Proc. Natl. Acad. Sci. USA
Vol. 75, No. 10, pp. 4704–4708, October 1978
Biochemistry

Sequence of a RNA templated by the 3′-OH RNA terminus of defective interfering particles of vesicular stomatitis virus

(RNA sequences/RNA polymerase/two-dimensional gel sequencing)

BERT L. SEMLER*, JACQUES PERRAULT†, JOHN ABELSON*, AND JOHN J. HOLLAND*

* Departments of Biology and Chemistry, University of California, San Diego, La Jolla, California 92093; and † Department of Microbiology and Immunology, Washington University School of Medicine, St. Louis, Missouri 63110

Communicated by Clifford Grobstein, July 5, 1978

ABSTRACT **We have sequenced the endogenous RNA polymerase product produced by disrupted purified virions of vesicular stomatitis virus defective interfering particles by using the newer one-dimensional rapid gel sequencing techniques and confirming this with a modified two-dimensional gel vectoring technique. The sequence of this 46-nucleotide RNA is:**
5′(pp)pACGAAGACCACAAAACCA-
GAUAAAAAAUAAAAACCACAAGAGGG(U)C_{OH}3′
We infer that this sequence is identical to the sequence at the 5′ end of infectious vesicular stomatitis virus RNA and is complementary to the sequence of the 3′-OH terminus of this defective interfering particle genome RNA.

Vesicular stomatitis virus (VSV) is a negative-strand enveloped RNA virus that has a transcriptase within purified virions (1). This negative strand of virion RNA acts as a template for the transcriptase and generates 5′ capped, methylated, and polyadenylated RNA species that code for five different proteins, each of which can be found in infected cells and in purified virions. Transcription initiates at the 3′-OH terminus of virion RNA and generates a 68-nucleotide leader RNA, followed sequentially by mRNA species for the N, NS, M, G, and L proteins, respectively (2–4). The *in vivo* amounts of each mRNA and of each protein show this same N > NS > M > G > L abundance order (5). All mRNA species contain G(5′)-ppp(5′)AACAG, whereas the leader initiates with ppACG at the 5′ end; in the virion RNA, the 5′ end is (p)ppACG . . . and the 3′-OH end is . . . PyG U 3′-OH (2, 6).

VSV and other animal RNA and DNA viruses generate defective interfering virus particles (DI) at high frequency (for recent review, see ref. 7). These are subgenomic deletion mutant virus particles lacking all or part of the standard virus protein-coding capacity and they are able to interfere with infectious virus replication, probably because they compete for virus-coded replicase and other viral proteins (8, 9). In addition, they contain inverted terminal complementary sequences about 60 bases in length which allow intramolecular circularization or intermolecular concatamer formation, whereas standard virus RNA lacks extensive 5′–3′-OH terminal sequence complementarity, but complementarity exists at the few terminal residues for which sequences are available (2, 6, 10). A similar situation exists with Sendai virus DI (11).

Although purified virions of most VSV DI lack transcriptase activity (8, 9), they do synthesize a small (2S) A-rich endogenous polymerase product RNA (12, 13). This is a discrete 46-base RNA molecule which has been shown, by annealing and gel sizing of the double-stranded RNA hybrids formed (14), to be completely complementary to the 3′-OH terminus of the genome RNA of a wide variety of VSV Indiana DI. Therefore, this DI polymerase product RNA is complementary in sequence to 46 bases of the 3′-OH portion of inverted terminal "stems" of DI RNA and identical in sequence to the first 46 of the 60 nucleotides in the 5′ strand of the inverted terminal "stems." Obviously, therefore, sequencing of this DI polymerase product RNA will provide sequence information 46 nucleotides in from each of the two ends of DI RNA. In addition, it will provide the sequence at the 5′ terminus of standard infectious VSV RNA because annealing and double-stranded RNA hybrid sizing studies show that the 3′-OH "stem" strand of DI is completely complementary to the 60-nucleotide sequence at the 5′ terminus of standard infectious virus RNA and identical in sequence to the 46-residue DI polymerase product RNA (14). Sequencing data should provide insight into mechanisms of DI generation and replication and of DI interference with standard virus replication. This has considerable biological significance because DI are increasingly being implicated in antiviral protection *in vivo* and in the capacity of virulent RNA viruses to cause long-term persistent infection (15–21).

