Proc. Natl. Acad. Sci. USA
Vol. 75, No. 7, pp. 3225–3229, July 1978
Biochemistry

Nucleotide sequence homology at the 3′ termini of RNA from vesicular stomatitis virus and its defective interfering particles

(replicase initiation sites/autointerference/generation of defective viruses)

JACK D. KEENE*, MANFRED SCHUBERT*, ROBERT A. LAZZARINI*, AND MARTIN ROSENBERG†

* Molecular Virology Section, Laboratory of Molecular Biology, National Institute of Neurological and Communicative Disorders and Stroke, National Institutes of Health, Bethesda, Maryland 20014; and † Laboratory of Molecular Biology, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20014

Communicated by Robert M. Chanock, May 5, 1978

ABSTRACT **Vesicular stomatitis virus (VSV) and defective interfering (DI) particle RNAs were labeled at their 3′ ends by using RNA ligase and cytidine 3′,5′-bis[^{32}P]phosphate. The RNAs were subjected to partial digestion with alkali and analyzed by oligonucleotide fingerprinting in two dimensions. VSV and DI particle RNAs have complete sequence homology for the first eight bases from the 3′ end. The following four positions contain three mismatched nucleotides in which guanosine residues in one strand are replaced by uridine residues in the other. There is again complete homology for the next five bases (positions 13–17). The locations of purine residues within the sequence were confirmed by partial digestion with RNase T1 and RNase U2 and separation by size on 20% acrylamide gels. The latter method also indicated that sequences of VSV and DI particle RNAs diverge beyond the 18th nucleotide from the 3′ termini.**

The replication of vesicular stomatitis virus (VSV) involves the production of a linear full-length (+) strand of RNA from the 10-kilobase-long genomic (−) strand. The (+) strand, in turn, serves as template for production of progeny (−) strands that are eventually encapsidated. This replicative process can be curtailed by a competing population of defective interfering (DI) particles that arise during high multiplicity or repeated undiluted passage of VSV in cultured cells (1). The mechanism of this interference of viral replication is poorly understood. However, because DI particles and VSV probably replicate by the same mechanism, it has been suggested that interference is the result of competition between VSV and DI particle templates for limited quantities of replicase (2).

This model requires that the replicase initiation sequences on VSV and its DI particle RNAs are located at the termini and are similar enough that both are recognized and copied by the same enzyme. We recently demonstrated (3) that VSV RNA and three dissimilar DI particle RNAs all possess the same 3′-terminal nucleotides (PypGpU-OH). The identity of this short stretch of nucleotides on four different particle RNAs suggests that they have a common function.

Recently, we analyzed additional sequences at the 3′ end of DI particle RNAs by annealing of the leader RNA of VSV, which is transcribed from the 3′ terminus of the genome (4), to the DI particle RNAs. The DI whose RNA is homologous to the 3′ portion of the genome (DI_{LT}) showed complete annealing to the full leader RNA (5), suggesting that the 3′ termini of VSV and DI_{LT} are identical. On the other hand, DI particles whose genomes were derived from the 5′ portion of the VSV genome formed only small and spurious RNA·RNA duplexes when annealed to the leader RNA (unpublished data). In addition, Perrault *et al.* (6) reported that RNA isolated from a DI particle self-anneals to a limited extent and that the short duplexes formed contained the 3′ and 5′ regions of the genome. However, the separated strands could only anneal to the 5′-terminal sequences of VSV RNA. They concluded that the 3′- and 5′-terminal sequences of the DI RNA are complementary to one another, but that the 3′ end of the DI particle RNA is not homologous to the 3′ end of VSV RNA. Leppert *et al.* (7) made similar observations with DI particle RNAs from Sendai virus-infected cells and concluded that only short stretches (10–20 nucleotides) of homology could exist between the 3′ ends of Sendai virus and its DI RNA and go undetected by their method of hybridization.

