JOURNAL OF VIROLOGY, Feb. 1980, p. 789-794
0022-538X/80/02-0789/06$02.00/0

Vol. 33, No. 2

Intervening Sequence Between the Leader Region and the Nucleocapsid Gene of Vesicular Stomatitis Virus RNA

JACK D. KEENE,[1]* MANFRED SCHUBERT,[2] AND ROBERT A LAZZARINI[2]

Department of Microbiology and Immunology, Duke University Medical Center, Durham, North Carolina 27710,[1] and Molecular Virology Section, Laboratory of Molecular Biology, National Institute of Neurological, Communicative Disorders and Stroke, National Institutes of Health, Bethesda, Maryland 20205[2]

The base sequence at the 3′ end of vesicular stomatitis virus RNA was determined by using terminal labels and chemical RNA sequencing. The leader RNA was complementary to 47 bases at the 3′ terminus, whereas the nucleocapsid gene (*N*) began 51 nucleotides from the 3′ end of the genomic RNA. The intervening bases were 3′...GAAA...5′ for the Indiana serotype and 3′...GAAAA...5′ for the New Jersey serotype. The complements of these bases did not appear in either the leader RNA or the *N* mRNA. This sequence may function as a stop signal or cleavage site during transcription. Furthermore, processing or termination at this sequence must be inhibited during the production of full-length RNA plus-sense strands (replication). We recently found similar sequences approximately 46 to 48 nucleotides from the 3′ ends of several defective interfering particle RNAs where the short defective interfering particle transcription products terminate. This sequence is present also at the end of the polymerase (*L*) gene.

Vesicular stomatitis virus (VSV) contains a single-stranded RNA genome of antimessage sense. Evidence from studies of UV inactivation and in vitro transcription suggests that the viral polymerase initiates at the 3′ end and continues in the 5′ direction, with the production of each mRNA being sequential and dependent upon synthesis of the previous transcript (1, 2). The gene order of the Indiana serotype of VSV (VSV_{IND}) was found to be 5′-L-G-M-NS-N-leader 3′ (2, 3). The leader RNA initiates at the precise 3′ end of the genomic RNA (4) and has been fully sequenced (5).

The replication of VSV RNA requires that a full-length plus-sense RNA be made from the genomic minus strand. Unlike the mRNA's, the full-length plus-sense RNA is not capped, methylated, or polyadenylated. The transition from primary transcription to replication during VSV growth correlates with cessation of these modifications. Therefore, the switch to the replicative process depends upon suppression of the termination or cleavage at the intergenomic junction.

To study the details of the transcriptive-replicative transition, we determined the nucleotide sequences at the first intervening junction in VSV RNA.

MATERIALS AND METHODS

Virus growth and RNA purification. VSV_{IND} of both the Mudd-Summers and heat-resistant strains and the New Jersey serotype of VSV (VSV_{NJ}) of the Hazelhurst strain were grown and purified as described (11, 14). RNA was extracted with sodium dodecyl sulfate-phenol and isolated on 10 to 30% sucrose gradients in 0.1% sodium dodecyl sulfate.

Terminal labeling of RNA. RNAs stored in water were precipitated with NaCl and ethanol immediately before being labeled. After rinsing twice with 80% ethanol, 10 μl of buffer containing 16 mM $MgCl_2$, 6 mM dithiothreitol, 100 μM ATP, 100 mM HEPES (*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid; pH 8.0), and 100 μCi of 5′-[^{32}P]cytidine diphosphate were added. After lyophilization to dryness, the residue was dissolved in 17 μl of 10% dimethyl sulfoxide–10% glycerol, and 3 μl of RNA ligase (P-L Biochemicals, Inc.) was added. The reaction was incubated for 18 h at 4°C and stopped by the addition of 100 μl of buffer containing 0.4 M NaCl, 0.01 M Tris (pH 7.4), and 0.1% sodium dodecyl sulfate. Labeled RNA was isolated by chromatography over Sephadex G-75 and was recovered by ethanol precipitation in the presence of carrier tRNA.

Sequencing procedure. Labeled RNAs were analyzed for base sequence by using the chemical RNA sequencing procedure of Peattie (17) with only minor modifications. After chain scission the fragments were separated by 12% acrylamide gel electrophoresis at 1,200 V for various times. Labeled fragments were identified by autoradiography with Kodak XR-2 film.