MATERIALS AND METHODS

Cells, Virus, and DI. Standard infectious virus was the indiana strain of VSV originally used by Mudd and Summers (22) for biochemical studies. The DI (DI_{MS}) used for DI RNA polymerase product synthesis was originally generated by the above standard virus during high-multiplicity passages. The very short DI (DI_{ST31}) was the DI originally described from the ts G_{31} mutant of VSV Indiana (23, 24). The carrier culture-derived DI (DI_{CAR4}) was obtained from BHK_{21} cells after years of persistent infection *in vitro* (16, 17). DI were purified as described (15, 25) with additional purification on density gradients of 15–45% potassium tartrate and on an additional 5–40% sucrose velocity gradient.

Endogenous RNA Polymerase Reaction with Purified DI of VSV. The endogenous polymerase reaction was carried out by a slight modification of the method used in our original study of this activity (12). The reaction was carried out in a total of 3 ml containing 0.15 M potassium acetate, 5.0 mM magnesium acetate, 3 mg of dithiothreitol, 1 mM each of ATP, GTP, CTP, and UTP, 0.05 M Tris acetate buffer (pH 7.6), 1 mg of purified VSV DI, and 0.1% Nonidet P40.

End-Labeling and Sequencing of RNA. DI polymerase product RNA was labeled at the 5′ end by using T4 polynucleotide kinase after alkaline phosphatase removal of terminal phosphate molecules as described (26, 27). The labeled RNA was purified on a 20% acrylamide gel and the RNA bands were visualized by autoradiography, cut out, and eluted from gel

The publication costs of this article were defrayed in part by page charge payment. 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 particles; DI_{MS}, DI generated from Indiana strain of VSV used by Mudd and Summers; DI_{ST31}, very short DI from ts G_{31} mutant of VSV Indiana; DI_{CAR4}, carrier culture-derived DI from BHK_{21} cells.

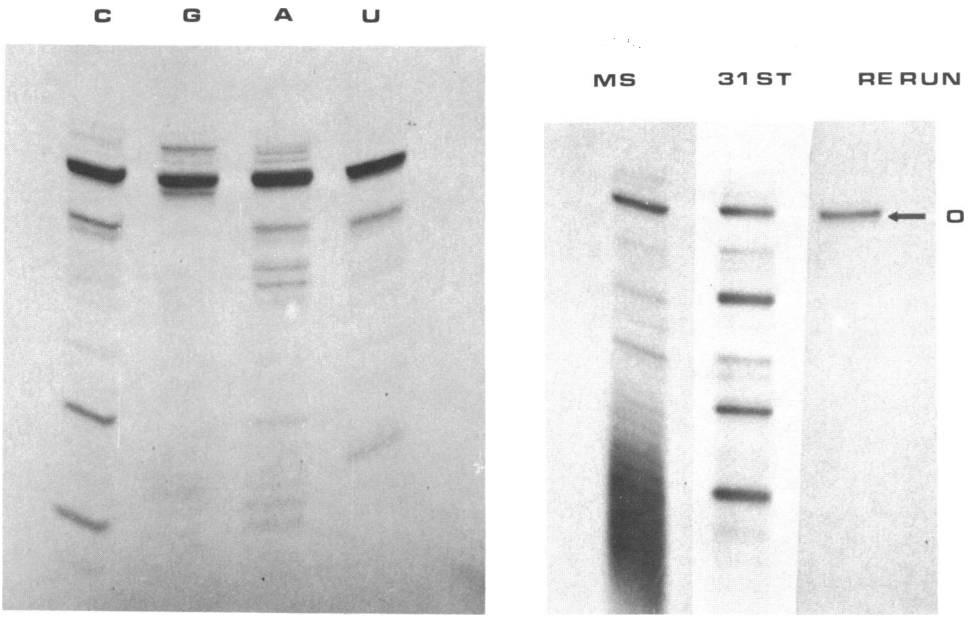

FIG. 1. (*A*) Autoradiograph of preparative 20% polyacrylamide gel of *in vitro* RNA polymerase products synthesized in the presence of the four [α-^{32}P]ribonucleotide triphosphates. Lanes: C, cytosine; G, guanosine; A, adenosine; U, uridine. (*B*) Composite autoradiograph of preparative 20% polyacrylamide gels fractionating the products of 5′-end-labeling of *in vitro* RNA polymerase products. End-labeling was done with [γ-^{32}P]ATP and T4 polynucleotide kinase (26, 27). Lanes: MS, product RNA derived from DI_{MS}; 31ST, product RNA derived from DI_{ST31}. Rerun, 5′-end-labeled 46-nucleotide RNA species from DI_{MS} (band O) (arrow) was cut out and eluted from the gel MS and run again (specific activity of this eluted band O was 1×10^8 cpm/μg in most preparations).

strips. The eluted RNA was repeatedly precipitated with ethanol and the final precipitate was dissolved and subjected to sequence analysis.