In this communication, we demonstrate that there is base sequence homology between the 3′ ends of VSV and DI particle RNAs with mismatched bases at defined points. The homologous regions span the first 17 nucleotides from the 3′ end. Beyond that point there is sharp divergence between the infectious and defective viral sequences. Thus, the terminal nucleotide sequences are different enough to exclude hybrid stability in the presence of single-strand-specific nuclease, yet similar enough to allow replicase initiation on both RNAs. These findings also indicate that the 3′ end of the DI particle RNAs are not derived directly from the 3′ end of VSV RNA but come either from somewhere else in the molecule or from an extragenomic source. The implications of these data with respect to the origin of DI particles and the replication of VSV are discussed.

MATERIALS AND METHODS

Virus Growth. BHK_{21} cells were grown as monolayers in Eagle's minimal essential medium containing nonessential amino acids and 10% fetal calf serum. Cells were infected with the Indiana serotype of VSV [Mudd–Summers strain (VSV_{MS})] alone or together with DI_{611} (8) or with the Indiana serotype of VSV [San Juan strain (VSV_{SJ})] alone or together with DI_T (9). The RNAs of these VSV strains have been shown by Clewley *et al.* (10) to differ significantly by oligonucleotide fingerprinting of complete RNase T1 products.

Virus and RNA Preparation. Products from the infected cells were clarified by centrifugation, precipitated with polyethelene glycol, and purified by isopycnic banding in linear gradients between 30% (wt/wt) glycerol and 50% (wt/wt) potassium tartrate with the SW 27 rotor (Beckman Instruments) at 20,000 rpm for 15 hr. Virus particles were further purified by isopycnic banding in 7–52% sucrose in the SW 27 rotor at 20,000 rpm for 15 hr and recovered by centrifugation in 1 mM

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviations: VSV, vesicular stomatitis virus; DI, defective interfering; DI_{LT}, DI whose RNA is homologous to the 3′ portion of the genome; VSV_{MS}, Mudd–Summers strain of VSV; VSV_{SJ}, San Juan strain of VSV.

Tris, pH 7.5/1 mM EDTA for 2 hr at 24,000 rpm in the SW 27 rotor. VSV RNA was prepared from purified standard virus; DI RNA was prepared from mixed infectious and defective particles. RNA species were resolved by sodium dodecyl sulfate/sucrose gradient centrifugation as described (9).

Preparation of 3′-Labeled RNA. Purified RNA was labeled at its 3′ end by using modifications of the methods of England *et. al.* (11). RNA ligase from T_4-infected *Escherichia coli* was incubated at a concentration of 830 units/ml in 50 mM *N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid (Hepes), pH 8.0/8 mM $MgCl_2$/3 mM dithiothreitol/5 μM ATP/67 nM RNA with 3 μM cytidine 3′,5′-bis[^{32}P]phosphate (New England Nuclear). After 20 min at 37°, the reaction was stopped with 0.4 M NaCl/10 mM Tris, pH 7.6/0.1% sodium dodecyl sulfate and chromatographed over Sephadex G-150 in the same buffer. Radioactive RNA was recovered from the excluded volume and precipitated with ethanol in the presence of 50 μg of *E. coli* tRNA.

Partial Digestion and Electrophoresis. Terminally labeled RNAs were subjected to partial digestion by using modifications of the method of Donis-Keller *et al.* (12). RNase T1 was added to a final ratio of 0.01 unit/μg of RNA in the presence of 20 mM sodium citrate, pH 5.0/1 mM EDTA/7 M urea/0.02% bromophenol blue and xylene cyanol FF. Samples were heated to 100° for 2 min, cooled before addition of enzyme, and then incubated with enzyme at 50° for 15 min. RNase U2 was used under the same conditions except that no urea was added and the ratio of enzyme to RNA was 0.1 unit/μg. Partial alkaline digestion was in 50 mM $NaHCO_3$, pH 9.0/1 mM EDTA at 90° for 30 min. The digests were subjected to electrophoresis in slab gels containing 20% acrylamide or analyzed in two dimensions by oligonucleotide fingerprinting according to the method of Sanger *et al.* (13). Fragments were identified by autoradiography using Kodak X-Omat film.

