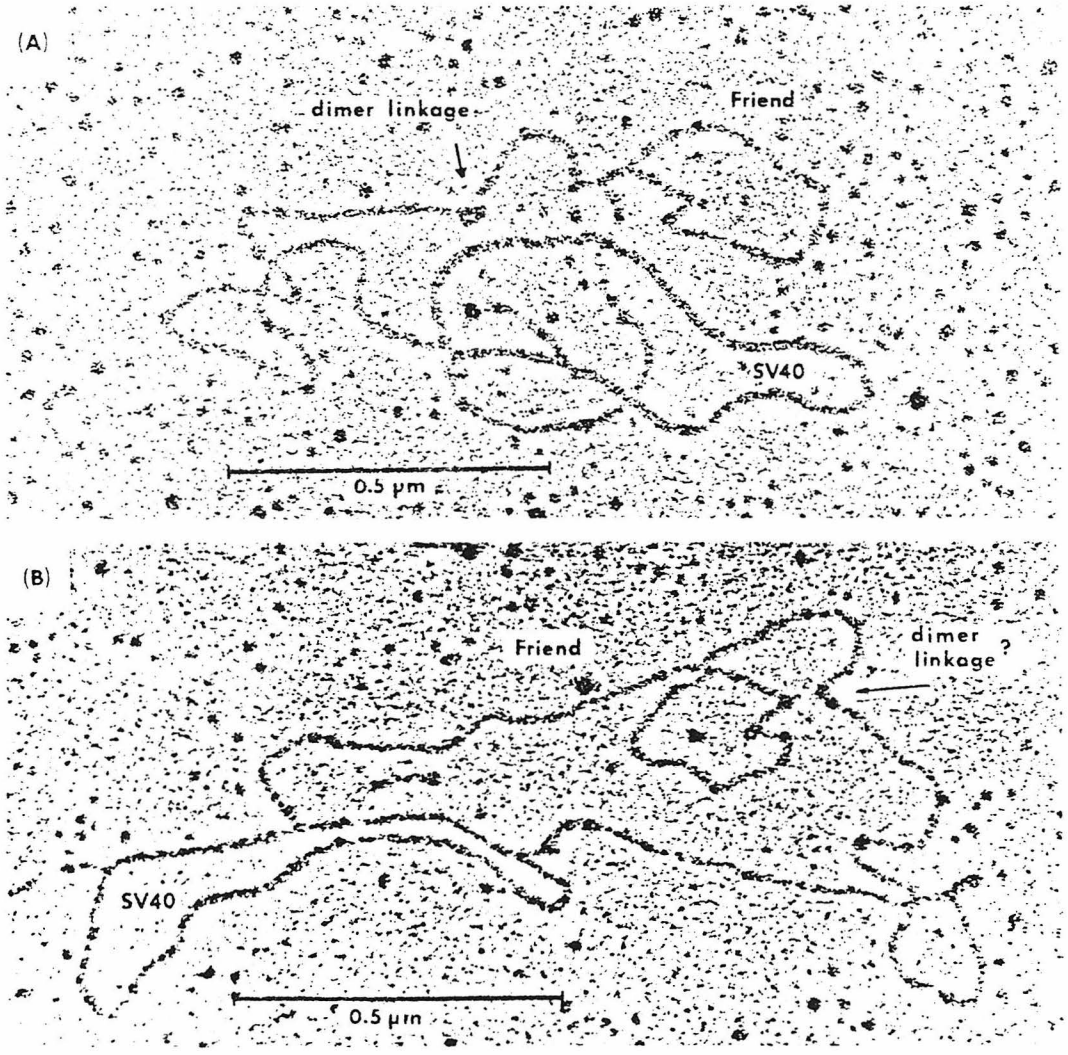



FIG. 4. *Poly(A) mapping on FV dimers. The poly(A) sequences at both outside ends of an FV 50 to 60S RNA dimer anneal to short poly(dT) sequences polymerized onto SV40 circles. Both dimers show the characteristic loops on each monomer half. The DLS in (B) is indistinct, as is true for many FV dimers. All subunits here are of medium size (7 to 8 kb).*

with one long strand with an end attached to SV40(dT), and with one short strand with no SV40(dT). The monomer constituent here and the two monomer constituents of a complete dimer with two attached SV40(dT)'s can be used with confidence for length measurements of full-length monomers. Figure 5 shows a histogram of the length distribution for all complete monomers, either from unbroken dimers or from dimers with one intact monomer unit. As explained in the legend to Fig. 5, the length calibration is based on measurements with BKD viral RNA treated with glyoxal and spread under the same conditions as used for FV RNA.

The distribution in Fig. 5 shows two partially resolved main components with an average molecular length of 6.2 ± 0.2 kb for the smaller (S) component and 7.4 ± 0.5 kb for the medium (M) component. These values are in reasonable agreement with the molecular weights of 6.7 ± 0.6 and 7.7 ± 0.6 kb for the S and M components measured by gel electrophoresis under denaturing conditions.

There are a few monomer units with molecular lengths in the broad range from 8.5 to 10.5 kb in Fig. 5. This range more or less agrees with the molecular length of the large (L) component in the gel electrophoresis experiment of 9.5 ± 0.6 kb. However, the molecules in the fraction labeled L in Fig. 5 have a peculiar skewed distribution. In view of the small sample of molecules, we cannot draw conclusions as to whether there is one or several size classes in this L component or whether the electron microscope and gel electrophoresis data are really in good agreement.

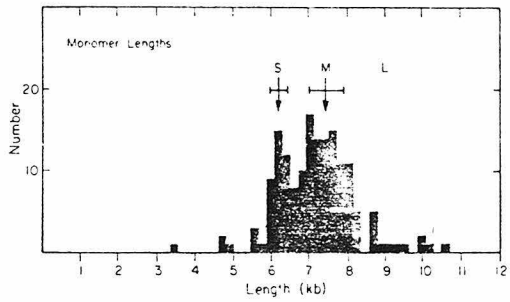



FIG. 5. *Distribution of monomer subunit lengths. All unbroken monomers that had their 3' poly(A) ends attached to SV40-poly(dT) and their 5' ends in a clearly recognizable DLS were measured, and their total contour lengths were plotted. Regions of the histogram are designated S, M, and L, corresponding to the small, medium, and large subunits seen on denaturing gels. The average sizes (arrow) and standard deviations (brackets) are indicated for the small (6.2 ± 0.2 kb) and medium subunit regions (7.4 ± 0.5 kb). For length calibration, E. coli 23S ribosomal RNA, treated by the standard glyoxal procedure, was spread and measured. The FV RNA for this figure was glyoxal treated in the presence of 50 mM NaCl. BKD viral RNA, after treatment with glyoxal in this salt medium, measured 4.4% shorter than after the standard glyoxal treatment, and this correction factor was included. This leads to the calibration that 1 μm = 4.1 kb.*

Given that there are S, M, and L monomer units, we may ask whether the dimers found are all homodimers SS, MM, or LL or whether the heterodimers, SM, SL, and ML are also present. To study this question, we plot in Fig. 6 the length of the long arm of each dimer versus the length of the short arm. Due solely to statistical fluctuations, the measured lengths of the two monomer arms of any one homodimer will, in general, be different. The nature of the plot in Fig. 6 is such that all points must lie below the straight line of slope 1.00 indicated on the figure. The figure also has a dotted line with a slope of 0.89 going through the origin. The usual standard deviation in length for a homogeneous RNA component is about 8%. As explained in the figure legend, if only homodimers are present, we would expect most of the points to lie within a vertical distance of about 1.4 standard deviations in length for a homogeneous component below the slope 1.00 line, that is, within the sector between the lines of slope 1.00 and 0.89.

Suppose the relative frequencies of the S, M, and L components are s, m, and l. Suppose dimers are formed by random association. Then the relative numbers of the several components would be $s^2$, $m^2$, $l^2$, 2 sm, 2 sl, and 2 ml. The random association model would predict that there would be a considerable density of points below the sector defined by the two straight lines of slopes 1.0 and 0.89. However, there are relatively few such points in Fig. 6.

To give a specific example, suppose there were the number of SM heterodimers expected for random assortment. There are about 30 points in Fig. 6 in the x interval of 6 to 7 kb within the two lines and about 30 points in the x interval of 7 to 8 kb. Then there should be about 60 points clustered in a Gaussian distribution around the coordinates x = 7.4 kb, y = 6.2 kb. Similarly, there are perhaps 6 LL dimers in Fig. 6. For random assortment, there ought to be about 27 SL and 27 ML dimers clustered around the points x = 9.5 kb, y = 6.2 kb and x = 9.5 kb, y = 7.4 kb. These three expected heterodimer positions are marked in the figure. It is clear that in no case is there this predicted frequency of heterodimers. In fact,

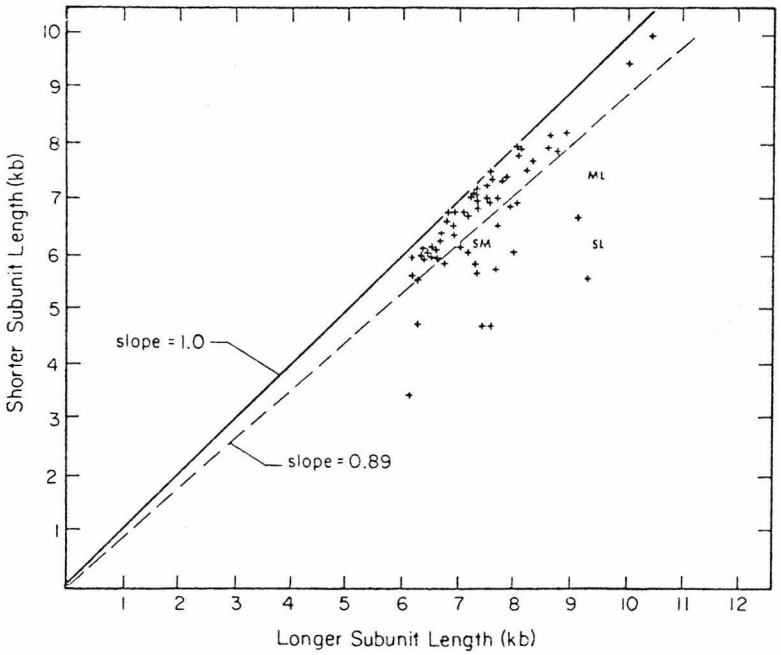



FIG. 6. *Correlation of the two monomer lengths in dimer molecules. All complete dimer RNA molecules with both poly(A) monomer ends joined to SV40-poly(dT) and with clear DLS were measured and plotted. Each cross represents one dimer molecule; its x and y coordinates are the contour lengths of the longer and shorter subunits of the dimer, respectively. It can be shown that, if dimers consists of subunits of identical length and if the standard deviation for the subunit lengths is σ, the distribution curve of the difference (x − y) is: f[(x − y)] = A exp [−(x − y)²/4σ²]. Thus, the standard deviation (σ') of the difference distribution curve is (2σ²)^1/2 = 1.4 σ. The standard deviation in length for tumor virus RNAs containing only a single monomer component is about 8% (9). If both halves of Friend dimers are always identical, and if the only length difference is due to measurement error, 68% of the dimers should have monomers differing by less than 1.4 × 8% = 11%, and these molecules will be plotted above the line with a slope of 0.89. If heterodimers are formed, they should be found centered at the positions marked SM, SL, and ML.*

the number of points below the 0.89 line is approximately the number expected due to statistical variations in length, assuming no heterodimers. Furthermore, a few of the outside points may be artifacts due, for example, to a broken monomer strand accidentally lying on an SV40 label. We can positively conclude that the frequency of heterodimers is much less than expected for random assortment of monomers, but we cannot decide whether heterodimers are completely absent or are present at a low, but nonzero, frequency.

Figure 7 and its legend summarize the appearance and length data for the several secondary structure features found in FV RNA. The DLS is often hard to identify; at best it looks as shown in Fig. 7d; at worst it is not discernible at all. The loops frequently have multiple crossover points (Fig. 7b and c), suggesting that there are multiple junctions therein. The histograms of loop length and of distance from the 5' end or from the 3' end (not shown) are broader than observed for RD-114, BKD, and WoMV, perhaps because different members of the multiple junction points are open or closed for different molecules. If there are multiple-loop junctions, some of which are open and some of which are closed in any one molecule, then the position of the center of the loop relative to an end of the molecule should be more constant than the loop length. This is indeed the case. The average distance from poly(A) end to the center of the loop for the several size classes of RNA are tabulated in the legend to the figure.

## DISCUSSION

The first and most basic point is that RNA from the murine virus, FV, shows the same dimer structure as seen previously in BKD, RD-114, and WoMV RNAs. This observation strengthens the inference that the dimer structure is a general characteristic of the RNAs of all mammalian type C viruses.

The three general features of the dimer structure are: the DLS joining the two monomer

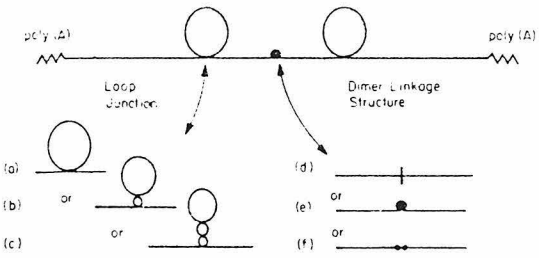



FIG. 7. *Diagram of the structure of FV RNA dimers. (a), (b), and (c) show alternate appearances of the large loop, and (d), (e), and (f) show different morphologies seen at the dimer linkage. For all monomer subunits with single loop junctions as shown in (a), the distance from poly(A) end to loop junction is 2.8 kb, the loop circumference is 2.7 kb, and the loop junction to dimer linkage distance is 1.5 kb. All these distances are about 20% longer in medium (between 6.9 and 8.4 kb) subunits than in small (between 5.9 and 6.5 kb) subunits, and the standard deviations as percentages for all these measurements are about 25%. The distance from poly(A) end to the midpoint of the loop, measured for all molecules with either single or multiple loop junctions, is 4.2 kb ± 17%. This distance for medium subunits only is 4.4 kb ± 10% and for small subunits is 3.6 kb ± 5%.*

units close to their 5′ ends, a loop feature close to the center of each monomer subunit, and a poly(A) segment at the two outside ends. Since each of these features occurs quite generally for type C viral RNAs, each has probably evolved for a biological function common to all type C viruses.

Different RNAs differ in the stability of the dimer linkage. In the case of FV, we observe that the midpoint of the transition occurs at about 25°C in 66% (U+F), 0.03 M cations. We estimate that this corresponds to a melting temperature of 60°C in aqueous 0.1 M $Na^+$. The corresponding melting temperatures of BKD, RD-114, and WoMV RNAs are 87, 87, and 66°C, respectively.

[The conversion of melting temperature in (U+F) solvents to that in aqueous solution with 0.1 M $Na^+$ is based on the following equation: $\Delta T_m$ = 0.4 × percent (U+F) + 16.6°C log ([cations]), where $T_m$ is the melting temperature. The predicted effect of cation concentration in the above equation is taken from Schildkraut and Lifson (15) and is probably fairly accurate. The estimated effect of (U+F) concentration is based on very limited data. Friedrich and Feix (7) show that the melting temperature of the double-stranded replicative form of MS2 RNA is depressed by 0.3 to 0.5°C per percentage of formamide. The melting point depression of DNA by (U+F) and by pure formamide are approximately the same per volume percent solvent (8). If we assume the same is true for RNA duplexes, we arrive at the equation above.]

These predicted melting temperatures in aqueous 0.1 M $Na^+$ from the equation may not be very accurate. However, the qualitative fact that the FV dimer is much less stable than that of RD-114 and BKD RNAs and slightly less stable than WoMV dimers is quite clear.

The dimer molecule also contains two loops, symmetrically disposed with respect to the DLS. The loops are also seen in monomers spread under moderately denaturing conditions (Fig. 3). Roughly speaking, these loops are of comparable stability in RD-114, BKD, WoMV, and FV RNAs.

We observe that the RNA from viruses derived from FSD-3 cells contains three monomer size classes, L, M, and S, of approximate molecular weights of 3.3 × $10^6$, 2.7 × $10^6$, and 2.3 × $10^6$. The M and S components are present in roughly equal amounts and represent about 80 to 90% of FV RNA. The electron microscope data on complete dimers show that either heterodimers are not present or present at a frequency much less than expected for a random assortment model.

By comparison of the RNA gel patterns of FV with those of other murine sarcoma-leukemia virus complex, Maisel et al. (11) have suggested that the M and L components may be assigned to the spleen focus-forming virus and lymphatic leukemia virus activities, respectively. The L and M subunit classes in relative amounts similar to that observed here have been consistently found in many other FV-transformed cell lines (S. Dube and W. Ostertag, private communication). The S component, however, has not been described previously. The induction of an endogenous virus in FV-transformed cells has been described (5, 14). Perhaps S RNA is from the endogenous virus. Alternatively, the S component may be a deletion mutant of the M component. Nucleotide sequence studies on resolved bands are underway to determine if the S and M components are related.

## ACKNOWLEDGMENTS

S.D. has been supported as a Sherman Fairchild Distinguished Scholar at the California Institute of Technology. W.B. has been the recipient of a National Science Foundation fellowship and of a training grant from the National Institutes of Health. This research has been supported by a contract, NO-1-CP-43306 within the Virus-Cancer Program of the National Cancer Institute.

## LITERATURE CITED

1. **Axelrad, A. A., and R. A. Steeves.** 1964. Assay for friend leukemia virus: rapid quantitative method based on enumeration of macroscopic spleen foci in mice. Virology 24:513-518.

2. **Bailey, J. M., and N. Davidson.** 1976. Methylmercury as a reversible denaturing agent for agarose gel electrophoresis. Anal. Biochem. 70:75–85.
3. **Bender, W., and N. Davidson.** 1976. Mapping of poly(A) sequences in the electron microscope reveals unusual structures of type C oncornavirus RNA molecules. Cell 7:595–607.
4. **Dube, S. K., G. Gaedicke, N. Kluge, B. J. Weimann, H. Melderis, G. Steinheider, T. Crozier, H. Beckmann, and W. Ostertag.** 1973. Hemoglobin-synthesizing mouse and human erythroleukemic cell lines as model systems for the study of differentiation and control of gene expression, p. 99–135. *In* W. Nakahara, T. Ono, T. Sugimuna, and H. Sugano (ed.), Differentiation and control of malignancy of tumor cells. University of Tokyo Press, Tokyo.
5. **Dube, S. K., I. B. Pragnell, N. Kluge, G. Gaedicke, G. Steinheider, and W. Ostertag.** 1975. Induction of endogenous and of spleen focus-forming viruses during dimethylsulfoxide-induced differentiation of mouse erythroleukemia cells transformed by spleen focus-forming virus. Proc. Natl. Acad. Sci. U.S.A. 72:1863–1867.
6. **Friedrich, R., and G. Feix.** 1972. RNA-RNA hybridization in aqueous solutions containing formamide. Anal. Biochem. 50:467–471.
7. **Friend, C.** 1957. Cell-free transmission in adult swiss mice of a disease having the character of a leukemia. J. Exp. Med. 105:307–318.
8. **Kung, H. J., J. M. Bailey, N. Davidson, M. O. Nicolson, and R. M. McAllister.** 1975. Structure, subunit composition, and molecular weight of RD-144 RNA. J. Virol. 16:397–411.
9. **Kung, H. J., S. Hu, W. Bender, J. M. Bailey, N. Davidson, M. O. Nicolson, and R. M. McAllister.** 1976. RD-114, baboon, and woolly monkey viral RNAs compared in size and structure. Cell 7:609–620.
10. **Lilly, F., and T. Pincus.** 1973. Genetic control of murine viral leukemogenesis. Adv. Cancer Res. 17:231–277.
11. **Maisel, J., V. Klement, M. M-C. Lai, W. Ostertag, and P. Duesberg.** 1973. Ribonucleic acid components of murine sarcoma and leukemia viruses. Proc. Natl. Acad. Sci. U.S.A. 70:3536–3540.
12. **Ostertag, W., T. Cole, T. Crozier, G. Gaedicke, J. Kind, N. Kluge, J. C. Krieg, G. Roesler, G. Steinheider, B. J. Weimann, and S. K. Dube.** 1973. Viral involvement in the differentiation of erythroleukemic mouse and human cells, p. 493–520. *In* W. Nakahara, T. Ono, T. Sugimuna, and H. Sugano (ed.), Differentiation and control of malignancy of tumor cells. University of Tokyo Press, Tokyo.
13. **Ostertag, W., H. Melderis, G. Steinheider, N. Kluge, and S. K. Dube.** 1972. Synthesis of mouse haemoglobin and globin mRNA in leukaemic cell cultures. Nature (London) New Biol. 239:231–234.
14. **Ostertag, W., G. Roesler, C. J. Krieg, J. Kind, T. Cole, T. Crozier, G. Gaedicke, G. Steinheider, N. Kluge, and S. K. Dube.** 1974. Induction of endogenous virus and of thymidine kinase by bromodeoxyuridine in cell cultures transformed by friend virus. Proc. Natl. Acad. Sci. U.S.A. 71:4980–4985.
15. **Schildkraut, C., and S. Lifson.** 1965. Dependence of the melting temperature of DNA on salt concentration. Biopolymers 3:195–208.
16. **Steeves, R. A., R. J. Eckner, M. Bennett, E. A. Mirand, and P. J. Trudel.** 1971. Isolation and characterization of a lymphatic leukemia virus in the friend virus complex. J. Natl. Cancer Inst. U.S.A. 46:1209–1217.

# Structure of 50-70S RNA from Moloney Sarcoma Viruses

by

Jan Maisel[1], Welcome Bender[2], Sylvia Hu[3], P.H. Duesberg,

and Norman Davidson*

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

and

Department of Chemistry
California Institute of Technology
Pasadena, California 91125*

1. Present address, Department of Pediatrics, University of California, San Francisco, California 94143

2. Present address, Department of Biochemistry, Stanford University, School of Medicine, Stanford, California 94305

3. Present address, Department of Microbiology, University of Southern California, School of Medicine, Los Angeles, California 90033

Contribution Number 5625

## ABSTRACT

The 50-70S RNAs of two clonal isolates of defective Moloney sarcoma-leukemia helper virus complex were analyzed by gel electrophoresis and electron microscopy. The RNAs extracted from both clone 3 and clone 124-5R Mo-MSV(Mo-MLV) contained some large monomer subunits, about 10,000 nucleotides in length (10 kb) which are believed to be the Mo-MLV subunits. Both had an excess of a smaller, sarcoma-specific subunit, 5 kb (clone 3) or 6 kb (clone 124-5R) in length. Electron microscopy of intact 50-70S dimer RNA molecules showed, for both clones, many dimers of two small subunits, some dimers of two large subunits, but few if any heterodimers with one large and one small subunit. This result is unexpected because the sequences near the 5' end of the RNA subunits, which are believed to be involved in the dimer linkage, are probably homologous between the large and small subunits. We also observed that some small-small dimers migrate anomalously slowly on nondenaturing gels. The nature of this slow-migrating complex is unknown; it could be a higher aggregate of the small-small dimer with additional small or large subunits, or it could be an extended conformation of the small-small dimer.

## INTRODUCTION

Cloned RNA tumor viruses contain a 60-70S RNA complex consisting of two very similar or identical 30-40S subunits (4, 5, 8, 22). In several mammalian type C RNA tumor viruses, the two 30-40S subunits are joined at their 5' ends (6, 18, 12, 17). The function of the dimer molecules is not understood but it has been proposed that viral heterozygotes, perhaps containing heterodimer RNA complexes consisting of two genetically distinct subunits, are obligate intermediates in recombination (28, 31, 4). However, such heterodimer RNAs have thus far not been detected in virus released by cells infected with two or more different viruses.

