Proc. Natl. Acad. Sci. USA
Vol. 82, pp. 3110–3114, May 1985
Biochemistry

Nucleotide sequence of the satellite of peanut stunt virus reveals structural homologies with viroids and certain nuclear and mitochondrial introns

(cucumoviruses/replication/conserved sequences in RNA splicing/RNA oligomers/origin of small RNAs)

CANDACE WHITMER COLLMER, A. HADIDI, AND J. M. KAPER

Plant Virology Laboratory, Plant Protection Institute, Agricultural Research Service, U.S. Department of Agriculture, Beltsville, MD 20705

Communicated by T. O. Diener, January 3, 1985

ABSTRACT **Peanut stunt virus-associated RNA 5 (PARNA 5), the satellite of a plant cucumovirus, is a linear RNA of 393 nucleotides with a 5′ cap and a 3′ hydroxyl group. Determination of its nucleotide sequence has revealed two consecutive open reading frames that together extend most of its length. Sequences at the 5′ and 3′ ends are homologous with those of the satellite of the related cucumber mosaic virus, and the double-stranded forms of both satellites contain an unpaired guanosine at the 3′ end of the minus strand. However, little other homology exists between the two satellites. In contrast, PARNA 5 has several regions of 90% sequence homology with various plant viroids, including sequences of the conserved central region of most viroids. Such homologies suggest a common origin with viroids coupled with specific adaptation as a linear RNA. The presence within PARNA 5 of conserved intron sequences essential to proper RNA processing suggests a possible origin from plant introns and/or involvement of such sequences in the processing of PARNA 5 multimers to monomers at some stage of replication.**

Two distantly related plant viruses of the cucumovirus group, cucumber mosaic virus (CMV) and peanut stunt virus (PSV) (1), each encapsidate a small, linear satellite RNA along with their own four major RNAs (2–4). Each satellite RNA is dependent upon its respective helper virus for replication, thus contrasting with plant viroids—circular molecules of similar size that replicate independently of a helper virus (5). The two satellite RNAs, CARNA 5 (CMV-Associated RNA 5) and PARNA 5 (PSV-Associated RNA 5), show little relationship to each other or to the RNAs of their respective helper viruses (3, 4, 6). This latter characteristic distinguishes satellite RNAs from defective interfering RNAs (7) and invites speculation on the question of their origin.

Studies on the satellites of different strains of CMV have unveiled the existence of many different sequence variants. The biological effects of their presence range from exacerbation to attenuation of normal CMV symptoms, the latter often accompanied by a decrease in virus yield presumably due to interference with the CMV replicative machinery (8–10). Because of their ability to perturb both virus replication and symptom development, their small size ($M_r \approx 110{,}000$), and their naturally occurring variability, the CMV satellites provide an excellent model system in which to explore structure–function relationships in RNA molecules. Accordingly, over a dozen already have been sequenced (refs. 11–14; unpublished data).

In contrast, PARNA 5 has received little attention beyond its characterization and a study of some of its replicative properties (15). However, its larger size ($M_r \approx 130{,}000$), lack of relationship to CARNA 5, and replication specificity for PSV but not CMV predicted that comparison of its primary structure with that of CARNA 5 and other small pathogenic RNAs might provide clues into the origin and replication strategy of cucumovirus satellite RNAs. As reported here, the nucleotide sequence of PARNA 5 shows homologies with various regions of different viroids, yet with only the termini of CARNA 5. In addition, striking homologies with intron sequences essential to proper RNA processing suggest a possible origin for PARNA 5 and/or intron-like processing of PARNA 5 multimers to monomers at some stage of replication.

MATERIALS AND METHODS

Purification of PARNA 5 and Double-Stranded (ds) PARNA 5. PARNA 5 isolated from a 1976 PSV isolate was propagated with a 1974 PSV isolate in tobacco (*Nicotiana tabacum* Linnaeus cultivar Xanthi nc) as described (4). Methods for isolating PARNA 5 from virions were the same as those for CARNA 5 (2, 16). The isolation of ds PARNA 5, the separation of the complementary strands in polyacrylamide gels, and the elution of the purified strands have been described (17).

