JOURNAL OF VIROLOGY, Feb. 1980, p. 807-817
0022-538X/80/02-0807/11$02.00/0

Vol. 33, No. 2

Oligonucleotide Sequence Analyses Indicate that Vesicular Stomatitis Virus Large Defective Interfering Virus Particle RNA Is Made by Internal Deletion: Evidence for Similar Transcription Polyadenylation Signals for the Synthesis of All Vesicular Stomatitis Virus mRNA Species

CORRIE M. CLERX-VAN HAASTER, JON P. CLEWLEY, AND DAVID H. L. BISHOP*

Department of Microbiology, The Medical Center, University of Alabama in Birmingham, Birmingham, Alabama 35294

RNase T_1 oligonucleotide fingerprint analyses of three vesicular stomatitis virus Indiana serotype small defective interfering (DI) particle RNA species indicate that they only have oligonucleotides derived from the 5′ region of the viral genome. These studies also indicate that these three DI RNAs have partial L gene sequences as well as two 5′ viral oligonucleotides (59 and 70) that are not transcribed into L (or other) mRNA species (J. P. Clewley and D. H. L. Bishop, J. Virol. **30**:116-123, 1979). Analyses of the large DI RNA (LT DI) reveal a different origin. The LT DI RNA has oligonucleotides derived from both the 3′ end of the genome (including all the large oligonucleotides identified for N, NS, M, and G genes), in addition to at least one of the 5′-proximal L gene oligonucleotides (47), as well as all seven oligonucleotides (3, 38, 42, 43, 44B, 59, and 70) that are not protected from nuclease digestion after the formation of mRNA-viral RNA duplexes (Clewley and Bishop). It appears therefore that the genesis of LT RNA involves a deletion of internal L gene sequences from the viral RNA. Oligonucleotide sequence analyses have been undertaken on several of the vesicular stomatitis viral RNA oligonucleotides, including all seven (3, 38, 42, 43, 44B, 59, and 70) that are not transcribed into mRNA. The analyses confirm that oligonucleotides 59 [3′...GAACACCAAAAAUAAAAAAUA(G)...5′] and 70 [3′...GACCAAAACACCA(G)...5′] are at the 5′-end region of the viral genome. Oligonucleotide 38 [3′...GAAAUUCAUACUUUUUU(U)(G)...5′] may represent the termination signal for L mRNA synthesis (R. A. Lazzarini, personal communication). Oligonucleotide 43 [3′...GUAUACUUUUUUU(G)...5′] corresponds to the sequence shown to be the N gene mRNA polyadenylation signal (D. J. McGeoch, Cell **17**:673-681, 1979). The other three oligonucleotides share a common feature with oligonucleotides 43 and 38, viz., a stretch of 6 or 7 U residues preceded by an AUAC sequence. Thus the sequence of oligonucleotide 3 is 3′...GAAUUAAUAUAAAAUUAAAAAUUAAAAAUACUUUUUU(U)(G)...5′, whereas that of oligonucleotide 42 is 3′...GAUACUUUUUUUCAU(U)(G)...5′, and that of oligonucleotide 44B is 3′...G(U)AUACUUUUUU(G)...5′. These sequence analyses suggest a common polyadenylation signal for the synthesis of all vesicular stomatitis virus mRNA species, i.e., the sequence (3′)...AUACUUUUUU(U)...(5′).

Studies on the origin of certain small defective interfering (DI) particle RNA species of vesicular stomatitis virus (VSV) Indiana serotype (VSI) have suggested that they contain the 5′-end sequence of the viral RNA genome, as well as a 3′ sequence which is a complementary copy of a terminal portion of the 5′-end sequence (17, 19a-21), supporting the model of DI generation proposed by Huang (12) and Leppert et al. (14), and referred to in this paper as the read-back model. Since the gene order of VSI is known (i.e., 3′..N,NS,M,G,L..5′; 1, 2) and RNase T_1-derived oligonucleotides representing the five viral genes have been identified (4, 10), it has been possible to demonstrate by fingerprinting studies that certain VSI small DI RNA species (DI-1, DI-2, DI-3; 13) have some (but not all) of the L gene oligonucleotides and lack all of the N, NS, M, and G gene oligonucleotides (4, 5, 13). In addition, these small DI RNA species have at least three viral oligonucleotides (38, 59, and 70) that are not protected from removal when four

nucleases are used to produce mRNA-viral RNA duplexes lacking 3'-polyadenylic acid tails (4). RNase T_1-derived oligonucleotides complementary to VSV leader RNA, which also should not be protected, are too small to be recovered in the unique region of the two-dimensional fingerprints. From the oligonucleotide sequence analyses described in this paper, compared to the deduced 5'-end sequence of the VSI genome (20, 21; M. Schubert, J. D. Keene, R. C. Herman, and R. A. Lazzarini, submitted for publication), it has been possible to show that oligonucleotides 59 and 70 are (5') distal to the L gene, whereas oligonucleotide 38 is the proposed L mRNA polyadenylation (termination) signal (Schubert et al., submitted for publication).

In this paper we describe analyses of the LT DI RNA isolated from the heat-resistant (HR) derivative of VSI Toronto strain (18) and show that it has a different origin from that of the small DIs studied previously, in that its genesis has involved internal deletion of L gene sequences. To support these conclusions we show that the LT DI RNA has N, NS, M, and G sequences and some 5'-region L gene sequences, including the L gene polyadenylation-termination signal (oligonucleotide 38) and distal 5' noncoding oligonucleotides (59 and 70).

Oligonucleotide sequence analyses of four other viral oligonucleotides (3, 42, 43, and 44B) that are not protected against nuclease removal when four nucleases are used to generate mRNA-viral RNA duplexes (4) have been found to be strikingly similar. Together with oligonucleotide 38, they all share the sequence 3'...AUACUUUUUU(U)...5'. In view of the fact that oligonucleotide 43 corresponds to the N gene mRNA polyadenylation signal, these data are interpreted as indicating that the five oligonucleotides represent mRNA polyadenylation signals corresponding to the five viral genes.

MATERIALS AND METHODS

Materials. RNases U_2 and T_1 came from Calbiochem, La Jolla, Calif. *Physarum polycephalum* RNase (Phy 1) and nuclease P1 were purchased from P.L. Biochemicals Inc, Milwaukee, Wis. T4 polynucleotide kinase (RNase free) came from Enzo Biochem Research Products, New York, N.Y. RNase A was obtained from Sigma Chemical Co., St. Louis, Mo., and DEAE-cellulose (9 parts cellulose to 1 part DEAE-cellulose) thin-layer glass plates (20 by 40 cm) came from Analtech, Newark, Del. High-specific-activity [γ-^{32}P]ATP (1,000 to 3,000 Ci/mmol) was purchased from New England Nuclear Corp., Boston, Mass.; other radioisotopes came from ICN, Irvine, Calif. Cellogel strips (3 by 75 cm) or cellulose acetate strips (3 by 100 cm) came from Kalex Science Corp, Manhasset, N.Y., and Schleicher and Schuell Inc, Keene, N.H., respectively. Yeast RNA was purchased from BDH Chemicals Ltd., Poole, England.