Linkage of RNA subunits in a 50-70S RNA complex is thought to result from base-pairing (13, 14, 23) and possibly involves a short self-complementary sequence (18) near the 5' end of the RNA (6). Accordingly heterodimer formation might be limited to closely related virus strains whose RNAs share nucleotide sequences involved in dimer formation.

In the Friend virus complex which has several different RNA components (12, 20, 26), it was observed that heterodimers were not common, but the length differences among the different RNA subunits are not much greater than the expected variation in length measurement for a homogenous species, and the possibility of rare heterodimers could not be excluded.

In this study, the RNA of two Moloney murine sarcoma-leukemia virus isolates (designated Mo-MSV(Mo-MLV) or here simply MSV) are examined for heterodimers. The system is advantageous for two reasons. (i) The MSV virus complex contains two classes of RNA subunits of significantly different size, a 10 kb leukemia

virus-specific RNA, termed the L subunit, and a 5-6 kb sarcoma virus-specific RNA, termed the S subunit (20, 5, 19). LS heterodimers should be readily distinguished from LL and SS homodimers; especially since dimer molecules containing broken subunits can be excluded visually in the electron microscope assay by labeling the 3' poly(A)-ends of intact RNAs (6). (ii) The sarcoma virus-specific S subunit shares extensive sequence homology with the L subunit of Moloney leukemia virus (Mo-MLV or here simply MLV) (7, 24, 10). Moreover, mapping of the homology between the S subunits of a related strain (clone 124) of MSV and L subunits of MLV by electron microscopy of heteroduplex molecules (17) and by biochemical methods (Beemon, Dina and Duesberg, unpublished) has indicated that the two molecules share sequences near their 5' ends, where the dimer linkage appears to be located.

We report here that, when examined by electron microscopy, the 50-70S RNAs of clone 3 and clone 124-5R MSV are found to contain both SS and LL homodimers, but that there are no detectable LS heterodimers. However, gel electrophoresis of viral RNA under non-denaturing conditions shows that while there are LL dimers and SS dimers with mobilities as expected for their respective lengths there is also another kind of complex which has an electrophoretic mobility close to that of LL dimers. The structure of this third component is not known, but it includes SS dimers.

## MATERIALS AND METHODS

Cells and viruses. MSV clone 3 (20) was isolated by infection of rat NRK cells with a tissue culture-derived stock (HT-1) of

MSV virions obtained from A. Hackett (9), followed by cloning in microtiter wells (20). A line of TB mouse cells infected with MSV clone 124 (3) was a gift of J.K. Ball; this virus was derived from an MSV stock provided by J.B. Moloney. Virus from these cells was plated at low multiplicity on rat NRK cells, a nonproducing focus was isolated and superinfected with MLV, and these cells produced virus (MSV, clone 124-5R) containing RNA subunits of two sizes equivalent to those of MLV and MSV clone 124 (19). 3T3 (NIH) mouse cells infected by MLV were obtained from D. Smotkin. All cells were grown in Eagle's minimal essential medium (MEM) containing 10% fetal bovine serum. 0.5 µg/ml of Fungizone, 100 U/ml of penicillin and 50 µg/ml of streptomycin.

The following observations show that the RNAs used are not substantially contaminated with other RNA components. A cDNA probe prepared from clone 1 MLV RNA protects 85% of clone 1 RNA, and 85% and 58% of gel purified L and S components of MSV clone 3 respectively in hybridization and ribonuclease digestion experiments (L. Evans, P. Duesberg; personal communication). The MLV RNA used has the expected complexity and fingerprint (5). The complexity of the S component of MSV clone 3 was determined by fingerprint analysis and gave the value expected for its molecular length (J. Maisel, personal communication). The S component of MSV clone 124-5R shows the same heteroduplex pattern with MLV clone 1 cDNA as was seen for MSV clone 124 RNA (17), showing that these RNA's are basically identical (Y.H. Chien, personal communication).

Gel electrophoresis. 2.1 or 1.9% polyacrylamide gels were prepared and analyzed as previously described (19).

Agarose gels were prepared in borate (2) or Tris-acetate buffer (20). For denaturing gels, 5 mM methylmercury hydroxide was added to the gel solution (2).

RNA was eluted from polyacrylamide gels as previously described (21), and was recovered from agarose gels by the freeze-squeeze method (25). The RNA recovery by freeze squeeze was lower than by elution (30% vs. approximately 80%), but the RNA could be obtained quickly in a small volume and it was essentially free of agarose. In some experiments, oligo(dT)-cellulose column fractionation (29) was used to remove from electrophoretically-purified RNA preparations any RNA species which had been nicked near the 3' end and therefore no longer contained poly(A).

Electron Microscopy. RNA samples were treated with glyoxal, dialyzed, incubated briefly with SV40-poly(dT) or SV40-poly(dBrU) and spread by the formamide cytochrome c method as described (6). (We now use SV40-poly(dBrU) instead of SV40-poly(dT) as a poly(A) label, because of the greater stability of poly(dBrU)·poly(A) duplexes.) The standard glyoxal treatment (1 M glyoxal, 0.01 M potassium phosphate, pH 7.0 37°, 1h) dissociated most of the dimers of these murine viruses. Therefore, less harsh glyoxal treatments (incubations for only 20' or in the presence of 0.1 M NaCl) were used in order to preserve the dimer linkage. The SV40 circles were used as an internal length standard and the sizes of RNA molecules were calibrated by comparison with E. coli 23 S ribosomal RNA (1μm = 4.1 kb).

## RESULTS

Gel Electrophoresis of Clone 3 MSV RNA. RNA extracted from clone 3 MSV virus contains two sizes of subunits. When 50-70S viral RNA is heat denatured and analyzed by electrophoresis on a 2% polyacrylamide gel, two well-separated peaks are seen (20). The larger species comigrates with heat denatured MLV and is identified as the leukemia (L) subunit. The smaller (S) component is presumed to be the monomer unit of the genome of the defective sarcoma virus. Molecular lengths of 10.1 $\pm$ 0.5 kb and 5.1 $\pm$ 0.3 kb for the L and S subunits were estimated by agarose gel electrophoresis in the presence of the strong denaturing agent $CH_3HgOH$ (2) (Fig. 1).

MSV 50-70S RNA which has not been denatured gives a single broad peak when sedimented in a sucrose gradient (9, 20). However the high molecular weight RNA can be resolved into two components, I and II, when run on a nondenaturing agarose or acrylamide gel (Fig. 2a). 50-70S RNA from MLV has an electrophoretic mobility on these gels nearly identical to that of the slower migrating MSV component I. RNA from the two peaks of the nondenaturing gels was extracted, heat denatured and applied to 2% acrylamide gels. The faster migrating peak II gave after denaturation only S monomers (Fig. 2b), but the slower peak I contained subunits of both the L and S sizes (Fig. 2c).

The results on the composition of the two peaks was confirmed by cutting peak I and peak II slices from a non-denaturing agarose gel, polymerizing them directly onto a denaturing $CH_3HgOH$ gel, and electrophoresing into the second gel. By this technique, over 85% of the material in the slices from the first gel is recovered

in the second gel, thus showing that the results in Fig. 2b and c are not distorted because of possible inefficient recovery from the non-denaturing gel of Fig. 2a.

Electron Microscopy of Clone 3 MSV RNA. The structures of the several components present in clone 3 MSV RNA were studied by electron microscopy. Because peak I had both L and S subunits, we were particularly interested in the question of whether any LS heterodimers existed.

Viral RNA was glyoxal treated and spread from formamide as previously described (18). SV40 double stranded circular DNA with poly(dT) tails was hybridized to the viral RNA prior to spreading in order to label the 3' poly(A) ends (6). The dimers of MSV were mostly dissociated by the standard glyoxal treatment (37°, 1h, 1.0 M. glyoxal, 0.01 M potassium phosphate, pH 7.0), therefore, 0.1 M NaCl was added to the solution to stabilize the dimer linkage structure during glyoxal treatment.

As already reported (17), the principal structure found in 50-70S RNA from MLV alone is the LL dimer. This dimer has the same general structural features previously observed for other mammalian viral RNAs (18, 12).

In contrast, the 50-70S RNA from MSV yielded two kinds of structures. The less abundant component had exactly the same secondary structure features as described previously for MLV RNA; such an LL dimer hybridized to SV40 poly(dT) is shown in Figure 3a. The more abundant molecule, which must represent the defective-sarcoma component, is shown in Figure 3b. Like the MLV RNA dimers, the SS dimer molecules displayed two-fold symmetry with a central dimer linkage structure. The SS dimers are only about half as

long as LL dimers. The secondary structure loops in the center of the S monomer units are also quite different. These loops are small figure-eight structures often with a short hairpin near the midpoint of the loop. Schematic representations of the structures of both components and the contour lengths of the several features are given in Fig. 4.

In the preparations from 50-70S RNA of MSV, we saw only LL and SS dimers but no LS heterodimers. In order to study this question critically, we photographed and measured about 50 randomly selected dimers which had clear dimer linkage structures and which had two unbroken subunits both ending in a poly(A) sequence hybridized to an SV40-poly(dT) label. Fig. 5 is a plot of the length of the shorter subunit of each dimer vs. the length of the longer subunit. The analysis of this kind of a plot has been discussed previously (12). A heterodimer molecule would be represented by a point that deviated significantly from the 45° line in the plot. The vast majority of molecules in Figure 5 are LL or SS dimers with subunits of almost equal length. There are a few molecules with one unusually short subunit; these presumably are broken subunits whose ends accidentally lie close to an SV40 circle, but there are none with the dimensions expected of a heterodimer.

In subsequent studies, we observed over 300 clone 3 MSV dimers; of these, three molecules had the structure expected for an LS heterodimer, but only one (Fig. 3c) was attached at both outside ends to SV40-poly(dT), thus showing that it is unbroken. No

certain conclusions can be drawn from a single rare molecule; for example, the molecule could possibly be a broken or partially collapsed LL dimer. However this molecule does appear to be a true heterodimer both by the contour lengths of the two subunits and by the positions and shapes of the two loops.

It seemed possible that heterodimers might exist but have a much lower stability than homodimers, and thus be disrupted by our spreading conditions and not be scored. We therefore separated 50-70S MSV RNA on a nondenaturing agarose gel as in Fig. 2a, isolated RNA from each peak, treated it with glyoxal in high salt, and then mounted it for microscopy. Samples from the fast-migrating peak II gave sarcoma dimers (SS) and sarcoma monomers (S) with occasionally broken MLV fragments observed. Extensive searching of the sample from the slow-migrating peak I yielded four types of molecules: LL-dimer, L-monomer, SS-dimer and S-monomer. The ratios obtained from two independent

Table 1

experiments are listed in Table 1. These data confirm that there are significant amounts of SS and S components migrating anomalously in the slow peak. If unstable LS heterodimers do exist, the number of L and S monomers seen here requires that these heterodimers are a minor component. Finally, we wish to note that there are a certain number of tangled uninterpretable molecules on grids prepared by the gentle glyoxal treatments used here, so that the numbers in Table 1 may not be strictly representative.

Gel Electrophoresis of Clone 124 MSV. In an attempt to resolve the question whether the S subunits of MSV can form, in addition to homodimers, larger complexes consisting of more than two S subunits or whether both S and L subunits are necessary to form such large complexes, the RNA of an MSV variant, namely clone 124, was investigated. The RNA of MSV clone 124, propagated in a mouse cell line, differs from that of MSV clone 3 in that it contains a large excess of S over L subunits, at an approximate mass ratio of 30:1 (19). Hence, if MSV clone 124 S subunits are able to form multimeric complexes as well as homodimers, the 50-70S RNA would consist of two electrophoretic species. If both L and S subunits are necessary to form complexes larger than S subunit homodimers, the 50-70S RNA of MSV clone 124 would consist of only one major 50-70S RNA species, namely SS homodimers, because insufficient L subunits are present in this virus stock to permit formation of larger complexes. It is shown in Fig. 6A that the 50-70S RNA of clone 124 MSV consists essentially of only one major species, which migrates faster than 60-70S MLV RNA and is similar to peak II of MSV clone 3 RNA. After heat-dissociation the RNA was converted predominantly into a single S subunit species, which migrated faster than the L subunits but a little slower than the S subunits of MSV clone 3. This electrophoretic difference between the S subunits of the two strains of MSV has been described previously (19, Fig. 4).

Fig.6a

Clone 124-5R MSV RNA. A possible interpretation of the above result is that the SS dimers of MSV clone 124 do not form any slowly migrating complexes because L subunits are necessary for

this phenomenon and are present in insufficient amount. We have therefore studied the RNA from a cell line, 124-5R, derived from 124; it releases a larger amount of the L component than does the parental cell line.

On a non-denaturing gel, the 50-70S RNA of this virus complex
Fig.7 was resolved into two peaks, I and II (Fig. 7a). After heat-dissociation, 50-70S clone 124-5R MSV RNA was resolved into two subunit species. In accord with a previous study (19) the larger L subunit comigrated with the L subunit of clone 3 MSV and the smaller S subunit migrated more slowly than the S subunit of MSV clone 3 (Fig. 7b). The RNA was also sized on 1% agarose gels in the presence of 5 mM $CH_3HgOH$; the L subunit ran as 9.7 $\pm$ 0.5 kb and the S subunit was 6.1 $\pm$ 0.3 kb (data not shown). As with clone 3 MSV, gel electrophoretic analysis of heat dissociated peak I RNA reveals both L and S subunits (Fig. 7c). The signal to noise ratio for these experiments is not as good as for the experiments with clone 3 MSV RNA in Figs. 2 and 3, because the ratio of peak I to peak II RNA is less and the separation between the two peaks is not as great; nevertheless, the results support the view that there are S as well as L subunits in peak I.

For electron microscope studies, 50-70S RNA from clone 124-5R MSV was glyoxal treated (in this case in 1 M glyoxal, 0.01 M $PO_4$, pH 7.0, for 20 min), hybridized with SV40-poly(dBrU), and spread from 50% formamide. As with the clone 3 RNA, there were dimers with two large (9.5 kb) subunits and dimers with two smaller (6 kb) subunits. These dimers are diagramed in Figure 4. The LL dimers

of clone 124-5R are indistinguishable from those of MLV or clone 3 MSV; an example is shown in figure 3d. The SS dimers, as shown in figure 3e are identical in size and secondary structure to those of clone 3 MSV, except that the distance from the poly(A) end to the figure eight loop is 2.7 kb in clone 124-5R instead of 1.8 kb for clone 3. In the region between the poly(A) and the loop on clone 124-5R S subunits there is often a small loop or hairpin about 0.6 kb in length and about 1.5 kb from the poly(A) end; it is conceivable that the clone 124-5R and clone 3 S subunits are identical except for an insertion or deletion of this 0.6 kb sequence.

As with clone 3 MSV, we did not find LS heterodimers in clone 124-5R. About 100 randomly selected dimers were photographed and measured, taking only molecules with a clear dimer linkage and both ends attached to SV40-poly(dBrU). The relative lengths of the two monomer units of each dimer pair is plotted in figure 5b. There are no molecules with the dimensions expected for heterodimers.

Further Experiments. It was not clear whether the SS dimers found in peak I of clone 3 MSV were there because of variation in conformation, because of nonspecific aggregation, or because of specific large complexes such as LSS trimers or $S_4$ tetramers. Several experiments were done to test these possibilities.

It seemed possible that S subunits artificially aggregated with MLV RNA after the RNAs had been extracted from the respective virions. To test this possibility MSV clone 124 culture fluid ([$^{32}$P] labeled) was mixed with MLV culture fluid ([$^{3}$H] uridine-

labeled), and their RNAs were extracted together. It can be seen in figure 6b that the 50-70S MLV [$^3$H] RNA formed a distinct more slowly migrating peak. The small amount of [$^{32}$P] RNA, which migrated with the slower [$^3$H] MLV RNA, was also present in clone 124 RNA extracted in the absence of MLV (Fig. 6a) and is thought to be helper MLV RNA present as a minor component (10, 19). We conclude that simultaneous extraction of MLV and MSV clone 124 RNA did not generate the type of hybrid RNA complex seen in clone 3 or clone 124-5R MSV. Therefore, the presence of S subunits in peak I in clone 3 and 124-5R RNAs is not an artifact of the RNA extraction and handling.

In an attempt to resolve possibly heterogeneous RNA complexes present in peak I, we subjected clone 3 MSV RNA to electrophoresis in agarose gels under partially denaturing conditions. It has been shown that 60-70S Rous sarcoma virus RNA is dissociated into 30-40S subunits at room temperature by 8 M urea at low ionic strength (11). 50-70S clone 3 RNA was not dissociated into subunits when electrophoresed in agarose gels containing up to 6 M urea. In the presence of 6 M urea, peak I appeared slightly broader than in the absence of urea, but RNA eluted from both sides of the 6 M urea peak I still contained S subunits (data not shown).

The stabilities of the several components of clone 3 MSV RNA were also tested by thermal melting. RNA was extracted from peaks I and II of a nondenaturing agarose gel; samples were

dialyzed against 0.1 M NaCl, 0.01 M Tris, 0.001 M EDTA, pH 8.5; the dialyzed samples were heated to various temperatures for two minutes; and then the samples were rerun on nondenaturing agarose gels. MLV dimers and the SS dimers of peak II both melt into the respective monomer units in the temperature interval 55-65°C. Peak I melts into subunits in the same temperature range, and there is no conversion of peak I material to any faster migrating components by heating at 50° (data not shown).

A Secondary Structure Feature in L Monomers. In the course of the electron microscopic examination we noted an unusual secondary structure feature of the L subunit RNA produced by heating and quenching. 50-70S RNA of clone 3 MSV was dissociated by heating (58% formamide, 0.006 M EDTA, 0.06 M Tris, pH 8.5 at 40° C for 1 min) and quenched on ice. Samples were mounted from 55% formamide solution onto an 18% formamide hypophase. After this treatment, the RNA was all of monomer length, and the subunits all lacked the usual loops. The S monomers appeared as simple linear molecules. However, the L monomers displayed a new loop and stem structure of contour length 1.3 ± 0.1 kb beginning 0.9 ± 0.1 kb from the 5' end (Fig. 8). The large loops typical of the L subunits in the dimer never reformed. In rare molecules, the dimer linkage is intact, with two of the new hairpins on either side. This shows that the hairpin feature is close to the 5' end of the L monomer. The large loop in the native molecule begins at 1.6 kb from the 5' end. The sequence to which it is evidently paired is at 5.9 kb. The RNA sequences of the secondary structure feature seen in heated and quenched RNA map 0.9 to 2.2 kb from the 5' end. Thus the loop structure and the hairpin in the

Fig. 8

monomer may be alternative base-paired structures for this particular region. This structural transition also occurs when other mammalian RNA tumor viruses such as RD-114 and baboon endogenous virus 50-70S RNAs are dissociated and rapidly quenched (19).

## DISCUSSION

The experiments described here indicate that the 50-70S RNA of MSV clone 3 and 124-5R consist of two classes of RNA complexes, i.e. sarcoma virus (SS) and leukemia virus (LL) specific homodimers. Sarcoma-leukemia heterodimers are absent or comprise less than 1% of the dimer complexes, as assayed by direct visualization in the electron microscope; whereas, on the basis of random assortment, one would expect for the sample of molecules in Fig. 5a (ratio of S to L subunits of 3:1) that the relative fractions of SS, LS, and LL dimers would be 0.56, 0.38 and 0.06, respectively.

The signal to noise ratios in the various gel electrophoresis experiments are such that we cannot rigorously exclude the possibility that there is a modest fraction of LS heterodimers in the sample, but that these are dissociated much more readily than LL and SS dimers and are not seen under the partially denaturing conditions used for electron microscopy. However, gel electrophoresis experiments did not detect any complexes significantly less stable than LL and SS dimers. As discussed in more detail

below, L and S monomer units probably have the same sequences at their 5' ends, so that LL, SS, and LS dimers should all involve the same base pairs and therefore have the same stability.

Our other principal finding is contradictory to the simple view that the sample of 50-70S MSV virion RNA consists only of SS and LL dimers. The gel electrophoresis studies show that under non-denaturing and mildly denaturing conditions, there are LL and SS dimers with mobilities as expected for their respective lengths; however there is also another kind of a complex which has an electrophoretic mobility close to that of LL dimers but which contains SS dimers. No evidence for these special structures has been observed by electron microscopy.

The interaction responsible for the slow migration of some SS dimers is not an artifact of the extraction process. It appears to depend on synthesis of both S and L genomes in the same cell. The interaction does not melt out at a temperature substantially lower than the dissociation temperature of SS and LL homodimers.

Several hypotheses may be suggested to explain these results.

a) There is a conformational isomer of the SS dimers which is more extended and more slowly migrating in non-denaturing and even in 6 M urea gels than the standard SS dimers.

b) There is a higher specific aggregate, say an LSS or an $S_4$ tetramer. These molecules form tangled structures which haven't been recognized and analyzed under the conditions used for electron microscopy.

c) There is a general interaction resulting in higher ggregates involving, say, SS and LL dimers. The interaction s slightly weaker than that responsible for dimerization, but he resolution of the urea gel and thermal melting experiments is nsufficient to reveal this.

d) A reviewer has suggested that higher aggregates of iral RNA may be formed when "double budded virions" are released rom cells (32).

We know of no decisive evidence to choose among these several ossibilities, but the ratio of L monomers to SS dimers in experi- ent 1 of Table 1 argues against hypothesis (b). We note that hatever the nature of the slowly migrating S containing species, ts formation may be dependent on the cell type in which the irus grows and may be different, for example, in mouse TB cells clone 124) than in rat NRK cells (clone 3 and 124-5R).

The findings that all known strains of MSV infect with double- it kinetics (15, 1) argue against the existence of heterodimers, ecause virions containing a sarcoma and leukemia virus-specific ubunit in the same RNA complex might be expected to plate with ingle-hit kinetics. Also, no recombination between a mammalian arcoma virus and its helper to give a non-defective sarcoma irus has been observed. This lack of recombination may be due o the absence of heterozygotes, if heterozygote formation is n obligatory step in RNA tumor virus recombination (27, 28, 31). ote however that in the particular case of clone 124 MSV, the ormation of a non-defective sarcoma virus would probably be mpossible by homologous recombination. Near the 3' end of lone 124 MSV, a new RNA sequence of 1.5 kb has been substituted

. the RNA for a 2.9 kb region of the helper. Assuming that the w 1.5 kb sequence is essential for transformation and that the 9 kb region of the helper contains information essential for rus replication , each possible recombinant would have either .e or the other sequence in the substitution region, and thus non-defective sarcoma viruses would be formed (see also 30).

From the appearance of the dimer linkage structure of several ammalian viruses in cytochrome films and from its appearance in ene 32 and other higher resolution spreading methods (S. Hu, . Wu, personal communications), we believe that the base paired equences involved in the dimer linkage structure probably lie omewhere within the first 500 nucleotides from the 5' end. eteroduplex analysis (17) indicates that clone 124 MSV RNA orms a duplex with MLV cDNA throughout the 2.25 kb region starting t the 5' end of the RNA. Analysis of ribonuclease T1 oligonucleo-ides reveals no sequence differences in this region (Beemon, ina and Duesberg, private communication). Heteroduplexes between lone 3 MSV and MLV have not yet been studied but we expect to ee the same homology regions, because both clone 3 and clone 124 re derivatives of J.B. Moloney's original MSV and because the wo types of MSV subunits are so similar in secondary structure. uch a heteroduplex analysis can not detect sequence mismatch elow the level of about 10% nor can a short single stranded region f RNA near the 5' end of the MSV RNA be seen if the MLV cDNA was ctually not quite full length. Therefore, it is probable, but ot absolutely proven, that the S and L subunits have identical r very similar sequences involved in the dimer linkage.