RESULTS

VSV RNAs were labeled at their 3′ ends with RNA ligase and 5′-[^{32}P]cytidine diphosphate, as previously described (13). Labeled RNAs were recovered by chromatography over Sephadex G-

75 and ethanol precipitation. RNAs were sequenced by using chemical cleavages (17). Specific bases were modified with dimethyl sulfate for G, diethyl pyrocarbonate for A, hydrazine in 3 M NaCl for C, and 50% hydrazine-water for U residues. RNA chains were broken with aniline treatment and analyzed on 12% sequencing gels.

We previously used partial nuclease digestion of 3′-labeled VSV RNA to determine the first 18 bases of VSV_{IND} (13). This sequence was confirmed by chemical sequencing methods. In Fig. 1, we show the sequence of VSV_{IND} RNA starting at position 13. The G, A, C, and U bands at the left of the figure (secondary loading) do not align precisely with bands in the alkali ladder (Fig. 1, L), presumably due to the presence of 5′ OH groups on ladder bands and 5′ phosphate groups after chain scission with aniline. At po-

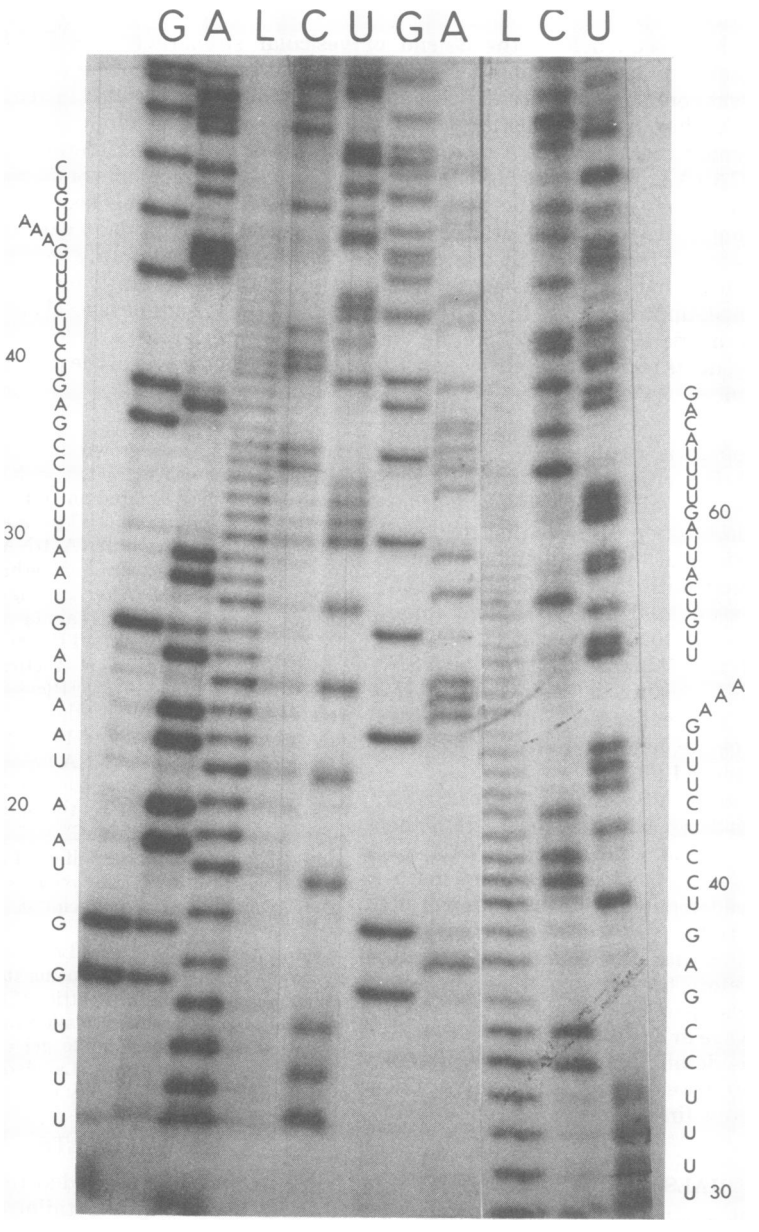

FIG. 1. *Rapid chemical RNA sequencing of 3′-labeled VSV_{IND} RNA on a 12% urea-acrylamide gel. The five positions on the right side of the autoradiogram constitute the primary loading, and the five positions on the left show the secondary loading. Each base cleavage is denoted at the top. L denotes the partial alkali ladder.*