RNA Sequencing of PARNA 5. End-labeling of RNA followed standard procedures as described (13). PARNA 5 purified from virions and (+)- and (−)-strands of PARNA 5 isolated from ds PARNA 5 were labeled at the 3′ terminus by using [5′-^{32}P]pCp (1000–3000 Ci/mmol, ICN; 1 Ci = 37 GBq) and T4 RNA ligase (P-L Biochemicals). Purified (+)- and (−)-strands from ds PARNA 5 were labeled at the 5′ terminus with [γ-^{32}P]ATP (4500 Ci/mmol, ICN) and T4 polynucleotide kinase (P-L Biochemicals) following treatment with calf intestinal alkaline phosphatase (Boehringer Mannheim). PARNA 5 purified from virions could not be labeled by this procedure and, alternatively, was decapped by periodate oxidation and β-elimination (18) and then was 5′-labeled by using [α-^{32}P]GTP (>400 Ci/mmol, ICN) and guanylyltransferase (Bethesda Research Laboratories) (19). Gel purification of labeled RNA, sequencing on polyacrylamide gels containing 98% formamide by using the partial enzymatic cleavage method (20), and determination of the 3′ terminal nucleotide have been described (13).

The cap structure at the 5′ terminus of PARNA 5 purified from virions was labeled with KB^{3}H$_4$ (21), the RNA was hydrolyzed with nuclease P1, and the digest was separated by two-dimensional chromatography on cellulose thin-layer sheets (22). Digests of CARNA 5 and PARNA 5 gave essentially identical patterns, with one labeled spot corresponding in position to m^{7}GpppG, the cap core previously determined for CARNA 5 (23).

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: PSV, peanut stunt virus; CMV, cucumber mosaic virus; PARNA 5, PSV-Associated RNA 5; CARNA 5, CMV-Associated RNA 5; ORF, open reading frame; ds, double-stranded; *box*, region involved in the splicing of certain introns.

Synthesis and Cloning of ds PARNA 5 cDNA. Plasmid pUC9 (24) was isolated from cultures of *Escherichia coli* strain JM83 (25). First-strand cDNA synthesis followed the method of Taylor *et al.* (26) using random priming on the template RNA for the initiation of cDNA synthesis. Reaction mixtures (50 μl) contained 50 mM Tris·HCl (pH 8.1); 50 mM KCl; 10 mM $Mg(OAc)_2$; 10 mM dithiothreitol; 0.1 mg of actinomycin D per ml; 0.5 mM dGTP, dTTP, and dCTP; 0.02 mM dATP; 1.5 μM [α-^{32}P]dATP (650 Ci/mmol; ICN); 3 μg of PARNA 5; 100 μg of calf thymus primer DNA; and 44 units of avian myeloblastosis virus reverse transcriptase (Life Sciences). After incubation for 2 hr at 37°C, reactions were stopped with EDTA, and the RNA was hydrolyzed. The cDNA was extracted with phenol/chloroform, separated from unincorporated dNTPs on Sephadex G-50, and precipitated with ethanol following standard procedures (27). Second-strand cDNA was synthesized for 4 hr at 37°C by using the Klenow fragment of *E. coli* DNA polymerase I (Bethesda Research Laboratories) (27). The resulting cDNA was treated with S1 nuclease, tailed with oligo(dC), hybridized with pUC9 that had been cleaved with *Pst* I and tailed with oligo(dG), and used to transform *E. coli* strain JM83 (28).

Isolation and Sequence Analysis of Cloned PARNA 5 cDNA. Use of the plasmid pUC9 allowed rapid one-step screening for hybrid transformants (24). *E. coli* clones with recombinant plasmids containing PARNA 5 cDNA were selected by standard colony hybridization techniques (27). Recombinant plasmids were isolated from JM83 by a rapid isolation procedure (29), and the size of the insert was determined by polyacrylamide gel electrophoresis after plasmid cleavage with *Pst* I. Inserts were 3′-labeled with [α-^{32}P]dATP (400 Ci/mmol; ICN) and the Klenow fragment of DNA polymerase I at either the *Hin*dIII or the *Eco*RI site of pUC9 (30). The inserts were then excised with the second restriction enzyme (*Eco*RI or *Hin*dIII), purified on a polyacrylamide gel, and sequenced by using the base-specific chemical cleavage method (31). Manipulations involving recombinant DNAs were performed under BL1 containment conditions in compliance with current National Institutes of Health guidelines.

