Volume 9 Number 12 1981 Nucleic Acids Research

Chrysanthemum stunt viroid: primary sequence and secondary structure

James Haseloff and Robert H. Symons*

Department of Biochemistry, University of Adelaide, Adelaide, SA 5001, Australia

Received 10 April 1981

ABSTRACT

The sequence of the 356 nucleotide residues of chrysanthemum stunt viroid (CSV) has been determined. Overlapping linear viroid fragments were obtained by partial ribonuclease digestion, radiolabelled *in vitro* at their 5'-ends, and sequenced using partial enzymic cleavage methods. Of the CSV sequence, 69% is contained in the published sequence of potato spindle tuber viroid (PSTV). Differences in the primary sequence of CSV and PSTV suggest that neither the positive nor putative negative strands of these two viroids code for functional polypeptide products. However, the two viroids can form similar secondary structures, implicating a role for viroid structure in replication.

INTRODUCTION

The causative agent of chrysanthemum stunt disease is a member of the unique group of plant pathogens known as viroids of which only eight have been described so far (1,2). Together with the other members of this group, chrysanthemum stunt viroid (CSV) consists of a single-stranded covalently closed circular RNA species existing as a highly base-paired rod-like structure (3,4), which is infectious and is not encapsidated (3,4). No viroid has been shown to code for a protein product either *in vivo* (5) or *in vitro* (6,7).

The primary and secondary structure of only one viroid, potato spindle tuber viroid (PSTV), has been determined and it contains 359 residues (8). The biological properties of PSTV differ from those of CSV; for example, the host range of CSV is mostly confined to some species in the Compositae family whereas PSTV will replicate in species from a number of plant families (1,9). Ribonuclease fingerprinting has also shown that the primary sequence of PSTV differs significantly from that of CSV (10). However, cDNA-RNA hybridization techniques have indicated that at least 20% of the PSTV sequence is common to that of CSV (11). It was of interest, therefore, to determine the

complete primary and secondary structure of CSV and to determine the common features shared by these two viroids.

MATERIALS AND METHODS

Materials

CSV, purified from infected chrysanthemums as described by Palukaitis and Symons (4), was kindly provided by Dr. P. Palukaitis. RNases T_1 and A were obtained from Sigma Chemical Co. and RNase U_2 from Sankyo. RNase PhyM was purified (12) from the culture supernatant of *Physarum polycephalum*, the inoculum of which was kindly provided by the School of Biological Sciences, Flinders University of South Australia. The extracellular RNase from *Bacillus cereus* was prepared as described by Lockard *et al.* (13) with a culture obtained from Dr. G. Brownlee and the phage T_4 polynucleotide kinase was purified essentially as described by Richardson (15). γ-^{32}P-ATP was prepared from carrier-free ^{32}Pi (14).

Isolation of 5'-^{32}P-RNA fragments of CSV

Partial RNase digestions. Purified circular CSV (5 µg) was dissolved in 40 µl of 600 mM NaCl, 10 mM $MgCl_2$, with either 20 mM Tris-HCl, pH 7.5, for RNases T_1 and A digestions or with 20 mM sodium citrate, pH 3.5, for RNase U_2 digestions. Incubation was at 0°C for 60 minutes with 150 units RNase T_1, 80 ng RNase A, or 2 units RNase U_2. Digestions were terminated by phenol extraction, ether washing and ethanol precipitation.

5'-^{32}P-labelling of CSV fragments. A dried mixture of CSV fragments and 200 µCi of γ-^{32}P-ATP was dissolved in 5 µl 2 mM spermidine, heated at 80°C for one minute and cooled on ice. The final reaction mixture of 10 µl contained 20 mM Tris-HCl, pH 9.0, 10 mM $MgCl_2$, 10 mM dithiothreitol, 5% (v/v) glycerol, 1 mM spermidine and 4 units T_4 polynucleotide kinase. After incubation at 37°C for 30 minutes, 10 µl of 95% (v/v) formamide, 10 mM EDTA, 0.01% bromophenol blue and xylene cyanol FF (formamide-dye mix) was added to terminate the reaction.