sition 27 and beyond, however, the bands are precisely coincident. In addition, the first group of samples loaded on the gel (right half on Fig. 1) are all precisely aligned with the respective ladder bands. The inclusion of a ladder in these gels offers an additional indexing method for rapid RNA sequencing. The RNA sequence is unambiguously read to well over 70 bases (Fig. 1). The sequence between positions 13 and 46 is 3′... UUUGGUAAUAAUAGUAAUUUUCCGAGUCCUCUUU... 5′ which is complementary to the leader RNA (5). The leader RNA contained three or four adenylate residues near its 3′ terminus, whereas the 3′-terminal base was an adenylate in 32% of the molecules and a cytidine in 68% of the molecules. These values were determined by 3′ in vitro labeling and may not accurately reflect the relative content of these terminal bases, since RNA ligase is known to preferentially label according to the terminal base sequence of the acceptor molecule. The template, as shown in Fig. 1 and 3, contains only three uridylates followed by a G. If the leader incorporates an additional A residue, it may do so by slippage or chattering on the short U stretch. Other stretches of U residues occur earlier in the template (positions 9 to 15 and 30 to 33), however, and no additional A residues were seen in the complementary leader sequence. The possibility of chattering on oligo(U) in the template is interesting because the *L* and *N* mRNA's of VSV appear to be polyadenylated by chattering (16; M. Schubert, J. D. Keene, R. C. Herman, and R. A. Lazzarini, submitted for publication).

Immediately following the three uridylates at the end of the leader region is the sequence 3′... GAAA... 5′. The C which is sometimes present at the 3′ end of leader is presumably copied from the G, and termination or cleavage occurs at this point during transcription.

After the short intervening purine sequence, the characteristic mRNA initiation sequence 3′... UUGUC... 5′ is found, followed by 26 bases that are each precisely complementary to the 5′ end of the *N* mRNA (20). The sequence is 3′... AUUAGUUUUACAGACAAUGUCAGUUC... 5′ (Fig. 3). Thus, it is clear that a short intervening sequence of just four nucleotides resides at the first intervening region of VSV_{IND} RNA.

The same analysis described above was applied to the VSV New Jersey serotype (VSV_{NJ}). We previously reported the sequence of the first 21 nucleotides at the 3′ end of VSV_{NJ} RNA and found the identical 21 bases at the 3′ end of VSV_{NJ} defective interfering (DI) particle RNAs (12). Also, Colonno and Banerjee (6) have reported the sequence of the leader RNA from VSV_{NJ}.

The results of the chemical sequencing studies of VSV_{NJ} RNA are shown in Fig. 2 and 3. In the first 48 bases, there are two differences from the sequence reported by Colonno and Banerjee (6). In positions 21 and 31, they reported A in the leader RNA and thus predicted U in the complementary genomic RNA. The sequence determined here is G in position 21 and A in position 31. This may be due to the use of different strains of VSV_{NJ}.

Another difference between the VSV_{NJ} leader sequence determined by Colonno and Banerjee and the sequence reported here is the presence of three adenylate residues near the 3′ end of the leader, whereas only two U residues are present in the genome (Fig. 2 and 3). This could again be due to chattering of the polymerase; however, it is necessary to involve additional flanking sequences in the chattering process since uridylates occur frequently in the VSV_{NJ} template where chattering is not observed (e.g., positions 9 to 15). Thus, chattering might occur when the U residues are followed by an adenylate-rich sequence.

Following the end of the leader region of VSV_{NJ} is the purine-rich sequence 3′... GAAAA... 5′ (Fig. 2 and 3). The commonality of this sequence with that between the leader region and first gene of VSV_{IND} is evident. The 5′ ribosomal binding sites of VSV_{NJ} mRNA's have not been reported; however, the VSV_{NJ} mRNA's, like the VSV_{IND} RNAs, initiate with the common sequence 5′... AACAG... 3′ (8). Consequently, it is very likely that the first gene after the leader RNA begins at position 52 with the sequence UUGUC, which immediately follows the intervening sequence 3′... GAAAA... 5′ (Fig. 2 and 3). The subsequent sequence (Fig. 2 and 3) is presumed to be the *N* gene of VSV_{NJ}; however, the gene order of VSV_{NJ} has not been reported.