Two sequencing methods were used. The first used the newer, rapid RNA sequencing, one-dimensional gel techniques of Simoncsits *et al.* (28) and Donis-Keller *et al.* (29). The second used the two-dimensional oligonucleotide mapping acrylamide gels developed by DeWachter and Fiers (30) and modified by Kennedy (31). These large gel separation techniques were utilized to provide a high-resolution modification of the mobility shift RNA sequencing technique originally designed on the basis of electrophoretic separation according to charge at pH 3.5 in the first dimension and according to DNA or RNA fragment size in the second dimension (32, 33). In our modification, the mobility shift vectors are essentially the same as in the older techniques even though charge separation in the first dimension is carried out in 10% acrylamide gels containing pH 3.5 citric acid and urea, and the size separation in the second dimension is carried out in 20% acrylamide gels containing pH 8.3 Tris borate buffer (31).

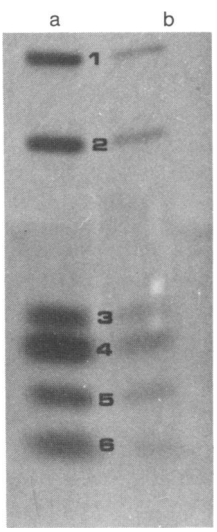

FIG. 2. Autoradiograph of a one-dimensional 20% polyacrylamide gel of the complete ribonuclease T1 digestion products of the 46-nucleotide polymerase product RNA synthesized *in vitro* in the presence of [α-^{32}P]GTP. Digestions were for 30 min at 37° with 100 units of ribonuclease T1 in 10 mM Tris (pH 7.6). Lanes: a, from DI_{ST31}; b, from DI_{MS}.

RESULTS

Synthesis of A-Rich 46-Base Polynucleotide by Purified Virions of DI of VSV. Reichmann *et al.* (12) originally reported synthesis of an A-rich 2S RNA, but no messenger-size RNA, by various VSV DI examined. Perrault *et al.* (14) showed that this RNA is a discrete species of approximately 45 bases and that it is complementary to one strand of VSV DI RNA inverted complementary termini and homologous to the other strand. It was also shown that the 3′-OH terminal sequence of DI virion RNA is not present in infectious virion RNA, but that it was derived by some "recombinational" event as a complementary copy of the 5′ terminus of infectious virion RNA during DI generation. Thus, the sequencing of this DI polymerase product RNA will simultaneously allow deduction of the 3′-OH terminal sequences of DI RNA, the 5′ inverted complementary sequence (stem) of the DI, and the 5′ terminus of the original infectious virion RNA. Because the inverted complementary sequences are conserved in a wide variety of VSV Indiana DI (14), this sequence will be common to the majority of DI.

Fig. 1*A* shows acrylamide gel electrophoretic analysis of the DI *in vitro* RNA polymerase products labeled with [α-^{32}P]triphosphates of cytosine, guanosine, adenosine, and uridine. The major RNA product (arrow) was 46 nucleotides long. In addition, smaller amounts of other shorter and longer discrete RNA bands also were synthesized [these minor bands are all identical in sequence from the 5′ end of the product RNA but represent early and late termination respectively (see below)]. Fig. 1*B* shows the same DI polymerase product RNA bands after 5′-end labeling of the completed product RNA with purified T4 polynucleotide kinase. This purified, gel-eluted, 5′-end-labeled 46-base product RNA of DI_{MS} was subjected to sequence analysis numerous times.