RESULTS

Determination of the Complete PARNA 5 Sequence. The identity of each of the 393 nucleotides of PARNA 5 was determined by at least two independent methods. Sequencing of end-labeled RNA by the enzymatic cleavage method was only partially successful because of band compression on the sequencing gels. Presumably this was due to persistence of secondary structure, which rendered regions of the molecule inaccessible to enzyme cleavage (32). However, sequencing both (+)- and (−)-strand RNA and using polyacrylamide gels containing 98% formamide allowed unambiguous determination of 100–150 nucleotides at each end of the molecule. In addition, long internal stretches were similarly sequenced. The sequence was completed from the longest cDNA clone, which included nucleotides 124–322. The sequence of this region was determined from each of the DNA strands and agreed with all previous RNA sequencing data.

Features of the PARNA 5 Sequence. Like CARNA 5, PARNA 5 is a linear molecule with a 5′ cap and a 3′ hydroxyl group (Fig. 1). It has a rather unusual nucleotide composition, containing over 33% guanosine with several long runs of guanosine (see positions 102–107, 115–123, 162–169, and 233–238) and several purine-rich regions (see especially positions 102–123, 162–189, and 323–334). It is of potential interest that the RNA of the *E. coli* RNase P, a tRNA-processing enzyme, has a very similar, and notably unusual, base composition and distribution (33).

Open Reading Frames (ORFs) in PARNA 5. PARNA 5 has the potential to code for polypeptides in each of its three reading frames. The satellites of CMV sequenced to date also contain ORFs of varying sizes. PARNA 5's first AUG codon at position 42 (followed by a second at position 48) could begin a polypeptide 31 amino acids long, ending at a UGA codon at position 138. This first AUG and its flanking sequences match the favored sequence for eukaryotic initiation sites, $\frac{A}{G}$-X-X-A-U-G-G (34). Shortly after the UGA, another AUG at position 144 begins a second ORF that continues the entire remaining length of the molecule. The two putative polypeptides encoded by these ORFs would have molecular weights of ≈3,500 and ≈9,100, respectively.

Short ORFs occur in each of the other two reading frames, one beginning at position 68 and encoding a putative polypeptide of 22 amino acids and the second beginning at position 76 and encoding a polypeptide 23 amino acids long. The translational capability of PARNA 5 has not been investigated.

An Unpaired Guanosine in ds PARNA 5. Determination of the terminal sequences of the separated (+)- and (−)-strands of ds PARNA 5 revealed the presence of an unpaired guanosine at the 3′ end of the (−)-strand (Fig. 2). Similar termini exist in ds CARNA 5 (35).

DISCUSSION

Comparison of PARNA 5 with Other Small Pathogenic RNA Molecules. Surprisingly, PARNA 5 shows little homology at the 90% level with several CARNA 5 molecules except at the termini where, for example, in two CARNA 5 molecules, 9 of the 10 5′- and 3′-terminal nucleotides match those of PARNA 5 (13, 14). One CARNA 5 (11) contains one other region of 10 nucleotides (positions 35–44) that is 90% homologous with sequences of PARNA 5 (positions 75–84). It seems likely that the recognition of PARNA 5 but not CARNA 5 by the PSV replication machinery involves more than terminal sequences.

In contrast, comparison with various plant viroid sequences revealed several regions of 90% sequence homology. For example, there are nine regions of chrysanthemum

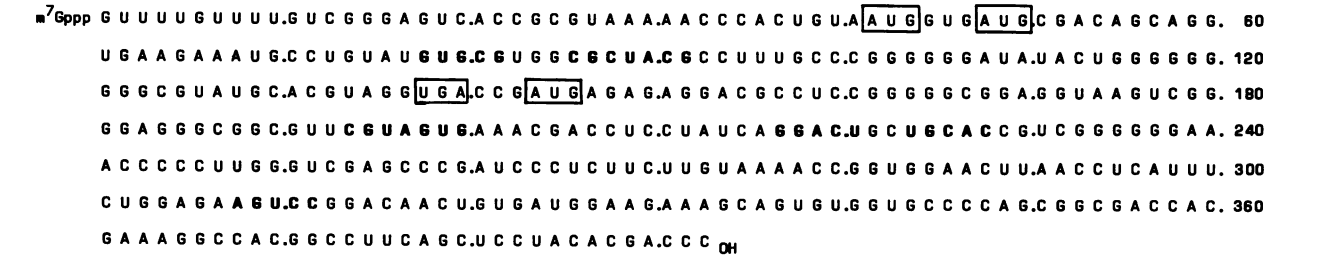

FIG. 1. The complete nucleotide sequence of PARNA 5. Putative initiation codons and corresponding termination codons are boxed. Conserved sequences of introns capable of base pairing (see Fig. 4) are in bold type.