RESULTS

Prior to use for sequence studies, a portion of each end-labeled RNA was digested for 18 hr in 0.5 M NaOH and analyzed by high-voltage electrophoresis on polyethyleneimine sheets (data not shown). In every case, >95% of the radioactivity was found to migrate with the optical density marker for Up, the terminal nucleotide of VSV and DI particle RNAs. This highly specific ligation indicates that the viral RNA was homogeneous and that extraneous 3′-hydroxyl termini were not generated during the incubation.

Fig. 1 shows the array of labeled fragments of RNA from VSV_{SJ} and the DI_T that arose from the San Juan parental stock. The first eight nucleotides yielded identical fingerprints, indicating complete homology over this tract.

When the same analysis was performed using VSV_{MS} and DI_{611} derived from this strain, an identical fingerprint was obtained (Fig. 2). We conclude that the RNAs from all four particles are identical for the first eight nucleotides from the 3′ end. The chromatograms shown in Figs. 1 and 2 were developed with solutions designed to maximally mobilize the labeled fragments. As a result, very small fragments at the terminus do not appear in these figures. However, other chromatograms that were developed with weaker homochromatographic solutions clearly demonstrated that the terminus was CpGpU in each case, as was shown previously (3).

The precise sequence of nucleotides on these fingerprints was determined by analysis of the relative position of each fragment. In the vertical dimension (homochromatography), each additional nucleotide moves a fragment's position toward the bottom of the chromatographic plate while the composition of the added base, in general, determines the relative position to the left- or right-hand side in the first dimension. Thus, addition of a C residue retards migration in the first dimension, resulting in an apparent shift to the left as shown in Figs. 1 and 2. Addition of a U residue, on the other hand, accelerates migration in