Purification of 5'-^{32}P-CSV fragments. Reaction mixtures were heated at 80°C for 30 seconds and loaded onto an 80 x 20 x 0.05 cm 6% polyacrylamide gel prepared in 90 mM Tris-borate, pH 8.3, 2 mM EDTA, 7 M urea (16). After electrophoresis for 6 hours at 25 mA, the gel was autoradiographed at room temperature for 15 minutes and the autoradiograph used as a template to locate and excise the labelled CSV fragments. Excised bands were eluted by soaking overnight at room temperature in 500 µl of 500 mM ammonium acetate, 1 mM EDTA, 0.1% SDS, containing 60 µg *E. coli* tRNA as carrier. Eluted frag-

ments were further purified by two ethanol precipitations.

Sequencing of 5'-labelled CSV fragments

This was done using partial enzymic cleavage methods. Partial digestions were carried out with RNase T_1 and alkali (17), RNase U_2 (18), RNase PhyM (12), and *Bacillus cereus* extracellular RNase (13), and the digests fractionated on either 80 x 20 x 0.05 cm 8% polyacrylamide gels or 40 x 20 x 0.05 cm 20% polyacrylamide gels in 90 mM Tris-borate, pH 8.3, 2 mM EDTA, 7 M urea.

Formamide gels. Some partial sequencing digests of viroid fragments were fractionated on polyacrylamide gels containing 98% formamide in order to eliminate band compression (19) arising from incomplete denaturation of the fragments during electrophoresis. Partial digests were ethanol precipitated, resuspended in formamide-dye mix (see above), heated at 80°C for one minute, cooled on ice and loaded onto a 40 x 20 x 0.05 cm 20% polyacrylamide gel containing 98% formamide buffered with 16 mM Na_2HPO_4, 4 mM NaH_2PO_4 (20).