If L and S subunits do have the same sequence at the 5' end, ınd if dimer formation is simply due to base-pairing in this region s proposed by Kung et al. (18) or Haseltine et al. (16), one ould expect to see LS heterodimers with the same stability as he homodimers.

As noted above, failure to form heterodimers may be due to nall differences between L and S subunits in the base pairing equences for dimerization. We think it is more likely that sociations between two RNA subunits in regions other than 5' ds are necessary for dimer formation and/or viral packaging. ternatively, dimer formation may require that two neighboring A chains link during transcription.

## ACKNOWLEDGMENTS

W.B. has been the recipient of a National Science Foundation lowship and of a training grant from the National Institute of eral Medical Sciences. This research has been supported Research Grant CA 11426 from the National Cancer Institute and contracts NO1 CP 61001 and NO1 CP 43306 with the Virus Cancer gram of the National Cancer Institute.

## LITERATURE CITED


1. Aaronson, S.A., and W.P. Rowe. 1970. Non-producer clones of murine sarcoma virus-transformed BALB/373 cells. Virology 42:9-19.
2. Bailey, J.M., and N. Davidson. 1976. Methylmercury as a Reversible Denaturing Agent for Agarose Gel Electrophoresis. Anal. Biochem. 70:75-85.
3. Ball, J.K., J.A. McCarter, and S.M. Sunderland. 1973. Evidence for helper independent murine sarcoma virus. I. Segregation of replication-defective and transformation-defective viruses. Virology 56:268-284.
4. Beemon, K., P. Duesberg, and P.K. Vogt. 1974. Evidence for crossing over between avian tumor viruses based on analysis of viral RNAs. Proc. Nat. Acad. Sci. USA. 71:4254-4258.
5. Beemon, K., A. Faras, J. Haase, P.H. Duesberg, and J. Maisel. 1976. Genomic complexities of murine leukemia and sarcoma, reticuloendotheliosis and visna viruses. J. Virol. 17:525-537.
6. 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.
7. Benveniste, R., and E. Scolnick. 1973. RNA in mammalian sarcoma virus transformed non-producer cells homologous to murine leukemia virus RNA. Virology 51:370-387.
8. Billeter, M.A., J.T. Parsons, and J.M. Coffin. 1974. The nucleotide sequence complexity of avian tumor virus RNA. Proc. Nat. Acad. Sci. USA. 71:3560-3564.
9. Bondurant, M.C., A.J. Hackett, and F.L. Schaffer. 1973. Infectivity and RNA patterns as functions of high- and low-dilution passage of murine sarcoma-leukemia virus: evidence

for autointerference within an oncornavirus population. J. Virol. 11:642-647.

10. Dina, D., K. Beemon, and P.H. Duesberg. 1976. The 30S Moloney sarcoma virus RNA contains leukemia virus nucleotide sequences. Cell 9:299-309.
11. Delius, H., P.H. Duesberg, and W.F. Mangel. 1975. Electron microscope measurements of Rous sarcoma virus RNA. Cold Spring Harbor Symp. Quant. Biol. 39:835-843.
12. Dube, S., H.J. Kung, W. Bender, N. Davidson, and W. Ostertag. 1976. Size, subunit composition, and secondary structure of the Friend virus genome. J. Virol. 20:264-272.
13. Duesberg, P.H. 1968. Physical properties of Rous sarcoma virus RNA. Proc. Nat. Acad. Sci. USA. 68:1511-1518.
14. Erikson, R.L. 1969. Studies on the RNA from avian myeloblastosis virus. Virology 37:454-459.
15. Hartley, J.W., and W.P. Rowe. 1966. Production of altered cell foci in tissue culture by defective Moloney sarcoma virus particles. Proc. Nat. Acad. Sci. USA. 55:780-786.
16. Haseltine, W.A., A.M. Maxam, and W. Gilbert. 1977. Rous sarcoma virus genome is terminally redundant: the 5' sequence. Proc. Nat. Acad. Sci. USA. 74:989-993
17. Hu, S., N. Davidson, and I. Verma. 1977. A heteroduplex study of the sequence relationship between the RNAs of M-MSV and M-MLV. Cell 10:469-477.

18. Kung, H.J., S. Hu, W. Bender, J.M. Bailey, and N. Davidson. 1976. RD-114, baboon, and woolly monkey viral RNAs compared in size and structure. Cell 7:609-620.

19. Maisel, J. D. Dina, and P.H. Duesberg. 1977. Murine sarcoma viruses: The helper-independence reported for a Moloney variant is unconfirmed; Distinct strains differ in size of their RNAs. Virology 76:295-312.

20. Maisel, J., V. Klement, M. M-C. Lai, W. Ostertag, and P. Duesberg. 1973. Ribonucleic acid components of murine sarcoma and leukemia viruses. Proc. Nat. Acad. Sci. USA. 70:3536-3540.

21. Maisel, J., E. Scolnick, and P.H. Duesberg. 1975. Base sequence differences between the RNA components of Harvey sarcoma virus. J. Virol. 16:749-753.

22. Mangel, W.F., H. Delius, and P.H. Duesberg. 1974. Structure and molecular weight of the 60-70S RNA and the 30-40S RNA of the Rous sarcoma virus. Proc. Nat. Acad. Sci. USA. 71:4541-4545.

23. Montagnier, L., A. Golde, and P. Vigier. 1969. A possible subunit structure of Rous Sarcoma Virus RNA. J. Gen. Virol. 4:449-452.

24. Scolnick, E.M., R.S. Howk, A. Anisowicz, P.T. Peebles, C.D. Scher, and W.P. Parks. 1975. Separation of sarcoma virus-specific and leukemia virus-specific genetic sequences of Moloney sarcoma virus. Proc. Nat. Acad. Sci. USA. 72:4650-4654.

25. Thuring, R., J.P.M. Sanders, and P. Borst. 1975. The freeze-squeeze method for recovering long DNA from agarose gels. Anal. Biochem. 66:213-220.

26. Troxler, D.H., J.K. Parks, W.P. Parks, and E.M. Scolnick. 1977. Friend strain of spleen focus-forming virus: A recombinant between mouse type C ecotropic viral sequences and sequences related to xenotropic virus. J. Virol. 22:361-372.

27. Vogt, P.K. 1977. The genetics of RNA tumor viruses. In "Comprehensive Virology" (H. Fraenkel-Conrat and R.R. Wagner, eds.), Plenum Press, New York. (in press).

28. Vogt, P.K., and P.H. Duesberg. 1973. On the mechanism of recombination between avian RNA tumor viruses. Second ICN-UCLA Symposium; In "Virus Research" (W.S. Robinson and F. Fox, eds.,) Academic Press, New York.

29. Wang, L.H. and P. Duesberg. 1974. Properties and locations of poly(A) in Rous sarcoma virus RNA. J. Virol. 14:1515-1529.

30. Wang, L.H., P.H. Duesberg, T. Robins, H. Yakota, and P.K. Vogt. 1977. The terminal oligonucleotides of avian tumor virus RNAs are genetically linked. Virology, (in press).

31. Weiss, R., W. Mason, and P.K. Vogt. 1973. Genetic recombination between endogenous and exogenous avian RNA tumor viruses. Virology 52:535-552.

32. Yuen, P.H., and P.K.Y. Wong. 1977. A morophological study on the ultrastructure and assembly of murine leukemia virus using a temperature-sensitive mutant restricted in assembly. Virology 80, 260-274.

Table 1. Analysis of peak I RNA by electron microscopy*

| Type / No. | LL | L | SS | S | L/S |
|---|---|---|---|---|---|
| Expt. 1 | 114 | 7 | 46 | 14 | 2.24 |
| Expt. 2 | 145 | 231 | 65 | 60 | 2.74 |

*Clone 3 MSV was electrophoresed in a 1% agarose gel, and the peak I gel slices were eluted by the freeze-squeeze method. The recovered RNA was glyoxal treated, (1 M glyoxal, 0.1 M NaCl, 37°, for 40 min in (expt. 1) or 60 min (expt. 2)), and spread for electron microscopy, the grids were then scanned for LL, SS and LS dimers and for L and S monomers. Two separate experiments were performed. Only three presumptive LS dimers were found in the whole search; the numbers of molecules in the other four classes are listed for the separate experiments. In experiment 2, the longer glyoxal treatment disrupted many of the dimers.

Figure 1. Molecular weight determination of viral RNAs by electrophoresis in 1% agarose gels in the presence of 6 mM $CH_3HgOH$. The left gel contained $^3H$ Sindbis viral RNA, HeLa 28S ribosomal RNA, and E. coli 23S and 16S ribosomal RNAs (molecular lengths 13.5, 5.5, 3.1, and 1.7 kb respectively). The right gel had the same markers plus $^{32}P$ clone 3 MSV RNA. After being stained with ethidium bromide and photographed, the gels were sliced in 2 mm sections and counted.

Figure 2. Gel electrophoresis of the two components of 50-70S RNA of clone 3 MSV. In (A) $^{32}P$ clone 3 RNA and $^3H$ MLV RNA were electrophoresed together in a 1.9% polyacrylamide gel. In (B) $^{32}P$ clone 3 RNA was electrophoresed as in (A), but the gel slices were counted by Cerenkov radiation. The slices indicated by the horizontal bars were pooled and eluted, and the recovered RNA was heat denatured and rerun on 2.1% polyacrylamide gels (C+D).

Figure 3. Electron micrographs of MSV RNA dimers. (A),(B),(C) show clone 3 RNA, glyoxal treated and spread from 40% formamide; in each case both poly(A) ends are hybridized to short poly(T) tails on an SV40 DNA circle. (A) is a dimer of large subunits (LL); an interpretive tracing is shown. (B) is a dimer of small subunits (SS). (C) is the single example of an apparent heterodimer (LS); the loop at the upper left is similar in size and position to those of MLV subunits, and the lower right loop looks like a partially collapsed figure-8 loop of an S subunit. (D) and (E) show clone 124-5R RNA dimers, glyoxal

treated, spread from 50% formamide, and attached to SV40 circles with poly(dBrU) tails. (D) is an LL dimer and (E) an SS dimer. The SV40-poly(dBrU) preparation used here had few tails per circle; in both cases shown, both viral poly(A) tails have hybridized to a single poly(dBrU) tail. In all pictures the dimer linkage is marked DLS and the characteristic large loops are marked ℓ. All pictures are at the same magnification.

Figure 4. Diagram of the structures of MLV and MSV dimers. The dimers of MLV and MSV clone 3 and clone 124-5R are all drawn to scale. The table lists contour lengths of the various subunits and their secondary structure features. The zig-zag lines at the ends of the molecules symbolize poly(A). The S subunits of clone 124-5R sometimes show a small loop or hairpin between the figure-8 loop and the poly(A) end.

Figure 5. Correlation of the two monomer lengths of MSV dimers. Randomly selected dimer molecules, each with a clear dimer linkage structure and with both poly(A) ends labeled, were photographed and measured. Each dot represents one dimer molecule; the horizontal coordinate is the length of its longer subunit, and the vertical coordinate the length of its shorter subunit. LL and SS homodimers lie near the diagonal, since both subunits are about the same length. LS heterodimers, had they been found, would be plotted around the position marked LS. (A) shows 41 clone 3 MSV dimers; (B) shows

98 clone 124-5R dimers.

Figure 6. Electrophoresis of clone 124 MSV RNA. (A) 50-70S RNA of clone 124 run on a 1.9% polyacrylamide gel. (B) $^{3}H$ labeled MLV (intact virus) was mixed with $^{32}P$ labeled clone 124 MSV (intact virus), and the mixture was phenol extracted, sedimented to prepare 50-70S RNA, and the mixed RNA was electrophoresed on a 1.9% polyacrylamide gel.

Figure 7. Electrophoresis of clone 124-5R MSV RNA. (A) 50-70S $^{3}H$ RNA of clone 124-5R MSV was mixed with heat dissociated $^{14}C$ RNA of clone 3 MSV and electrophoresed in a 1.6% agarose gel in Tris acetate buffer. (B) 50-70S $^{3}H$ RNA of clone 124-5R was mixed with 50-70S $^{14}C$ RNA of clone 3 MSV; the mixture was heat dissociated and electrophoresed in a 2% polyacrylamide gel. (C) Each 1 mm gel slice of Figure 7a was melted (1 min, 100°) in 100 µl 0.2 x Tris acetate buffer and its radioactivity determined by counting a 10 µl aliquot. Peak I (slices 50-53) was pooled, mixed with 50-70S $^{14}C$ RNA of clone 3 MSV, heated at 100° for 45 sec, and rerun on a 2% polyacrylamide gel.

Figure 8. A secondary structure feature in L monomers produced by heating and quenching. 50-70S clone 3 MSV RNA was dissociated by heating in 58% formamide, 60 mM Tris, 6 mM EDTA, pH 8.5 at 40° for 1 min, quenched on ice, and spread from 55% formamide. A new secondary structure feature (arrow) is observed on the L subunits.

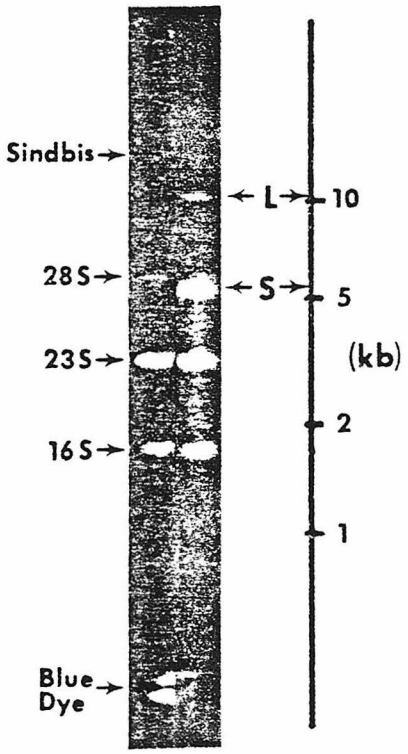



Fig. 1

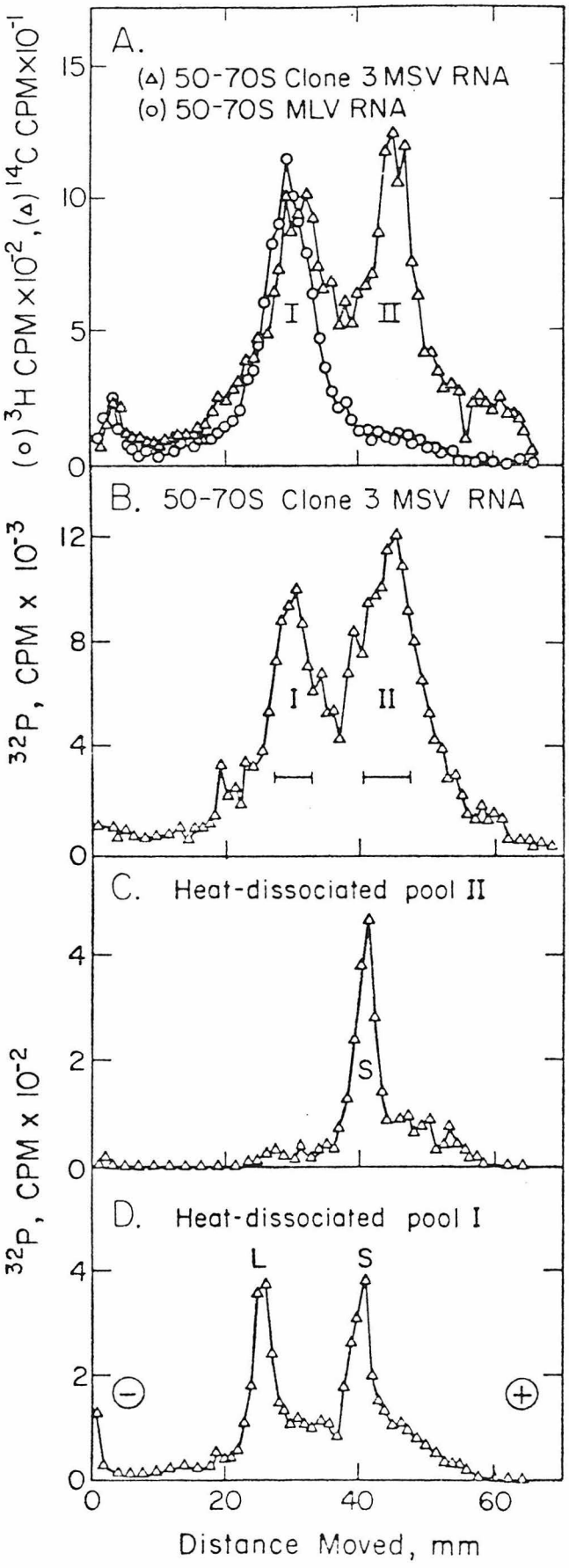
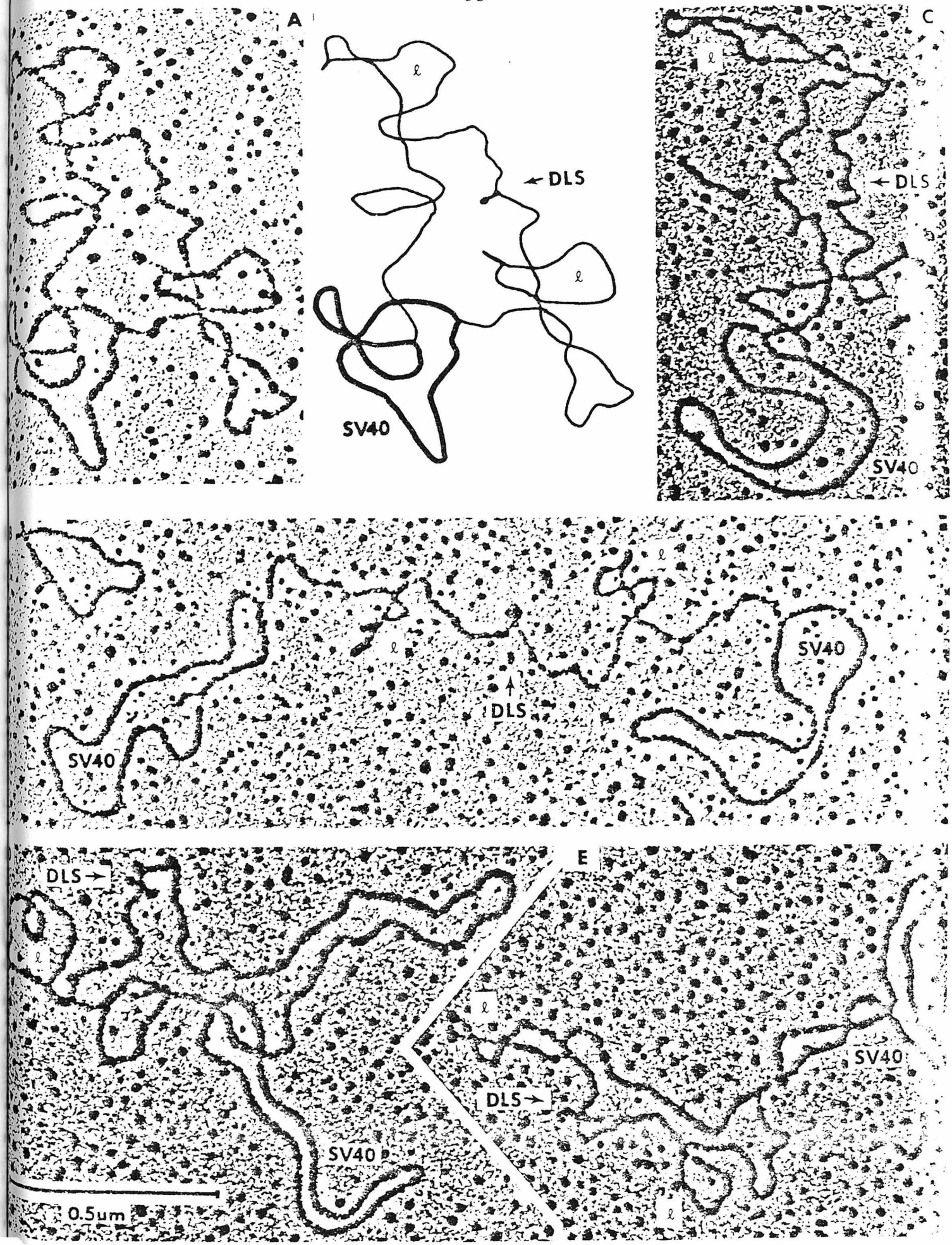
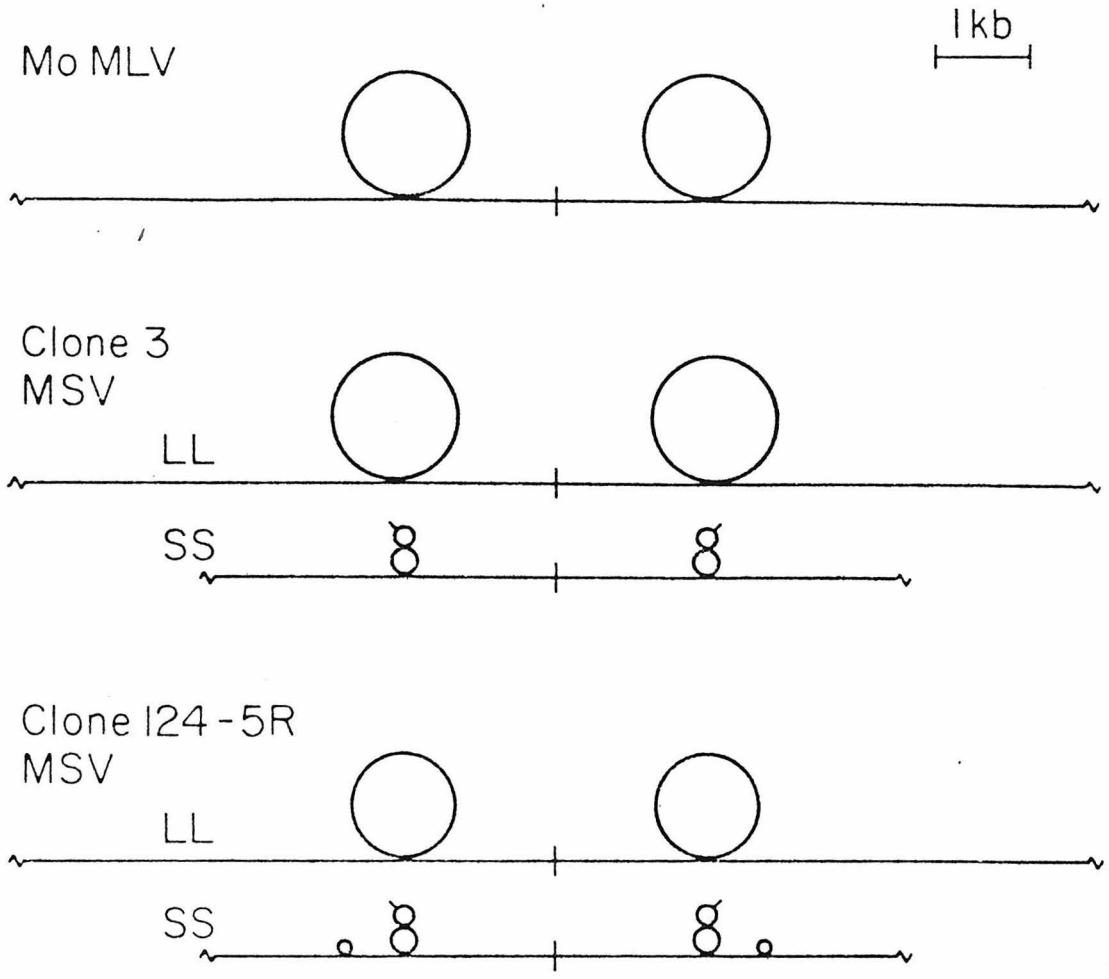