DISCUSSION

The leader RNAs of both VSV_{IND} and VSV_{NJ} are 48% adenylate residues (5, 6), and the templates for leader transcription have relatively few adenylate residues. The first adenylate residue in the genome RNA is 19 bases from the 3′ terminus of both VSV_{IND} (5, 13) and VSV_{NJ} (6, 12). The work described here shows that an adenylate-rich sequence resides at the end of the VSV leader RNA template. This sequence may represent a polymerase termination site or a processing site at which endonucleolytic cleavage takes place (9).

In a recent study we have shown that the base sequence at the end of the *L* gene is 5′... AAAGUUUUUUU... 3′ (Schubert et al., submitted for publication). The polyadenylate

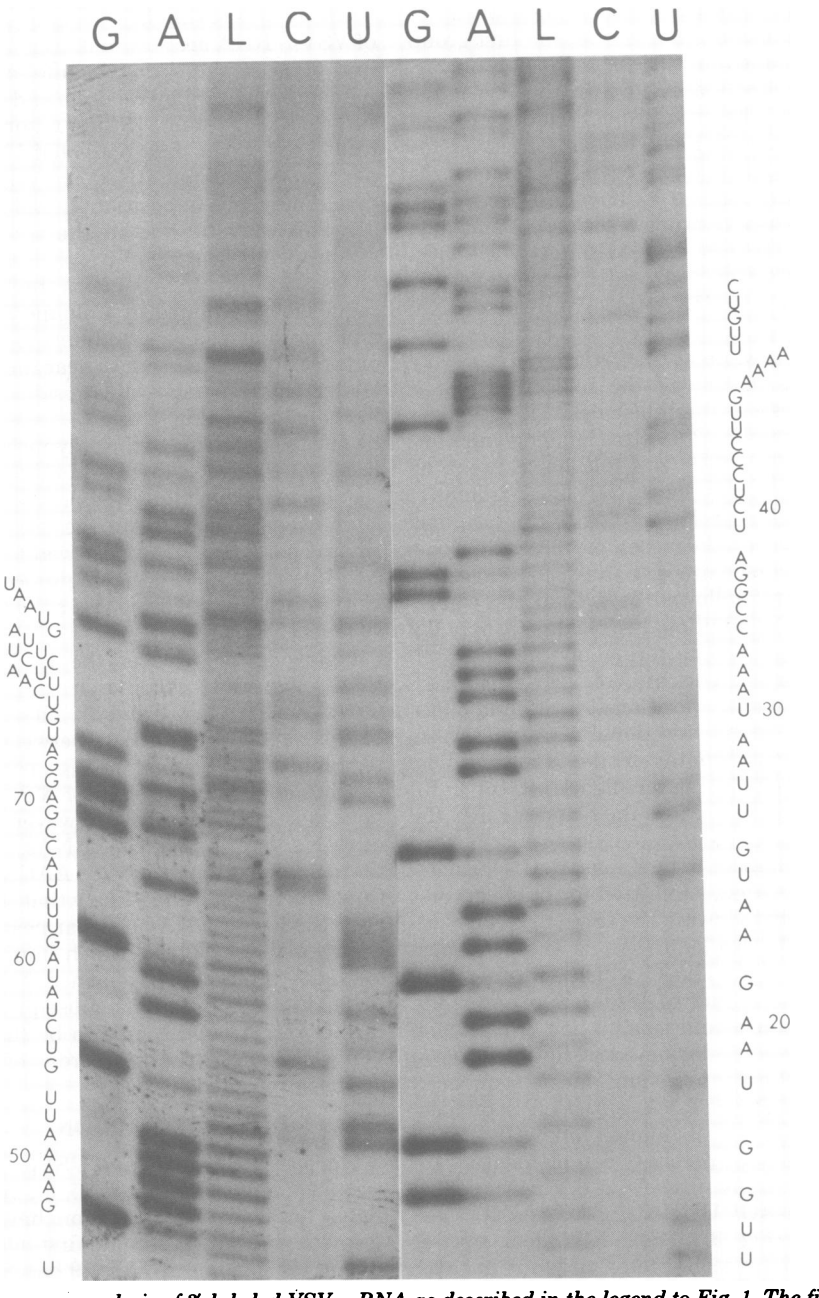

FIG. 2. *Sequence analysis of 3′-labeled VSV_{NJ} RNA as described in the legend to Fig. 1. The five positions on the left side show the primary loading, and the five positions on the right show the secondary loading.*

at the 3′ end of the *L* mRNA is presumably coded by a chattering mechanism on the tract of seven uridylates. The sequence immediately after this adenylation site is identical to the sequence between the leader region and the *N* gene described here. These findings are consistent with the hypothesis that adenylates can specify the end of a transcript either by termination of VSV polymerase or processing of the transcript (Fig. 3).