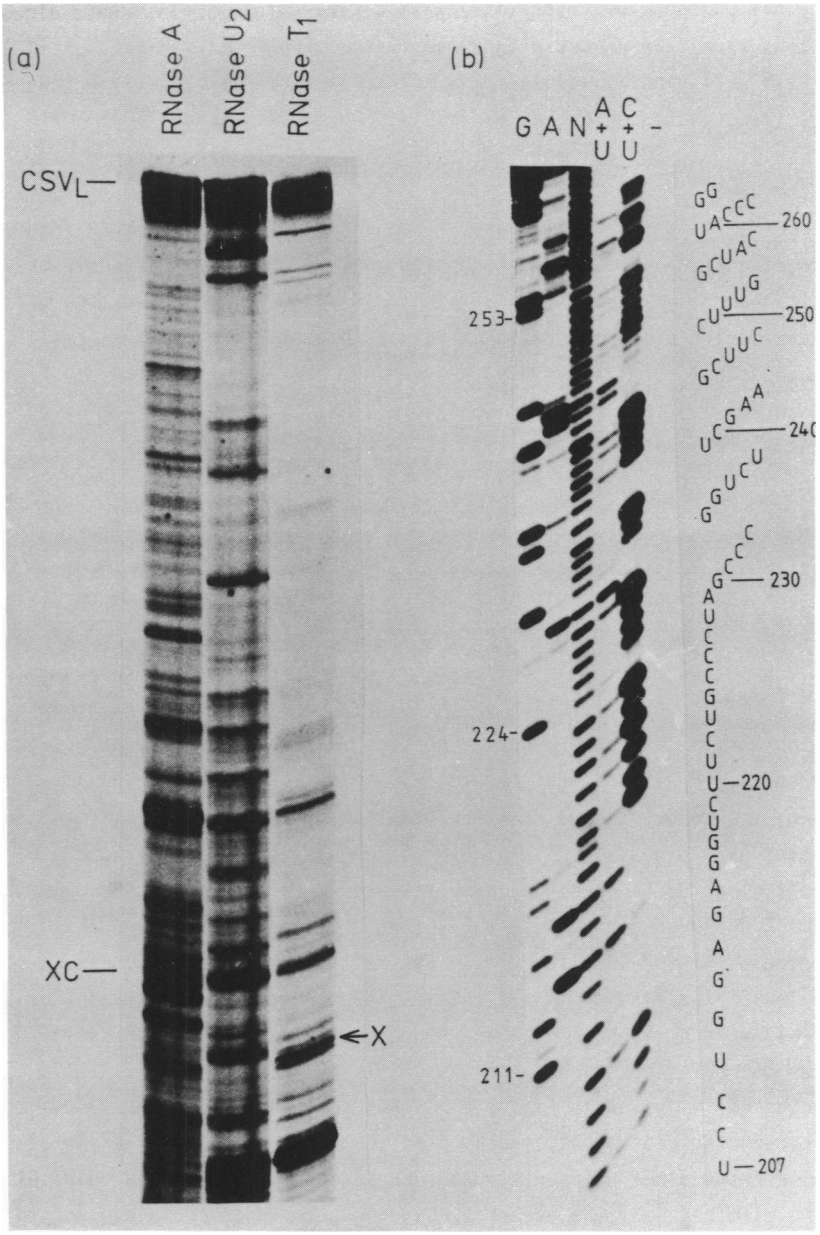
RESULTS

Preparation of 5'-^{32}P-Labelled Viroid Fragments for Sequencing

Partial digests of circular CSV were carried out with RNases T_1, U_2 and A under conditions of high salt concentration (600 mM NaCl, 10 mM $MgCl_2$) and at 0°C in order to limit cleavage by the single-strand specific RNases to relatively few accessible sites on the highly base-paired RNA molecule. The resulting viroid fragments were 5'-labelled *in vitro* with polynucleotide kinase and γ-^{32}P-ATP and fractionated by size on a denaturing polyacrylamide gel; Fig. 1(a) shows the gel pattern obtained for partial digests with RNases T_1, U_2 and A. Digestion with either of the single base specific RNases T_1 (G specific) or U_2 (A specific) gave rise to fewer fragments than digestion with the C and U specific RNase A.

Sequencing of 5'-^{32}P-Labelled Viroid Fragments

The 5'-labelled fragments obtained by partial RNase digestion of CSV (Fig. 1a) were excised, eluted and sequenced using the partial enzymic cleavage method as described under Materials and Methods. An example of one sequencing gel is given in Fig. 1(b). For some regions of the viroid molecule, sequencing was complicated by band compression (19) due to the presence of stable base-paired hairpin structures. However, these band compressions could be eliminated by the use of sequencing gels which contained 98% formamide rather than 7 M urea in order to ensure complete denaturation of the RNA fragments.


(a)
RNase A
RNase U2
RNase T1
CSVL
XC
X
(b)
G A N
A C
+ +
U U
-
253
224
211
260
250
240
230
220
207

Primary Sequence and Secondary Structure of CSV

The complete base sequence of CSV (Fig. 2) was assembled from sequence data of a large number of RNA fragments obtained from partial RNase digestion. The 356 residues are numbered according to the scheme of Gross *et al.* (8) for PSTV while the main overlapping sequences used for the primary structure determination are given in Fig. 2.

A possible secondary structure model was constructed from the CSV sequence using the methods of Tinoco *et al.* (21,22) and is compared with the published structure for PSTV (23) (Fig. 3). The relative number of G.C base pairs in the predicted CSV structure (64 G.C, 44 A.U, 16 G.U) is lower than that of PSTV (73 G.C, 37 A.U, 16 G.U) and the calculated free energies (22) for native CSV and the more stable PSTV are ΔG (25°C) = -140 Kcal.mol^{-1} and -161 Kcal.mol^{-1}, respectively.