A.
(Δ) 50-70S Clone 3 MSV RNA
(o) 50-70S MLV RNA
(o) $^{3}H$ CPM x $10^{-2}$, (Δ) $^{14}C$ CPM x $10^{-1}$
I
II
B. 50-70S Clone 3 MSV RNA
$^{32}P$, CPM x $10^{-3}$
C. Heat-dissociated pool II
S
$^{32}P$, CPM x $10^{-2}$
D. Heat-dissociated pool I
L
S
⊖
⊕
Distance Moved, mm

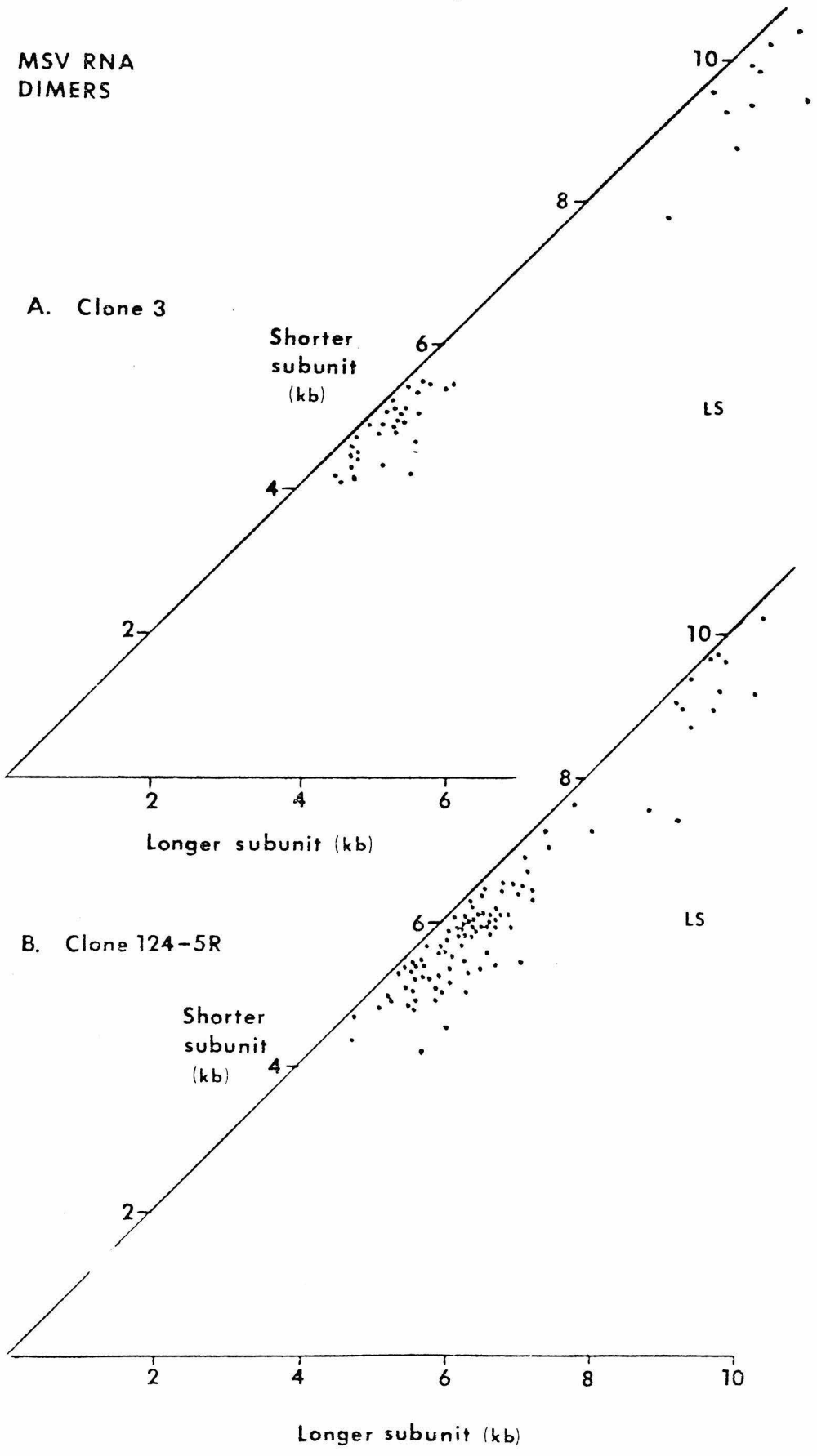
Fig. 2.


A
DLS
SV40
C
DLS
SV40
SV40
DLS
SV40
DLS
E
SV40
DLS
SV40
0.5um




Contour Lengths (kb)

| | | monomer | poly(A) to loop | loop | loop to DLS |
|---|---|---|---|---|---|
| Mo MLV | | 9.73 ±.60 | 3.88 ±.21 | 4.27 ±.21 | 1.57 ±.12 |
| Clone 3 MSV | L | 9.87 ±.65 | 3.96 ±.27 | 4.24 ±.27 | 1.67 ±.21 |
| | S | 5.15 ±.43 | 1.77 ±.16 | 1.78 ±.16 | 1.60 ±.15 |
| Clone 124-5R MSV | L | 9.50 ±.47 | 3.91 ±.39 | 3.68 ±.54 | 1.60 ±.17 |
| | S | 6.06 ±.70 | 2.74 ±.44 | 1.71 ±.31 | 1.59 ±.33 |


MSV RNA
DIMERS
A. Clone 3
Shorter subunit (kb)
LS
2
4
6
8
10
Longer subunit (kb)
B. Clone 124-5R
Shorter subunit (kb)
LS
2
4
6
8
10
Longer subunit (kb)

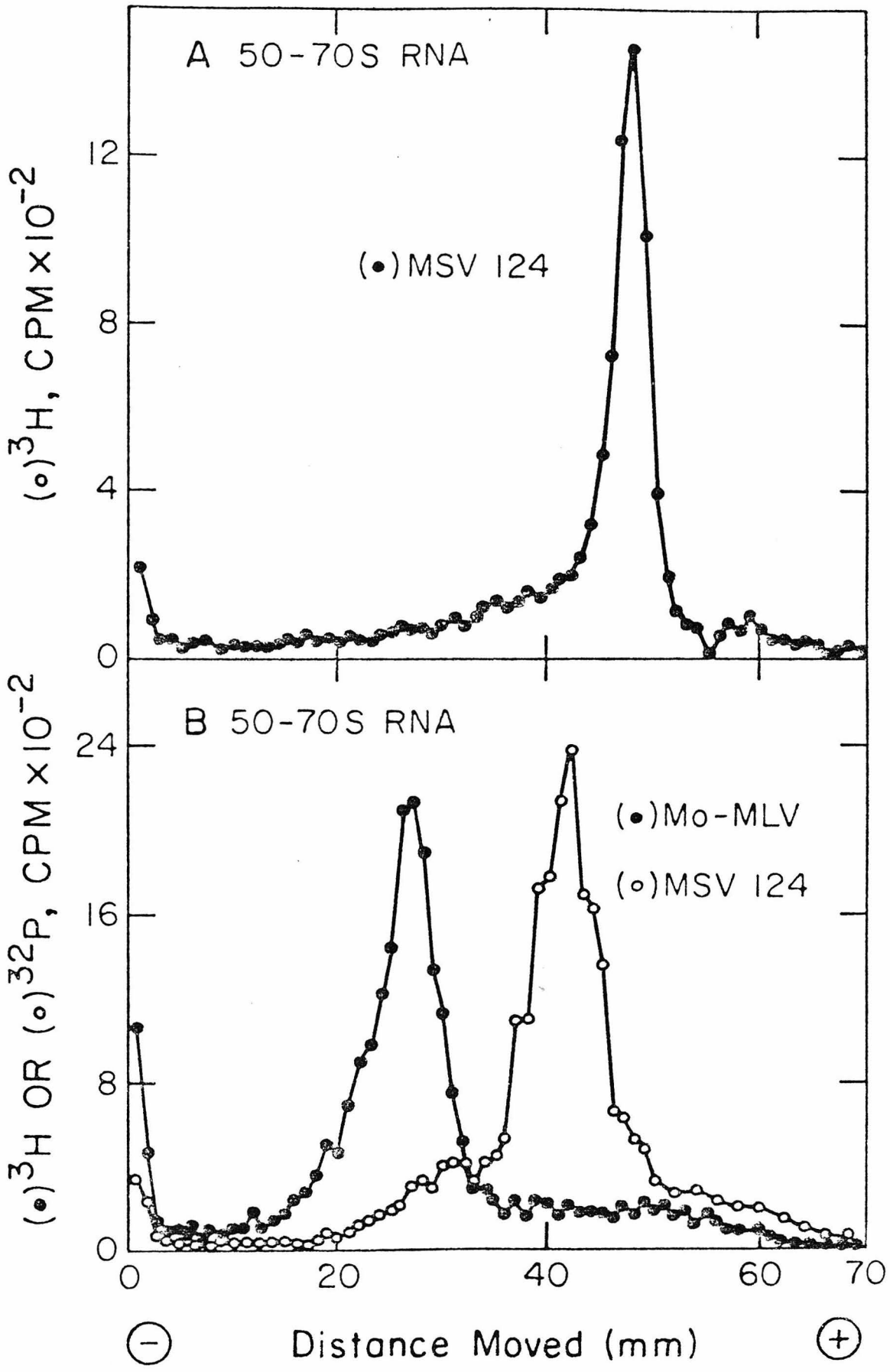

A 50-70S RNA
(•)MSV 124
(•)$^3$H, CPM×10$^{-2}$
12
8
4
0
B 50-70S RNA
(•)Mo-MLV
(○)MSV 124
(•)$^3$H OR (○)$^{32}$P, CPM×10$^{-2}$
24
16
8
0
0
20
40
60
70
⊖
Distance Moved (mm)
⊕

Fig. 6

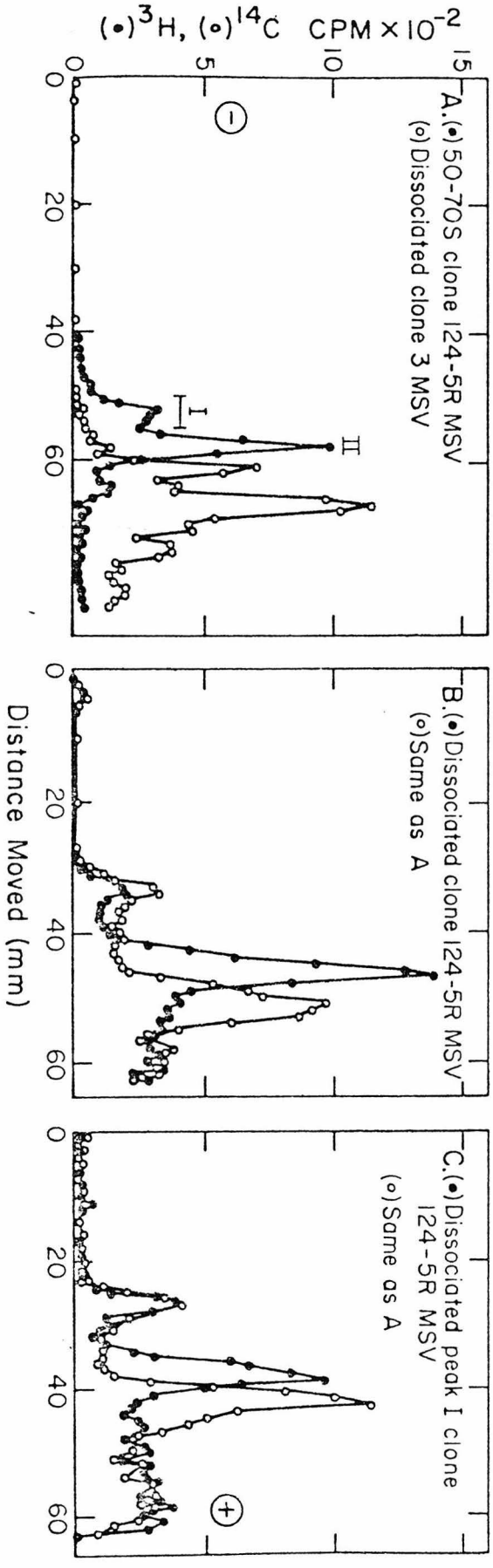

(•)$^3$H, (○)$^{14}$C CPM × 10$^{-2}$
A.(•) 50-70S clone 124-5R MSV
(○) Dissociated clone 3 MSV
I
II
B.(•) Dissociated clone 124-5R MSV
(○) Same as A
C.(•) Dissociated peak I clone 124-5R MSV
(○) Same as A
Distance Moved (mm)

Fig. 7

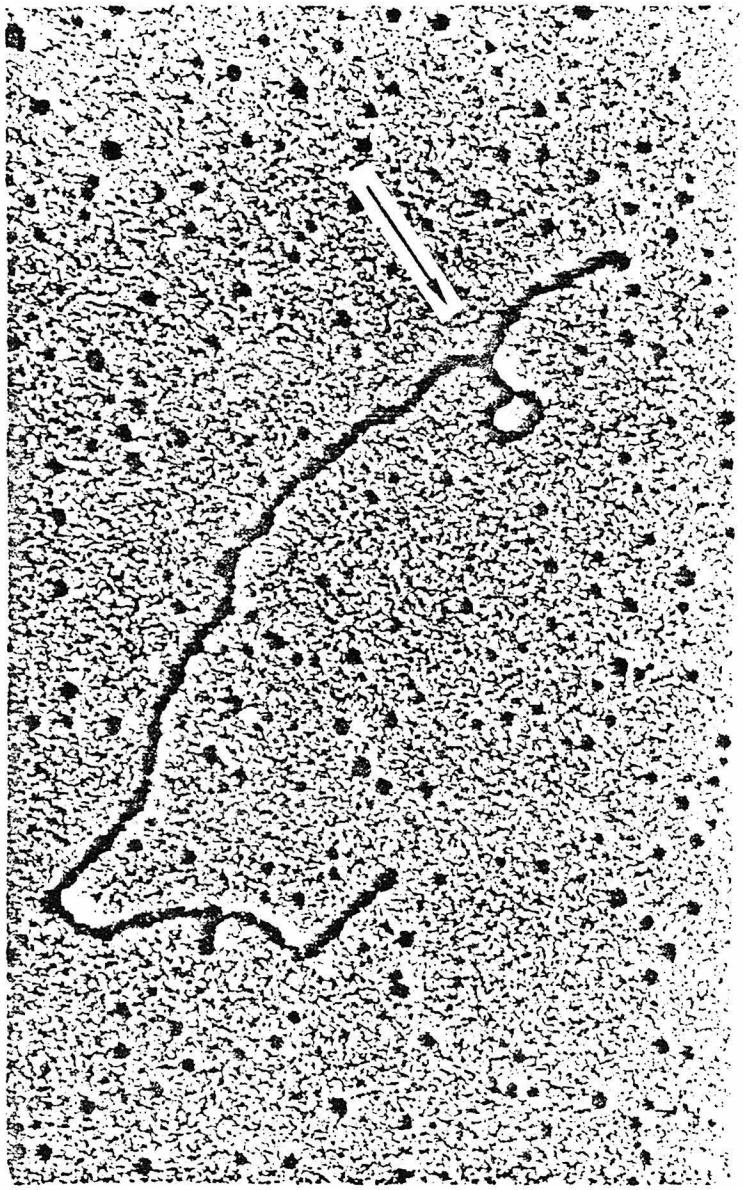


Fig. 8

# High Molecular Weight RNAs of AKR, NZB, and Wild Mouse Viruses, and Avian REV A11 have Similar Dimer Structures

Welcome Bender, Yueh-hsiu Chien,
Sisir Chattopadhyay, Peter K. Vogt,
Murray B. Gardner, and Norman Davidson

Division of Biology and Department of Chemistry
California Institute of Technology
Pasadena, California 91125

Laboratory of Viral Diseases,
National Institute of Allergy and Infectious Diseases
National Institutes of Health
Bethesda, Maryland 20014

Department of Microbiology and Department of Pathology
University of Southern California School of Medicine
Los Angeles, California 90033

Contribution number 5672

## ABSTRACT

Several 50-70S tumor viral RNA's have previously been shown by electron microscopy to be dimers, with the two monomers subunits joined near their 5' ends. Five additional RNA tumor viruses have now been examined: AKR, an endogenous murine ecotropic virus; NZB, an endogenous murine xenotrope; an ecotropic and an amphitropic virus isolated from a wild mouse; and the avian reticuloendotheliosis virus (REV). All five 50-70S RNA's show a similar 5' to 5' dimer structures. REV is the first example of an avian virus with a recognizable 5' to 5' dimer linkage; these observations suggest that the structure may be common to all RNA tumor viruses. All of the mammalian viral RNAs, but not REV, show symmetrically placed loops in each subunit of the dimer. Possible molecular structures and biological functions of the dimer linkages and loops are discussed.

## INTRODUCTION

The first 50-70S RNA from a tumor virus to be studied by electron microscopy in this laboratory was that of the endogenous cat virus, RD114. The RNA was observed to be dimer of two monomer subunits, each 10 kilobases (kb) in length, and the two monomers were joined together close to their 5' ends in a structure termed the dimer linkage structure (22, 2). The choice of RD114 was a fortunate historical accident in that the dimer linkage of RD114 RNA is more stable to dissociation than most other viruses we have studied, and so the dimer structure is usually maintained under the partially denaturing conditions needed to extend the RNA for spreading in a basic protein film.

Subsequent studies demonstrated very similar RNA dimers for BKD, an endogenous baboon virus (24); for WoMV, a sarcoma virus complex isolated from a woolly monkey (24); for the Friend murine virus mixture (8); and for Moloney murine leukemia virus and two variants of Moloney sarcoma virus (19, 25). The dimer linkage of BKD is as stable to dissociation as that of RD114, but for all the other viruses, the linkage is considerably less stable and can be preserved only by relatively mild denaturation conditions of spreading for electron microscopy, such that many RNA molecules are not fully extended and are difficult to interpret. Nevertheless, many dimers with the same general structural features as RD114 are seen on the grids. By contrast, avian sarcoma virus RNA (Prague C, Prague B, and B77 strains of ASV have been examined), treated with the same mild denaturation conditions, do not show any reproducible dimer structures (2).

It was of interest to examine the RNAs of other tumor viruses to see how universal the dimer structure is and to attempt to correlate the variations in secondary structure features of the RNA's with other differences among the viruses in their sequences or biological properties. The RNAs from four murine and one avian virus are described here; they all have the 5' to 5' dimer linkage.

## MATERIALS AND METHODS

### Virus

The AKR-L1 strain of AKR virus, originally isolated from a leukemic AKR mouse (11) was propagated on secondary NIH Swiss mouse embryo cells (7). NZB virus was originally isolated by treating cells derived from a New Zealand Black mouse embryo with iododeoxyuridine, and then co-cultivating these cells with mink lung fibroblast cells (13, 26). The NZB virus used here was propagated on these mink cells, the CCL64 line (17). The wild mouse viruses were produced by cultured embryo cells from a wild mouse trapped near Los Angeles, California (strain no. 1504). The ecotropic (grows only on cells from species of origin) and amphotropic (grows on species of origin and on other mammalian cells) viral strains were cloned by limiting dilutions of the original virus stock onto mouse, rabbit, mink, or human cells (10). Both wild mouse viruses were grown on SC-1 cells, a line obtained from a wild mouse embryo (12). Reticuloendotheliosis virus (REV), strain T, (26) was propagated on quail embryo cells.

Preparation of viral RNA

Viruses were pelletted from the cell supernatant, in some cases banded in sucrose, and then extracted with phenol or phenol-SDS as previously described (24). The RNA was sedimented through a sucrose gradient (10-30% in 0.1 M NaCl .01 M Tris .001 M EDTA) and the 50-70S peak was used for microscopy.

Electron microscopy

50-70S viral RNA was treated with glyoxal (1 M glyoxal, 0.01 M $PO_4$, pH 7.0) for various times, dialyzed, hybridized with SV40-poly(dT), and spread from 40% formamide as previously described (2). Alternately, the treated RNA was hybridized to SV40·poly(dBrU) and spread from 50% or 55% formamide (25). Molecules were photographed and measured; the magnification was calibrated using the SV40 circles as internal length standards (SV40 contour length taken as 1.76 μm). RNA lengths were computed assuming 1 μm = 3.9 kb for glyoxal treatment of 1 h at 37° or 30' at 42° (18), or 1 μm = 4.1 kb for glyoxal treatment of 30' at 37° or 1 h at 37° in the presence of 0.1 M NaCl. A less stringent glyoxal treatment gives slightly less extended RNA molecules; the exact length conversion depends on the particular treatment and the particular RNA. This factor is thus impossible to calibrate accurately for the mild glyoxal reaction conditions used here. The second calibration factor quoted above was obtained for BKD viral RNA treated with glyoxal in the presence of 50 mM NaCl (8).

Denaturing Gels

Molecular lengths of viral RNAs were measured by agarose gel electrophoresis in the presence of 5 mM $CH_3HgOH$ as previously described (1), except that the EDTA in the ER buffer was reduced from 1 mM to 0.1 mM. We believe that the reduction in EDTA concentration decreases the rate of transport of the $CH_3HgOH$ out of the gel as a negative $CH_3Hg$-EDTA complex. This should ensure that the slowly electrophoresing RNA is always in the presence of 5 mM $CH_3HgOH$.

$T_m$ Determinations

Aliquots of 50-70S viral RNA in 0.1 M NaCl, 0.01 M Tris, 1 mM EDTA, pH 7.5 were heated to various temperatures for 60-90 sec., and then chilled to 0°. The samples were then electrophoresed on a nondenaturing agarose gel (0.8% agarose) in ER borate buffer (1). The gels were stained with ethidium bromide (0.1 µg/ml) and photographed, or sliced and counted.

## RESULTS

Monomer Molecular Lengths

The RNAs of AKR virus, NZB virus, the ecotropic wild mouse virus and REV each showed a single band on a denaturing $CH_3HgOH$ gel with mobilities corresponding to estimated molecular lengths of 9.6, 9.5, 9.3., and 9.2 kb respectively (Figure 1). (Amphitropic wild mouse virus RNA was not available in sufficient quantity for gel electrophoresis.)

AKR Virus

50-70S RNA from AKR virus was treated with glyoxal (37°, 30 min), dialyzed, mixed with SV40-poly(dBrU) and spread for

Fig 2 microscopy from 50% formamide. Figure 2A shows a typical dimer molecule with both poly(A) ends hybridized to short poly(dBrU) tails on an SV40 DNA circle. The AKR dimer structure is shown
Fig 3 schematically in Figure 3 together with the structures of all the other viral RNA dimers studied here or previously. Like the other mammalian viral RNAs, AKR RNA has a 5' to 5' dimer linkage and a loop near the middle of each monomer subunit. The lengths and positions of the various secondary structure features are given in Table 1.