Viruses. The sources of the prototype strain of VSI used in this study (VSI B, i.e., VSI Birmingham derivative), and three VSI (Toronto)-derived small DI particles have been described (5, 12). These three DI particles are the DI-1, DI-2, and DI-3 described by Kang and associates (13). The large DI was obtained from the HR strain of VSI (Toronto, 18).

Infection of cells, virus purification, extraction of RNA, and oligonucleotide fingerprinting of RNA species. The methods used for the preparation and purification of ^{32}P-labeled virus and the extraction and purification of RNA have been described (5). Digestion of RNA by RNase T_1 and resolution of the resulting oligonucleotides by two-dimensional polyacrylamide gel electrophoresis utilizing the method of De Wachter and Fiers (7), as modified by Clewley and associates (6), have also been described (5). Preparations of DI particle RNA were made for us by C.-Yong Kang, Southwestern Medical School, Dallas, Tex.

Extraction and RNase A analyses of T_1-derived oligonucleotides. Oligonucleotides in polyacrylamide gels were located by autoradiography, and gel plugs containing a resolved oligonucleotide were punched out with a plastic straw (6-mm diameter). Each gel plug was placed in a sterile siliconized glass minivial and crushed, and 0.8 ml of 1 M NaCl was added. After 1 to 2 days at 4°C, the mixture was filtered and precipitated with an equal volume of isopropanol in the presence of 0.04 mg of yeast RNA, previously repurified by phenol-*m*-cresol-8-hydroxyquinoline extraction. Usually between 60 and 90% of the radioactivity was recovered. The precipitated materials were lyophilized and digested with 0.005 ml of a solution of 2 mg of RNase A per ml of 0.1 M ammonium bicarbonate (pH 8.0). The resulting products were then resolved by a modification of the thin-layer chromatography procedure described by Volckaert and Fiers (24). Each nuclease digest was applied to a cellulose MN 300 polyethyleneimine-impregnated plate (10 by 10 cm by 0.1 mm; Brinkmann Instruments, Inc., Westbury, N.Y.), and the plate was washed with 70% aqueous ethanol to remove any residual salt. After drying, an initial chromatography was run with deionized water (to 1 cm above the sample), followed by chromatography with 5 M urea containing 7.4 M formic acid. The plate was then dried to remove formic acid, washed three times with water to remove the urea, and again dried. Chromatography in the second dimension involved an initial 1-cm development with deionized water, followed by chromatography with 4 M urea containing 1 M formic acid, adjusted to pH 4.3 with pyridine. The plate was dried, washed three times with deionized water to remove urea, and again dried. The locations of the radiolabeled compounds were determined by autoradiography, and the individual labeled nucleotides were recovered and counted to obtain an estimate of the relative label in each nucleotide.

Recovery and 5'-end labeling of T_1 oligonucleotides. For end labeling of RNase T_1-derived oligonucleotides resolved by two-dimensional polyacrylamide gel electrophoresis, the gel plugs containing the desired oligonucleotides were eluted as described above, except that 0.5 ml of 1 M NaCl was employed and the oligonucleotides were recovered by precipitation with 2 volumes of ethanol at −70°C in the absence

of carrier RNA. After recovery of the precipitated oligonucleotides by centrifugation, each sample (containing 100 to 1,500 cpm of Cerenkov radiation) was dissolved in 0.01 ml of water, and 0.005 ml of a solution containing 0.25 M Tris-hydrochloride (pH 8.0), 0.1 M $MgCl_2$, and 0.1 M dithiothreitol was added. A sample of 10 pmol (20 to 30 μCi) of [γ-^{32}P]ATP and 3 U of T4 polynucleotide kinase were added, and the mixture was adjusted with water to a final volume of 0.05 ml and incubated at 37°C for 15 min. After addition of 0.05 ml of 4 M ammonium acetate, 0.02 mg of carrier RNA, and 0.3 ml of ethanol, the 5'-end-labeled oligonucleotide was precipitated at −70°C for 10 min. The precipitate was then recovered by centrifugation, dried, and dissolved in 0.01 ml of 7 M urea–5 mM Tris-borate (pH 8.3)–1 mM EDTA, containing 0.01% bromophenol blue and 0.01% xylene cyanol FF. Resolution of the labeled material by electrophoresis in a thin 20% polyacrylamide gel (15 by 30 cm by 0.4 mm) was carried out as described by Donis-Keller and associates (8). After autoradiography, the end-labeled oligonucleotide was eluted from the gel as described above. Usually 10^4 to 3×10^5 cpm of Cerenkov radiation was recovered.

Partial digestion of 5'-end-labeled RNase T_1 oligonucleotides by specific nucleases. The reaction conditions employed for the partial digestion of 5'-end-labeled oligonucleotides by RNase A (0.01 ng/μg of RNA) and RNase U_2 (0.001 U/μg of RNA for the 1× concentration, and 0.01 U/μg of RNA for the 10× concentration) used the same reaction mixtures and incubation times described by Donis-Keller and associates (8). Partial digestion of the labeled nucleotides with nuclease P1 (0.005 U/μg of RNA) or Phy 1 (0.1 U/μg of RNA) was accomplished by incubating the samples in 25 mM ammonium acetate buffer (pH 5.0) for 15 min at room temperature. The latter two enzymatic digestions were terminated by heating to dryness at 100°C in an open tube (usually taking 10 min). The dried samples were then dissolved in 0.003 ml of 7 M urea–5 mM Tris-borate (pH 8.3)–1 mM EDTA, containing 0.01% bromophenol blue and 0.01% xylene cyanol FF. Partial hydrolysis of oligonucleotides was achieved by boiling a mixture of 0.001 ml (0.001 mg of RNA) with 0.01 ml of formamide for 30 min. On occasion, oligonucleotide ladders were generated using nuclease P1, even though this resulted in minor mobility differences by comparison to oligonucleotides with 3'-phosphates. Before loading on a gel, 0.001 ml of a dye mixture was added (0.1% bromophenol blue, 0.1% xylene cyanol FF). Digests were resolved by electrophoresis in thin polyacrylamide gels as described above (8).

Homochromatography of partial digests of 5'-end-labeled oligonucleotides. The P1 partial enzymatic digests of 5'-end-labeled oligonucleotides were resolved by electrophoresis and homochromatography as described by Silberklang and associates (22) using Homomix C prepared as described by Barrell (3).