During the replication of VSV RNA the minus strand is copied fully into a single contiguous plus-sense strand. The termination or cleavage that occurs at the sequence 3′ . . . GAAA . . . 5′

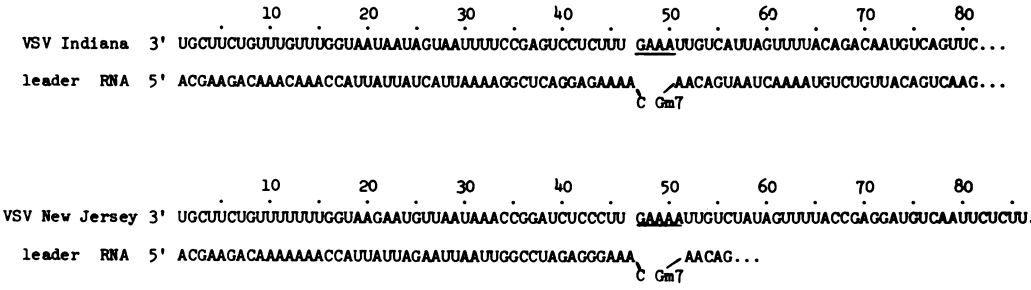

FIG. 3. *Alignment of 3'-terminal base sequences of VSV_{IND} and VSV_{NJ} with published data of RNA transcripts from this region. The intervening base sequence is underlined.*

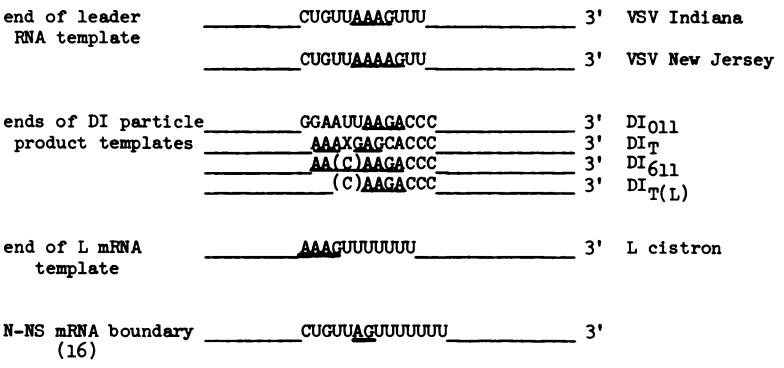

FIG. 4. *Comparison of purine-rich base sequences about the template termination or cleavage sites of various RNA transcripts from VSV and DI particles.*

must be suppressed so that the viral replicase can copy this sequence. Therefore, events which frequently occur at this site are suppressible and may be part of the regulatory mechanism balancing transcription and replication. Whether this involves interaction of this sequence with other proteins is not known.

Studies from several laboratories have suggested that DI particles of VSV originate by premature termination of nascent minus strands and copy-back synthesis of its own 5' end (10, 12, 13, 15). Such a mechanism results in the production of panhandled duplexes (18, 19) and the retention of competent 3'-terminal polymerase initiation sites in DI particle RNAs. Recently, we reported the presence of a specific, internal polymerase initiation site 43 to 48 nucleotides from the 5' end of VSV RNA which allows the nascent minus strand to copy back on itself (21). Careful examination of the 3' base sequence immediately following the panhandle region reveals adenylate-rich stretches (Fig. 4) which are 45 to 48 nucleotides from the 3' terminus. Since these particles produce transcription products in the range of 46 bases long (7, 22, 23; Lazzarini, unpublished data), it is possible that regulatory elements which interact at these A-rich sequences result in the termination of transcription of DI particle RNAs. The inhibition of this termination process could result in the production of full-length DI particle RNAs that are the necessary intermediates in DI particle replication.