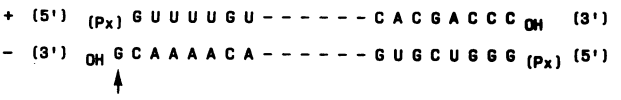

FIG. 2. Terminal sequences of ds PARNA 5 The number of phosphates (p_x) at the 5′ termini of the plus and minus strands has not been determined. The arrow notes the unpaired guanosine at the 3′ end of the (−)-strand.

stunt viroid (CSV) (36) in which 9 of 10 nucleotides match a sequence within PARNA 5, whereas in randomly generated sequences of the same size and base composition as CSV, only 1 or 2 such matches occur. Starting with the 5′ end of PARNA 5, the nine regions are designated by the position numbers of their first nucleotides (CSV in parentheses) as follows: 3 (340), 58 (65), 88 (282), 235 (92), 279 (260), 279 (329), 301 (207), 303 (69), and 338 (21). Although these and other homologous regions are not all in the same relative positions within the two molecules, several are of particular interest. First, two of the PARNA 5 regions listed above (positions 235–243 and 279–287) perfectly match regions of the upper and lower parts of the central, conserved region of most viroids (37). Second, the 22 5′-terminal nucleotides of PARNA 5 (G-U-U-U-U-G-U-U-U-U-G-U-C-G-G-G-A-G-U-C-A-C) show 77% homology with nucleotides 343–10 of the circular CSV (A-G-U-U-U-U-G-U-U-C-C-C-U-C-G-G-G-A-C-U-U-A-C). This region of CSV corresponds in position to the 5′ end of one of the two observed "natural linear" forms of potato spindle tuber viroid (38).

Although their biological significance has not yet been established, these linear forms of potato spindle tuber viroid may be compared with the linear forms of certain viral satellite RNAs that also exist in circular configurations (39–41). In the case of the satellite of tobacco ringspot virus specifically, evidence is accumulating that both its linear and circular forms could represent different stages in its life cycle (39, 42).

Whatever the significance of linear forms of viroids and circular satellites, it seems important to note that the termini of PARNA 5 appear unique. Whereas these other linear molecules apparently have 3′ phosphate and 5′ hydroxyl termini (see ref. 43), PARNA 5 has a 5′ cap and a 3′ hydroxyl group, termini which give no hint of circular precursors or products. Therefore, while the homologies with plant viroids may be suggestive of a common origin for PARNA 5 and viroids, it seems that the two kinds of molecules are specifically adapted for existence as linear and circular molecules, respectively.

Homologies with Conserved Sequences of Introns. The possible origin of viroids from introns (intervening sequences) of eukaryotic RNAs has been discussed previously (36, 44–49). Cech (50) has delineated three general classes of introns and corresponding RNA splicing strategies—one for tRNA, one for nuclear-encoded mRNA, and, finally, one encompassing nuclear rRNA and mitochondrial mRNA and rRNA. Splicing in this last class alone appears to be directed by the intron itself, whose integrity is absolutely essential to the splicing reaction. A combination of phylogenetic, genetic, and biochemical evidence has established that the base pairing of three particular pairs of sequences in these introns most probably plays a crucial role in forming a secondary structure essential to splicing (51–56). Two of these pairs (designated as regions A and B, regions *box9L* and *box2*) are conserved in nucleotide sequence across introns of nuclear rRNAs as well as those of mitochondrial mRNAs and rRNAs (53, 57–59); the third pair (regions 9R′ and *box9R*) is conserved in the different introns in location and ability to base pair only (52, 59, 60).

PARNA 5 contains sequences that are related to all three pairs of the highly conserved sequences of these introns. In addition, the sequences are located in the same relative positions within the molecule as those of nuclear rRNA introns and mitochondrial introns (Fig. 3). Whereas actual nucleotide sequence is not always conserved in the PARNA 5 sequence, what is particularly striking is that the six regions of PARNA 5 can base-pair within a hypothetical secondary structure that is very similar to that proposed for group I introns (51). Fig. 4 shows these base-pairings of PARNA 5 compared to noticeably similar ones of intron 4 of the yeast mitochondrial gene for cytochrome b, *cob*. The A–B region comprises six base pairs in the yeast intron and seven in the comparable region of PARNA 5. The five base pairs in the *box9R*–9R′ region of the yeast intron are maintained in PARNA 5 through compensation for the loss of an A·U pair at one end of the region by acquisition of a G·U pair at the other end. And, a similar compensation in the tertiary interaction between the *box9L* and *box2* regions of PARNA 5 results in maintaining both a spacer region between *box9L* and *box9R* and the five base pairs present in that region of the yeast intron (see Fig. 4 *Right*).