Oligonucleotide Analysis of 46-Nucleotide DI_{MS} Polymerase Product RNA. Preliminary analysis of α-ATP- or α-GTP-labeled 46-base product RNA by one- and two-dimensional electrophoretic gel analyses showed that this RNA contained six oligonucleotides after digestion by ribonuclease T1 (Fig. 2) and eight oligonucleotides after digestion by pancreatic ribonuclease (not shown). The shorter, minor band, RNA species contained an increasing number of the same oligonucleotides with increasing size of the RNA band. The number of sizes of T1 oligonucleotides seen here agree with those reported by

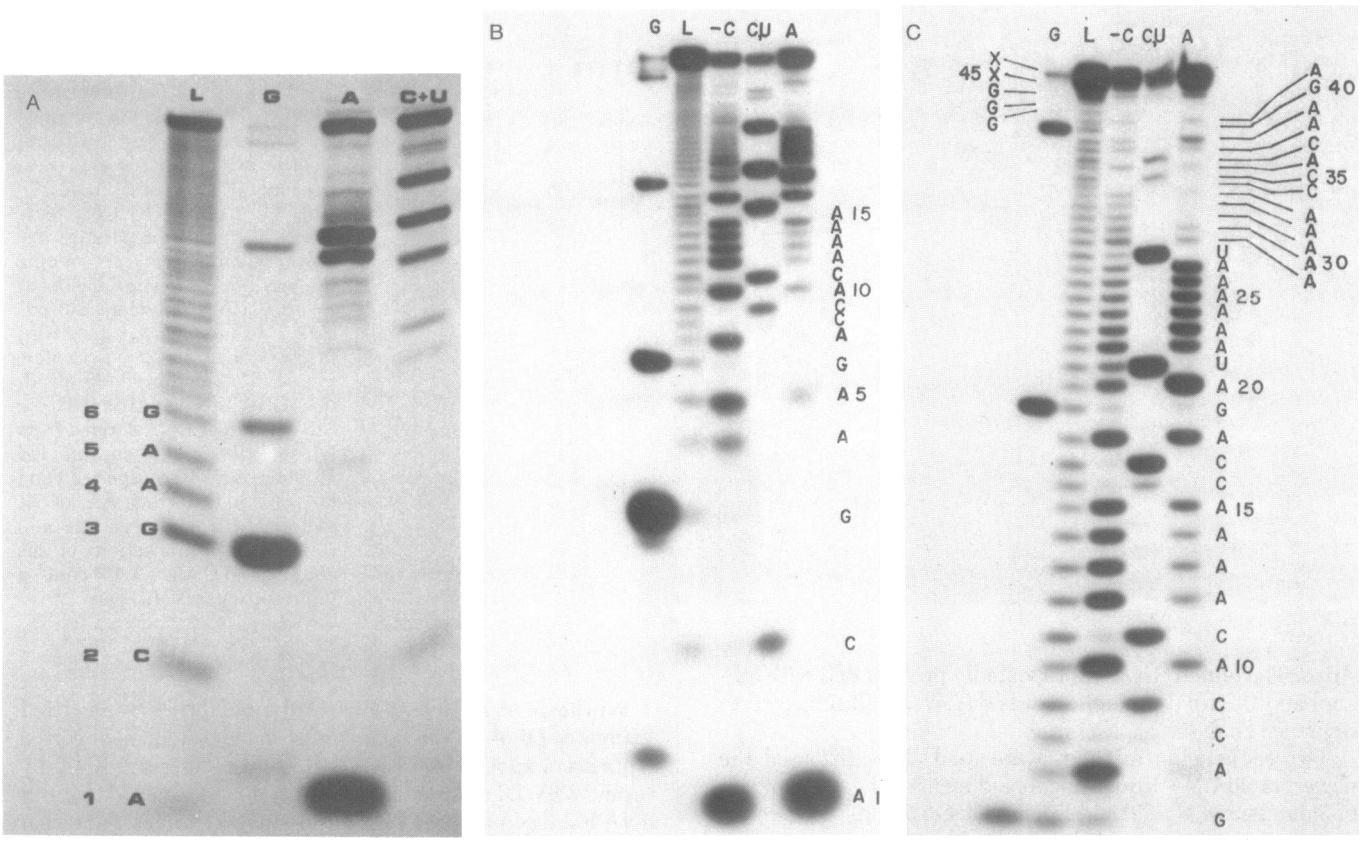

FIG. 3. Autoradiographs of one-dimensional 20% polyacrylamide RNA sequencing gels of partial enzymatic digests. Lanes: G, RNase T1 digestion; L, formamide degradation for identification of every possible nucleotide; −C, Phy I RNase digestion; C + U, pancreatic RNase digestion; A, RNase U2 digestion. Digestion conditions were as described (28) for 1–2 μg of carrier tRNA included in each digestion with the following modifications: L (ladder) samples were boiled at 100° in freshly deionized formamide for 2½ hr; −C samples were digested with 5 μl of solution containing 10 units of *Phy* I RNase per ml in 40% glycerol and 5 μl of 10 mM NaOAc, pH 5.0/1 m*M* EDTA buffer. (*A*) Sequence analysis of 5′-end-labeled product RNA from DI_{MS}, resolving the nucleotides near the 5′ terminus. (*B* and *C*) Sequence analysis of 5′-end-labeled RNA from DI_{MS}; these gels were run longer to resolve nucleotides distal to the 5′ terminus. Note that the assignment of nucleotides 45 and 46 (here denoted X) was made on the basis of two-dimensional gel analysis (see Fig. 4*B*) and analysis of many other one-dimensional sequencing gels (data not shown).

Emerson *et al.* (13) and are as predicted by the derived sequence below (see Fig. 5).

One-Dimensional Gel Electrophoresis Sequence Analysis of 46-Nucleotide DI_{MS} Polymerase Product RNA. Fig. 3*A* shows a typical gel providing resolution of the nucleotides near the 5′ (labeled) end of the DI product RNA. Fig. 3 *B* and *C* show similar sequencing gels run longer to resolve the nucleotide sequence more distal to the 5′ terminus.