DISCUSSION

Rapid Sequencing of Viroids

The approach used in this paper for the sequencing of CSV proved to be rapid and required only small amounts of purified viroid. A large number of 5'-labelled fragments can be rapidly obtained from partial RNase digests of only a few µg of purified viroid while the sequencing of these fragments by the partial enzymic cleavage method is rapid and reliable. Further, the problem of band compression found on sequencing gels of some labelled fragments can be resolved by the use of polyacrylamide gels containing 98% formamide. The final sequence of the circular viroid is then obtained from the sequence of overlapping fragments.

Homology Between CSV and PSTV

The striking feature of both the primary and postulated secondary

Figure 1. Purification and nucleotide sequence determination of CSV fragments. (a) Autoradiogram of the 5'-^{32}P-labelled products of the partial digestion of CSV by RNases A, U_2 and T_1 after fractionation by electrophoresis on a polyacrylamide slab gel as described in Materials and Methods. The largest labelled fragment is full length linear CSV (CSV_L) which migrated about 30 cm from the origin. XC is the position of the xylene cyanol FF dye marker which corresponds to fragments about 80 residues long. A number of the shorter CSV fragments (including band X) were excised and eluted for sequencing by partial enzymic digestion. (b) Autoradiogram of part of a sequencing gel (8% acrylamide) containing the various partial enzymic digests of fragment X. Digestions, as described in Materials and Methods, were with RNase T_1 (G), RNase U_2 (A), alkali (N) to produce the reference ladder, RNase PhyM (A + U), and *Bacillus cereus* RNase (C + U). Part of the sequence of fragment X from residues 207 to 265 is given.