The large loops in the AKR monomers often had several cross-over points that could be interpreted as the loop junction, especially if the RNA had been treated lightly with glyoxal (i.e. for 30' instead of 1 h). A dimer with unusually many loop
Fig 4 junctions is shown in Figure 4A. For the diagram of Figure 3, the average position for the loop junction is given, but the length measurements for this feature have very large standard deviations (see Table 1) which reflect these multiple or alternate loop junction points. The position of the midpoint in the loop, however, is a much more reproducible measurement (Table 1). Thus it appears that there are multiple pairs of sites which can be bound together, each pair being symmetrically placed relative to a fixed center point of the loop. Any one pair may or may not be joined in a particular molecule as mounted for electron microscopy.

NZB Virus

The dimer linkage and the other secondary structure features of NZB dimers are much more stable to denaturation by glyoxal than those of any other murine viral RNA we have examined. After

treatment with glyoxal at 37° for 30' or 1 h, the 50-70S RNA molecules were mostly too condensed and tangled to interpret. Therefore, the NZB RNA was treated at higher temperature (½ h, 42°) before spreading. A dimer molecule with both poly(A) ends hybridized to SV40 (dBrU) is shown in figure 2B. Even after the harsher glyoxal treatment, the RNA was still a dimer and showed somewhat more secondary structure, in the form of small loops and hairpins, than the other murine viral RNAs. NZB RNA dimers have the large loops in each subunit, and as in the loops in AKR RNA; there appear to be multiple cohesive sites for loop formation, all defining approximately the same loop midpoint (Figure 4B). The NZB dimer structure is diagramed in figure 3 and quantitative data on contour lengths is presented in Table 1.

Wild Mouse Virus

The RNA's of two different types of RNA tumor virus isolated from wild mice, designated ecotropic and amphotropic to indicate their host ranges, were also examined. The 50-70S RNA of each virus was spread for microscopy after treatment with glyoxal at 37° for 1 h in the presence of 0.1 M NaCl. An example of each dimer is shown in figure 2. The monomer subunits of the dimer usually had the characteristic large loop. These loops did not show the multiple cohesion sites as seen in AKR or NZB although they frequently had two pairing sites (see the left loop of figure 2c and the right loop of figure 2d) to form a small asymmetric loop at the base of the larger one. The wild mouse dimer structures

are diagramed in figure 3 and measurements are included in Table 1.

Reticuloendotheliosis Virus

REV 50-70S RNA was treated with glyoxal (37°, 30') and spread from 50% formamide. A dimer with both poly(A) ends attached to SV40-poly(dBrU) is shown in figure 2E. The monomer subunits frequently had small loops (about 1 kb in circumference) or short hairpins but these features were not reproducible in position. In particular, there was no evidence of the larger symmetric loops in each subunit as seen in all mammalian viruses.

Thermal Melting of Dimers

It was of interest to measure the temperature for dissociation of 50-70S ASV RNA into 35S subunits to see if the lack of a reproducible dimer linkage in preparations for electron microscopy might be due to lower stability. $^{3}H$ labeled 50-70S ASV RNA was heated briefly to various temperatures and then analyzed on a nondenaturing agarose gel. At 55°, about 90% of the RNA still migrated as 50-70S; at 60°, only about 40% was still 50-70S; and at 65°, the RNA was 90% dissociated to 35S. This $T_m$ of about 60° is equivalent to those of Friend and Moloney viral RNAs (8, 25).

The NZB dimers were also examined because of their unexpected stability. Heating NZB RNA to 80° did not affect the migration of 50-70S RNA; 90° disrupted about half of the complexes, and 100° disrupted them all. The approximate $T_m$ of 90° is about the same as those of RD114 and BKD viral RNA's (23, 24).

The RNA's of AKR and the wild mouse viruses were not available in sufficient quantities to do such melts, but for all three RNAs, the extent of denaturation seen by electron microscopy,

after treatment with glyoxal or high concentrations of urea and formamide, was roughly equivalent to that seen for Friend or Moloney viral RNA's.

## DISCUSSION

Figure 3 shows the basic structures of the high molecular weight RNA's of all mammalian RNA tumor viruses so far examined. RD114 and BKD have 10-20% sequence homology to each other but not to any other virus in the list (27, 20). WoMV is not related to RD114 or BKD, but has 15 to 20% sequence homology with several murine viruses (3). All murine viruses are partially related to one another, but the viruses of Figure 3 are quite diverse in their degree of sequence homology. The AKR virus and the NZB virus both have about 60% sequence homology with Moloney MLV (Y. Chien and S. Chattopadhyay, unpublished observations) and there is about 50% homology between AKR and NZB (S. Chattopadhyay, unpublished observations). The defective spleen-sarcoma component of Friend virus is closely related to NZB but has only a little homology with the nondefective Friend lymphatic leukemia component or AKR virus (28). Wild mouse amphotropic and ecotropic viruses are closely related to each other, but hybridize only 40-50% with AKR and 20-30% with NZB (5).

In view of the sequence diversity of these mammalian viruses, the similarity in their structures is particularly striking. All nondefective viral RNA's have monomers of between 8 and 10 kb; all have two monomers joined at their 5' ends in a dimer linkage, and all have a large loop near the middle of each monomer subunit.

The smaller defective viral RNA's retain the same basic structure, with proportionately smaller loops, although the distance from dimer linkage to the base of the loop is not greatly reduced in the small Friend and Moloney viral components.

The dimers of the nondefective mammalian viruses are not identical in structure however. They show minor differences in the several secondary structure features. There are large differences in the stability of the dimer linkage towards dissociation; the endogenous xenotropic virus, RD-114, BKD, and NZB all have a stable dimer linkage with a $T_m$ measured (or estimated from measurements in other solvents) as 85-90° in 0.1 M $Na^+$ aqueous solvents. All of the other viruses we have studied are amphotropic or ecotropic, and, where measured, they have $T_m$'s for dimer dissociation in the range of 60 to 70° in aqueous solvents with 0.1 M $Na^+$.

We have not been able to recognize any reproducible dimer linkage structure in the 50-70S RNA of avian sarcoma viruses (2). However, B77 ASV has about the same melting temperature as do Moloney-MLV and MSV, and the Friend virus dimers, all of which have the characteristic 5' to 5' dimer structure. REV was examined because it is another avian virus unrelated to the Rous family of sarcoma viruses (21). The ultrastructure of the viral particle is reported to be closer to that of mammalian viruses than the avian sarcoma viruses (29). REV RNA shows a dimer linkage structure almost identical in position and morphology to those of the mammalian viruses, although it does not have the characteristic large loops. This dimer of avian viral RNA

suggests that the dimer linkage is a structural feature common to all RNA tumor viruses.

The apparent universality of the dimer structure of tumor viruses implies a biological function, but as yet, the function is unknown. It is likely that the structure of the 50-70S RNA is important for the copying of the RNA into DNA, since this happens quickly after penetration of the virus; stable 50-70S RNA does not reappear in the viral life cycle until after the virus buds off of the cell (6). Two monomers per virus might be required if reverse transcription proceeded by jumping from the 5' end of one monomer to the 3' end of another or if the initial DNA minus strand were a complete copy of one monomer plus a portion of the second monomer (i.e. slightly greater than the 10 kb unit length). The dimer linkage might also serve as a specific binding site for reverse transcriptase.

The physical nature of the dimer linkage is unknown. The viral RNA has been treated with phenol and SDS, but protein linkers have not been excluded. The dimer linkage of RD114 does not reform after melting and annealing (H.J. Kung, personal communication). Several models have been proposed to explain how such structures could be formed using only antiparallel base pairing of RNA (24). We calculate that an RNA duplex of only 10-15 base pairs (50% G + C) would have a $T_m$ of 60° ( the $T_m$ of Moloney or Friend dimers) and 30-40 base pairs would melt at about 90° (the $T_m$ for RD114, BKD and NZB) (4, 14). It has been suggested from sequence information that the primer tRNA may be base paired to both monomers and thus help to hold the dimer

together (16). By measuring the size of the initial DNA transcript, the 3' OH terminus of the primer is known to be bound to the RNA at 100 or 135 bases from the 5' end, for ASV or MoMLV viruses respectively (15). Figure 5 is a gallery of dimer linkage structures as found in several different tumor viruses. There is often considerable variation in appearance among examples on the same grid, presumably due to variation in the extent of glyoxal-induced denaturation. (The four examples of NZB are typical of the range of morphologies of the dimer linkage structure seen in murine viruses.) The dimer linkages of RD114 and BKD, and the most extended forms of the linkages seen with murine viruses and REV appear to involve at least 200 bases at the 5' end of each mononer (Figure 5), so it is probable that viral RNA sequences on the 3' side of the primer binding site are involved in the dimer linkage.

The loops near the middle of each monomer are also a reproducible feature in the mammalian RNA tumor viruses. The loops are not quite centered in the monomer; the midpoint of the loop is usually about 60% of the monomer length from the poly(A) end.

The molecular structure of the cohesive sites or sequences that cause loop formation is not known. The most obvious assumption is that they are due to base pairing between inverted repeat sequences, similar to the clover leaf structure for tRNA and the base-paired secondary structure features proposed for other RNAs (9). However, the loop structures do not reform after thermal melting and cooling. They could be held together by proteins or by small RNA linker molecules, or perhaps there are RNA sequences

base paired in the native dimer which after melting form alternate pairing arrangements.

In most murine viral RNA's, the base of the loop is held together at two or more points; this was most obvious for Friend virus (8), and for AKR and NZB as reported here. In BKD, the monomer RNA was often held together again between a point close to the 3' end and one close to the 5' end (24). In NZB, such multiple associations between the 3' region and the 5' region were often seen (as in Figure 4b) although the junction points were not reproducible. This multiple pairing of the monomer with itself suggests that in the virion each monomer is folded back on itself. Such a configuration would bring the 3' ends close to the dimer linkage, and transcription initiated at the primer near the 5' end could then more easily jump to a 3' end. Alternatively, the loop structures we see could, if they also exist in viral mRNA, be important for translation control.

Finally, it is important to emphasize that the RNA molecules studied here have been partially denatured to extend them for microscopy. There are probably other association in the virion, weaker than the dimer linkage or loop junctions, which are always disrupted by these techniques. Indeed, the absence of a dimer linkage structure in ASV may be because the structure is slightly more sensitive to glyoxal than those of mammalian viruses. A more detailed picture of the native structure of viral RNA may possibly be obtained by chemically or photochemically crosslinking the RNA in the virion, extracting and spreading the RNA for electron microscopy under highly denaturing conditions, and then systematically mapping the points of crosslinking.

ACKNOWLEDGMENTS

We are grateful to Madeline R. Lander and Wallace P. Rowe for assistance and advice. W.B. has been the recipient of a National Science Foundation fellowship and of a training grant from the National Institute of General Medical Sciences. This research has been supported by contract NO1 CP 43306 with the Virus Cancer Program of the National Cancer Institute.

1. Bailey, J.M. and N. Davidson. 1976. Methylmercury as a reversible denaturing agent for agarose gel electrophoresis. Anal. Biochem. 70:75-85.

2. 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.

3. Benveniste, R.E., and G.J. Todaro. 1973. Homology between type-C viruses of various species as determined by molecular hybridization. Proc. Nat. Acad. Sci. USA 70, 3316-3320.

4. Billeter, M.A., C. Weissmann, and R.C. Warner. 1966. Replication of viral ribonucleic acid. IX. Properties of double-stranded RNA from Escherichia coli infected with bacteriophage MS2. J. Mol. Biol. 17:145-173.

5. Bryant, M.L., P. Roy-Burmon, M.B. Gardner, and B.K. Pol. 1977. Evolutionary relationships between ecotropic, xenotropic, and amphotropic murine type C RNA viruses. Nature (Lond.) in press.

6. Canaani, E., K. Vander Helm, and P. Duesberg. 1973. Evidence for 30-40S RNA as precursor of the 60-70S RNA of Rous sarcoma virus. Proc. Nat. Acad. Sci. USA 72:401-405.

7. Chattopadhyay, S.K., D.R. Lowy, N.M. Teich, A.S. Levine, and W.P. Rowe. 1974. Qualitative and quantitative studies of AKR-type murine leukemia virus sequences in mouse DNA. Cold Spring Harbor Symp. Quant. Biol. 39:1085-1101.

8. Dube, S., H.J. Kung, W. Bender, N. Davidson, and W. Ostertag. 1976. Size, subunit composition, and secondary structure

of the Friend virus genome. J. Virol. 20:264-272.

9. Fiers, W., R. Contreras, F. Duerinck, G. Haegeman, D. Iserentant, J. Merregaert, W. Min Jou, F. Molemans, A. Raeymaekers, A. VandenBerghe, G. Volckaert, and M. Ysebaert. 1976. Complete nucleotide sequence of bacteriophage MS2 RNA: primary and secondary structure of the replicase gene. Nature (Lond.) 260:500-507.

10. Gardner, M.B. 1977. Amphotropic and ecotropic type C viruses of wild mice. Current Topics in Microbiology and Immunology. in press.

11. Hartley, J.W., W.P. Rowe and R.J. Huebner. 1970. Host range restrictions of murine leukemia viruses in mouse embryo cell cultures. J. Virol. 5:221-225.

12. Hartley, J.W., and W.P. Rowe. 1975. Clonal cell lines from a feral mouse embryo which lack host-range restrictions for murine leukemia viruses. Virology 65:128-134.

13. Hartley, J.W., and W.P. Rowe. 1976. Naturally occurring murine leukemia viruses in wild mice. Characterization of a new "amphotropic" class. J. Virol. 19:19-25.

14. Hayes, F.N., E.H. Lilly, R.L. Ratliff, D.A. Smith, and D.L. Williams. 1970. Thermal transitions in mixtures of polydeoxyribonucleotides. Biopolymers 9:1105-1117.

15. Haseltine, W.A., D.G. Kleid, A. Panet, E. Rothenberg, and D. Baltimore. 1976. Ordered transcription of RNA tumor virus genomes. J. Mol. Biol. 106:109-131.

16. Haseltine, W.A., A.M. Maxim, and W. Gilbert. 1977. Rous sarcoma virus genome is terminally redundant: the 5' sequence. Proc. Nat. Acad. Sci. USA. 74:989-993.

17. Henderson, I.C., M.M. Lieber, and G.J. Todaro. 1974. Mink cell line Ngv-1Lu (CCL 64). Focus formation and generation of "nonproducer" transformed cell lines with murine and feline sarcoma viruses. Virology 60:282-287.

18. Hsu, M.T., H.J. Kung, and N. Davidson. 1973. An electron microscope study of Sindbis virus RNA. Cold Spring Harbor Symp. Quant. Biol. 38:943-950.

19. Hu, S., N. Davidson, and I.M. Verman. 1977. A heteroduplex study of the sequence relationships between the RNAs of M-MSV and M-MLV. Cell 10:469-477.

20. Hu, S., N. Davidson, M.O. Nicolson, and R.M. McAllister. 1977. A heteroduplex study of the sequence relations between RD-114 and Baboon viral RNAs. J. Virol. 23:345-352.

21. Kong, C.Y., and H.M. Temin. 1973. Lack of sequence homology among RNAs of avian leukosis-sarcoma viruses, reticuloendotheliosis viruses, and chicken endogenous RNA directed DNA polymerase activity. J. Virol. 12:1314-1324.

22. Kung, H.J., J.M. Bailey, N. Davidson, P.K. Vogt, M.O. Nicolson, and R.M. McAllister. 1974. Electron microscope studies of tumor virus RNA. Cold Spring Harbor Symp. Quant. Biol. 39:827-834.

23. Kung, H.J., J. Bailey, N. Davidson, M.O. Nicolson, and R.M. McAllister. 1975. Structure, subunit composition, and molecular weight of RD-114 RNA. J. Virol. 16:397-411.

24. Kung, H.J., S. Hu, W. Bender, J. Bailey, N. Davidson, M.O. Nicolson, and R.M. McAllister. 1976. RD-114, Baboon and woolly monkey viral RNAs compared in size and structure. Cell 7:609-620.

25. Maisel, J., W. Bender, S. Hu, P.H. Duesberg, and N. Davidson. 1977. Structure of the 50-70S RNA from Moloney sarcoma viruses. J. Virol. in press.

26. Theilen, G.H., R.J. Ziegel, and M.J. Twiehaus. 1966. Biological studies with RE virus (strain T) that induces reticuloendotheliosis in turkeys, chickens, and Japanese quail. J. Nat. Cancer. Inst. 37:731-743.

27. Todaro, G.J., C.J. Sherr, R.E. Benveniste, M.M. Lieber, and J.L. Melnick. 1974. Type C virus of baboons: isolation from normal cell cultures. Cell 2:55-61

28. Troxler, D.H., J.K. Boyars, W.P. Parks, and E.M. Scolnick. 1977. Friend strain of spleen focus-forming virus: a recombinant between mouse type C ecotropic viral sequences and sequences related to xenotropic virus. J. Virol. 22: 361-372.

29. Zeigel, R.F., G.H. Theilsen and M.J. Twiehaus. 1966. Electron microscopic observations on RE virus (strain T) that induces reticuloendotheliosis in turkeys, chickens, and Japanese quail. J. Nat. Cancer Inst. 37:709-729.

Table 1

Contour lengths of viral RNA structural features (in kb).

| | Monomer Length | poly(A) to loop | loop | loop to dls | loop midpoint to dls |
|---|---|---|---|---|---|
| AKR | 8.49 ±8% | 3.39 ±17% | 3.58 ±24% | 1.64 ±30% | 3.39 ±10% |
| NZB | 8.67 ±7% | 3.22 ±14% | 3.44 ±18% | 1.94 ±15% | 3.68 ±6% |
| WME | 9.29 8% | 3.69 11% | 3.88 16% | 2.23 9% | 4.17 8% |
| WMA | 9.55 ±8% | 3.68 14% | 3.94 16% | 2.03 9% | 4.00 9% |
| REV | 8.63 ±8% | | | | |

## FIGURE LEGENDS

Figure 1. Molecular weight determinations of viral RNAs on $CH_3HgOH$ agarose gels. Lanes A and H contain molecular weight standards which include Sindbis RNA, HeLa 28S ribosomal RNA, E. coli 23S and 16S ribosomal RNAs and E. coli tRNA (molecular lengths of 13.5, 5.5, 3.1, 1.7, and 0.1 kb respectively). Lane B has Prague B ASV; the two bands are from transformation competent and transformation defective virions. Lane C has AKR RNA, lane D has NZB RNA, lane E has wild mouse ecotropic RNA, lane F has avian reticulotheliosis virus, and lane G has RNA from clone 3 Moloney MSV grown with Moloney MLV as helper. All viral RNAs were 50-70S RNA selected after sucrose gradient sedimentation except for the REV (lane F) which was total RNA extracted from pelletted virus. The REV RNA is contaminated with 18 and 28S ribosomal RNAs from the quail cells on which the virus was grown.

Figure 2. Electron micrographs of viral RNA dimers. All dimers are hybridized by their poly(A) tails to SV40 duplex circles with short tails of poly(T) or poly(dBrU). (A) is AKR RNA, (B) is NZB RNA, (C) is wild mouse ecotropic viral RNA, (D) is wild mouse amphotropic viral RNA, and (E) is avian reticuloendotheliosis viral RNA. In (C) only one poly(A) end appears

attached to the SV40. The dimer linkage structures are marked DLS and the large loops are labeled ℓ. All micrographs were printed at the same magnification.

Figure 3. Diagram of the structures of all tumor virus RNA dimers so far examined in the laboratory. The dimer linkage structure is shown in the middle of each molecule; the zig-zag lines at the ends of the molecules symbolize poly(A). Next to each structure is listed the approximate melting temperature of the dimer linkage, where it has been determined. The viruses are: RD114, the cat endogenous xenotrope (23); BKD, the baboon endogenous xenotrope (24); WoMV, the simian sarcoma virus complex (24); Friend murine virus complex, including the lymphatic leukemia virus and the spleen focus forming virus (three different viral RNAs were present in the mixture studied) (8); MoMLV, the Moloney murine leukemia virus (19, 25); two isolates, clone 3 and clone 124 of MoMSV, Moloney murine sarcoma virus (19, 25); AKR murine endogenous ocotrope; NZB murine endogenous xenotrope; WME, a wild mouse ecotrope; WMA, a wild mouse amphotrope; and REV, the avian reticuloendotheliosis virus.

Figure 4. RNA dimers showing multiple pairings in each subunit. (A) is AKR RNA and (B) is NZB RNA. Both molecules have more possible pairing sites than average dimers, but

such multiple pairings are common for both viruses. The dimer linkage structures are indicated.

Figure 5. Gallery of dimer linkage structures. The dimer linkages are from RD 114 (A), REV (B), NZB (C-F), and BKD (G-I). The NZB structures (C-F) were all found on the same grid, and were chosen to show the range of variation typically seen in all the murine viruses studied. The Y shaped linkage (A) is distinctive to RD 114, and BKD apparently has two linkage sites, either one or both of which may be paired in any individual dimer.

