Avocado sunblotch viroid: primary sequence and proposed secondary structure

Robert H. Symons

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

Received 13 October 1981

ABSTRACT

The sequence of the 247 nucleotide residues of the single strand circular RNA of avocado sunblotch viroid (ASBV) was determined using partial enzymic cleavage methods on overlapping viroid fragments obtained by partial ribonuclease digestion followed by ^{32}P -labelling *in vitro* at their 5'-ends. ASBV is much smaller than potato spindle tuber viroid (PSTV; 359 residues) and chrysanthemum stunt viroid (CSV; 356 residues). A secondary structure model for ASBV is proposed and contains 67% of its residues base paired. In contrast to the extensive (69%) sequence homology of CSV with PSTV, only 18% of the ASBV sequence is homologous to PSTV and CSV. There are eight potential polypeptide translation products with chain lengths from 4 to 63 amino acid residues coded for by the plus (infectious) strand and four potential translation products (2 to 60 residues) coded for by the minus strand. An improved method is described for the synthesis of γ - ^{32}P -ATP of high specific activity.

INTRODUCTION

Sunblotch is a serious disease affecting avocados (*Persea americana*). It can be transmitted readily by mechanical means to healthy avocados, with symptoms of the disease appearing about 6 months to two years after inoculation (1-3). It appears to have a very limited host range since it has only been transmitted to one other species, cinnamon (*Cinnamomum zelanicum*), which, like the avocado, is a member of the family Lauraceae (4). We recently purified and characterized a viroid, named avocado sunblotch viroid (ASBV), from sunblotch-infected avocado leaves (5) and infectivity studies have indicated that highly purified ASBV is the causative agent of the disease (6). Since the sequences of only two other viroids have been reported, potato spindle tuber viroid (PSTV) (7) and chrysanthemum stunt viroid (8), it is of considerable importance to determine the primary sequence and secondary structure of ASBV and to determine the common features shared by these viroids.

MATERIALS AND METHODS

Materials

ASBV (strain SB-1) was purified (5,9,10) from sunblotch infected avocado leaves kindly provided by D. Alexander, CSIRO Division of Horticultural Research, Merbein, Victoria. For most of the sequencing work, partly purified ASBV obtained after one preparative polyacrylamide slab gel (5,9,10) was used; this material consisted mostly of a mixture of the circular and linear forms of the viroid. Identical sequencing results were always obtained when the partly purified ASBV was compared with the more highly purified material (5,9,10). RNases T₁ and A were from Sigma Chemical Co. and RNase U₂ was from Sankyo. RNase PhyM (11), *Bacillus cereus* RNase (12) and phage T4 polynucleotide kinase (13) were purified as described.

Preparation of γ -³²P-ATP of high specific activity

A simple, routine procedure was modified from that of Johnson and Walseth (14). Carrier-free ³²Pi (10 mCi, New England Nuclear NEX054, in 200 μ l of 0.02 N HCl) in a sterile plastic tube (1.3 x 9 cm) was taken to dryness once with an oil pump in the presence of NaOH pellets. A reaction mixture was prepared from stock solutions, all stored at -20°C; to 100 μ l of 100 mM Tris-HCl, pH 9.0, 25 mM MgCl₂, 12 mM DTT, and 3.6 mM spermine, was added 80 μ l water, 2 μ l 10 mM L- α -glycerophosphate, 10 μ l 10 mM NAD, 2 μ l 5 mM ADP and 2 μ l 100 mM pyruvic acid. The dried down ³²Pi was dissolved in 180 μ l of this reaction mixture; the remaining 20 μ l was used to dissolve the precipitate obtained after centrifugation for 10 minutes at 4°C in an Eppendorf centrifuge of 20 μ l of the ammonium sulphate suspension of the Stage 1 enzymes of Johnson and Walseth (14). Four μ l of the enzyme solution was added to the dissolved ³²Pi and the mixture incubated at room temperature. After 20-30 minutes, a very small sample was taken with a drawn out capillary tube and spotted onto a 1.5 x 5.0 cm strip of PEI-cellulose (Macharey-Nagel Polygram CEL 300 PEI/UV₂₅₄) on which markers of AMP, ADP and ATP (each at 2 mg/ml) were previously dried at the origin. Ascending chromatography was carried out for 5 minutes in 0.5 M KH₂PO₄, the markers located under UV-light, and the strip cut through the ADP spot. Approximate conversion of ³²Pi (runs with AMP) to γ -³²P-ATP was determined with a radiation monitor and was always greater than 90%.