In an attempt to assess the significance of the intron-like sequences within PARNA 5, we calculated the probability of finding by chance the four remaining sequences, given the existence of the *box9L*–9R-like sequence near the middle of PARNA 5. To take into account the importance of the relative positions of the conserved sequences, we demanded that each sequence in PARNA 5 be found within 35 nucleotides to either side of the position corresponding to that of the respective sequence in the similarly-sized *Tetrahymena* intron (ref. 59; see Fig. 3). Our calculation included the probability of finding within the appropriate 70-nucleotide windows a *box2*-like sequence with four of five bases complementary to those of the PARNA 5 *box9L* sequence and a 9R′-like sequence with 4 of 5 bases complementary to those of the PARNA 5 *box9R* sequence; an A region-like sequence with 3 of 6 bases identical to the A sequence of the yeast *cob* intron 4 (Fig. 4) and in the same position relative to the 9R′ sequence (52); and a B region-like sequence with 5 of 6 bases complementary to the PARNA 5 A-like sequence within the

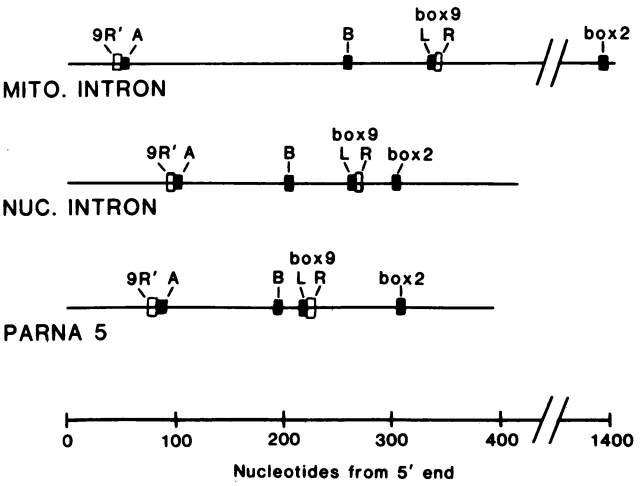

FIG. 3. Relative positions of conserved sequences in the yeast mitochondrial (MITO) *cob* gene intron 4, the *Tetrahymena* nuclear (NUC) rRNA intron, and PARNA 5. Positions of the sequences in the yeast and *Tetrahymena* intron, as well as terminology and organization of the figure, are as described by Cech *et al.* (59). Solid boxes indicate conserved intron sequences that can interact by base pairing (region A with B, region *box9L* with *box2*); open boxes indicate regions conserved only in their location and potential to base pair (region 9R′ with *box9R*); all are somewhat enlarged for clarity. The precise locations of the conserved sequences within PARNA 5 are: 9R′–A, nucleotides 78–92; B, 194–200; *box9L*–9R, 217–228; *box2*, 308–312 (see Fig. 1).