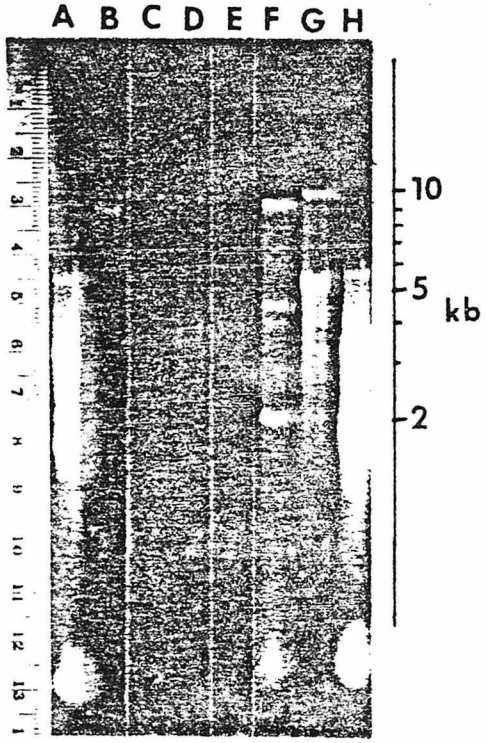



Fig. 1

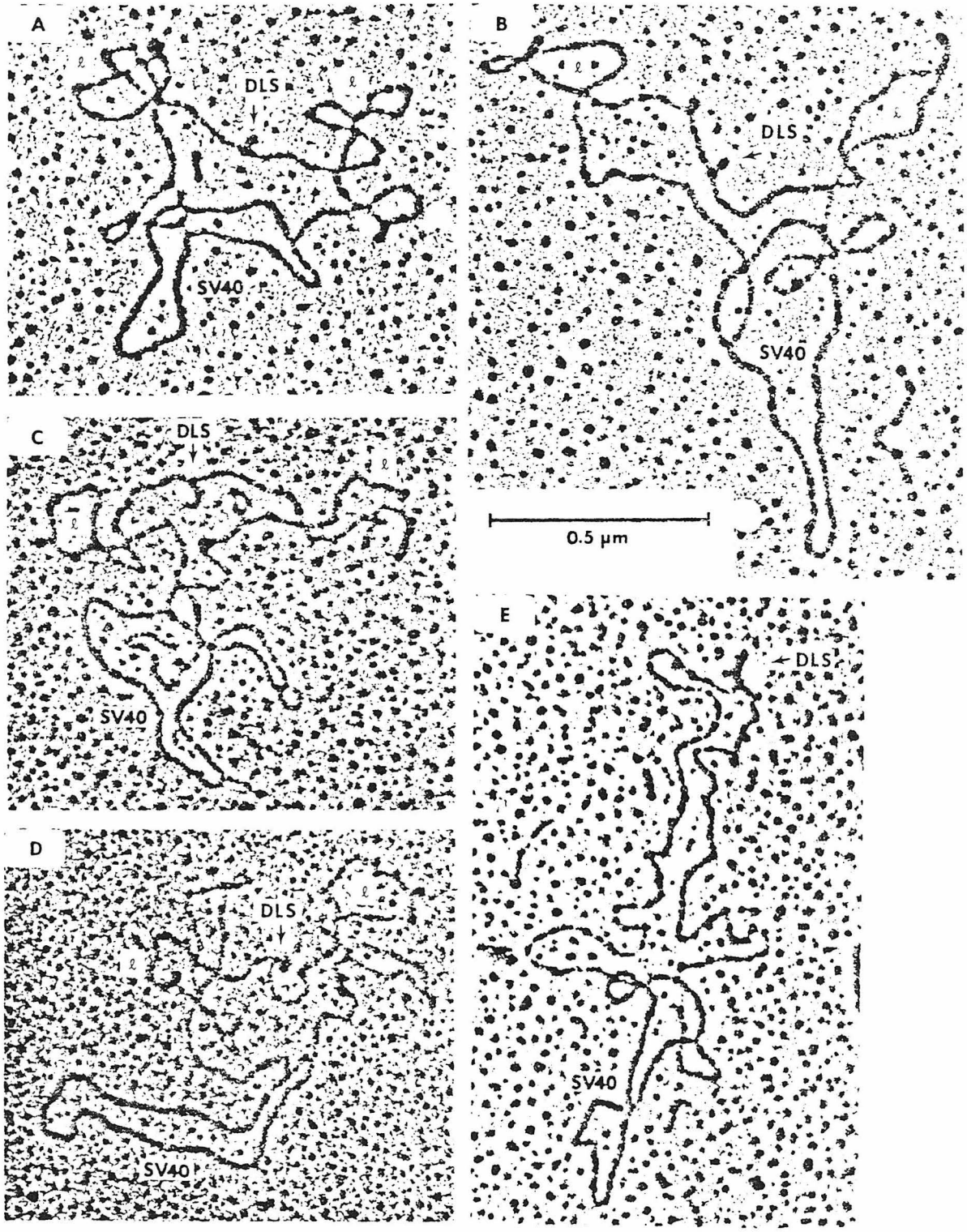



Fig. 2

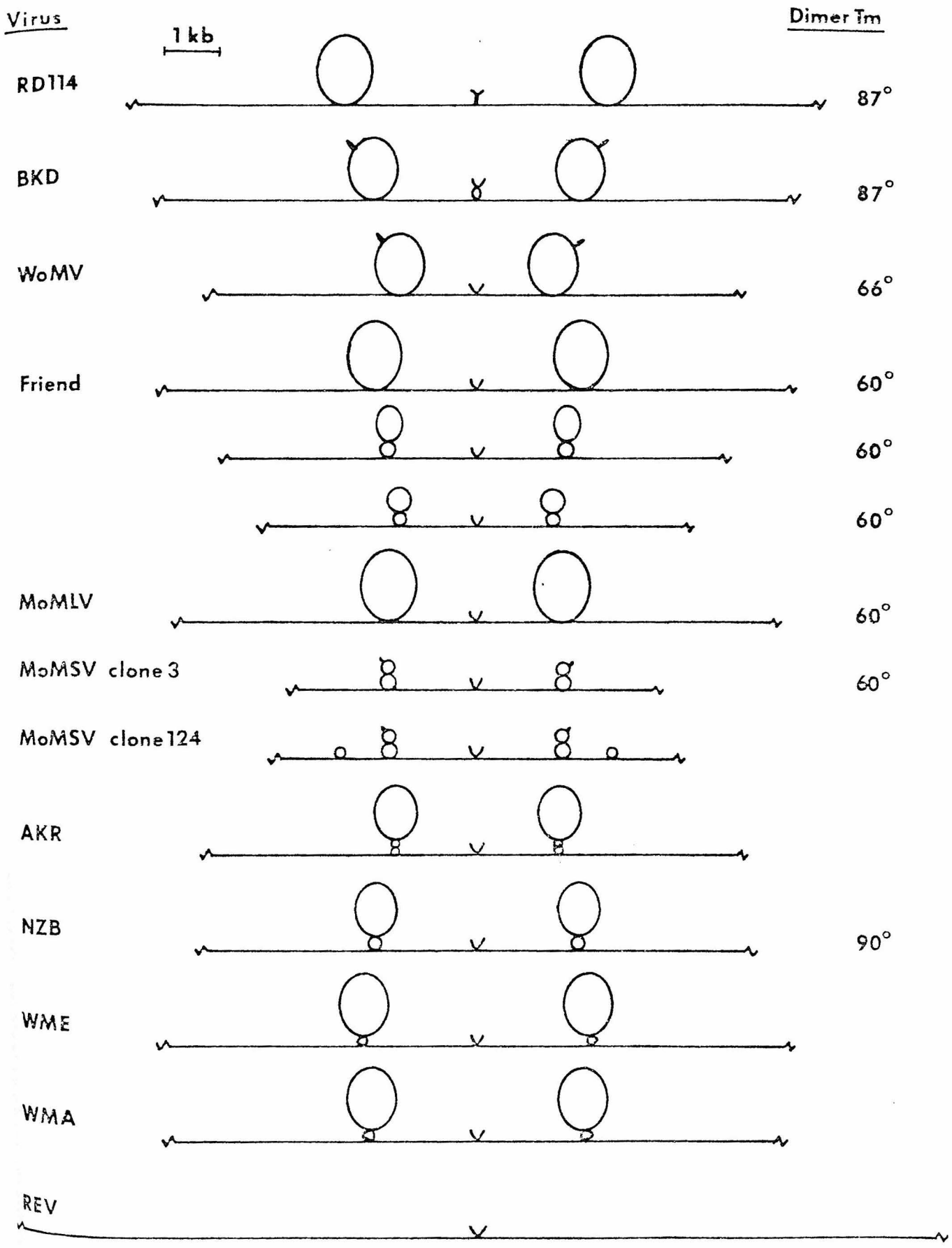

Virus
1 kb
Dimer Tm
RD114
87°
BKD
87°
WoMV
66°
Friend
60°
60°
60°
MoMLV
60°
MoMSV clone 3
60°
MoMSV clone 124
AKR
NZB
90°
WME
WMA
REV

Fig. 3

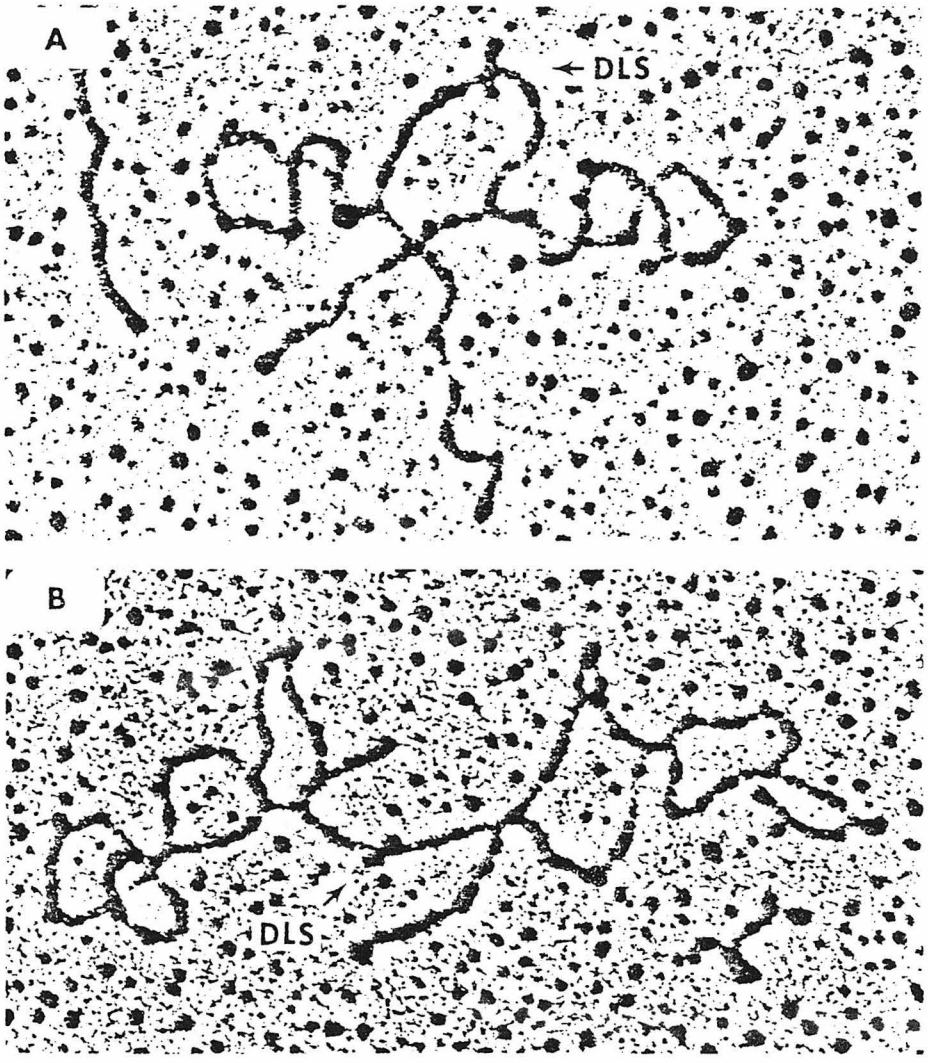



Fig. 4

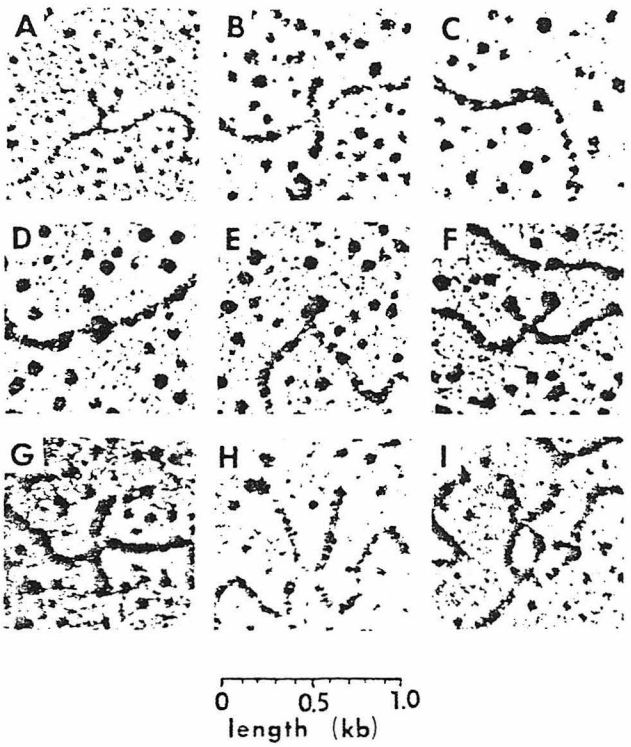



Fig. 5

# The Structure of M6, a Recombinant Plasmid Containing *Dictyostelium* DNA Homologous to Actin Messenger RNA.

Welcome Bender and Norman Davidson

Division of Biology and Division of Chemistry
California Institute of Technology
Pasadena, California 91125

Karen Kindle, William Taylor
Michael Silverman, and Richard A. Firtel

Department of Biology
University of California, San Diego
La Jolla, California 92093

Contribution number 5670

## ABSTRACT

The recombinant plasmid, M6, contains a DNA sequence from the cellular slime mold Dictyostelium discoidium which hybridizes to actin messenger RNA. The plasmid contains 6 kilobase pairs (kb) of Dictyostelium DNA inserted into a pMB9 vector. Ten cleavage sites for four different restriction enzymes have been mapped. Other work has shown that a central restriction fragment, 1.8 kb in length, contains sequences repeated about twenty times in the genome, and this fragment hybridizes to actin mRNA. Heteroduplexes between M6 and H6, another plasmid which contains two copies of the actin repeat, were used to define the position of this repeat in M6. Two plasmids with inserts of cDNA made from actin mRNA were heteroduplexed to M6 to define the position and orientation of the message complementary region. An improved electron microscope technique was developed for labeling poly(dA) sequences with dBrU polymers attached to suitable markers, and the technique was used to map several short poly(dA) tracts in the M6 insert, relative to the actin coding region.

## INTRODUCTION

Kindle et al. (preceding paper) constructed the recombinant plasmid, M6, and showed that the Dictyostelium DNA insert hybridizes to actin mRNA. Actin is a well characterized protein, highly conserved in evolution, whose function is at least partly understood and whose expression is developmentally regulated in slime molds (Kindle et al., preceding paper). In addition, there are multiple forms of actin in slime molds and other eukaryotes, and there is a family of related actin genes (Kindle et al., preceding paper, Garrels and Gibson, 1976). Because of the sequence homology among the different actin genes, M6 can be used as a probe to isolate the other family members for examination and comparison. (McKeowan et al., following paper).

The initial characterization of the M6 plasmid is reported here. We have mapped the cleavage sites of several restriction endonucleases on M6 DNA, located the repeated sequence, and determined the position and orientation of the actin mRNA complementary sequence.

Many eukaryotes have short homopolymeric tracts of dA:dT in genomic DNA (Mol and Borst, 1976); such short poly(dA) tracts are particularly prevalent in Dictyostelium DNA (Jacobson et al. 1974). An improved electron microscope technique is described here for identifying these poly(dA) sequences, and these sites are mapped on M6 relative to the actin gene.

## RESULTS

The construction of M6, and its characterization as clone hybridizing to actin mRNA is described in the accompanying paper

(Kindle and Firtel). Briefly, nuclear DNA from Dictyostelium discoidium was sheared to about 8kb, it was resected with λ exonuclease and extended with short poly(dA) sequences. The vector plasmid, pMB9, was cut at its single R1 site, resected with λ exonuclease and extended with short poly(dT) sequences. The insert and vector pieces were then annealed together and transformed into E. coli. M6 was one of several plasmids initially selected because they hybridized to Dictyostelium mRNA. mRNA complementary to M6 was isolated by hybridization to M6 DNA cellulose, and was translated in vitro to make a protein identical to Dictyostelium actin.

Figure 1 presents the map of M6 derived from the studies reported below. Positions on the map are designated by their distances in kilobase units from the left hand junction between plasmid and insert.

Restriction mapping

Closed circular M6 DNA was cleaved with the restriction endonucleases Bam HI, Sal I, Eco RI, Hind III, Hae III, and Hap II, and the lengths of the fragments were analyzed by electrophoresis in 1% agarose gels. Overlaps were defined by sizing the products of double and triple digests, and restriction sites were positioned relative to the known restriction map of the vector plasmid, PMB9 (Rodrigeuez et al. 1976). The positions of the Bam HI, Eco RI, Hind III, and Hae III sites were confirmed by electron microscopy in the course of the heteroduplex and oligo(dA) mapping experiments discussed below. None of the restriction endonucleases

Pst I, Sma I, Hga I, and Hha I cleaves within the Dictyostelium insert.

The initial restriction mapping was confused by the apparent absence of the Alu I site which was expected to lie between the Eco RI site (the site of insertion) and the Hind III site of pMB 9. Furthermore, poly(dA) mapping (see below) indicated that the insert vector junction was unexpectedly close to the Bam HI site. Heteroduplexes between M6 (cut at the Bam HI site) and pMB 9 (cut at either Rl or Bam sites) confirmed that approximately 300 bases of pMB 9 to the right of the Rl site had been lost in the construction of M6 (see Figure 2 for details of the heteroduplex). Apparently equivalent deletions occurred in the construction of both M4 and H6 (see preceding and following papers, respectively), two other plasmids made in the same way as M6. The reason for these deletions is unknown but we note that the endpoint of the deletions is at or near the edge of the tetracycline resistance gene in pMB 9 (Rodrigeuez et al. 1976).

Previous work has shown that the central Hae III - Hap II DNA fragment of M6 (2.1 to 4.0 kb) contains sequences which are repeated 15-20 times in the Dictyostelium genome and that this fragment hybridizes to actin message (Kindle et al. preceding paper). The region of the repeated DNA in M6 and the region of actin messenger homology were defined more precisely by heteroduplex analysis.

The plasmid H6 has been characterized as having two repeats of the actin gene on its 4.3 kb insert; its restriction map is

given in figure 1 (McKeown et al., following paper). The repeated regions of H6 proved to be in the opposite orientation to that of M6. Thus, the majority of M6/H6 heteroduplex molecules showed complete pairing between the pMB 9 vector portions of M6 and H6, but no hybridization at all between the two inserts. Therefore, one plasmid was first heteroduplexed with pMB 9, and then single strands of a restriction digest of the second plasmid were added to the hybridization mixture. Figure 2A shows an M6/pMB 9 heteroduplex hybridized in the region of the actin repeat with the EcoRI-BamHI fragment of H6 (0.5 to 4.6 kb on H6 map). Figure 2B shows an H6/pMB 9 heteroduplex with two EcoRI fragments (1.0 to 5.7 kb) of M6 hybridized to the two copies of the actin repeat of H6. Both H6 actin sequences hybridize to the same region from 3.0 to 4.1 kb on M6. The positions of the repeats on M6 and H6 are shown in Figure 1.

Position and orientation of actin coding sequence

Total Dyctyostelium cytoplasmic poly(A)-containing mRNA was used to make clones of cDNA by the technique of Efstratiadis et al., (1977) (Rowenkemp and Firtel, manuscript in preparation). The RNA was copied with reverse transcriptase from an oligo(dT) primer, the RNA was removed by base hydrolysis, the DNA was made double stranded with E. coli DNA polymerase I, the ends of the cDNA duplex were trimmed with S1 nuclease, and the 3' ends were extended with short poly(dC) sequences. This was annealed with the vector plasmid, pBR 322, cut at its single Pst I restriction site and extended at its 3' ends with short poly(dG)

sequences, and the mixture was used to transform E. coli. In the resulting plasmids, Pst I restriction sites are regenerated on either side of the cDNA insert. cDNA plasmids coding for actin mRNA were identified by colony filter hybridization using the M6 HaeIII-HapII fragment (2.1-4.0 kb) as a probe (Rowenkamp and Firtel, manuscript in preparation). Two actin cDNA plasmids, A1 and B1, have been studied; they have cDNA inserts of 0.8 and 1.1 kb respectively. Figure 3A shows a heteroduplex between B1 and pMB 9 to illustrate the position of the cDNA insert (which is often held in a small loop due to pairing of the poly(G)-poly(C) tails) and the regions of non-homology between the two vectors, pMB9 and pBR322. A B1/M6 heteroduplex is shown in Figure 3B. The B1 insert is hybridized to M6 in the region of the actin repeat; the exact region of homology is shown in Figure 1. The poly(dA) mapping of M6 described below shows that there are poly(dA) sequences at 0.0 and 0.3 kb and a poly(T) sequence at 0.9 kb (on one strand, with the complementary T and A sequences on the other strand). The single stranded M6 segments of the M6/B1 heteroduplex often show an additional loop due to base pairing of these A and T sequences. This feature locates in the heteroduplex molecule the left hand junction of M6 between vector and insert. The region of M6 homologous to B1 ends at or near the end of the repeated region, as defined by the M6/H6 heteroduplexes. There is a Hind III restriction site close to the right end of the B1 cDNA insert (Figure 1) which is presumably equivalent to the Hind III site at 4.1 kb in M6.

The cDNA plasmid A1 forms heteroduplexes with pMB9

have been copied from the poly(rA) sequence of the actin mRNA, and thus it defines the orientation of the actin mRNA sequence in M6 (3' end on the left in Figure 1). The 3' end of the message is at or near the left end of the repeated region, as defined by the H6 heteroduplexes.

In order to confirm the orientation of the M6 actin region, the EcoR1 fragment (1.0-5.7 kb) was isolated and inserted into the bacteriophage λ gt vector (Thomas et al. 1974) in both orientations. The orientation of the actin region in individual λ clones was determined by digesting the recombinant λ DNA with Hap II or Hind III and measuring the size of the products. Intact DNA of recombinant λ clones was then strand separated on $Cs_2SO_4$ gradients in the presence of poly(U,G) (Szybalski et al., 1971), and the separated strands were hybridized to Dictyostelium cytoplasmic mRNA. The mRNA hybridization was highly strand specific, and in both orientations of the M6 fragment, the hybridization confirmed the mRNA orientation defined above.

Poly(dA) mapping

Poly(A) sequences in DNA or RNA can be visulaized in the electron microscope by hybridization with homopolymers of T or dBrU. Poly(dBrU) forms more stable hybrids with poly(A) than does poly(T), and so dBrU labeling can be done in more denaturing spreading solutions or can be used to mark very short poly(A) sequences. (See appendix for detailed discussion of A:dBrU duplex stability as a function of length.) Short homopolymers of T or dBrU are difficult to visualize by themselves, but they can easily be polymerized with the enzyme terminal transferase onto any

recognizable duplex DNA molecule, such as nicked circular SV40 DNA (Bender and Davidson 1976), linear SV40 DNA (Carbon et al. 1975), or the other duplex DNAs described below.

As explained above, the A1 actin cDNA plasmid was thought to contain a long poly(dA) sequence copied from the poly(rA) of the message. Duplexes of rA:dBrU of 100 or more base pairs are stable in spreading solutions of up to 60% formamide, and so the site in A1 could be labeled under normal heteroduplex spreading conditions. A1 was heteroduplexed with pMB9; a dBrU label was added; the mixture was incubated briefly and then was spread from 50% formamide. The dBrU label in this case was an 800 base pair circular duplex DNA isolated from mitochrondria of the trypanisome Leishmania torentolae (Wesley and Simpson, 1973) onto which tails of poly(dBrU) of about 200 bases had been synthesized. Figure 4 shows a labeled A1/pMB9 heteroduplex; the poly(dA) site maps at the left hand edge of the actin cDNA insert (0.0 to 0.2 kb in the A1 map of Figure 1).

For locating the short poly(dA) sequences in the Dictyostelium DNA of M6, much less denaturing spreading conditions were required. In order to keep the single stranded M6 DNA extended, it was first reacted with glyoxal (which should react with adenine, cytosine, and guanine) and then dialyzed to remove the glyoxal from all but the guanine residues (Bender and Davidson 1976, Broude and Budowsky, 1971). Glyoxal treated DNA was mixed with a dBrU label and spread from 20% formamide at 25° or from 40% formamide at 4°. The glyoxal treated DNA remained extended and traceable under

such spreading conditions, although the contrast was usually poor and the grids often did not pick up any DNA molecules or cytochrome film. The dBrU labels used in these studies were linear duplex DNA molecules of about 3 kb (R1 cut pMB8 (Rodrigeuez et al., 1976)) with a poly(dBrU) tail of about 200 bases polymerized at one or both ends. Figure 5A shows a pMB8-poly(dBrU) label attached to a short poly(dA) site in single stranded M6 DNA. In Figure 5B, the label is attached by poly(dBrU) at both of its ends to two adjacent poly(dA) sites in M6.