Although the γ -³²P-ATP can be used without purification (15), improved phosphorylation with polynucleotide kinase was obtained after further purification by step-wise chromatography on DEAE-cellulose (16). The reaction mixture was diluted with 2 ml water and run onto a 0.6 x 2.0 cm column of

DEAE-cellulose (HCO_3^-). The column was washed with 3.5 ml freshly prepared 0.07 M NH_4HCO_3 and the γ - ^{32}P -ATP eluted with 3.5 ml 0.25 M NH_4HCO_3 . The solution was desalted after the addition of distilled triethylamine and ethanol by taking it to dryness on a rotary evaporator (16,17). The product was dissolved in 2 ml 50% ethanol, 0.05 mM EDTA, pH 8, 5 mM 2-mercapto-ethanol, and stored at -20°C . The yield of γ - ^{32}P -ATP relative to starting $^{32}\text{P}_i$ was always greater than 90%.

Preparation of 5'- ^{32}P -RNA fragments of ASBV

Procedures used were similar to those described in detail elsewhere (8). In summary, purified or partly purified ASBV was partially digested with RNase T $_1$ or A (pH 7.5) or RNase U $_2$ (pH 3.5) in 600 mM NaCl, 10 mM MgCl_2 , at 0°C or 37°C . Reactions were terminated by phenol extraction, ether washing and ethanol precipitation. RNA fragments were then 5'-labelled with γ - ^{32}P -ATP and phage T4 polynucleotide kinase and the mixtures fractionated by electrophoresis on an 80 x 20 x 0.05 cm 6% polyacrylamide gel in 90 mM Tris-borate, pH 8.3, 2 mM EDTA, 7 M urea, at 20-25 mA for 4-6 hours. RNA bands were located by autoradiography at room temperature for 30 minutes, excised and eluted by soaking overnight at room temperature in 0.5 ml 500 mM ammonium acetate, 1 mM EDTA, 0.1% sodium dodecyl sulphate, containing 60 μg *E. coli* tRNA as carrier. Eluted fragments were precipitated twice with ethanol.

Sequencing of 5'-labelled ASBV fragments

Partial enzymic cleavage methods were used. Digestions were carried out with alkali to produce a reference ladder and with RNase T $_1$ (18), RNase U $_2$ (19), RNase PhyM (11), and *Bacillus cereus* RNase (12). Reaction mixtures were fractionated on either 80 x 20 x 0.05 cm 6% polyacrylamide gels (20) or on 40 x 20 x 0.05 cm 20% polyacrylamide gels in 90 mM Tris-borate, pH 8.3, 2 mM EDTA, 7 M urea, and the radioactive bands were detected by autoradiography.

Dideoxynucleotide sequencing on ASBV with cloned DNA primer

ASBV (1 μg) plus purified double-strand DNA restriction fragment (residues 101 to 153) obtained from a recombinant DNA clone of ASBV (kindly provided by Dr. P.J. Murphy) in 50 μl 0.2 M NaCl, 10 mM Tris-HCl, pH 8.0, 1 mM EDTA, 0.05% sodium dodecyl sulphate, were heated at 95°C for 2.0 minutes then incubated at 65°C for 30 minutes. The DNA:RNA hybrid was recovered by ethanol precipitation, the precipitate rinsed twice with cold 70% ethanol, dried, and the ASBV sequence on the 3'-side of the primer determined by the dideoxynucleotide chain termination procedure (20,35).

RESULTS

Approach Used for the Sequencing of ASBV

Since ASBV exists mostly as a covalently closed single-strand circular RNA (5), the approach used successfully for the sequencing of chrysanthemum stunt viroid (CSV) (8) was also used here. Basically, the viroid was partially digested with RNases T₁, U₂ and A under conditions of high salt concentration (600 mM NaCl, 10 mM MgCl₂) in order to limit cleavage by these enzymes to the limited number of exposed single-strand regions of the molecule. After labelling the exposed 5'-ends of the fragments with γ -³²P-ATP and phage T4 polynucleotide kinase, the labelled fragments were fractionated by size on a long (80 cm) denaturing polyacrylamide slab gel. Twenty to 30 bands from each gel were eluted and sequenced using the partial enzymic cleavage method as described under Materials and Methods. The complete sequence of the circular viroid was eventually obtained by overlaps from a large number of fragment sequences. Further details of this sequencing approach and the type of results obtained can be found in Haseloff and Symons (8). Although the complete sequence was obtained by this method, overlapping fragments corresponding to the left hand end of the secondary structure model considered below (Fig. 2) were difficult to obtain. That the sequence in this region was correct was confirmed by the dideoxynucleotide chain termination technique for sequencing RNA (20,35) using a specific DNA primer (corresponding to residues 101 to 153; Fig. 1) isolated from a recombinant DNA clone of ASBV in the plasmid vector pBR322 and provided by Dr. P.J. Murphy.