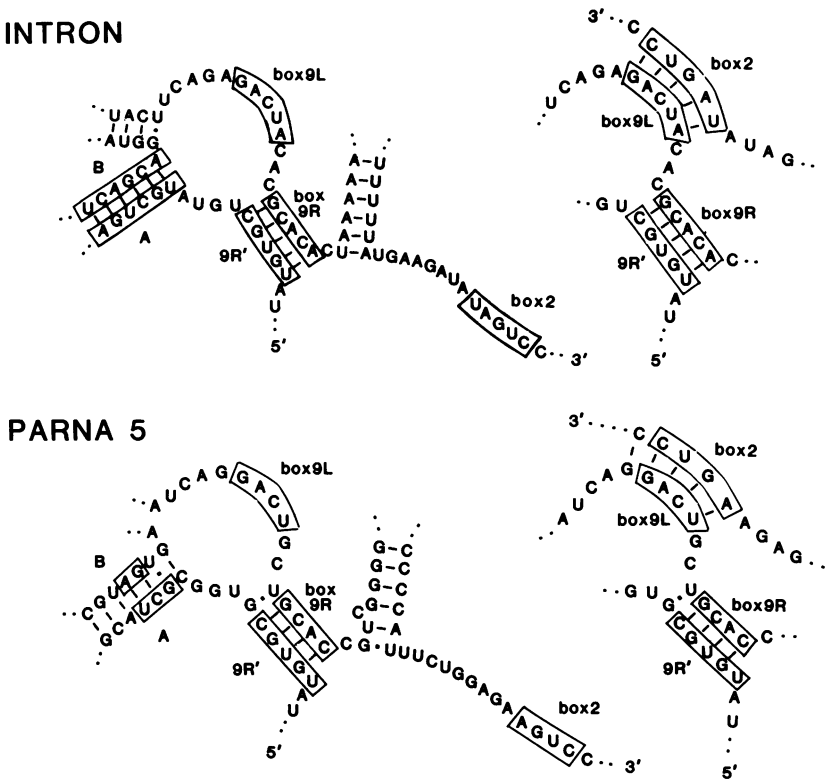

FIG. 4. Schematic presentation of possible base-pairings of the conserved intron sequences of PARNA 5 as compared to those in the mitochondrial *cob* intron 4 (drawn from figure 4 in ref. 56). Boxed sequences *box9L, box2*, A, and B in the mitochondrial intron are the sequences conserved in most mitochondrial and certain nuclear introns. Boxed sequences *box9R* and 9R′ of the intron are sequences conserved only in their position and ability to base-pair. Boxed sequences in PARNA 5 are those homologous to conserved sequences of the yeast intron. (*Left*) Possible secondary structure interactions. (*Right*) Possible tertiary interactions.

appropriate 45-nucleotide window (reduced from 70 by the presence of the *box9L* sequence at the 3′ border). The occurrence of the four sequences at the appropriate relative positions within PARNA 5 is significant at the 0.05 level.

Divergent or Convergent Evolution? Sequences related to six conserved sequences whose base-pairing is essential to proper splicing in certain nuclear and mitochondrial introns have been found not only in PARNA 5 but also in plant viroids (Dinter-Gottlieb and Cech, personal communication; Hadidi and Diener, unpublished data). Given that the unexpected resemblance between fungal mitochondrial introns and *Tetrahymena* nuclear introns is believed more likely to reflect movement of sequences between nucleus and mitochondrion than convergent evolution (57, 59), it seems possible that PARNA 5 and viroids could represent escaped introns of either origin. As might be expected from their nucleolar location (61), viroids are similar to nuclear rRNA introns (62) in lacking mRNA activity and circularizing after being processed from larger molecules (43). PARNA 5, on the other hand, more resembles certain mitochondrial introns that contain long ORFs and remain linear after excision, probably to function as mRNAs (53, 63). In addition, the regions *box9R* and 9R′, conserved between nuclear and mitochondrial introns in position and ability to base-pair only, are nearly identical in PARNA 5 and one mitochondrial intron (see Fig. 4). The origin of one plant virus satellite, that of turnip crinkle virus, from within the plant genome has been suggested by apparent homology between cloned satellite cDNA and sequences in the DNA of uninfected plants (64). And, observed homologies between proteins encoded by yeast mitochondrial introns and by two strains of a plant RNA virus are suggestive of an evolutionary connection between introns and pathogenic RNAs (65).

However, whether the intron-like sequences within PARNA 5 originated from introns or, alternatively, by convergent evolution, the existence of conserved sequences in rapidly evolving RNA genomes (66) implies selective pressure for their survival. The presence of both (+)- and (−)-stranded oligomers of PARNA 5 in infected tissue (15) requires mechanisms for both the formation of oligomers and their processing back to monomers. Perhaps the intron-like sequences within oligomeric (+)-stranded PARNA 5 form a secondary structure that, like the *Tetrahymena* intron (59), directs protein-independent processing. Capping of the resulting PARNA 5 monomers may be analogous to cyclization of the *Tetrahymena* intron, both preventing the excised RNA from driving the splicing reaction backwards (67).

Although a rolling circle form of replication could generate oligomers, no circular forms of PARNA 5 have been detected in infected tissue (15). In addition, the termini of both single-stranded and ds PARNA 5 match those of the linear cucumovirus CMV and hint of a linear RNA replicative strategy (35). An alternative mechanism for oligomerization may involve the unpaired guanosine at the 3′ end of the (−)-strand of monomeric PARNA 5 (Fig. 2), whose attack between the 5′-terminal guanosine and the penultimate guanosine of another (−)-stranded molecule could form a dimer through the same kind of cleavage-ligation reaction catalyzed by the 3′-terminal guanosine of the *Tetrahymena* intron during autocatalytic cyclization (67). In fact, oligomerization of *Tetrahymena* introns via such a mechanism has been observed (A. Zaug and T. Cech, personal communication). Transcription of (−)-stranded PARNA 5 oligomers to (+)-stranded oligomers followed by intron-like processing could restore the monomeric size perhaps necessary for biological activity. By manipulations of (−)-stranded monomeric PARNA 5 *in vitro* as well as determination of the sequence of dimeric PARNA 5 isolated from infected tissue, the occurrence of some of these putative steps in the PARNA 5 life cycle can be tested.