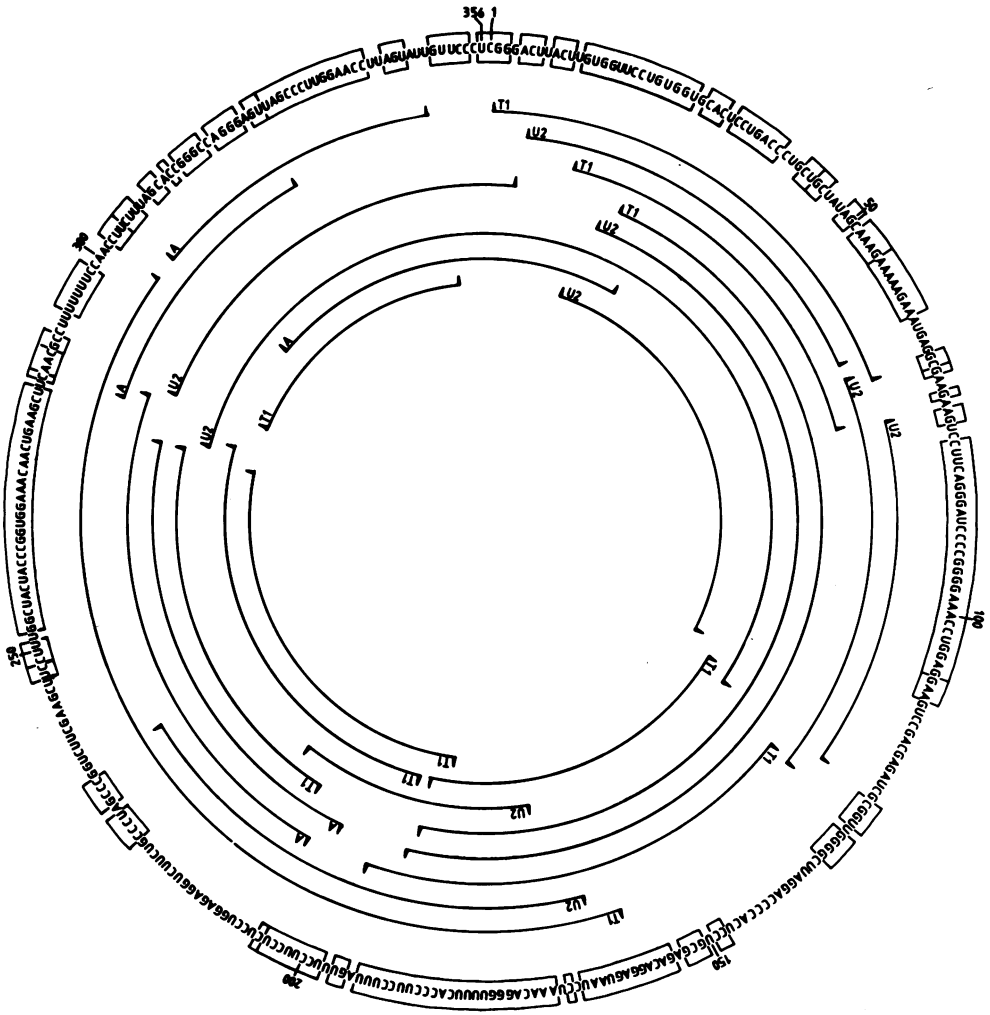


Figure 2. Primary sequence of CSV. The sequence of the 356 residues of CSV is given and the residues numbered according to the published sequence of PSTV (8). The 247 residues homologous with PSTV are boxed. Within the circular sequence are given the locations of overlapping sequences obtained from RNase fragments of CSV; these sequences do not represent the entire length of these fragments. Each sequence is labelled with the RNase (A, T_1, or U_2) which gave the fragment from which that sequence was derived.

structure of CSV is the extent of homology with the previously sequenced PSTV molecule (8). The 247 residues of CSV (69%) which are homologous with PSTV occur in two main areas in the primary structure (Fig. 2) extending from residues 247 to 110 and 148 to 206. These areas are separated by two regions

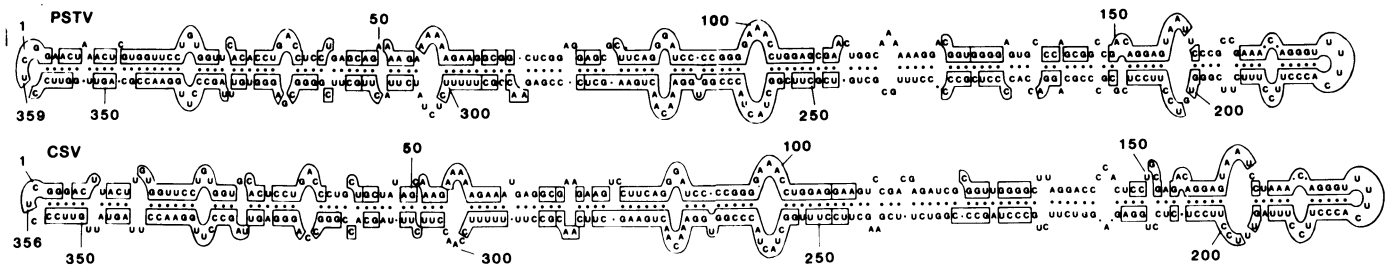


Figure 3. The predicted secondary structures of CSV and PSTV (23). The boxed areas contain residues homologous between the two viroids.

of about 40 residues each containing only two small areas of homology. The postulated secondary structure model for CSV (Fig. 3) shows that the two main areas of homology each correspond to one base-paired end of the native molecule. These are separated by the two regions of lesser homology which are positioned almost exactly opposite each other in the native molecule and are predominantly base-paired. Thus, the conservative arrangement and base-pairing of such non-conserved regions in the primary sequence allows the CSV molecule to form a stable secondary structure similar to that of PSTV.

Replication of CSV and PSTV

Although the host ranges of PSTV and our isolate of CSV differ significantly (see Introduction), they do overlap in such plants as the composite *Gynura aurantiaca* (1,4,24). It is feasible that replication of the two viroids in these plants will occur by similar mechanisms in view of their similarities in size and sequence. Matthews (25) has suggested that translation of the putative negative strand of PSTV may give rise to polypeptide(s) functional in replication.