RESULTS

Oligonucleotide fingerprinting analyses of VSI and four DI RNA species. Studies involving the generation and, after digestion with four nucleases, the recovery of VSI mRNA-^{32}P-labeled viral RNA duplexes (4) have demonstrated that: (i) N mRNA protects the viral oligonucleotides 11, 13, 58, and 65; (ii) NS mRNA protects 15, 28, 68, and 69; (iii) M mRNA protects 32, 35, 52, and 71; and (iv) G mRNA protects 31, 57, and 61. Other than seven oligonucleotides that are not protected under conditions of high nuclease by any of the mRNA species (3, 38, 42, 43, 44B, 59, and 70), all of the rest of the viral large oligonucleotides are protected by L mRNA.

In Fig. 1 are shown the oligonucleotide fingerprints of VSI (Fig. 1B), a schematic diagram of the same (Fig. 1A) showing the nomenclature system we have used to refer to its RNase T_1-resistant oligonucleotides (5), and the fingerprints of the LT DI RNA (Fig. 1C) derived from the HR strain of VSI Toronto (18), as well as three small DI RNA species (Fig. 1D to F), namely, DI-1, DI-2, and DI-3, recovered by Kang and associates from the Toronto strain of VSI (13). Certain of the large oligonucleotides that are found in the DI RNA digests are numbered according to the schematic diagram shown in Fig. 1A. The identities of these oligonucleotides were obtained by coelectrophoresis of the DI RNA digests with a digest of VSI RNA (data not shown). Oligonucleotides that are common to both the small DIs and the LT DI RNA species are underlined.

With regard to the LT DI RNA fingerprint (13; Fig 1C), it was observed that it lacked most of the oligonucleotides found in the small DI patterns, and therefore it was concluded that it was not grossly contaminated with small DI species. Gel electrophoresis of LT DI RNA preparations have failed to detect small DI-type RNA species (data not shown). The digest of LT DI RNA was found to have all seven of the viral oligonucleotides that were not protected against nuclease removal, as well as all of the N, NS, M, and G mRNA-protected oligonucleotides identified in a previous study (4). In addition it was found to have one oligonucleotide (47) that is protected by L mRNA (4), and which is found in all three small DI RNA fingerprints (Fig. 1D to F). Thus these results indicate that, whereas the LT DI RNA has mostly viral 3'-region sequences, it also has a limited amount of L gene and putative 5'-end sequences (see below).

The three small DI RNA fingerprints were found to have (i) three oligonucleotides that are not protected by the viral mRNA species (namely, 38, 59, and 70), and (ii) only L mRNA-protected oligonucleotides (4). These results therefore confirm the 5' origin of the three small DIs and suggest that oligonucleotides 38, 59, and 70 may represent 5' L mRNA termination and/

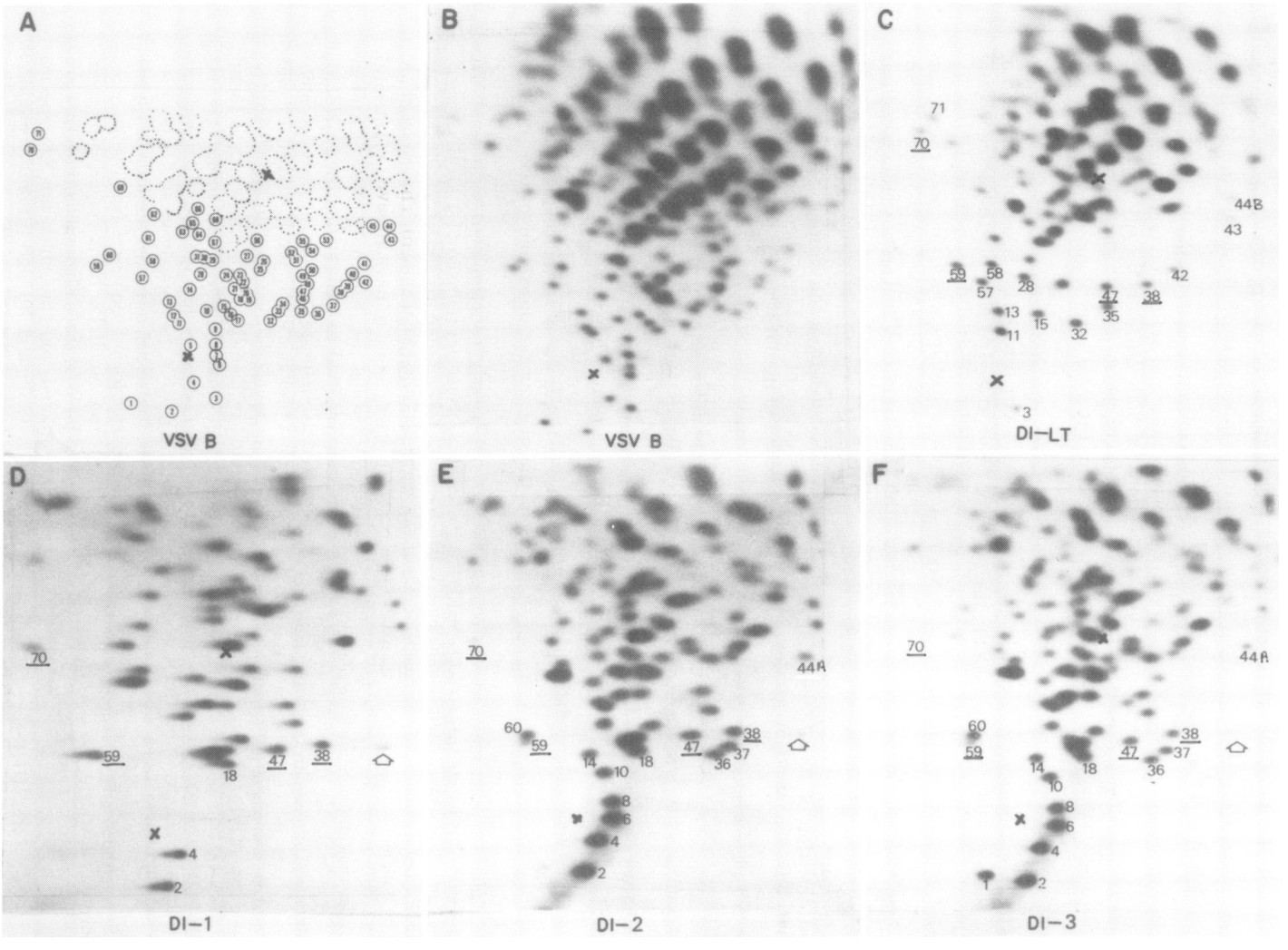

FIG. 1. *Oligonucleotide fingerprint analyses of the genome RNA of VSI and four of its DI RNA species. (B) Fingerprints of the RNase T_1-resistant oligonucleotides obtained from ^{32}P-labeled 40S RNA of VSI; (A) a schematic diagram of the same showing the oligonucleotide nomenclature system we have adopted (5); and fingerprints of (C) LT and (D-F) three small DI RNAs. Particular viral oligonucleotides present in the four DI species are numbered. The large oligonucleotides that are present in both the LT and the DI-1, DI-2, and DI-3 patterns are underlined. The arrowed oligonucleotide was irregularly found in DI fingerprints. It has not been analyzed further.*

or distal RNA sequences.