The final map of short poly(dA) tracts in M6 is shown in Figure 1 below the M6 restriction map. There are easily labeled poly(dA) tracts at the vector-insert boundaries (0.0 and 6.0 kb) which are derived from the A:T tails synthesized in vitro in the process of the plasmid construction. Since the Dictyostelium DNA inserts were extended with poly(dA) at their 3' ends and the vector with poly(T), the plus strand must have A's at 0.0 kb and T's at 6.0 kb, and these tracts can be used to orient the strand specificity of the Dictyostelium poly(dA) sites. There are three easily labeled Dictyostelium poly(dA) tracts at 0.3, 0.9, and 2.5 kb on M6. The 0.3 kb site is frequently labeled in the same molecule as the 0.0 kb site (these are the two sites labeled in Figure 5B) and so this site must also have A's in the plus strand. The site at 0.9 kb must have T's in the plus strand because it is often seen paired with the 0.0 or the 0.2 kb sites to form a small loop (as seen near the middle of the M6 DNA in Figure 5A). This and other secondary structure

features in these molecules must be due to A:T base pairs since the G residues are blocked by glyoxal. The 0.9 kb site also pairs with the 2.5 kb site (the one labeled in Figure 5A) and so the latter must be poly(dA) in the plus strand. A fourth internal site was found at 4.2 kb but it is probably very short because it is only very rarely labeled and does not readily pair with any of the other sites. It was occasionally seen labeled in molecules where the 2.5 kb site was also labeled; thus the plus strand has T's. These poly(dA) sites were observed on DNA from M6 cut once at the Bam HI site or from isolated EcoRI and Hind III restriction fragments. Other cases of pairing between sites or double labeling on single M6 molecules confirm the positions and strand assignments given above.

Spontaneous deletions

The bacterial strain carrying the M6 plasmid spontaneously generated deletions in the M6 DNA, even though the HB101 strain into which M6 had been transformed was recA. The strain was therefore frequently subcloned to eliminate these deletions from the M6 DNA preparations, but in a few cases where the deleted DNA was encountered, it was examined to locate the deletion. A preparation of normal M6 DNA contaminated with DNA having a small deletion (about 100 bases) gave a doublet of the largest band after digestion with either EcoRI or Hind III. There was not a doublet of the largest Hae III band (-0.35 to 2.1 kb) so the deletion is presumably in the pMB9 portion of M6. A second deletion was detected as an anomalously large band in a Hind III digest. It was mapped by heteroduplexing with the normal

Hind III fragments and shown to be missing the region from 0.0 to 2.5 kb on the M6 insert. These coordinates are both at or near poly(dA) sites on the plus strand, and so the deletion probably occurred by recombination within these short poly(dA) sequences.

## DISCUSSION

It should be emphasized that there are several different but related actin-complementary DNA sequences in Dictyostelium and there are probably related but distinct sequences of actin mRNA (McKeown el al., following paper). There is no evidence that the messages represented in the cDNA plasmids A1 and B1 were actually transcribed from the M6 actin gene, or, indeed, that any message is ever transcribed from this particular sequence. Any of the observations made below concerning the sequence organization of M6 may be irrelevant if M6 is not an actual coding region.

It is somewhat surprising that the message-complementary region of M6 (defined by the cDNA plasmids) is co-terminal with the whole actin repeat (defined by H6 homology) within experimental error. Since the left end of the plus strand of the A1 cDNA ends in poly(A), it defines the 3' end of the message, but the B1 cDNA sequence does not necessarily include the 5' end of the message. Actin mRNA electrophoresed on denaturing agarose gels, migrates as two bands of 1.3 and 1.4 kb (Kindle et al. preceding paper). Since the M6 actin repeat measures only 1.1 kb, either there is error in the length measurements or the M6 actin repeat is shorter, by 200 to 300 bases at the right end, than the actin gene that makes the mRNA. In any case there is no evidence for

repeated DNA sequences on either side of the mRNA complementary region which might be included in a larger nuclear precursor to actin message or which might be used as sites of transcriptional control (Britten and Davidson, 1969). Also there are no large insertions of non-message DNA within the message-complementary region, as have been reported for adenovirus DNA (Berget el al., 1977), SV40 DNA (Celma et al., 1977) and Drosophila ribosomal DNA (Glover and Hogness, 1977).

Short poly(dA) tracts occur in a wide variety of eukaryotes (Mol and Borst, 1976). The average length of the poly(dA) tracts in Dictyostelium is about 25 base pairs and they occur about once per 3 kb of DNA duplex (Jacobson et al., 1974); the slime mold poly(dA) sequences are slightly longer and much more frequent than in most other organisms (Mol and Borst, 1976). Their function is unknown, but it has been suggested that nuclear transcripts might end in a short poly(dA) tract which is then elongated before transport to the cytoplasm (Jacobson et al., 1974). Sequence data on the Drosophila 5S RNA precursor (Rubin and Hogness, 1975) and the yeast 5S RNA and tRNA-phe coding regions (W.J. Rutter, personal communication) suggests that poly(dA) tracts in the coding (minus) strand are transcription termination signals.

In M6 there is a poly(dA) tract in the plus strand approximately 400 bases beyond the 3' end of the message. If it is included in a nuclear transcript, it must be cleaved off before the long poly(A) is added. It is in the opposite strand orientation (T's in the coding strand) as the poly(dA) tract at the site

of 5S RNA termination. There is a second poly(dA) site near the right end of the actin repeat. This poly(dA) appears to be shorter, or to have more non-adenine bases within the A rich tract than the other sites mapped in M6. Because it was rarely labeled, it is not very accurately mapped relative to the right end of the actin repeat, and its possible significance is further confused by the aforementioned uncertainty about the 5' end of the actin message.

It will be important to define now which actin message comes from which actin repeat. Direct DNA sequencing will allow rapid and sensitive comparisons. The Hind III site at 4.0 kb in M6 and the comparable sites in H6 and B1 are the logical places to start.

## Appendix: Theory and Practice of Electron Microscopic Labels for Short Poly(dA) Sequences

There are two major difficulties in labeling poly(A) sequences in DNA and RNA for visualization in the electron microscope. First hybrids between poly(A) and poly(T) or poly(U) are disrupted by the denaturating conditions usually used to extend single stranded nucleic acids for microscopy, and second, polymers of T or U are not easy to see, particularly if they are short. The second problem was overcome by polymerizing poly(T) onto some easily recognized duplex DNA, using the enzyme terminal transferase (Carbon et al., 1975, Bender and Davidson, 1976). The first, problem could be solved for long A:T hybrids by pretreating the molecule of interest with glyoxal to keep it extended in low

formamide spreading conditions.

The stability problem is much more severe for labeling short poly(dA) tracts in eukaryotic DNA, which are only 20 to 25 bases long (Mol and Borst, 1976). The difference in $T_m$ between hybrids of poly$(dA)_{20}$:poly$(T)_{300}$ and poly$(dA)_{300}$:poly$(T)_{300}$ has been measured to be about 25° (in 0.15 M NaCl, 0.015 M sodium citrate) (Cassani and Bollum, 1967). Since long A:T hybrids are not stable in spreading solutions with more than about 40% formamide (Bender and Davidson, 1976), poly$(dA)_{20}$:poly(T) hybrids would not be stable in a 20% formamide spreading solution, which we take as the operational lower limit of the formamide Kleinschmidt technique. It was suggested by Tod Miles (personal communication) that poly(rBrU) or poly(dBrU) could be used to make more stable hybrids with the short poly(dA)'s. Long poly(dBrU):poly(dA) duplexes have a $T_m$ about 10° higher than that for poly(T):poly(dA) at pH 7.0 (Riley and Paul, 1971). It was found that the calf thymus terminal transferase will polymerize dBrUTP almost as rapidly as TTP onto an oligo $(dT)_6$ primer or a nicked duplex circular SV40 molecule. Poly(rBrU):poly(rA) hybrids have an even higher $T_m$ (Riley and Paul, 1970), but rBrU cannot be so easily polymerized onto duplex DNA.

The stabilities of hybrids between short A polymers and long dBrU polymers were measured directly by optical melting experiments. Long poly(dBrU) chains were polymerized onto oligo$(dT)_6$ primers with terminal transferase. The reaction products from 10-20 h reactions were chromatographed on Sepharose 4B, and the excluded poly(dBrU) was collected. Oligomers of $(rA)_n$ (actually $rA_nC$, generously provided by Olke Ullenbeck) with chain

lengths n=12, 17, 23 or 28 were mixed with an equal mass of poly(dBrU), and the mixtures were dialized against 0.1 M NaCl, 0.01 M TES, 0.001 M EDTA at various pH's. The mixtures were melted in a Gilford model 2000 recording spectrophotometer fitted with a linear temperature programmer. The sample was heated at 0.5°/min and the absorbance was measured at 260 nm. Figure 6 is a plot of $T_m$ against poly(rA) chain length, for poly(dBrU) hybrids at pH 7.0 and 8.0 and for poly(T) at pH 7.5. (The $T_m$'s for rA:dT hybrids were equivalent at pH 6.5 and pH 8.5, where measured.) The $T_m$ for long rA:dT hybrids was 61.5°; the $T_m$'s for shorter hybrids agree well with the $T_m$ depressions predicted by the approximate relationship $\Delta T_m = 500/n$ where n is the duplex length (Hayes et al., 1970). The $T_m$ of the dBrU hybrids at pH 7.0 is about 15° higher than that for the dT hybrids for three different duplex lengths. The effect of formamide on rA:dBrU hybrid stability was measured by melting the short hybrids in a simulated spreading solution (30% formamide, 0.1 M TES, 0.01 M EDTA pH 7.3); the $T_m$'s for duplex lengths greater than 20 base pairs are still well above room temperature. Hybrids of T or dBrU with short poly(dA)'s should be slightly more stable than those with poly(rA)'s as measured here (Riley et al., 1966).

The above analysis neglects the dependence of $T_m$ on concentration, which becomes important for short duplex lengths. As predicted and measured by Applequist and Damle (1965), the relationship between the melting temperatures $T_m$ and $T_m'$ for

corresponding polymer concentrations C and C' is

$$\frac{1}{T_m'} = \frac{1}{T_m} + \frac{R}{n\Delta H} \ln \frac{C'}{C}$$

where n is the duplex length and $\Delta H$ is the enthalpy per base pair. For an A:dBrU duplex of 20 base pairs (assuming $\Delta H$ is 8 kcal/base pair), the $T_m$ would decrease by about 2.6° for every 10-fold decrease in polymer concentration. In the spreading solutions the concentrations of poly(A) and poly(dBrU) are only about 10 ng/ml or about 1000-fold less than the concentrations used for the optical melting experiments. Thus, the $T_m$ for a 20 base A:dBrU duplex would be about 8° lower than that predicted by Figure 6. In a spread from 20% formamide at room temperature (as used in the poly(dA) mapping experiments with M6) we would predict a minimum stable A:dBrU duplex length of about 15 base pairs. This estimate may be optimistic because there is a further dilution in polymer concentration when the 50 µl spreading solution is spread onto the 100 ml of hypophase buffer. The molecules of interest are then trapped at the air-water interface, and so concentration effects become impossible to predict.

In practice, four sites in the 6 kb Dictyostelium insert can be labeled by this technique. The sizes of these poly(dA) tracts have not yet been measured, but the average size of poly(dA) tracts in Dictyostelium is about 25 bases (Jacobson et al., 1974). The labeling efficiency is somewhat low; we estimate each individual poly(dA) site has a 10-20% probability

of being labeled. The linear duplex pMB8 molecules with dBrU tails are considerably better for labeling internal poly(dA) sites than are tailed SV40 or trypanisome mitochondrial DNA circles because the dBrU tails are only at the ends of the pMB8 molecules, and so accidental crossovers between pMB8 linears and the molecules of interest are not confused with sites of labeling.

## EXPERIMENTAL PROCEDURES

### Materials

The construction of the recombinant plasmids and the preparation of plasmid DNA are described elsewhere (Kindle et al., preceding paper, Rowenkamp and Firtel, manuscript in preparation). pMB9 DNA was a gift of J. Douglas Engel, Leishmania tarentolae mitochondrial DNA was a gift of Larry Simpson, and pMB8 DNA was a gift of Mavis Shure. EcoR1 restriction endonuclease was purified according to Greene et al. (1974). Other restriction enzymes were purchased from New England Biolabs, and digestions were routinely done for 1 h at 37° in the assay buffer recommended by the supplier. Agarose (ME grade) was purchased from Seakem. Terminal deoxynucleotidyl transferase and 5-bromo deoxyuridine triphosphate (dBrUTP) were purchased from P.L. Biochemicals, and $^{3}H$ dBrUTP was from New England Nuclear.

### Gel Electrophoresis

Restriction fragments were resolved by electrophoresis on 0.8 or 1.0% agarose gels run in E buffer (0.04 M Tris, 0.005 M Na acetate, 0.001 M $Na_2$ EDTA, pH 7.4). The dimensions of the gels

were 18 x 14 x 0.4 cm and they were run at constant voltage (2 to 3 volts per centimeter). Gels were stained with 0.1 μg/ml ethidium bromide. Molecular lengths of various fragments were estimated by comparison to restriction fragments of bacteriophages λ and pM2; the sizes are accurate to within about 5% in the 1 - 5 kb range. For preparative work, samples were loaded in wide (3.5 cm) lanes at the top of the gel. After electrophoresis, only the edges of each lane were cut off and stained so that the bands of interest could be located without UV-induced nicking of the DNA. The bands were cut out and eluted by electrophoresing the DNA into a dialysis bag or by the freeze-squeeze method (Thuring et al. 1975).

Heteroduplex Mapping

Heteroduplexes were prepared and spread from a hyperphase containing 55% formamide according to Davis et al. (1971). In the triple heteroduplexes as in Figure 2, the two full length plasmid molecules were first annealed together, and then a restriction digest containing the third DNA fragment of interest (and other fragments) was heat denatured, quenched, and added to the heteroduplex mixture, and the mixture was annealed a second time before spreading. ϕX single and double stranded DNAs were used as internal length standards to calibrate length measurements of heteroduplex molecules.

Poly(dA) Mapping

Duplex DNA molecules were extended with short tails of poly(dBrU) by incubation with terminal transferase at 37° for

1h (in 50 mM TES, 40 mM kCl, 3 mM $CoCl_2$, and 3 mM dBrUTP, with 1000 enzyme units/ml and up to 0.5 mg/ml of duplex DNA). Tail length could be measured by incorporation at $^3H$ dBrUTP or by direct visualization of the tails in the electron microscope; a one hour reaction produced average tail lengths of 100 to 200 bases. Closed circular duplex DNAs could be used as primers without prior nicking; they are converted to nicked circular molecules in the course of the reaction, presumably by nuclease contamination in the commercial terminal transferase.

Duplex DNA molecules with poly(dA) tracts of interest were melted at 80° in 1 M glyoxal, 0.01 M potassium phosphate, pH 7.0; the solution was allowed to cool slowly to 37°, and then it was incubated at 37° for 1h. The glyoxal-treated DNA was then dialyzed overnight against 10 mM TES, 1 mM EDTA pH 7.0. Hybridization of the dBrU labels to the poly(dA) tracts and spreading of the hybrids were performed as previously described (Bender and Davidson, 1976) except that everywhere TES buffer at pH 7.0 was substituted for Tris buffer at pH 8.5. In spreads from a 20% formamide hyperphase, the hypophase buffer concentration was raised to 0.02 M TES, 0.002 M EDTA to maintain isodenaturing conditions.

## ACKNOWLEDGMENTS

We are grateful to Walter Rowenkamp, Doug Engel, Mavis Shure, Larry Simpson, and Olke Ullenbeck for providing materials essential for these studies.

FIGURE LEGENDS

Figure 1. Maps of the Recombinant Chromosomal DNA Plasmids M6 and H6, and the cDNA Plasmids A1 and B1. In each case the plasmid is shown aligned with the homologous portions of its vector. Only the portions of the vector sequences adjacent to the Dictyostelium inserts are shown; the non-paired portions of the vector on the right hand side of the M6, H6, and A1 maps indicate sequences of the vector which were deleted in the process of plasmid construction. The cleavage sites for various restriction endonucleases are marked and regions of homology among plasmids are indicated by horizontal bars below each map. Zig-zag lines indicate regions of poly(dA):poly(T) sequences at the vector-insert junctions and in the region of the A1 cDNA copied from the actin mRNA poly(A). The insert regions of all plasmids are drawn with thicker lines. The second M6 map shows the short poly(dA) tracts in that plasmid; the strand to which the actin message hybridizes is defined as the minus strand.

Figure 2. M6/H6/pMB9 Triple Heteroduplexes. In (A), M6 and pMB9 (both cut at the single Bam HI site) were heteroduplexed together, and then the molecules were hybridized with the EcoRI-BamHI fragment of H6. In (B) a H6/pMB9 heteroduplex (again, both cut at Bam HI site) was paired with two EcoR1 fragments of M6, one at each actin repeat. Both micrographs are at the same magnification. Each heteroduplex is diagrammed schematically below the micrograph. The

region marked Δ in each case is the 300 base sequence of pMB9 (extending from the EcoR1 site almost to the Hind III site) which was deleted in the construction of M6 and H6.

Figure 3. Heteroduplexes between M6 and Actin cDNA plasmids. (A) shows a B1/pMB9 heteroduplex. The G:C tails flanking the cDNA insert are paired in this molecule and thus hold the insert in a loop, as is indicated in the schematic drawing. (B) shows a B1/M6 heteroduplex. The B1 cDNA insert is completely paired with M6 in the actin repeat region. The small loop in the unpaired M6 DNA is due to pairing between the A:T tail region of the vector-insert junction with a short internal poly(dA):poly(T) site. (C) and (D) are A1/M6 heteroduplexes. (C) is exactly analogous to (B) except that the actin homology region is shorter. In (D), a poly(dA) sequence just to the left of the actin hybrid region is paired to a poly(T) sequence of M6 at its left vector-insert junction (assuming that the A1 and M6 strands are plus and minus strands, respectively). All plasmids were cut at their single Bam HI sites before heteroduplexing; all micrographs are at the same magnification.

Figure 4. Labeling of Poly(dA) site in A1. DNAs from the A1 and pMB9 plasmids (both cut at Bam HI site) were heteroduplexed, 800 base pair duplex circles with

poly(dBrU) tails were added to the mix and annealed briefly, and the mixture was spread from 50% formamide The dBrU label is attached to the A1 DNA at a site near the left end of the A1 cDNA insert. The position of the A1 cDNA insert was measured independently from A1/pMB9 heteroduplex molecules like the B1/pMB9 molecule of figure 3A.

Figure 5. Labeling Short Poly(dA) Tracts in M6. M6 DNA was cleaved at the Bam HI site, glyoxal-treated, mixed with linear duplex pMB8 molecules with dBrU tails at one or both ends, and spread from 20% formamide. In (A), one end of a pMB8 label is attached to a M6 poly(dA) site near the actin gene (at 2.5 kb on the M6 map). The small loop in the M6 DNA is equivalent to the small loop in figure 3B. In (B), both ends of pMB8 are attached to two adjacent poly(dA) tracts in M6 (at 0.0 and 0.3 kb). Both micrographs are at the same magnification.

Figure 6. Melting Temperatures of Short rA:dT or rA:dBrU Hybrids. Short defined polymers of rA were mixed with an equal mass of long poly(T) or poly(dBrU) in 0.1 M NaCl, 0.01 M TES, 0.001 M EDTA, and the $T_m$'s of the hybrids were measured optically. Two rA:dBrU hybrid lengths were also melted in a buffer equivalent to a low formamide spreading solution (30% formamide, 0.1 M TES, 0.01 M EDTA, pH 7.3). The $T_m$ for hybrids between long poly(rA) and long poly(dT) was 61.5°. The predicted

$T_m$'s for shorter rA:dT hybrids are indicated by the dashed line (using the approximate equation for the melting point depression $\Delta T_m = 500/n$, where n is the A:T duplex length.)

REFERENCES

1. Applequist, J., and Damle, V. (1965). Thermodynamics of the Helix-Coil Equilibrium in Oligoadenylic Acid from Hyperchromicity Studies. J. Am. Chem. Soc. 87, 1450-1458.

2. Bender, W., and Davidson, N. (1976). Mapping of Poly(A) Sequences in the Electron Microscope Reveals Unusual Structure of Type C Oncornavirus RNA Molecules. Cell 7, 595-607.

3. Berget, S.M., Moore, C. and Sharp, P.A. (1977). A Non-Complementary Sequence at the 5'-Terminus of Late Adenovirus 2 mRNA. Proc. Nat. Acad. Sci. USA, in press.

4. Britten, R.J., and Davidson, E.H. (1969). Gene Regulation for Higher Cells: a Theory. Science 165, 349-357.

5. Broker, T.R., Soll, L., and Chow, L.T. (1977). Underwound Loops in Self-renatured DNA Can Be Diagnostic of Inverted Duplications and Translocated Sequences. J. Mol. Biol. 113, 579-589.

6. Broude, N.E., and Budowsky, E.I. (1971). The Reaction of Glyoxal with Nucleic Acid Components III. Kinetics of the Reaction with Monomers. Biochim. Biophys. Acta 254, 380-388.

7. Carbon, J., Shenk, T.E., and Berg, P. (1975). Construction *in vitro* of mutants of Simian Virus 40: Insertion of a Poly(dA·dT) Segment at the *Hemophilus Parainfluenza* II Restriction Endonuclease Cleavage Site. J. Mol. Biol. 98, 1-15.

8. Cassani, G., and Bollum, F.J. (1967). Oligodeoxynucleotide-Polynucleotide Interactions. Adenine-Thymine Base Pairs. J.Am.Chem.Soc. 89, 4798-4799.

9. Celma, M.L. Dhar, R., Pan, J. and Weissman, S.M. (1977). Comparison of the Nucleotide Sequence of the Messenger RNA for the Major Structural Protein of SV40 with the DNA Sequence Encoding the Amino Acids of the Protein. Nucl. Acid Res. 4, 2549-2559.

10. Davis, R.W., Simon, M., and Davidson, N. (1971). Electron Microscope Heteroduplex Methods for Mapping Regions of Base Sequence Homology in Nucleic Acids. Meth in Enzymol. 21, 413-428.

11. Efstratiadis, A., Kafatos, F.C., and Maniatis, T. (1977). The Primary Structure of Rabbit β-Globin mRNA as Determined from Cloned DNA. Cell 10, 571-586.

12. Garrels, J.I., and Gibson, W. (1976). Identification and Characterization of Multiple Forms of Actin. Cell 9, 793-805.

13. Glover, D.M., and Hogness, D.S. (1977). A Novel Arrangement of the 18S and the 28S Sequences in a Repeating Unit of Drosophila melanogaster rDNA. Cell 10, 167-176.

14. Greene, P.J., Betlach, M.D., Goodman, H.M. and Boyer, H.W. (1974). In Methods in Molecular Biology: DNA Replication, 7, R.B. Wickner, ed. (New York: Marcel Dekker), p. 87.

15. Hayes, F.N., Lilly, E.H., Ratliff, R.L., Smith, D.A., and Williams, D.L. (1970). Thermal Transitions in Mixtures of Polydeoxyribonucleotides. Biopolymers 9, 1105-1117.

16. Jacobson, A., Firtel, R.A., and Lodish, H.F. (1974). Transcription of Polydeoxythymidylate Sequences in the Genome of the Cellular Slime Mold, Dictyostelium discoidium. Proc. Nat. Acad. Sci. USA 71, 1607-1611.

24. Thomas, M., Cameron, J.R., and Davis R.W. (1974). Viable Molecular Hybrids of Bacteriophage Lambda and Eukaryotic DNA. Proc. Nat. Acad. Sci. USA 71, 4579-4583.

25. Thuring, R., Sanders, J.P.M., and Borst, P. (1975). The Freeze-Squeeze Method for Recovering Long DNA from Agarose Gels. Anal. Biochem. 66: 213-220.

26. Wesley, R.D., and Simpson L. (1973). Studies on Kinetoplast DNA II. Biophysical Properties of Minicircular DNA from Leishmania tarentalae. Biochim. Biophys. Acta 319, 254-266.