Possible translation products of both the positive and putative negative strands of CSV and PSTV are given in Fig. 4. If there is read-through of the weak termination codon UGA then much longer products would be produced with more than one round of translation in the case of the PSTV positive strand and the CSV negative strand. Whether or not UGA codons are translated, the major differences between the possible polypeptide translation products of CSV and PSTV suggest that neither viroid codes for proteins involved in their replication. This is consistent with the lack of evidence for any viroid translation *in vivo* (5) and *in vitro* (6,7).

On the other hand, the overall secondary structures of CSV and PSTV are conserved despite differences in sequence. Thus, sequence and structural features common to both CSV and PSTV may play some role in the recognition by host enzymes which are capable of RNA-dependent RNA synthesis and are presumably responsible for viroid replication. An example of such a conserved feature is situated at the centre of the native molecules in Fig. 3 (CSV residues 74-110, 247-284; PSTV residues 76-112, 247-284) and includes two relatively large single-stranded regions which are completely conserved between CSV and PSTV. Also of interest are the stretches of purines (CSV residues 52-64; PSTV residues 49-65) and of uracil residues (CSV residues 291-297; PSTV residues 296-300). We have also found similar sequences (unpublished) in the encapsidated viroid-like RNAs of *Solanum nodiflorum* mottle virus and velvet tobacco mottle virus (26-28) and the avocado sun-