We are grateful to Marie Tousignant for providing PARNA 5 and ds PARNA 5, Dr. Michael Kiefer for assistance in the initial cloning work, Dr. Huub Linthorst for stimulating discussions, Tish Lysinger for cheerful and expert assistance in the manuscript preparation, and Drs. R. A. Owens and T. O. Diener for critically reading the manuscript. Financial support for this research was provided in part by U.S. Department of Agriculture Competitive Research Grant 82-CRCR-1-1136.

1. Devergne, J. C. & Cardin, L. (1975) *Ann. Phytopathol.* **7,** 255–276.
2. Kaper, J. M., Tousignant, M. E. & Lot, H. (1976) *Biochem. Biophys. Res. Commun.* **72,** 1237–1243.
3. Gould, A. R., Palukaitis, P., Symons, R. H. & Mossop, D. W. (1978) *Virology* **84,** 443–455.
4. Kaper, J. M., Tousignant, M. E., Diaz-Ruiz, J. R. & Tolin, S. A. (1978) *Virology* **88,** 166–170.
5. Diener, T. O. (1983) *Adv. Virus Res.* **28,** 241–283.
6. Diaz-Ruiz, J. R. & Kaper, J. M. (1977) *Virology* **80,** 204–213.
7. Huang, A. S. (1973) *Annu. Rev. Microbiol.* **27,** 101–117.
8. Murant, A. F. & Mayo, M. A. (1982) *Annu. Rev. Phytopathol.* **20,** 49–70.
9. Piazzolla, P., Tousignant, M. E. & Kaper, J. M. (1982) *Virology* **122,** 147–157.
10. Kaper, J. M. (1984) in *Control of Virus Diseases*, eds. Kurstak, E. & Marusyk, R. (Dekker, New York), pp. 317–343.
11. Richards, K. E., Jonard, G., Jacquemond, M. & Lot, H. (1978) *Virology* **89,** 395–408.
12. Gordon, K. H. J. & Symons, R. H. (1983) *Nucleic Acids Res.* **11,** 947–960.
13. Collmer, C. W., Tousignant, M. E. & Kaper, J. M. (1983) *Virology* **127,** 230–234.
14. Hidaka, S., Ishikawa, K., Takanami, Y., Kubo, S. & Miura, K. (1984) *FEBS Lett.* **174,** 38–42.
15. Linthorst, H. J. M. & Kaper, J M. (1984) *Virology* **139,** 317–329.
16. Kaper, J. M. & Tousignant, M. E. (1977) *Virology* **80,** 186–195.
17. Kaper, J. M. & Tousignant, M. E. (1983) *Biochem. Biophys. Res. Commun.* **116,** 1168–1175.
18. Dasgupta, K., Harada, F. & Kaesberg, P. (1976) *J. Virol.* **18,** 260–267.
19. Monroy, G., Spencer, E. & Hurwitz, J. (1978) *J. Biol. Chem.* **253,** 4490–4498.
20. Donis-Keller, H., Maxam, A. M. & Gilbert, W. (1977) *Nucleic Acids Res.* **4,** 2527–2538.
21. Symons, R. H. (1975) *Mol. Biol. Rep.* **2,** 277–285.
22. Marcu, K. B., Schibler, U. & Perry, R. P. (1978) *J. Mol. Biol.* **120,** 381–400.
23. Lot, H., Jonard, G. & Richards, K. (1977) *FEBS Lett.* **80,** 395–400.
24. Vieira, J. & Messing, J. (1982) *Gene* **19,** 259–268.
25. Holmes, D. D. & Quigley, M. (1981) *Anal. Biochem.* **114,** 193–197.
26. Taylor, J. M., Illmensee, R. & Summers, J. (1976) *Biochim. Biophys. Acta* **442,** 324–330.
27. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
28. Morrison, D. A. (1979) *Methods Enzymol.* **68,** 326–328.
29. Davis, R. W., Thomas, M., Cameron, J., St. John, T. P., Scherer, S. & Pagett, R. A. (1980) *Methods Enzymol.* **65,** 406–407.
30. Rüther, V., Koenen, M., Otto, K. & Müller-Hill, B. (1981) *Nucleic Acids Res.* **9,** 4087–4098.
31. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65,** 499–560.