17. Mol, J.N.M., and Borst, P. (1976). The Binding of Poly(rA) and Poly(rU) to Denatured DNA. II Studies with Natural DNAs. Nucl. Acids Res. 3, 1029-1051.

18. Riley, M., Maling, B. and Chamberlin, M.J. (1966) Physical and Chemical Characterization of Two- and Three-stranded Adenine-Thymine and Adenine-Uracil Homopolymer Complexes. J. Mol. Biol. 20, 359-389.

19. Riley, M., and Paul, A.V. (1970). Two- and Three-stranded Complexes containing Homopolymers Polyriboadenylic Acid and Polyribobromouridylic Acid. J. Mol. Biol. 50, 439-455.

20. Riley, M., and Paul, A.V. (1971) Properties of Synthetic Polydeoxyribonucleotide Complexes Containing Adenine and Bromouracil. Biochemistry 10, 3819-3825.

21. Rodrigeuez, R.L., Bolivar, F., Goodman, H.M., Boyer, H.W. and Betlach, M.C. (1976). Construction and Characterization of Cloning Vehicles. in Molecular Mechanisms in the Control of Gene Expression ICN-UCLA Symposium on Molecular and Cell Biology 5, D.P. Nierlich, W.J. Rudder and C.F. Fox, ed. (Academic Press, New York) p. 471.

22. Rubin, G.M., and Hogness, D.S. (1975). Effect of Heat Shock on the Synthesis of Low Molecular Weight RNAs in Drosophila: Accumulation of a Novel Form of 5S RNA. Cell 6, 207-213.

23. Szybalski, W., Kubinski, H., Hradecna, Z. and Summers, W.C. (1971) Analytical and Preparative Separation of the Complementary DNA Strands. in Methods in Enzymology vol. 21, part D. L. Grossman and K. Moldave, ed., Academic Press, New York.

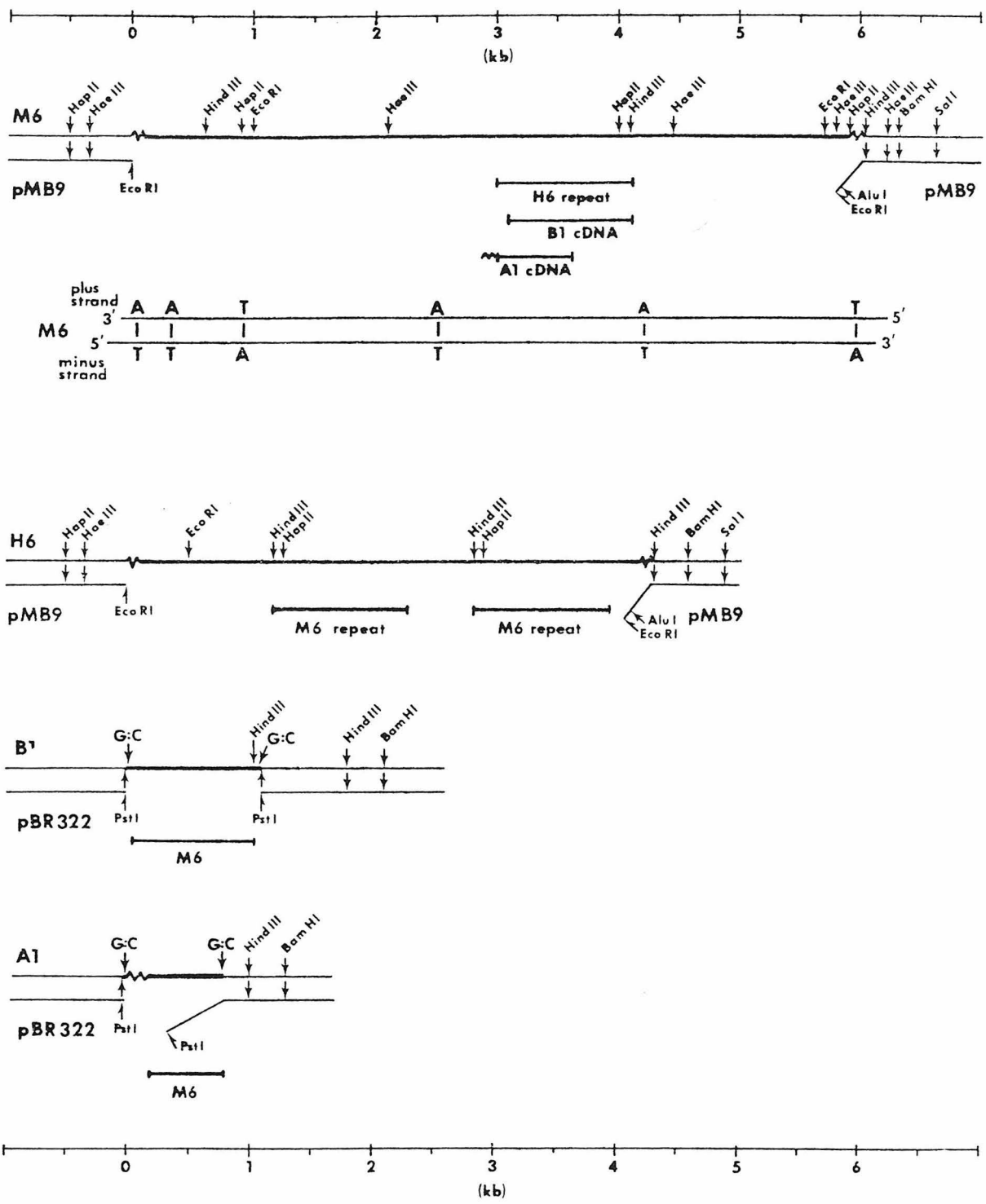

0
1
2
3
4
5
6
(kb)
M6
Hap II
Hae III
Hind III
Hap II
Eco RI
Hae III
Hap II
Hind III
Hae III
Eco RI
Hae III
Hap II
Hind III
Hae III
Bam HI
Sal I
pMB9
Eco RI
H6 repeat
B1 cDNA
A1 cDNA
Alu I
Eco RI
pMB9
plus strand
3'
5'
M6
5'
3'
minus strand
A A T A A T
T T A T T A
H6
Hap II
Hae III
Eco RI
Hind III
Hap II
Hind III
Hap II
Hind III
Bam HI
Sal I
pMB9
Eco RI
M6 repeat
M6 repeat
Alu I
Eco RI
pMB9
B1
G:C
Hind III
G:C
Hind III
Bam HI
pBR322
Pst I
Pst I
M6
A1
G:C
G:C
Hind III
Bam HI
pBR322
Pst I
Pst I
M6
0
1
2
3
4
5
6
(kb)

Fig. 1

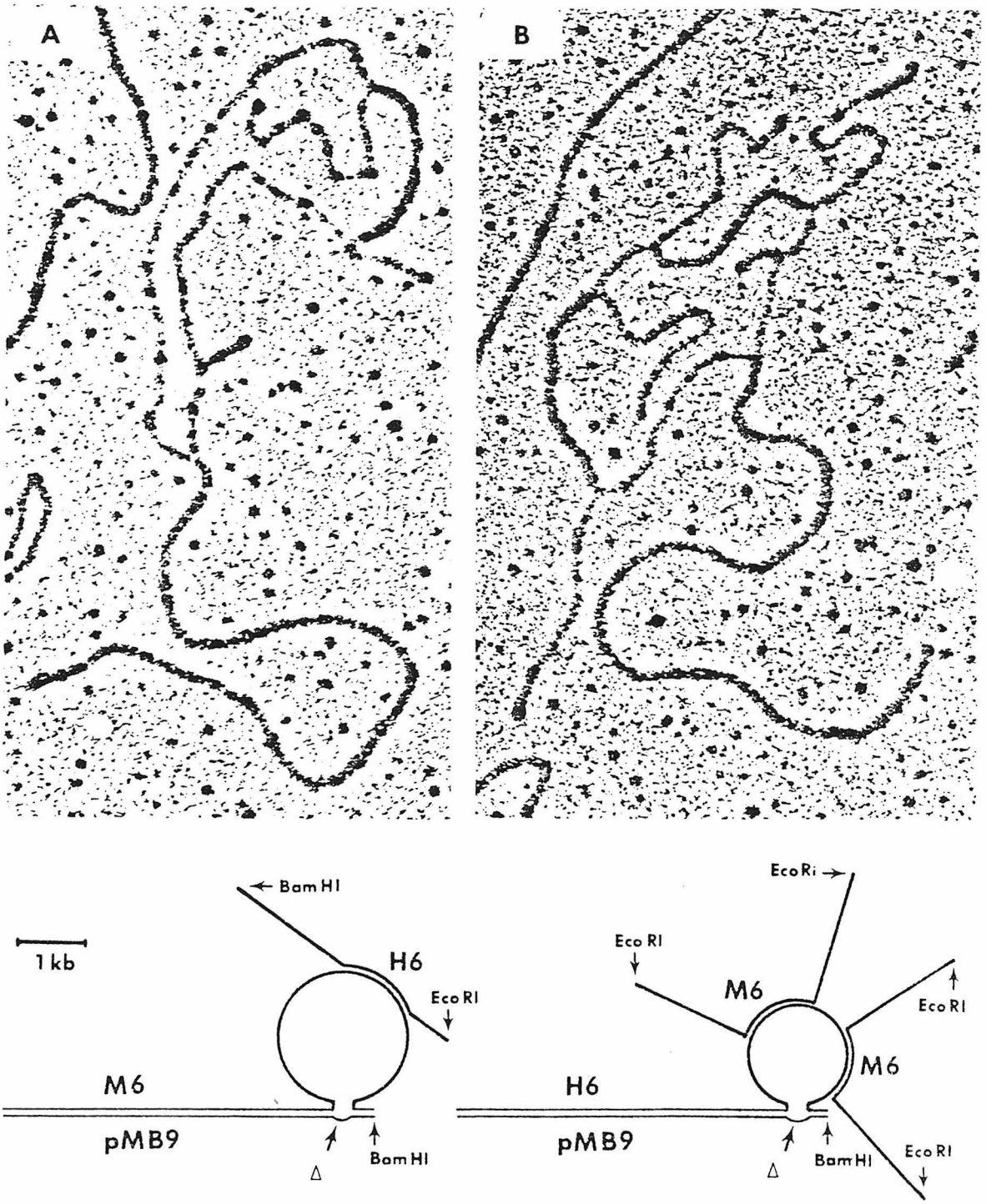

A
B
1 kb
BamHI
H6
EcoRI
M6
pMB9
BamHI
Δ
EcoRi
EcoRI
M6
EcoRI
M6
H6
pMB9
BamHI
EcoRI
Δ

Fig. 2

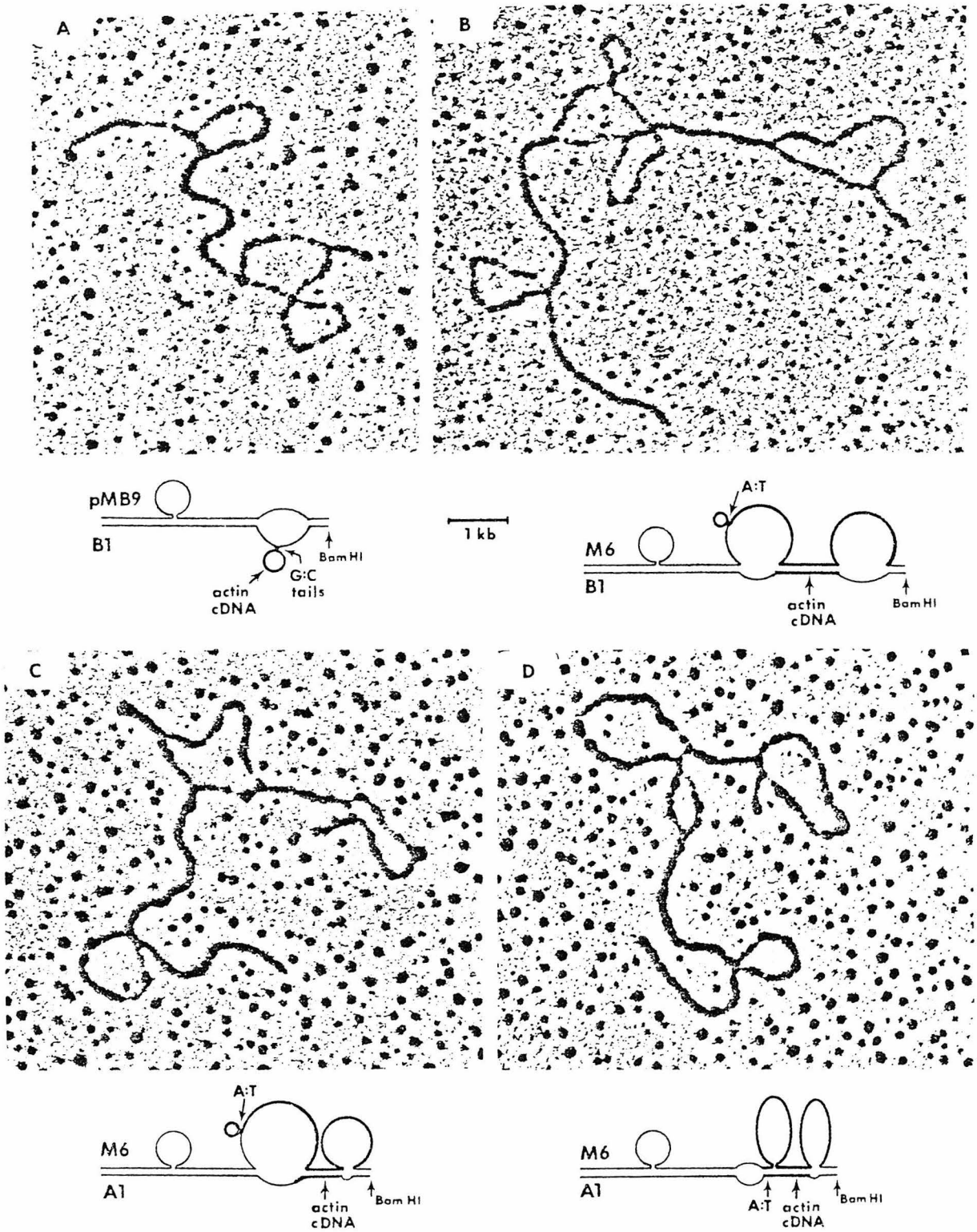

A
B
pMB9
B1
Bam HI
G:C
tails
actin
cDNA
1 kb
A:T
M6
B1
actin
cDNA
Bam HI
C
D
A:T
M6
A1
actin
cDNA
Bam HI
M6
A1
A:T
actin
cDNA
Bam HI

Fig. 3

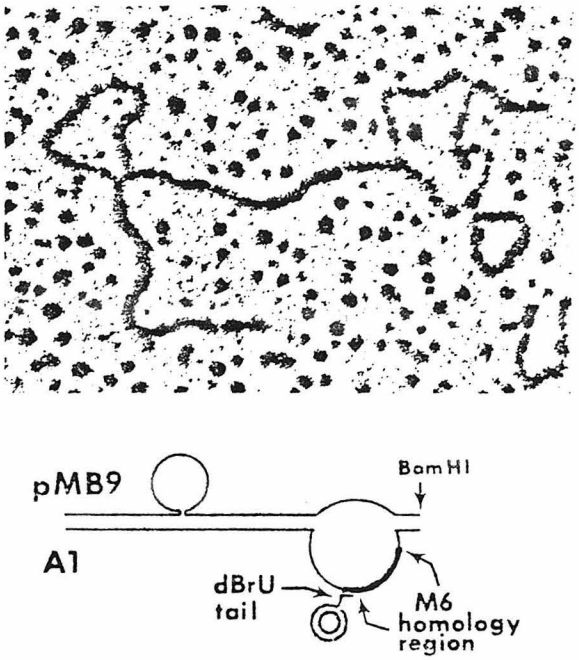



Fig. 4

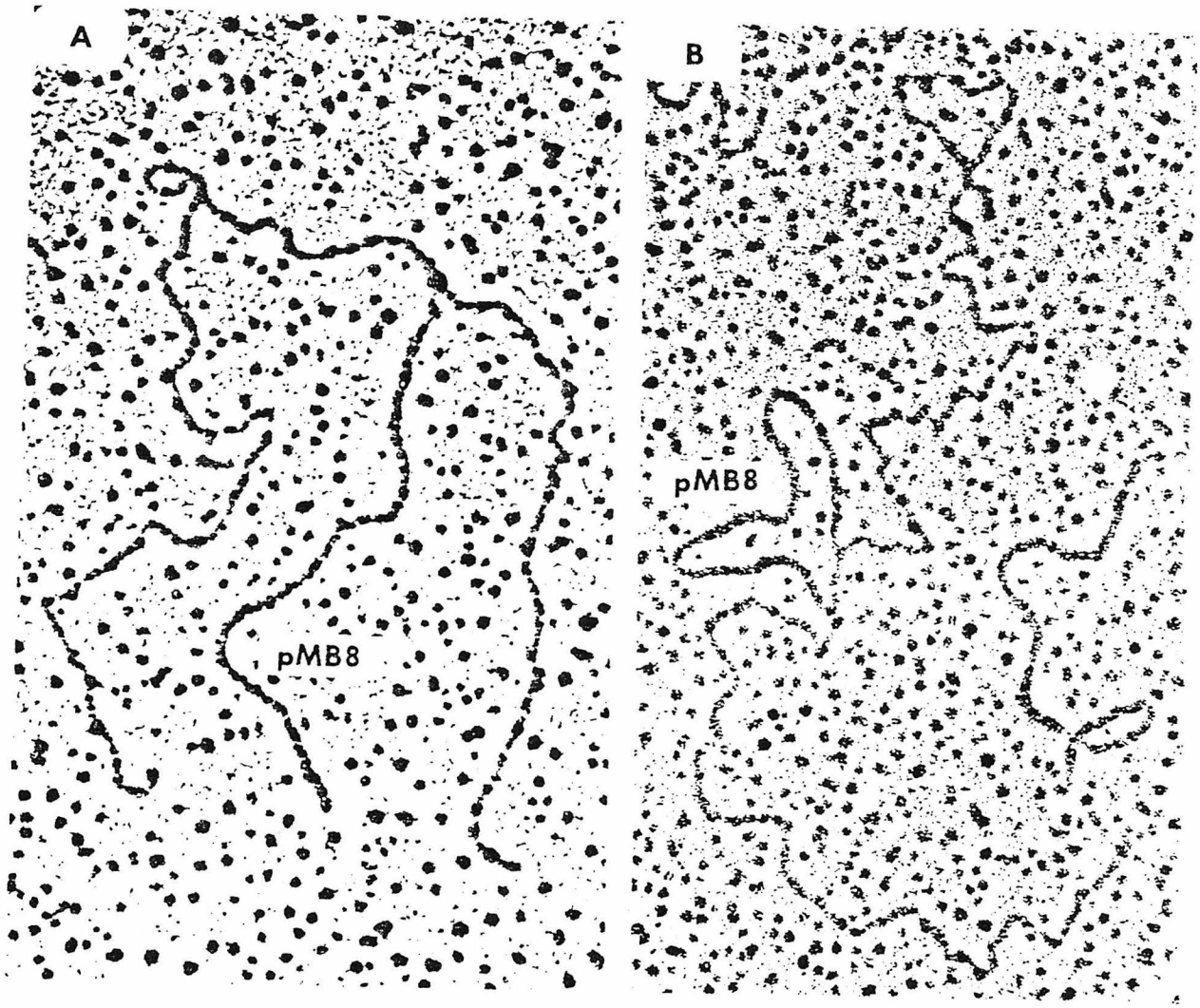



Fig. 5

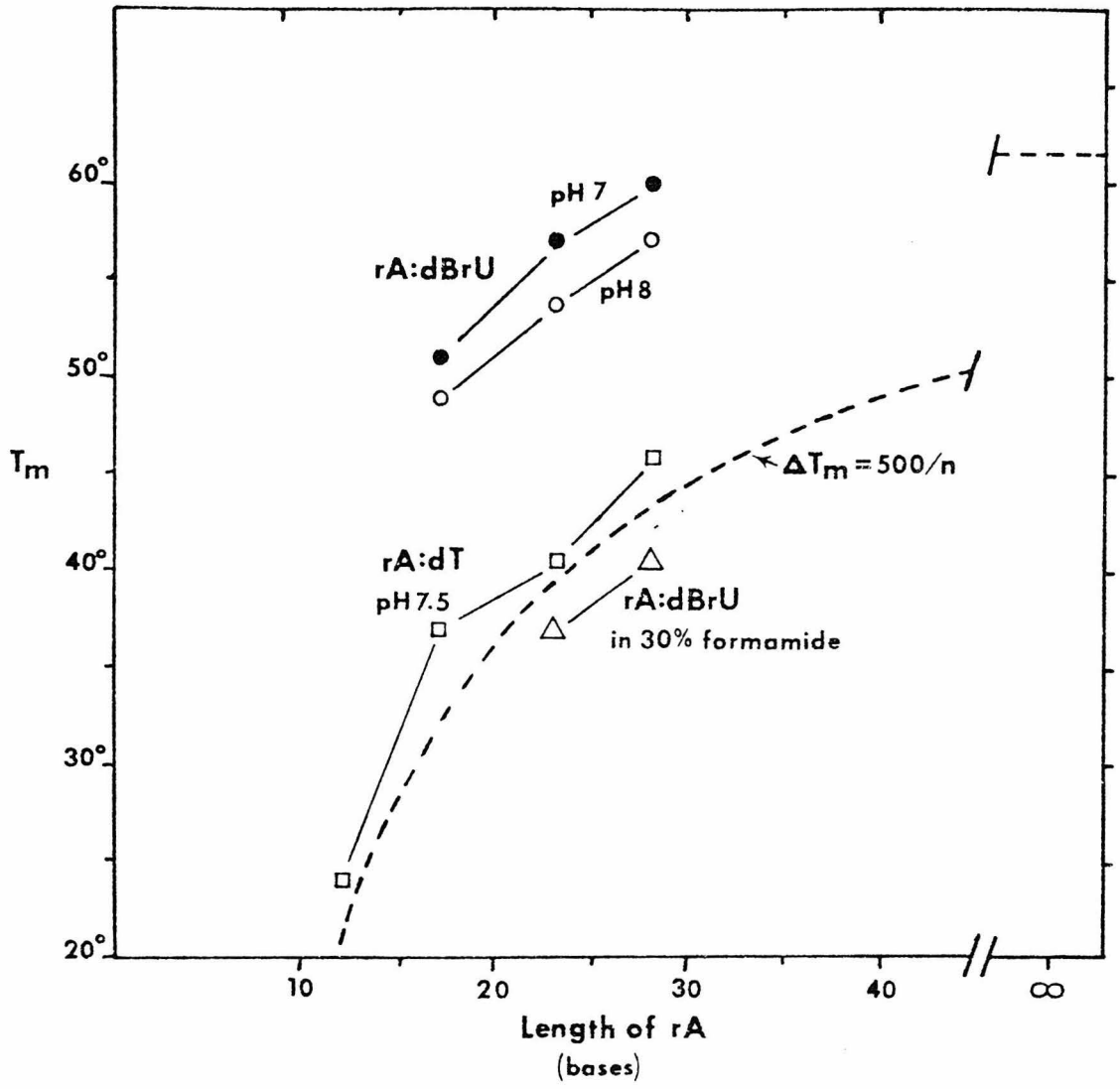

60°
50°
40°
30°
20°
$T_m$
pH 7
rA:dBrU
pH 8
$\Delta T_m = 500/n$
rA:dT
pH 7.5
rA:dBrU
in 30% formamide
10
20
30
40
∞
Length of rA
(bases)

Fig. 6