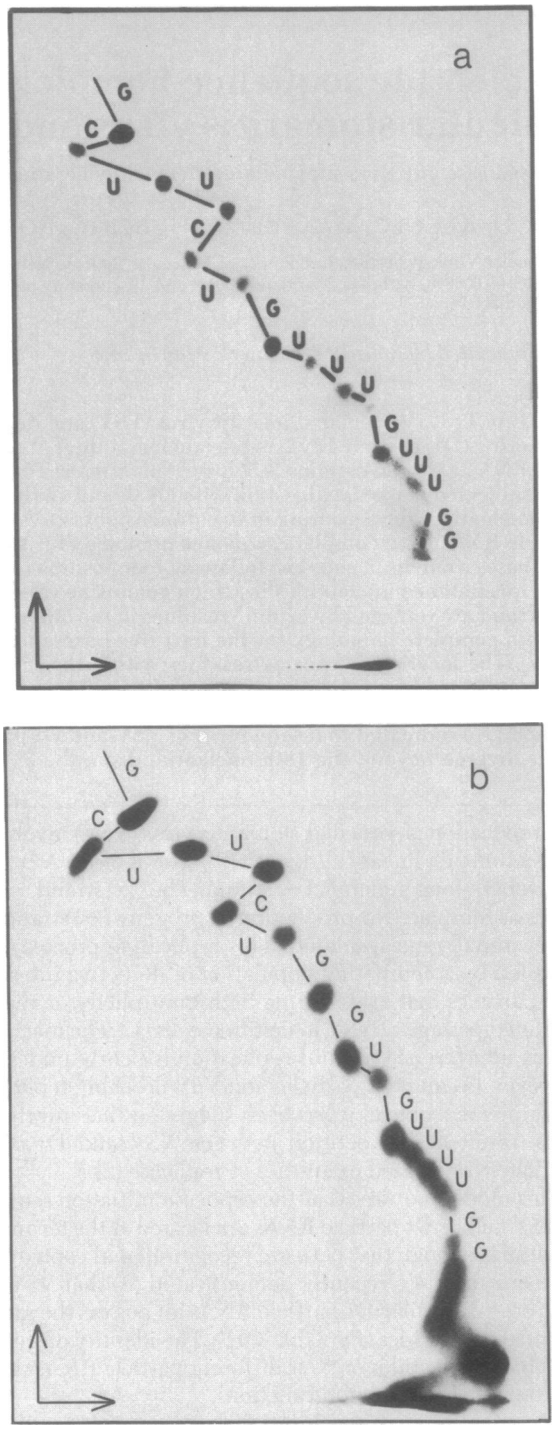

FIG. 1. Oligonucleotide fingerprint analysis of 3′-^{32}P-labeled RNA after partial digestion in alkali. Samples were adsorbed directly onto cellulose acetate strips and subjected to high-voltage electrophoresis in the first dimension (horizontal arrow). After transfer to DEAE-cellulose plates, fragments were resolved by homochromatography in the second dimension (vertical arrow) and identified by autoradiography. (*a*) VSV_{SJ}, (*b*) DI_T.

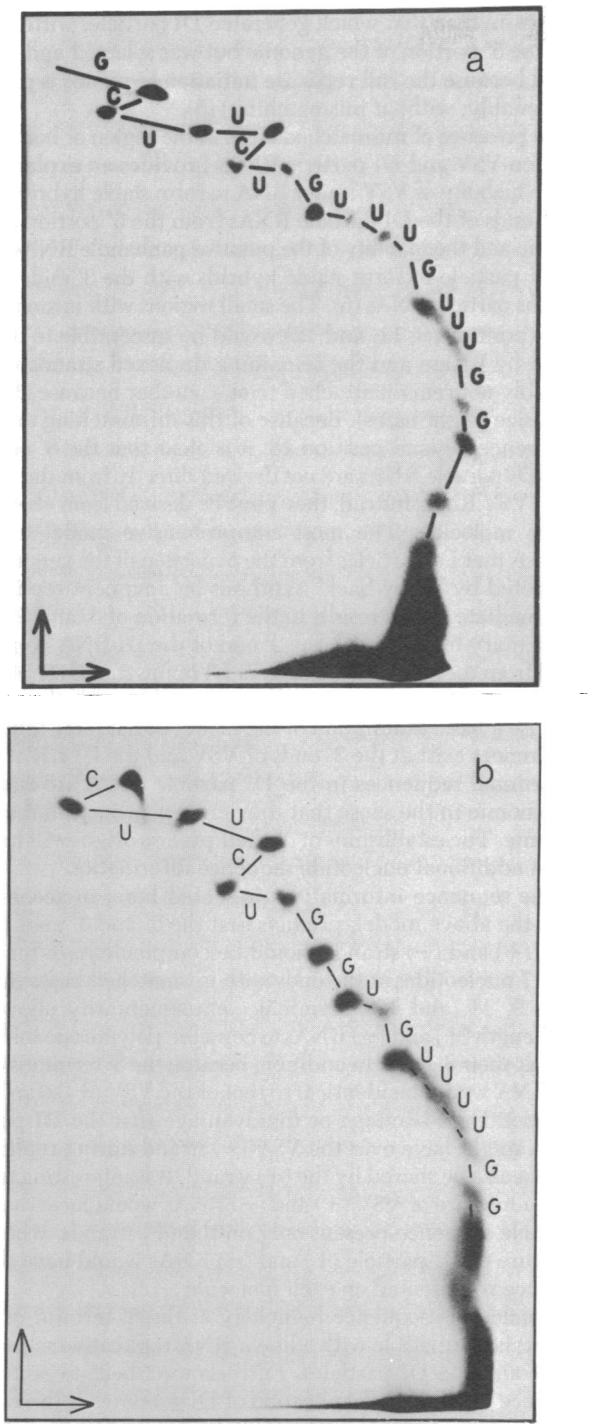

FIG. 2. Oligonucleotide fingerprint analysis of 3′-^{32}P-labeled RNA after partial digestion in alkali as described in the legend to Fig. 1. (*a*) VSV_{MS}, (*b*) DI_{611}.

this dimension and produces an apparent shift to the right. A and G residues are intermediate between C and U in their effect on the rate of migration of an oligonucleotide. The addition of an A residue has little affect on the mobility of an RNA fragment in the first dimension, whereas G slightly accelerates the rate of migration. As shown in Figs. 1 and 2, the first eight nucleotides at the 3′ termini of these VSV and DI particle RNAs are 5′ . . . pGpUpCpUpUpCpGpU-OH 3′.