RNA Sequencing by Base-Specific Vectoring Techniques Modified to Utilize Large Two-Dimensional Gel Electrophoresis Apparatus. Sanger *et al.* (32) developed a sequencing technique for short oligonucleotides involving electrophoretic separation of partial digestion fragments in the first dimension and homochromatography in the second (33). We have adapted this technique to large two-dimensional gel acrylamide gel electrophoresis separations. Fig. 4 shows the typical gel results with a partial snake venom digest of 5′-end-labeled DI product RNA. Individual base assignments are made according to the mobility shift resulting from each nucleotide loss (32, 33). It can be seen that the resulting sequence is in agreement with the sequence derived from the one-dimensional ladder technique (Fig. 3). Finally, we found that this same technique can be applied to partial digests of RNA degraded by boiling in formamide (28), and vectoring characteristics were nearly identical to those seen with partial snake venom digests.

Application of these techniques to the polymerase product RNA of a different VSV DI particle, DI_{ST31}, showed a nearly identical sequence, except that C is substituted for U at position 21. Preliminary data on a third DI, DI_{CAR4}, also showed a nearly identical sequence. It must be concluded that this sequence is common to many, if not most, DI [and hence that the 3′-OH end of many DI must share sequence identity, as was shown by annealing of complementary stems from 3′ and 5′ ends of various DI (14)].

Assignment of Nucleotide Sequences for RNA Termini of VSV DI and of Infectious Virion RNAs. It was shown previously (10, 14) that nearly all VSV DI differ from the infectious virions that generated them in having inverted complementary termini which can give double-stranded RNA stems upon self-annealing. These stems are about 60 nucleotides in length (120 bases in the double-stranded stem), and the stem component derived from the 3′-OH and of DI RNA anneals completely to (i.e., completely protects and is partially protected by) labeled DI polymerase product RNA 46 bases in length. This same stem component RNA molecule derived from the 3′-OH end of DI also anneals to (and is completely complementary to) the 5′ end of infectious virion RNA. This allows inference of the nucleotide sequences at the ends of DI virion RNA and of infectious virion RNA (Fig. 5), as well as for the sequence of the DI polymerase product RNA of 46 bases. The 46-nucleotide RNA polymerase product has an unusual composition with short tracts of A residues interspersed with other nucleotides. Conversely, the DI 3′-OH terminus template RNA contains corresponding tracts of U residues.