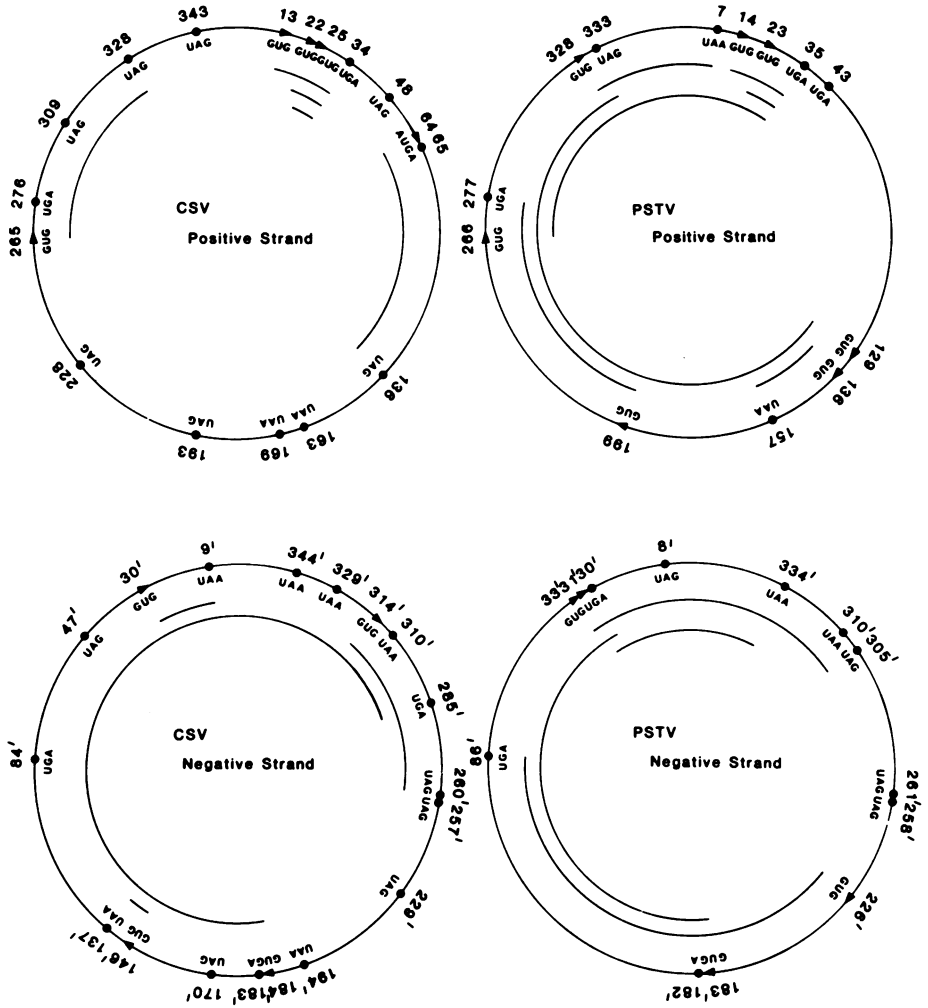


Figure 4. Possible polypeptide products of both the positive and negative strands of CSV and PSTV. The possible initiation codons (AUG, GUG; ▲) and termination codons (UAA, UAG, UGA; ●) are numbered at the first residue of the codon. Possible translation products are shown as solid lines within each circle.

blotch viroid (2). It will be of considerable interest to extend this comparison to other isolates of CSV and PSTV and to other viroids.

Relationship of Our Isolate of CSV to Other Viroid Isolates

Both Gross *et al.* (10) and Owens *et al.* (11) have used isolates of CSV,

of different origins from ours, for comparative studies with a PSTV isolate obtained from Dr. T.O. Diener and which has been sequenced (8). Gross *et al.* (10) have compared the oligonucleotide fingerprints of an isolate of CSV, obtained from Dr. M. Hollings, with those of PSTV and obtained distinctly different patterns. They concluded that PSTV and CSV differed significantly in sequence.

The complete catalogue of the fragments derived from complete RNase A and T_1 digestion of our isolate of CSV can be derived from the sequence of Fig. 2. A number of these fragments, up to 10 residues long, are common to PSTV (8), and the positions of these common oligonucleotides on the fingerprints of PSTV are known (8). However, several of these common oligonucleotides are not shared by PSTV and the CSV isolate of Gross *et al.* (10) as judged by comparative fingerprints. For example, CUACUACCG and AAACAACUG are common to PSTV and our isolate of CSV. These two oligonucleotides are closely positioned as two spots on the RNase T_1 fingerprints of PSTV (8,10) but only one spot is seen in the corresponding position of the RNase T_1 fingerprint of their isolate of CSV (10). This, together with other oligonucleotide differences, indicates non-identity of our isolate of CSV and that used by Gross *et al.* (8,10).

Owens *et al.* (11) used DNA complementary to PSTV to show by hybridization analysis that yet another isolate of CSV contained about 20% sequence homology with PSTV, whereas a viroid isolated from *Columnea erythrophae* contained about 40% sequence homology. In addition, their CSV isolate and the *Columnea* viroid had electrophoretic mobilities in a non-denaturing polyacrylamide gel which were markedly faster than that of PSTV, indicating appreciably different sizes and/or secondary structures. In contrast, our isolate of CSV has 69% sequence homology with PSTV (8), is only three residues shorter, and possesses a similar secondary structure.

Overall, the data indicate that our isolate of CSV differs from the CSV isolates of Gross *et al.* (10) and Owens *et al.* (11) and from the *Columnea* viroid, and is more closely related to PSTV, even allowing for some errors in estimates of sequence homology determined by cDNA.RNA hybridization analysis (29,30). It therefore appears likely that there exists a group of viroids, including PSTV, which share common sequences and possibly secondary structures, and which may be derived from a common ancestral viroid.

ACKNOWLEDGMENTS

We thank Dr. R.I.B. Francki for glasshouse facilities, Miss J.L. Rosey

for assistance, and Dr. P. Palukaitis for purified CSV and helpful discussions. This work was supported by the Australian Research Grants Committee.