Pancreatic RNase compositional analyses of T_1 oligonucleotides. To obtain some sequence data on the large oligonucleotides obtained from the viral genome, compositional analyses have been undertaken. ^{32}P-labeled oligonucleotides were eluted from gel plugs, recovered, and digested with pancreatic RNase, and the products were resolved by thin-layer polyethyleneimine two-dimensional chromatography as described in Materials and Methods. For each digest the phosphate specific activity was deduced from the label recovered in the guanine-containing nucleotide. Thus, for an oligonucleotide yielding A_2G, one-third of the recovered A_2G label was taken to be the phosphate specific activity. The moles of each nucleotide in a digest were then calculated by dividing their counts by the phosphate specific activity × the nucleotide size. These calculations assume that the α-phosphates of the precursor triphosphates used in vivo to make the RNA had essentially equivalent specific activities, an assumption which was borne out by obtaining essentially equivalent counts (±10%) for RNase T_2 (giving 3′-mononucleotides) and nuclease P1 (giving 5′-mononucleotides) digests of ^{32}P-labeled VSV RNA (data not shown).

The results obtained for 59 oligonucleotide spots, representing the average of several determinations, are shown in Table 1. The 12 other oligonucleotides were too poorly resolved for meaningful results to be obtained. Other than oligonucleotide 44, every oligonucleotide examined gave only one nucleotide containing guanine (G, AG, or A_2G, etc.), indicating that they were in all probability unique and not a mixture of two or more oligonucleotides. Oligonucleotide 44 gave both G and AG residues. As mentioned previously (4), we have observed that this oligonucleotide frequently has approximately twice the radioactivity of neighboring oligonucleotides (e.g., 43 or 45). For these reasons we conclude that the oligonucleotide spot 44 has two components, which we designate 44A and 44B.

The compositional analyses indicate that the four oligonucleotides that migrate the slowest in the second-dimension gel (1 to 4) have from 28 to 38 nucleotides. Although in general the compositional analyses indicated that oligonucleotides that moved slowly in the second dimension were larger than the faster-migrating oligonucleotides, this was not always true (e.g., compare 10 and 59), suggesting that the nucleotide sequences and possibly their secondary structures influenced the relative mobilities of the various oligonucleotides.

Sequence analyses of selected RNase T_1 oligonucleotides. In view of the fact that, when a battery of nucleases is used to prepare mRNA-viral RNA duplexes (4, 11), certain viral oligonucleotides are not recovered in any hybrid (i.e., oligonucleotides 3, 38, 42, 43, 44B, 59, and 70), the possibility was raised that some of these oligonucleotides may represent mRNA terminal or intercistronic (or other) noncoding sequences. Taking into account the published sequencing data relating to the genome of VSI (15–17, 19a–21), we have sequenced some of these oligonucleotides to determine, where possible, (i) the location of the oligonucleotides on the viral genome and (ii) whether there are any similarities between their sequences.

Oligonucleotide 70. Compositional analyses of oligonucleotide 70 indicated that it lacked U residues and contained one AG, two C, two AC, and one A_4C residue. When oligonucleotide 70 was 5′-end labeled and subjected to partial digestion with various nucleases (Fig. 2), its sequence was deduced, based on the results obtained and its compositional analyses (Table 1), to be:

3′...GACCAAAACACCA(G)...5′

Oligonucleotide 59. The compositional analyses of oligonucleotide 59 indicated that it contained one A_2G, one AU, one A_6U, one C, one AC, and one A_5C residue. Based on these results, and the results obtained by partial digestion with various nucleases of 5′-end-labeled material (Fig. 3), the following sequence was deduced:

3′...GAACACCAAAAAUAAAAAAUA(G)...5′

The genome location of oligonucleotides 59 and 70. It has been shown that neither oligonucleotide 59 nor 70 was protected when mRNA-viral RNA duplexes were formed (4) and that both are found in the small DI RNA species (Fig. 1D to F). Schubert and associates (20) and Semler and collaborators (21) have deduced the 5′-end sequence of VSI by analyzing the in vitro transcription product of a small DI particle which has a 3′ sequence complementary to the viral 5′ end (as well as its own 5′ sequence) and makes a small transcript in vitro identical to the viral 5′-end sequence. Direct sequence analyses of the 5′ ends of VSI DIs agree with the findings with the DI in vitro transcripts (19a; Schubert et al., submitted for publication). When the deduced sequence for the 5′ end of VSI (19a–21) was compared to the results obtained for oligonucleotides 59 and 70, it was found that their sequences exactly corresponded to two oligonucleotide sequences located near the 5′ terminus of the VSI genome:

TABLE 1. *Pancreatic RNase digests of VSI oligonucleotides*

	Moles of nucleotide per mole of oligonucleotide															
Oligo	G	AG	A_2G	A_3G	A_nG	U	AU	A_2U	A_3U	A_4U	A_nU	C	AC	A_2C	A_3C	A_nC
1	1	0	0	0	0	4-5	4	2	1	0	0	6-7	3	0	0	0
2	0	1	0	0	0	9-11	5	0	1	0	0	2-3	3	0	1	0
3	0	0	1	0	0	8-10	2	1	0	1	$2A_5U$	1	0	0	0	0
4	1	0	0	0	0	6-7	4-5	2-3	1	0	0	1-2	1	1	0	0
5	1	0	0	0	0	6-7	5-6	0	0	0	0	4-5	1-2	2-3	0	0
6	0	1	0	0	0	6-7	2-3	0	0	0	0	0-1	0	1	0	0
7	1	0	0	0	0	8	1-2	2	0	0	0	5-6	1	0	0	0
8	0	0	0	1	0	7	0	1	1	0	0	2	0	0	0	0
9	1	0	0	0	0	6-7	4-5	2	0	0	0	4	1-2	0	0	0
10	0	0	1	0	0	3-4	2	2-3	0	0	0	1-2	1-2	0	0	0
11	1	0	0	0	0	6-7	1	1-2	0	0	0	3-4	3	1-2	0	0
12	1	0	0	0	0	7-8	0	2	0	0	0	6-7	3-4	2	0	0
13	1	0	0	0	0	2-3	2	0	0	0	$1A_5U$	2	2	0	0	0
14	1	0	0	0	0	4-5	2-3	0	0	0	0	6-7	0	1-2	0	0
15	1	0	0	0	0	7-8	4-5	0	0	0	0	6-7	1-2	0	0	0
16	0	0	1	0	0	8-10	2	2-3	0	0	0	2-3	0	0	0	0
17	1	0	0	0	0	9-10	3-4	1	0	0	0	5-6	2	0	0	0
18	1	0	0	0	0	8	3	0	0	0-1	0	2-3	2-3	0	0	0
19	0	1	0	0	0	5-6	3	1-2	0	0	0	1-2	0	0	0	$1A_4C$
20	1	0	0	0	0	4-5	2	0-1	0	0	$0\text{-}1A_6U$	4-5	1	0	0	0
21	0	1	0	0	0	7-8	4	2	0	0	0	2-3	2	0-1	0-1	0
22,23	Not Analysed															
24	0	1	0	0	0	3-5	0	1-2	0	2	0	0-2	1-2	0	0	0
25	0	0	0	1	0	5	3	0	0	0	0	1-2	0-1	0	0	0
26	1	0	0	0	0	1-2	2	0	0	0	0	0-1	1	0	0	0
27	0	1	0	0	0	9-10	1	1	0	0	0	2-3	2	1	0	0
28	1	0	0	0	0	4-6	0	0	0	0	0	3	1-2	0	1-2	0
29,30,31	Not Analysed															
32	0	1	0	0	0	9-10	1-2	0	0	0	0	3-5	1	0	0	0
33	0	0	1	0	0	9-10	0	0-1	0	0	0	0-1	0	0	0	0
34	1	0	0	0	0	7-8	2	2	0	0	0	2	0-1	1	0	0
35	Not Analysed															
36	0	1	0	0	0	9-10	1-2	1	0	0	0	3-4	0	0	0	0
37	0	1	0	0	0	5-9	2-4	0	0-1	0	0	0-1	0-1	0	0-1	0
38	0	0	0	1	0	8-10	1	0	0	0	0	1	1	0	0	0
39	Not Analysed															
40	1	0	0	0	0	6	2	0	0	0	0	1	0	1	0	0
41	1	0	0	0	0	7-8	3	0	0	0	0	2	0	0	0	0
42	0	1	0	0	0	6-8	1	0	0	0	0	1	1	0	0	0
43	1	0	0	0	0	7-8	2	0	0	0	0	1	0	0	0	0
44A+B	1	1	0	0	0	8-12	4	0	0	0	0	2	0	0	0	0
45	0	1	0	0	0	9-10	2-3	1	0	0	0	2	0	0	0	0
46	Not Analysed															
47	0	0	1	0	0	7-8	2	1	0	0	0	1-2	0	0	0	0
48	Not Analysed															
49	1	0	0	0	0	3-4	2	0	0	0	0	2	1	0	0	0
50	1	0	0	0	0	6-7	2	1	0	0	0	2-3	1	0	0	0
51,52	Not Analysed															
53	1	0	0	0	0	4-6	1-2	0	0	0	0	0-1	0	1	0	0
54	1	0	0	0	0	2-3	0-1	0	0	0	0	2	2	2	0	0
55	0	1	0	0	0	4-5	1	1	0	0	0	0	0	0	0	0
56	0	0	1	0	0	9	0	0	1	0	0	4-5	1	1	0	0
57	1	0	0	0	0	4	0-1	0	0	0	0	0-1	1	2	1	0
58	1	0	0	0	0	2-3	1	0	0	0	0	0-2	1	1	0	$1A_4C$
59	0	0	1	0	0	0	1	0	0	0	$1A_6U$	1	1	0	0	$1A_5C$
60	1	0	0	0	0	0	2	1	0	0	0	5-6	3	0	0	0
61	0	0	0	1	0	0-1	1	0	0	0	$1A_6U$	2-3	1-2	0	0	0
62	0	1	0	0	0	4-5	5	2	0	0	0	3-4	2	0	0	0
63	1	0	0	0	0	2	0	1	0	0	0	2-3	0-1	0	0	$1A_4C$
64	Not Analysed															
65	1	0	0	0	0	2	2-3	0	0	0	0	1-2	0	0	1	0
66	0	0	1	0	0	2	2	0	0	1	0	2	2	0	0	0
67	0	1	0	0	0	6	2-3	0	0	0	0	4	2	1	0	0
68	1	0	0	0	0	6	3	2	0	0	0	2-3	2	3	0	0
69	1	0	0	0	0	0	1	1	0	0	0	1	1-2	1	0	0
70	0	1	0	0	0	0	0	0	0	0	0	2	2	0	0	$1A_4C$
71	0	0	0	0	$1A_6G$	0	0	0	0	0	0	2	0	1	0	0

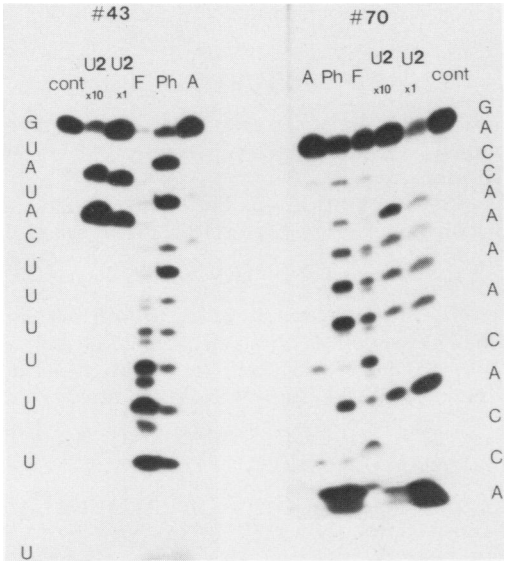

FIG. 2. *Sequence analyses of oligonucleotides 43 and 70. Undigested control (cont) 5'-end-labeled oligonucleotides, and samples incubated with formamide (F) and the enzymes RNase U_2 (two concentrations), Phy 1 (Ph), and RNase A (A), were electrophoresed in 20% polyacrylamide gels as described elsewhere (8). Although it has been shown that Phy 1 enzyme leaves A-U bonds (15) and most C-U, C-C, and C-G bonds intact (11, 19, 23), under the conditions used here C-A bonds were sometimes digested. The poor digestion by RNase A is probably caused by using too low an enzyme concentration. Under the conditions used, RNase A exhibited a preference for cutting pyrimidines next to A residues (see reference 11). The reason why RNase U_2 did not cut between A and G in oligonucleotide 70 is not known. However, compositional analyses indicate that 70 has an AG residue (Table 1). The deduced sequences shown at the sides were correlated with the compositional analyses given in Table 1 to allow decisions to be made on the positions of pyrimidine residues.*

3' ... GAAAUUCAUACUUUUUUU/
38
GAAACUAGGAAUUCUGGGA-
GAACACCAAAAAUAAAAAAUA/
59
GACCAAAACACCA/GAAGCAppp ... 5'
70

Thus it appears that oligonucleotides 59 and 70 come from the 5'-end region of the viral genome that does not code for mRNA.