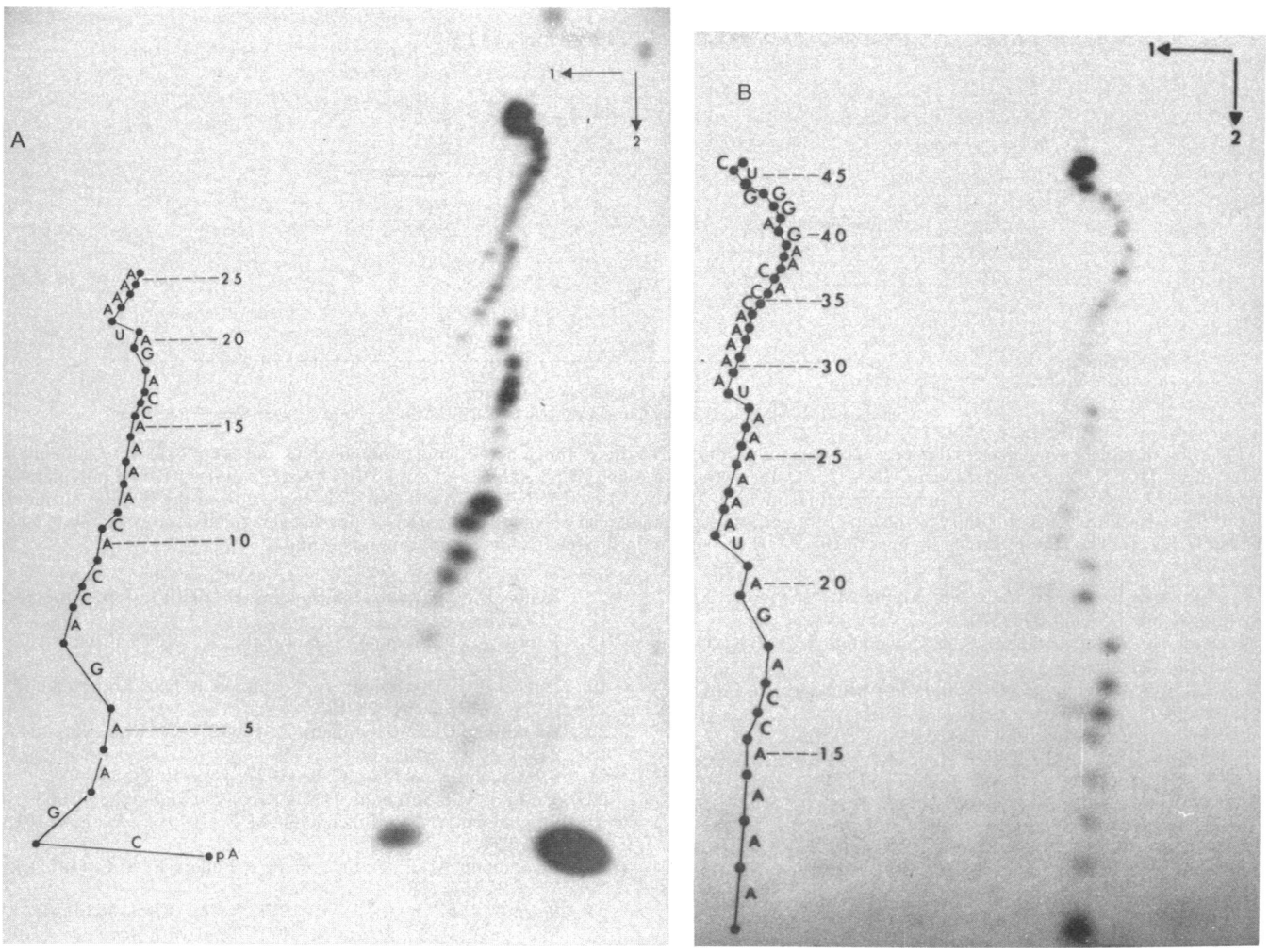

FIG. 4. Autoradiograph showing two-dimensional polyacrylamide gel analysis of partial snake venom phosphodiesterase digests of 5′-end-labeled 46-nucleotide product RNA from DI_{MS}. Enzyme reactions were carried out at 37° for 0, 5, 10, 20, and 30 min, each in the presence of 50 μg of tRNA carrier and 5 μl of snake venom phosphodiesterase (1 mg/ml in 25 mM Tris·HCl, pH 8.8/10 mM MgOAc). All samples were then pooled (at 0°), diluted into a urea sample buffer (31), heated at 70° for 30 sec, and then subjected to two-dimensional gel electrophoresis. (*A*) Sequence analysis resolving the 5′ end of the product RNA. (*B*) Sequence analysis resolving the nucleotides distal to the 5′ terminus. The minor spots (not depicted in the schematic tracing) represent a very small percentage of molecules that become labeled at position 2 (a C residue) because the 5′-end A residue was lost (hence they migrate differently in both dimensions). The vector at position 45 is either U or G. We rule out G because there are no RNase T1 cuts at this position in one-dimensional gels (although the −C and U + C slots yield equivocal results at position 45 because of its proximity to the 46-nucleotide starting material).