32. Kramer, F. R. & Mills, D. R. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 5334–5338.
33. Reed, R. E., Baer, M. F., Guerrier-Takada, C., Donis-Keller, H. & Altman, S. (1982) *Cell* **30,** 627–636.
34. Kozak, M. (1981) *Nucleic Acids Res.* **9,** 5233–5252.
35. Collmer, C. W. & Kaper, J. M. (1984) *Phytopathology* **74,** 808 (abstr.).
36. Gross, H. J., Krupp, G., Domdey, H., Raba, M., Jank, P., Lossow, C., Alberty, H., Ramm, K. & Sänger, H. L. (1982) *Eur. J. Biochem.* **121,** 249–257.
37. Haseloff, J., Mohamed, N. A. & Symons, R. H. (1982) *Nature (London)* **299,** 316–321.
38. Kikuchi, Y., Tyc, K., Filipowicz, W., Sänger, H. L. & Gross, H. J. (1982) *Nucleic Acids Res.* **10,** 7521–7529.
39. Kiefer, M. C., Daubert, S. D., Schneider, I. R. & Bruening, G. (1982) *Virology* **121,** 262–273.
40. Jones, A. T. & Mayo, M. A. (1984) *J. Gen. Virol.* **65,** 1713–1721.
41. Gould, A. R. & Hatta, T. (1981) *Virology* **109,** 137–147.
42. Linthorst, H. J. M. & Kaper, J. M. (1984) *Virology* **137,** 206–210.
43. Branch, A. D. & Robertson, H. D. (1984) *Science* **223,** 450–455.
44. Roberts, R. J. (1978) *Nature (London)* **274,** 530.
45. Crick, F. (1979) *Science* **204,** 264–271.
46. Diener, T. O. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 5014–5015.
47. Dickson, E. (1981) *Virology* **115,** 216–221.
48. Zimmern, D. (1982) *Trends Biochem. Sci.* **7,** 205–207.
49. Kiefer, M. C., Owens, R. A. & Diener, T. O. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 6234–6238.
50. Cech, T. R. (1983) *Cell* **34,** 713–716.
51. Michel, F., Jacquier, A. & Dujon, B. (1982) *Biochimie* **64,** 867–881.
52. Davies, R. W., Waring, R. B., Ray, J. A., Brown, T. A. & Scazzocchio, C. (1982) *Nature (London)* **300,** 719–724.
53. Burke, J. M. & RajBhandary, U. L. (1982) *Cell* **31,** 509–520.
54. De La Salle, H., Jacq, C. & Slonimski, P. P. (1982) *Cell* **28,** 721–732.
55. Netter, P., Jacq, C., Carignani, G. & Slonimski, P. P. (1982) *Cell* **28,** 733–738.
56. Weiss-Brummer, B., Holl, J., Schweyen, R. J., Rodel, G. & Kaudewitz, F. (1983) *Cell* **33,** 195–202.
57. Waring, R. B., Scazzocchio, C., Brown, T. A. & Davies, R. W. (1983) *J. Mol. Biol.* **167,** 595–605.
58. Michel, F. & Dujon, B. (1983) *EMBO J.* **2,** 33–38.
59. Cech, T. R., Tanner, N. K., Tinoco, I., Jr., Weir, B. R., Zuker, M. & Perlman, P. S. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 3903–3907.
60. Waring, R. B. & Davies, R. W. (1984) *Gene* **28,** 277–291.
61. Schumacher, J., Sänger, H. L. & Riesner, D. (1983) *EMBO J.* **2,** 1549–1555.
62. Kruger, K., Grabowski, P. J., Zaug, A. J., Sands, J., Gottschling, D. E. & Cech, T. R. (1982) *Cell* **31,** 147–157.
63. Garriga, G. & Lambowitz, A. M. (1983) *J. Biol. Chem.* **258,** 14745–14748.
64. Altenbach, S. B. & Howell, S. H. (1984) *Virology* **134,** 72–77.
65. Zimmern, D. (1984) *J. Mol. Biol.* **171,** 345–352.
66. Holland, J., Spindler, K., Horodyski, F., Grabau, E., Nichol, S. & VandePol, S. (1982) *Science* **215,** 1577–1585.
67. Zaug, A. J., Grabowski, P. J. & Cech, T. R. (1983) *Nature (London)* **301,** 578–583.