The 3'-terminal sequences of VSV RNA described here have also been determined by D. Rowlands and J. Holland (personal communication) and by D. McGeoch (personal communication) with different methods.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grant AI 16099 to J.D.K. from the National Institute of Allergy and Infectious Diseases and a grant from the North Carolina United Community Services.

We thank Barbara T. Hoffman for excellent technical assistance.

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. **Banerjee, A. K., G. Abraham, and R. J. Colonno.** 1977. Vesicular stomatitis virus: mode of transcription. J. Gen. Virol. **34**:1-8.

4. **Colonno, R. J., and A. K. Banerjee.** 1976. A unique RNA species involved in initiation of vesicular stomatitis virus RNA transcription *in vitro*. Cell **8**:197–204.
5. **Colonno, R. J., and A. K. Banerjee.** 1978. Complete nucleotide sequence of the leader RNA synthesized *in vitro* by vesicular stomatitis virus. Cell **15**:93–101.
6. **Colonno, R. J., and A. K. Banerjee.** 1978. Nucleotide sequence of the leader RNA of the New Jersey serotype of VSV. Nucleic Acids Res. **5**:4165–4176.
7. **Emerson, S. U., P. M. Dierks, and J. T. Parsons.** 1977. In vitro synthesis of a unique RNA species by a T particle of vesicular stomatitis virus. J. Virol. **23**:708–716.
8. **Franze-Fernandez, M. T., and A. K. Banerjee.** 1978. In vitro RNA transcription by the New Jersey serotype of vesicular stomatitis virus. I. Characterization of the mRNA species. J. Virol. **26**:179–187.
9. **Herman, R. C., S. Adler, R. A. Lazzarini, R. J. Colonno, A. K. Banerjee, and H. Westphal.** 1978. Intervening polyadenylate sequences in RNA transcripts of vesicular stomatitis virus. Cell **15**:587–596.
10. **Huang, A. S.** 1977. Viral pathogenesis and molecular biology. Bacteriol. Rev. **41**:811–821.
11. **Keene, J. D., M. Rosenberg, and R. A. Lazzarini.** 1977. Characterization of the 3′ termini of RNA isolated from vesicular stomatitis virus and from its defective interfering particles. Proc. Natl. Acad. Sci. U.S.A. **74**: 1353–1357.
12. **Keene, J. D., M. Schubert, and R. A. Lazzarini.** 1979. Terminal sequences of vesicular stomatitis virus RNA are both complementary and conserved. J. Virol. **32**: 167–174.
13. **Keene, J. D., M. Schubert, R. A. Lazzarini, and M. Rosenberg.** 1978. Nucleotide sequence homology at the 3′ termini of RNA from vesicular stomatitis virus and its defective interfering particles. Proc. Natl. Acad. Sci. U.S.A. **75**:3225–3229.
14. **Lazzarini, R. A., G. H. Weber, and G. M. Stamminger.** 1975. Covalently linked message and anti message (genomic) RNA from a defective vesicular stomatitis virus particle. J. Mol. Biol. **97**:289–307.
15. **Leppert, M., L. Kort, and D. Kolakofsy.** 1977. Further characterization of Sendai virus DI-RNAs: a model for their generation. Cell **12**:539–552.
16. **McGeoch, D.** 1979. Structure of the Gene N: Gene NS intercistronic junction in the genome of VSV. Cell **17**: 673–681.
17. **Peattie, D. A.** 1979. Direct chemical method for sequencing RNA. Proc. Natl. Acad. Sci. U.S.A. **76**:1760–1764.
18. **Perrault, J., and R. W. Leavitt.** 1977. Characterization of snap-back RNAs in vesicular stomatitis defective interfering virus particles. J. Gen. Virol. **38**:21–34.
19. **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.
20. **Rose, J. K.** 1978. Complete sequence of the ribosome recognition sites in vesicular stomatitis virus mRNAs: recognition by the 40s and 80s complexes. Cell **14**:345–353.
21. **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.
22. **Schubert, M., J. D. Keene, R. A. Lazzarini, and S. U. Emerson.** 1978. The complete sequence of the *in vitro* transcription product from a defective interfering particle of vesicular stomatitis virus. Cell **15**:102–112.
23. **Semler, B. L., J. Perrault, J. Abelson, and J. J. Holland.** 1978. Sequence of an RNA templated by the 3′-OH RNA terminus of a DI particle of VSV. Proc. Natl. Acad. Sci. U.S.A. **75**:4704–4708.