DISCUSSION

From the above data, terminal RNA sequences are now known for each end of the DI genome and for the 5′ end of the infectious virion genome but not for the 3′-OH end of the infectious virion genome. Annealing of DI stems and DI polymerase product RNA (14) clearly shows that the 3′-OH ends of several DI RNAs are nearly identical to each other but do not have *extensive* sequence homology to the 3′-OH end of virion RNA (14). Nevertheless, there is at least partial homology at the extreme 3′-OH ends of virion and DI RNAs, because the virion RNA 3′-OH end has been determined to be . . . PyG U 3′-OH by sequencing and to be complementary to the 5′ end of a leader RNA with the sequence (p)ppACG . . . (2, 6). The base position at which this 3′-OH sequence of infectious virion RNA deviates from the 3′-OH sequence reported here for several different VSV Indiana DI is the point at which the 5′ and 3′-OH ends of infectious virion RNA cease to show inverted terminal complementarity. This point is unlikely to be more than 10–20 bases in from the 3′-OH U of virion RNA because, if otherwise, annealing would show complementarity to DI 5′ stems and to DI product RNA, which it does not (14).

Leppert *et al.* (11) have demonstrated that Sendai virus DI contain inverted complementary sequences and, as with VSV (14), these show homology and complementarity to the 5′ end of infectious virus but no extensive (annealing) homology to the 3′-OH terminus of the infectious virus. They (11) and Huang (7) proposed a "back-copying" mechanism for DI generation involving replicase strand switching from template to nascent chain after partial replication from the 3′-OH terminus of the template RNA (11). All of the findings reported above are compatible with this strand-switching model but many other mechanisms are possible. Further sequencing of virus, DI, and polymerase products may provide greater insight into DI generation and interference and roles of DI in infection and persistence.

We are grateful to Martinez Hewlett for sequencing suggestions and to Dr. Ian Kennedy for many helpful suggestions. Also, we thank Michael Heggeness for photographic help. This work was supported by National Institutes of Health Grant AI 14627 to J.H. and CA 10894 to J.A.; B.L.S. was supported by U.S. Public Health Service Predoctoral Traineeship CA 05274.

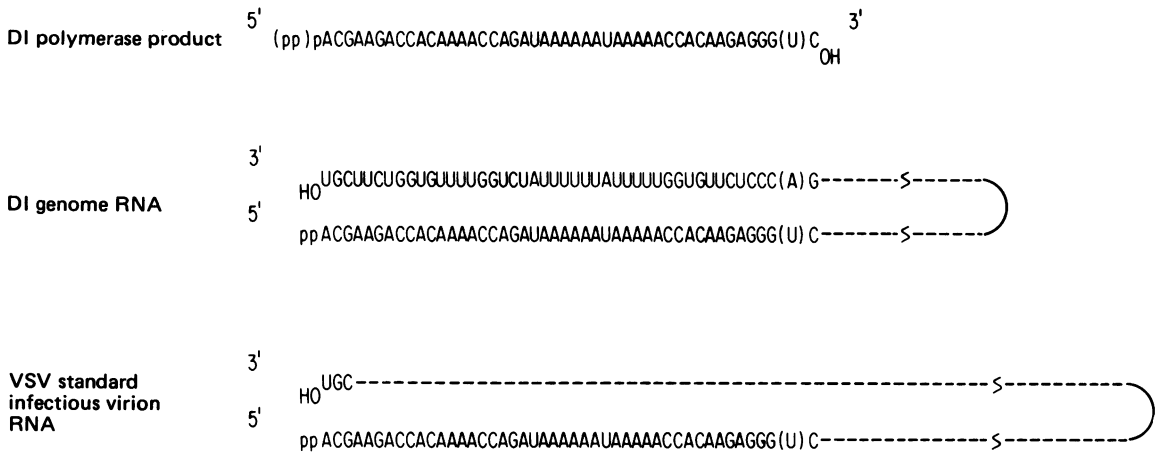

FIG. 5. Complete sequence of DI *in vitro* polymerase product RNA (determined herein) and of homologous and complementary sequences inferred for DI RNA and standard virus RNA. The latter were determined by annealing and duplex RNA sizing studies with DI polymerase product RNA and inverted complementary terminal stems of DI RNA (14). Residue (U) in position 45 is derived from the two-dimensional gels (Fig. 4). Because of its proximity to the top of the gel, its vector might also be compatible with G. However, G is ruled out by other data (see legend to Fig. 4). The trinucleotide sequence at the 3'-OH end of standard virus RNA is taken from refs. 2 and 6.