Oligonucleotide 38. The compositional analyses of oligonucleotide 38 (Table 1) indicated that it had one A_3G, 8 to 10 U, one AU, one C, and one AC residue. From these results and the data obtained by 5'-end labeling and partial nuclease digestions (Fig. 4), the following sequence was deduced:

3' ... GAAAUUCAUACUUUUUU(U)(G) ... 5'

This sequence is comparable to that proposed by Schubert and associates (submitted for publication) as the L mRNA termination and polyadenylation site (see above). The fact that oligonucleotide 38 can be recovered in mRNA-viral RNA duplexes (Schubert et al., submitted for publication) when they are prepared by RNase A and T_1 digestion, but not (4) when nucleases S1, T_2, T_1, and A were used, suggests that with the latter battery of nucleases there is paring down of the L mRNA-viral RNA duplex and consequent loss of oligonucleotide 38.

Oligonucleotide 43. Compositional (Table

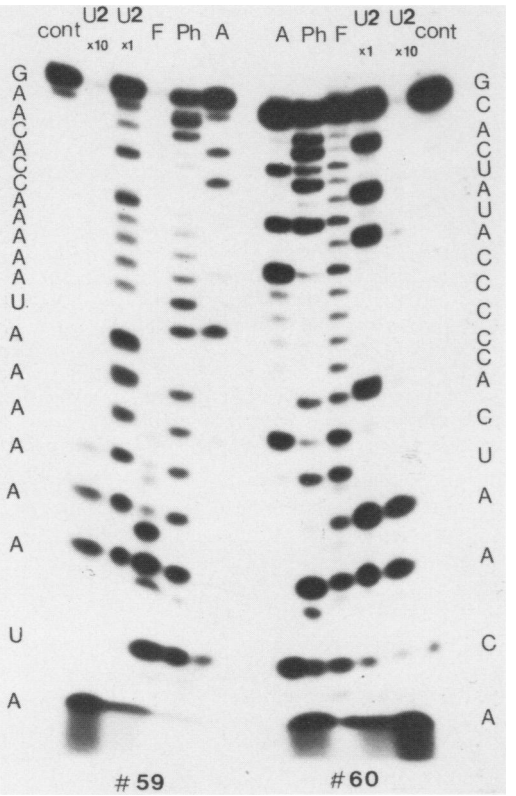

FIG. 3. *Sequence analyses of oligonucleotides 59 and 60. The electrophoretic separation is shown of control (cont) lanes of untreated 5'-end-labeled samples and samples treated with formamide (F) and the enzymes RNase U_2 (two concentrations), Phy 1 (Ph), and RNase A (see the text and legend to Fig. 2). The deduced sequences shown at the sides were correlated with the compositional analyses given in Table 1 to allow decisions to be made on the positions of pyrimidine residues.*

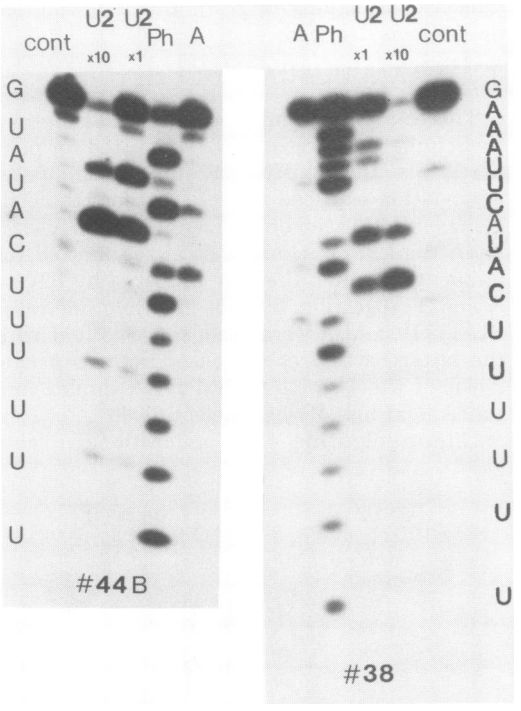

FIG. 4. *Sequence analyses of oligonucleotides 38 and 44B. The electrophoretic separation is shown of control (cont) lanes of untreated 5′-end-labeled samples and samples treated with the enzymes RNase U_2 (two concentrations), Phy 1 (Ph), and RNase A (see the text and legend to Fig. 2). The deduced sequences shown at the sides were correlated with the compositional analyses given in Table 1 to allow decisions to be made on the position of pyrimidine residues. A seventh U may have run off the bottom of the gels.*

1) and 5′-end sequencing (Fig. 2) analyses of oligonucleotide 43 indicate that it has the sequence:

3′ ... GUAUACUUUUUUU(G) ... 5′

This sequence exactly corresponds to the sequence recently described by McGeoch (15) at the N-NS intercistronic (N polyadenylation-termination) sequence.

Oligonucleotide 3. The sequence deduced for oligonucleotide 3 (see Fig. 5) was:

3′ ... GAAUUAAUAUAAAAUUAAAAAUU-AAAAAUACUUUUUU(U)(G) ... 5′

The position of the only C residue was determined by homochromatography (see Fig. 5), and the 3′-end nucleotide sequence was determined from more extensive electrophoreses than those shown in Fig. 5.

Oligonucleotide 44B. The sequence deduced for oligonucleotide 44B (Table 1, Fig. 4) was:

3′ ... GUAUACUUUUUU(U)(G) ... 5′

The position of the only C residue was determined by homochromatography (data not shown).

Oligonucleotide 42. The sequence deduced for oligonucleotide 42 (Table 1, Fig. 6) was:

3′ ... GAUACUUUUUUUCAU(U)(G) ... 5′

Oligonucleotides 1, 2, 13, 39, 58, 60, and 65. The sequences obtained for oligonucleotides 1, 2, 13, 39, 58, 65 (data for all of these, not shown), and 60 (Fig. 3) were as follows:

1: 3′ ... GUACUCUAAUCUACUAAAUU-AAUACACCUUACCACA(G) ... 5′

2: 3′ ... GAUUUUAUAUAUACCAUAC-AUUUUUAAAUUUCCACAAA(G) ... 5′

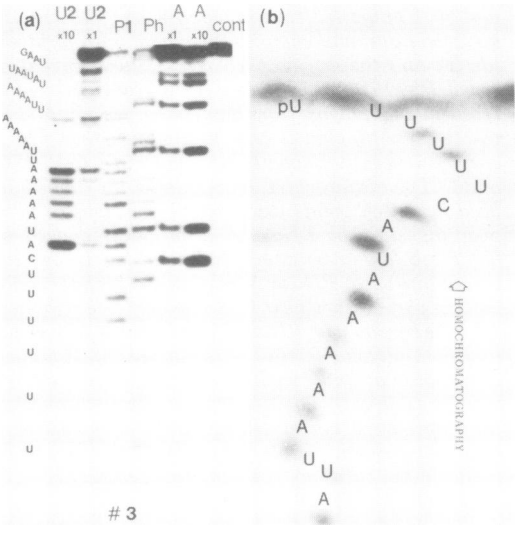

FIG. 5. *Sequence analyses of oligonucleotide 3. (a) Electrophoretic separation of a control (cont) lane of an untreated 5′-end-labeled sample and samples treated with the enzymes P1, Phy 1 (Ph), RNase A, and two concentrations of RNase U_2 (see Fig. 2, legend). The deduced sequence shown at the side was obtained from this and more extensive electrophoreses (data not shown) and correlated with the compositional analysis given in Table 1 to allow decisions to be made on the positions of the pyrimidine residues. A seventh U may have run off the bottom of the gel. (b) Top portion of an electrophoretic-homochromatography separation of a partial P1 enzyme digest of end-labeled oligonucleotide 3 (22). The resolved nucleotides allowed a decision to be made on the position of the only C residue in oligonucleotide 3 (see the text).*

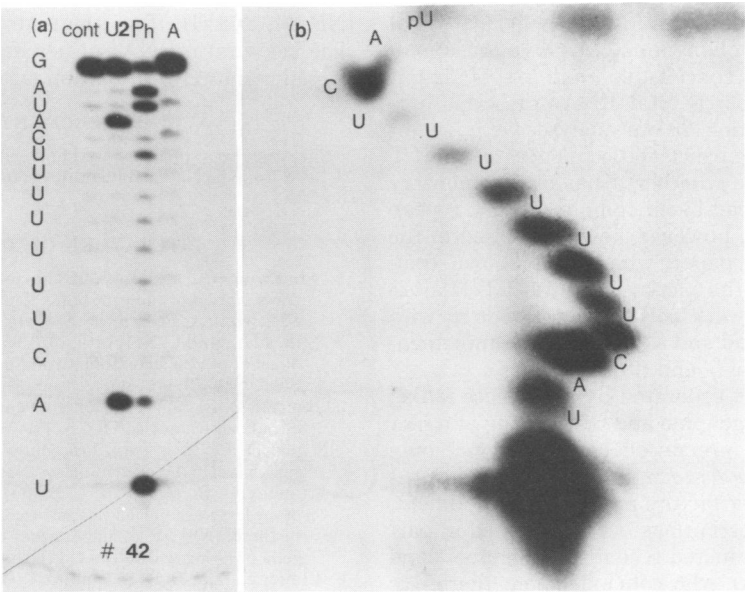

FIG. 6. *Sequence analyses of oligonucleotide 42. (a) Electrophoretic separation of a control (cont) lane of an untreated 5'-end-labeled sample of oligonucleotide 42, and samples treated with RNases U_2, Phy 1 (Ph), and A (see the text and legend to Fig. 2). The deduced sequence shown at the side was correlated with the compositional analysis given in Table 1 to allow decisions to be made on the position of the pyrimidine residues. A terminal second U may have run off the bottom of the gel. (b) Electrophoretic-homochromatography separation of a partial P1 enzyme digest of end-labeled oligonucleotide 42 (24). The resolved nucleotides allowed a decision to be made on the position of the two C residues in oligonucleotide 42.*

13: 3′ ... GUAAAAAC(or U)UACACAC(or U)CCCUU(or C)U(or C)AU(G) ... 5′

39: 3′ ... GUUUUCCCUUUCUAUU(G) ... 5′

58: 3′ ... GCAAAACACAACCUCUAU(G) ... 5′

60: 3′ ... GCACUAUACCCCCACUAACA-(G) ... 5′

65: 3′ ... GCCUACCAAACUUCUA(G) ... 5′

The sequence obtained for oligonucleotide 65 exactly corresponds to an N gene 3′ sequence recently characterized by McGeoch and Turnbull (16). This oligonucleotide has been shown to be protected by N mRNA (4).

DISCUSSION

The origin of the LT DI. The fingerprint analyses of LT DI RNA (Fig. 1C), plus the sequence analyses of oligonucleotides 38, 59, and 70, clearly indicate that the LT DI RNA has a sequence containing information derived from both the 5′ and 3′ ends of the viral genome. It is possible that other LT types are present in the LT preparations we have analyzed (see accompanying paper, reference 9) also having 3′-derived sequences. If so, then any differences in the oligonucleotide fingerprints caused by their presence have been subliminal to our analyses. When the recoveries of oligonucleotides 38, 59, and 70 have been compared to 3′-end nucleotides (e.g., 11, 13, 58, 65), no significant reduction in the amounts of the former three oligonucleotides have been noted, indicating that if any 3′ readback LT DI types are present they have not been a significant fraction of the preparations.

Although LT DI RNA lacks most of the L gene information, some has been retained, including the putative L mRNA polyadenylation (termination) signal and an L mRNA oligonucleotide (47) which may well be the nearest internal L gene large T_1 oligonucleotide to oligonucleotide 38. Studies reported in the accompanying paper (9), using alternate analytical techniques, agree with the conclusions drawn here on the origin of the LT DI we have analyzed. Our studies realize an alternate mechanism involved in DI genesis, i.e., that of internal deletion.

The mechanism of genesis of LT RNA by internal deletion may occur by the RNA polymerase enzyme falling off a template and reattaching elsewhere. Whether the same template or another is used for the continuation of RNA

synthesis is not known. Alternatively, instead of the genesis involving an RNA polymerase malfunction, it is conceivable that a splicing-out mechanism from a total RNA transcript produces the deletion. In our analyses we have not determined the exact end sequences of the LT DI RNA to see whether it has complementary ends (i.e., 3′ read-back, complementary types). This issue has, however, been addressed in the accompanying paper, where it is shown that, depending on the passage history, LT DIs are present either with both parental ends or with a parental 5′ end and a 3′ sequence complementary to the viral 5′ end (9).

If DIs can be generated by a replicase falling off a template genome and reattaching at a new site, questions are raised concerning not only what promotes dissociation, but also what sequence or other factors promote reattachment. This latter question has been addressed in part by a study conducted recently by Schubert and associates (19a), who concluded that there was a specific RNA polymerase recognition site involved in the generation of DI particles. Whether there are preferred detachment sites is an open question.