In order to identify unequivocally the major spots on the homochromatograms, spots 1–6 were eluted and analyzed by high-voltage electrophoresis at pH 1.7 and pH 3.5 on DEAE-paper (data not shown). From this analysis, it was determined that fragment 1 was UpCp, fragment 2 was GpUpCp, etc. In each case, this sequence corresponded to the predicted position on the chromatograms.

In some fingerprints a faint shadow for each major spot appeared at the right and followed the array of spots down the plate. This shift is predicted if a single C residue were removed from each fragment. These faint spots result from the fact that RNA ligated to p*Cp is labeled in the penultimate phosphate. After partial degradation with alkali or any nuclease that cleaves the sequence Up*Cp, two sets of fragments can be generated, one set ending in . . . Up* and the other ending in Up*Cp. Fragments terminating in Up* are generated by two cleavages, whereas the major fragments (terminal p*Cp) were each derived by a single cleavage. For the study reported here, this second array of fragments was not apparent on oligonucleotide fingerprints unless the alkaline digestions were extreme.

Beginning with fragment 9 in the DI particle RNAs, there was sequence divergence from the infectious viral RNAs (Figs. 1 *b* and 2 *b*). In position 9, the DI RNAs had a G residue and the VSV RNAs had a U. This divergence continued for the next three bases: the DI_T and DI_{611} RNAs contained the sequence 5′ . . . UpGpUpG . . . 3′ and the VSV parental RNAs had the sequence 5′ . . . GpUpUpU . . . 3′. Thus, within these four positions, there are three mismatched bases, all of which are G → U or U → G conversions.

Following the short stretch of mismatched bases, there was complete homology between the RNAs of the DI particles and their infectious parental viruses in the next five positions with the sequence 5′ . . . GpGpUpUpUp . . . 3′ (Figs. 1 and 2).

The sequence for the 3′ termini of VSV and the DI particle RNAs was in part confirmed by 20% polyacrylamide/urea gel electrophoresis of partial digests with RNases T1 and U2. This gel system allows the separation of these digestion products by size. RNase T1 makes specific cleavages next to G residues, whereas RNase U2 cleaves next to G and A residues under the conditions used in these experiments. Fig. 3 shows autoradiograms of acrylamide gels after electrophoresis of partially digested [3′-^{32}P]RNA. Alkaline digestion produced a complete array of fragment sizes and provided an index for the assignment of positions in the sequence of the RNase digestion products. Band 2′ in Fig. 3 is the product of a double cleavage due to extensive digestion (see above) and has the sequence GpUp*. Other doublet bands resulting from secondary cleavages of the terminal Cp were not observed. Cleavage with RNase T1 produced fragments that terminated at positions 2, 8, 12, 16, and 17 for the infectious particle RNA and at 2, 8, 9, 11, 16, and 17 for the DI RNA. Thus, the positions of G residues found by partial RNase digestion correspond precisely with those determined by homochromatography. In addition, the presence of a G doublet in positions 16 and 17 of all four RNAs confirms the sequence deduced from Figs. 1 and 2 in a region where precise compositional shifts are frequently difficult to interpret.