*Correspondence: Please send to Dr. R.H.Symons

REFERENCES

1. Diener, T.O. (1979) Viroids and Viroid Diseases, John Wiley and Sons, New York.
2. Palukaitis, P., Hatta, T., Alexander, D.McE. and Symons, R.H. (1979) Virology *99*, 145-151.
3. Sänger, H.L., Klotz, G., Riesner, D., Gross, H.J. and Kleinschmidt, A.K. (1976) Proc. Natl. Acad. Sci. USA *73*, 3852-3856.
4. Palukaitis, P. and Symons, R.H. (1980) J. Gen. Virol. *46*, 477-489.
5. Conejero, V. and Semancik, J.S. (1977) Virology *77*, 221-232.
6. Davies, J.W., Kaesberg, P. and Diener, T.O. (1974) Virology *61*, 281-286.
7. Semancik, J.S., Conejero, V. and Gerhart, J. (1977) Virology *80*, 218-221.
8. Gross, H.J., Domdey, H., Lossow, C., Jank, P., Raba, M., Alberty, H. and Sänger, H.L. (1978) Nature *273*, 203-208.
9. Hollings, M. and Stone, O.M. (1973) Ann. Appl. Biol. *74*, 333-348.
10. Gross, H.J., Domdey, H. and Sänger, H.L. (1977) Nucleic Acids Res. *4*, 2021-2028.
11. Owens, R.A., Smith, D.R. and Diener, T.O. (1978) Virology *89*, 388-394.
12. Donis-Keller, H. (1980) Nucleic Acids Res. *8*, 3133-3142.
13. Lockard, R.E., Alzner-Deweerd, B., Heckman, J.E., MacGee, J., Tabor, M.W. and RajBhandary, U.L. (1978) Nucleic Acids Res. *5*, 37-55.
14. Johnson, R.A. and Walseth, T.F. (1979) in Advances in Cyclic Nucleotide Research, Brooker, G., Greengard, P. and Robison, G.A., Eds., Vol. 10, pp. 135-167, Raven Press, New York.
15. Richardson, C.C. (1971) in Procedures in Nucleic Acid Research, Cantoni, G.L. and Davies, D.R., Eds., Vol. 2, pp. 815-828, Harper and Row, New York.
16. Sanger, F. and Coulson, A.R. (1978) FEBS Lett. *87*, 107-110.
17. Donis-Keller, H., Maxam, A.M. and Gilbert, W. (1977) Nucleic Acids Res. *4*, 2527-2538.
18. Krupp, G. and Gross, H.J. (1979) Nucleic Acids Res. *6*, 3481-3490.
19. Kramer, F.R. and Mills, D.R. (1978) Proc. Natl. Acad. Sci. USA *75*, 5334-5338.
20. Maniatis, T. and Efstratiadis, A. (1980) Methods in Enzymol. *65*, 299-305.
21. Tinoco, I., Uhlenbeck, O.C. and Levine, M.D. (1971) Nature *230*, 362-367.
22. Tinoco, I., Borer, P.N., Dengler, B., Levine, M.D., Uhlenbeck, O.C., Crothers, D.M. and Gralla, J. (1973) Nature *246*, 40-41.
23. Riesner, D., Henco, K., Rokohl, U., Klotz, G., Kleinschmidt, A.K., Domdey, H., Jank, P., Gross, H.J. and Sänger, H.L. (1979) J. Mol. Biol. *133*, 85-115.
24. Niblett, C.L., Dickson, E., Horst, R.K. and Romaine, C.P. (1980) Phytopathology *70*, 610-615.
25. Matthews, R.E.F. (1978) Nature *276*, 850.

26. Randles, J.W., Davies, C., Hatta, T., Gould, A.R. and Francki, R.I.B. (1981) Virology *108*, 111-122.
27. Gould, A.R. (1981) Virology *108*, 123-133.
28. Gould, A.R. and Hatta, T. (1981) Virology *108*, 499-509.
29. Palukaitis, P. and Symons, R.H. (1979) Virology *98*, 238-245.
30. Gonda, T.J. and Symons, R.H. (1978) Virology *88*, 361-370.