Oligonucleotide sequence analyses that suggest that each mRNA may have common termination and/or polyadenylation signals. The analyses reported here indicate that oligonucleotides 38 and 43 correspond to the L and N (respectively) mRNA polyadenylate-termination sites (15; Schubert et al., submitted for publication), with oligonucleotides 59 and 70 (5′) distal to the end of L mRNA. Both oligonucleotides 38 and 43 have a sequence (3′) ... AUACUUUUUU(U)(G) ... (5′). These two oligonucleotides are not protected from nuclease digestion when a battery of nucleases is used to derive mRNA-viral RNA hybrids (4). Two other oligonucleotides (3 and 44B) that are also not protected in the derivation of duplexes (4) have sequences that are comparable to oligonucleotides 38 and 43, i.e., they have a nucleotide stretch:

(3′) ... AUACUUUUUU(U)(G) ... 5′

Oligonucleotide 42, the only other oligonucleotide that is not protected from high nuclease digestion of mRNA-viral RNA duplexes, is slightly different at its 5′ end although it has a similar sequence to the other oligonucleotides, i.e.:

(3′) ... AUACUUUUUUUCAU(U)(G) ... 5′

Although not proven by direct analyses, these results suggest that oligonucleotides 3, 42, and 44B may represent the polyadenylation-termination sites of the other three mRNA species (NS, M, and G). If so, then our results suggest that the viral mRNA species have comparable polyadenylation-termination signals.

ACKNOWLEDGMENTS

This study was supported by Public Health Service grant AI-13402 from the National Institute of Allergy and Infectious Diseases.

LITERATURE CITED

1. **Abraham, G., and A. K. Banerjee.** 1976. Sequential transcription of the genes of vesicular stomatitis virus. Proc. Natl. Acad. Sci. U.S.A. **73**:1504-1508.
2. **Ball, L. A., and C. N. White.** 1976. Order of transcription of genes of vesicular stomatitis virus. Proc. Natl. Acad. Sci. U.S.A. **73**:442-446.
3. **Barrell, B. G.** 1971. Fractionation and sequence analysis of radioactive nucleotides, p. 751-779. *In* S. L. Cantoni and D. R. Davis (ed.), Procedures in nucleic acid research, vol. 2. Harper and Row, New York.
4. **Clewley, J. P., and D. H. L. Bishop.** 1979. Assignment of the large oligonucleotides of vesicular stomatitis virus to the N, NS, M, G, and L genes and oligonucleotide gene ordering within the L gene. J. Virol. **30**:116-123.
5. **Clewley, J. P., D. H. L. Bishop, C.-Y. Kang, J. Coffin, W. M. Schnitzlein, M. Reichmann, and R. E. Shope.** 1977. Oligonucleotide fingerprints of RNA species obtained from rhabdoviruses belonging to the vesicular stomatitis virus subgroup. J. Virol. **23**:152-166.
6. **Clewley, J. P., J. Gentsch, and D. H. L. Bishop.** 1977. Three unique viral RNA species of snowshoe hare and La Crosse bunyaviruses. J. Virol. **22**:459-468.
7. **de Wachter, R., and W. Fiers.** 1972. Preparative two dimensional polyacrylamide gel electrophoresis of ^{32}P-labeled RNA. Anal. Biochem. **49**:184-197.
8. **Donis-Keller, H., A. M. Maxam, and W. Gilbert.** 1977. Mapping adenines, guanines and pyrimidines in RNA. Nucleic Acids Res. **4**:2527-2538.
9. **Epstein, D. A., R. C. Herman, I. Chien, and R. A. Lazzarini.** 1980. Defective interfering particle generated by internal deletion of the vesicular stomatitis virus genome. J. Virol. **33**:818-829.
10. **Freeman, G. J., D. D. Rao, and A. S. Huang.** 1979. RNA synthesis of vesicular stomatitis virus. VIII. Oligonucleotides of the structural genes and mRNA. Gene **5**:141-157.
11. **Gupta, R. C., and K. Randerath.** 1977. Use of specific endonuclease cleavage in RNA sequencing—an enzymic method for distinguishing between cytidine and uridine residues. Nucleic Acids Res. **4**:3441-3453.
12. **Huang, A. S.** 1977. Viral pathogenesis and molecular biology. Bacteriol. Rev. **41**:811-821.
13. **Kang, C. Y., T. Glimp, J. P. Clewley, and D. H. L. Bishop.** 1978. Studies on the generation of vesicular stomatitis virus (Indiana Serotype) defective interfering particles. Virology **84**:142-152.
14. **Leppert, M., L. Kort, and D. Kolakofsky.** 1977. Further characterization of Sendai virus DI-RNA's: a model for their generation. Cell **12**:539-552.
15. **McGeoch, D. J.** 1979. Structure of the gene N: gene NS intercistronic junction in the genome of vesicular stomatitis virus. Cell **17**:673-681.
16. **McGeoch, D. J., and N. T. Turnbull.** 1978. Analysis of the 3′-terminal nucleotide sequence of vesicular stomatitis virus N protein mRNA. Nucleic Acids Res. **5**: 4007-4024.
17. **Perrault, J., and R. W. Leavitt.** 1977. Inverted complementary terminal sequences in single-stranded RNAs and snap-back RNAs from vesicular stomatitis defective interfering particles. J. Gen. Virol. **38**:35-50.
18. **Petric, M., and L. Prevec.** 1970. Vesicular stomatitis virus—a new interfering particle, intracellular struc-

tures, and virus specific RNA. Virology **41**:615–630.

19. **Pilly, D., A. Niemeyer, M. Schmidt, and J. P. Bargetzi.** 1978. Enzymes for RNA sequence analysis. J. Biol. Chem. **253**:437–445.

19a. **Schubert, M., J. D. Keene, and R. A. Lazzarini.** 1979. A specific internal RNA polymerase recognition site of VSV RNA is involved in the generation of DI particles. Cell **18**:749–757.

20. **Schubert, M., J. D. Keene, R. A. Lazzarini, and S. U. Emerson.** 1978. The complete sequence of a unique RNA species synthesized by a DI particle of VSV. Cell **15**:103–112.

21. **Semler, B. L., J. Perrault, J. Abelson, and J. J. Holland.** 1978. Sequence of a RNA templated by the 3'-OH RNA terminus of defective interfering particles of vesicular stomatitis virus. Proc. Natl. Acad. Sci. U.S.A. **75**:4704–4708.

22. **Silberklang, M., A. M. Gillum, and U. L. RajBhandary.** 1977. The use of nuclease P1 in sequence analysis of end group labeled RNA. Nucleic Acids Res. **4**:4091–4108.

23. **Simoncsits, A., G. G. Brownlee, R. S. Brown, J. R. Rubin, and H. Guilley.** 1977. New rapid gel sequencing method for RNA. Nature (London) **269**:833–836.

24. **Volckaert, G., and W. Fiers.** 1977. Micro thin-layer techniques for rapid sequence analysis of ^{32}P-labeled RNA: double digestion and pancreatic ribonuclease analyses. Anal. Biochem. **83**:228–239.