From the data presented in Fig. 3 and the nucleotide shifts shown in Figs. 1 and 2, the 18th nucleotide appears to be a pyrimidine in VSV and the DI particle RNAs, but we can not conclude at this point that there is sequence homology beyond position 17. It is apparent, however, that there is extensive divergence of sequences in the infectious and DI particle RNAs beyond position 18. As shown in Fig. 3, digestion with RNase U2 indicates that the first A residue in the DI particle RNAs is

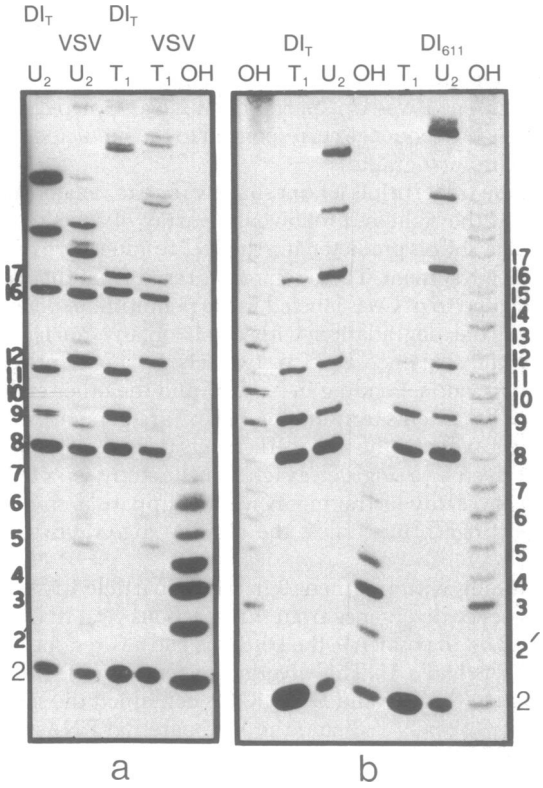

FIG. 3. Polyacrylamide gel electrophoresis of partial digestion products from 3′-^{32}P-labeled RNA. VSV or DI particle RNAs were digested with RNase T1, RNase U2, or alkali and analyzed on 40-cm-long slab gels of 20% acrylamide and identified by autoradiography. (*a*) Parallel electrophoresis of VSV$_{SJ}$ and DI$_T$ RNA digests. (*b*) Electrophoresis of DI$_T$ and DI$_{611}$ RNA digests on separate gels.

at position 21 whereas the first A in the VSV RNAs appears in position 19. From the data in Fig. 3 and other experiments, there is little apparent homology in the next 50 nucleotides between VSV and DI particle RNAs.

We conclude that the 3′-terminal sequences of DI particle and VSV RNA are:

YpXpApY[pGpGpUpUpU]pGpU[pU]pU[pGpUpCpUpUpCpGpU-OH] VSV
ApYpYpY[pGpGpUpUpU]pUpG[pU]pG[pGpUpCpUpUpCpGpU-OH] DI

Regions of homology are indicated by rectangles. Within this stretch of 21 nucleotides there are no extensive palindromic or hairpin structures. There is, however, an overlapping repeated sequence of six bases in the VSV RNAs that is not present in the DI particle RNAs: 5′ . . . pGpUpUpUpGpU . . . 3′.

DISCUSSION

During viral replication, several kinds of events might result in the generation of deletion mutants. The selection and amplification that is required for a particle to emerge in significant numbers restricts the kinds of particles to those with RNAs that have a competent replicase binding and initiation site. The 3′ termini of the RNAs of the infectious and DI particles represent one region in which to expect a common origin of replication because available evidence suggests that these replicative processes occur on linear ribonucleoprotein molecules. Therefore, the DI$_{LT}$ whose RNA is homologous to the 3′ portion of the VSV genome could have originated by a different mechanism than that which generated DI particles with RNAs from the 5′ portion of the genome but was selected and propagated because the full replicase initiation sequence is present (presumably, without mismatching) (5).