1. Baltimore, D., Huang, A. S. & Stampfer, M. (1970) *Proc. Natl. Acad. Sci. USA* **66**, 572–576.
2. Banerjee, A. K., Abraham, G. & Colonno, R. J. (1977) *J. Gen. Virol.* **34**, 1–8.
3. Ball, L. A. & White, C. N. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 442–446.
4. Colonno, R. J. & Banerjee, A. K. (1976) *Cell* **8**, 197–204.
5. Villarreal, L. P., Breindl, M. & Holland, J. J. (1976) *Biochemistry* **15**, 1663–1667.
6. Keene, J. D., Rosenberg, M. & Lazzarini, R. A. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 1353–1357.
7. Huang, A. S. (1977) *Bacteriol. Rev.* **41**, 811–821.
8. Perrault, J. & Holland, J. J. (1972) *Virology* **50**, 159–170.
9. Huang, A. S. & Manders, E. (1972) *J. Virol.* **9**, 909–916.
10. Perrault, J. & Leavitt, R. W. (1977) *J. Gen. Virol.* **38**, 35–50.
11. Leppert, M., Kort, L. & Kolakofsky, D. (1977) *Cell* **12**, 539–552.
12. Reichmann, M. E., Villarreal, L. P., Kohne, D., Lesnaw, J. & Holland, J. J. (1974) *Virology* **58**, 240–249.
13. Emerson, S. U., Dierks, P. M. & Parsons, J. T. (1977) *J. Virol.* **23**, 708–716.
14. Perrault, J., Semler, B. L., Leavitt, R. W. & Holland, J. J. (1978) in *Negative Strand Viruses*, eds. Mahy, B. W. J. & Barry, R. D. (Academic, New York), Vol. 10, in press.
15. Doyle, M. & Holland, J. J. (1973) *Proc. Natl. Acad. Sci. USA* **70**, 2105–2108.
16. Holland, J. J. & Villarreal, L. P. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 2956–2960.
17. Holland, J. J., Villarreal, L. P., Welsh, R. M., Oldstone, M. B. A., Kohne, D., Lazzarini, R. & Scolnick, E. (1976) *J. Gen. Virol.* **33**, 193–211.
18. Kawai, A., Matsumoto, S. & Tanabe, K. (1975) *Virology* **67**, 520–533.
19. Wiktor, T. J., Dietzschold, B., Leamnson, R. N. & Koprowski, H. (1977) *J. Virol.* **21**, 626–635.
20. Spandidos, D. A. & Graham, A. F. (1976) *J. Virol.* **20**, 234–247.
21. Schmaljohn, C. & Blair, C. D. (1977) *J. Virol.* **24**, 580–589.
22. Mudd, J. A. & Summers, D. F. (1970) *Virology* **42**, 958–968.
23. Schnitzlein, W. M. & Reichmann, M. E. (1976) *J. Mol. Biol.* **101**, 307–325.
24. Reichmann, M. E., Pringle, C. R. & Follett, E. A. C. (1971) *J. Virol.* **8**, 154–160.
25. Perrault, J. & Leavitt, R. W. (1977) *J. Gen. Virol.* **38**, 21–34.
26. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 560–564.
27. Richardson, C. C. (1965) *Proc. Natl. Acad. Sci. USA* **54**, 158–165.
28. Simoncsits, A., Brownlee, G. G., Brown, R. S., Rubin, J. R. & Guilley, H. (1977) *Nature (London)* **269**, 833–836.
29. Donis-Keller, H., Maxam, A. M. & Gilbert, W. (1977) *Nucleic Acids Res.* **4**, 2527–2538.
30. DeWachter, R. & Fiers, W. (1972) *Anal. Biochem.* **49**, 184–197.
31. Kennedy, S. I. T. (1976) *J. Mol. Biol.* **108**, 491–511.
32. Sanger, F., Brownlee, G. G. & Barrell, B. G. (1965) *J. Mol. Biol.* **13**, 373–398.
33. Tu, C. D., Jay, E., Bahl, C. P. & Wu, R. (1976) *Anal. Biochem.* **74**, 73–93.