The presence of mismatched bases in the region of homology between VSV and DI particle RNAs provides an explanation for the inability of VSV leader RNA to form stable hybrids with the 3′ ends of the DI particle RNAs from the 5′ portion of the genome and the inability of the putative panhandle RNAs from the DI particle to form stable hybrids with the 3′ ends of infectious particle RNAs (6). The small regions with mismatched bases (positions 9, 11, and 12) would be susceptible to degradation by RNase and the remaining duplexed strands would probably not remain attached to one another because of their small size (eight bases). Because of this mismatching and the divergence beyond position 18, it is clear that the 3′ ends of these DI particle RNAs are not derived directly from the 3′ end of the VSV RNA. Instead, they must be derived from elsewhere in the molecule. The most comprehensive model to date suggests that DI particles from the 5′ portion of the genome are generated by "copy-back" synthesis on a nascent replicative intermediate which results in the formation of 3′ and 5′ complementary termini with the 3′ end of the DI RNA being derived from the 5′ end of the nascent DI minus strand (7, 14). The data presented here are totally compatible with this model because a basic assumption of the model is that some sequence differences exist at the 3′ ends of VSV and the DI RNAs. The 3′-terminal sequences in the DI particle RNAs are thus extragenomic in the sense that they are not present in the VSV genome. The establishment of their precise origin will have to await additional nucleotide sequence information.

The sequence information presented here, in accordance with the above model, predicts that the 5′ and 3′ ends of the VSV (+) and (−) strands should be complementary for 14 of the 17 nucleotides at the ends with mismatched bases in positions 9, 11, and 12. Terminal complementarity allows the full-length (+) and (−) RNAs to conserve polymerase initiation sites at their 3′ ends. In addition, because the 3′ terminus of the DI RNA would be identical to that of the VSV (+) strand, any competitive advantage or disadvantage that the DI particle RNA might have over the VSV (−) strand during replication would also be shared by the (+) strand. It is interesting to note that a mixture of VSV (+) and (−) RNAs would have the 3′ DI particle sequence present only on the (+) strands, whereas a mixture of DI particle (+) and (−) RNAs would have this sequence represented on each molecule.

Nucleotide sequence homology at the 3′ termini of these RNAs is compatible with a linear RNA replicative scheme for VSV and the DI particles. Furthermore, because autointerference involves the replication of DI genomes at the expense of infectious viral genomes, these homologous 3′-terminal sequences are candidates for a site of competition, perhaps involving the replicase enzyme. One might speculate that the three mismatched bases in the terminal 17 nucleotides of the DI RNA are either tolerated by the replicase or perhaps confer slight advantages for replicase binding or initiation.

We thank Drs. Olke Uhlenbeck and Tom England for communicating their ligation method prior to publication. A preliminary report of these findings was presented at the ICN-UCLA Symposium on Persistent Viruses at Keystone, CO.

1. Huang, A. S. (1973) *Annu. Rev. Microbiol.* **27**, 101–117.
2. Perrault, J. & Holland, J. J. (1972) *Virology* **50**, 159–170.
3. Keene, J. D., Rosenberg, M. & Lazzarini, R. A. (1977) *Proc. Natl.*

Acad. Sci. USA **74,** 1353–1357.
4. Colonno, R. J. & Banerjee, A. K. (1976) *Cell* **8,** 197–204.
5. Colonno, R. J., Lazzarini, R. A., Keene, J. D. & Banerjee, A. K. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 1884–1888.
6. Perrault, J. & Leavitt, R. (1978) *J. Gen. Virol.* **38,** 35–50.
7. Leppert, M., Kort, L. & Kolakofsky, D. (1977) *Cell* **12,** 539–552.
8. Lazzarini, R. A., Weber, G. A., Johnson, L. D. & Stamminger, G. M. (1975) *J. Mol. Biol.* **97,** 289–308.
9. Stamminger, G. & Lazzarini, R. A. (1974) *Cell* **3,** 85–93.
10. Clewley, J. P., Bishop, D. H. L., Kang, C. Y., Coffin, J., Schnitzlein, W. M., Reichmann, M. E. & Shope, R. E. (1977) *J. Virol.* **23,** 152–166.
11. England, T. E., Gumport, R. I. & Uhlenbeck, O. C. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 4839–4842.
12. Donis-Keller, H., Maxam, A. & Gilbert, W. (1977) *Nucleic Acids Res.* **4,** 2527–2538.
13. Sanger, F., Brownlee, G. G. & Barrell, B. G. (1965) *J. Mol. Biol.* **13,** 373–398.
14. Huang, A. S. (1977) *Bacteriol. Rev.* **41,** 811–822.