Primary Sequence and Proposed Secondary Structure of ASBV

The complete base sequence of the 247 residues of ASBV is given in Fig. 1; although the viroid is a covalently closed circular molecule (5), the sequence is presented in linear form for convenience and for discussion of possible translation products (see below). A possible secondary structure model (Fig. 2) was constructed from the ASBV sequence following the base pairing matrix procedure of Tinoco *et al.* (21) and is compared with the published secondary structure models for PSTV (7) and CSV (8). The residues of ASBV are numbered following the scheme of Gross *et al.* (7) for PSTV in which residue 1 is in the single-strand loop at the left hand end of the secondary structure model for PSTV (Fig. 2).

The number of the three types of base pairs for the proposed ASBV structure are compared with the corresponding values for CSV and PSTV (Table 1). The much higher proportion of A:U plus G:U base pairs relative to G:C

Figure 1. Primary sequence of ASBV. The circular RNA is presented as a linear sequence with residue 1 corresponding to the left hand end of the secondary structure model of Figure 2. A bar beneath three residues indicates an initiation triplet and a bar above a sequence indicates a termination triplet.

base pairs in ASBV as compared with CSV is consistent with the lower T_m of the thermal denaturation profile of ASBV (5); PSTV was not available for direct comparison. The percentage of residues base paired in the three structures is approximately the same.

The free energies (ΔG) of the ASBV and CSV structures were calculated using up-dated values of Tinoco *et al.* (22) and kindly supplied by Dr. D. Riesner (G. Steger, H. Gross, H.L. Sanger and D. Riesner, in preparation), as was the ΔG value for PSTV. The results are consistent with the lower relative stability of ASBV as compared with CSV and indicate that PSTV is more stable than our isolate of CSV.

The homologous sequences shared by the three viroids are boxed in the structures of Fig. 2. In contrast to the extensive sequence homology of CSV (69%) with PSTV (8), the boxed regions of ASBV represent only 18% of the sequence. The most notable feature of the homologous regions are the conserved GAAACC sequence in the central loop of each molecule and the oligo(U) plus associated sequences at the left hand end. An unusual feature is the sequence CUUCAGGG which occurs on the top of the PSTV and CSV structures (residues 81-88 and 80-87 respectively) but is found on the bottom of the ASBV structure (residues 167-173).

Possible polypeptide translation products of ASBV

PSTV and citrus exocortis viroid (CEV) do not act as messenger RNAs in cell-free translation systems (23-25) and this may be related to the

Figure 2. Proposed secondary structure of ASBV together with those of PSTV (7) and CSV (8). The boxed areas contain residues in ASBV which are homologous in the other two viroids.

TABLE 1

Properties of proposed secondary structures for ASBV, CSV and PSTV

Viroid	No. of residues	No. of base pairs			G:C base pairs % of total	Residues base paired %	ΔG (KJ/mol at 25°C in 1 M NaCl)
		A:U	G:C	G:U			
ASBV	247	43	28	12	34	67	-280
CSV	356	44	64	16	52	70	-570
PSTV	359	37	73	16	58	70	-680

Data for CSV was obtained from (8) and for PSTV from (7). Values used for the calculation of the ΔG for ASBV and CSV were provided by Dr. D. Riesner (Steger, Gross, Sanger and Riesner, in preparation). The ΔG value for PSTV was provided by Dr. Riesner.

inability of eukaryotic ribosomes to initiate translation on circular RNA molecules (26). The possibility cannot be excluded at present that sub-genomic linear fragments of viroids may act as mRNAs. Hence, the number of start and stop codons in both the plus strand (the infectious strand as isolated) and the minus strand (27) are given in Table 2; all possible start and stop codons are present. The length of the potential polypeptide translation products from both strands are listed in Table 3 together with the number of the first residue of the start and stop codons. For the plus strand, six of the seven possible polypeptides terminate at the same UAA

TABLE 2

Number of start and stop codons in plus and minus sequences of ASBV

	Start codons		Stop codons		
	AUG	GUG	UAA	UAG	UGA
Plus strand	3	4	2	2	9
Minus strand	1	3	8	1	5

The plus strand is taken as the isolated viroid and the minus strand as its complementary RNA sequence.

TABLE 3

Potential polypeptide translation products from plus and minus strands of ASBV - positions of start and stop codons and length of products

Residue number at 1st residue					Polypeptide product (amino acid residues)
Start codons		Stop codons			
AUG	GUG	UAA	UAG	UGA	
<u>Plus strand</u>					
25		121			32
179		121			63
194		121			58
	17			41	8
	88	121			11
	106	121			5
	109	121			4
<u>Minus strand</u>					
227		212			5
	191			158	11
	159			226	60
	137	131			2

Residue numbers are at the first residue of the start or stop codon. The same residue numbers are retained in the minus strand and therefore run in the 3' to 5' direction.

stop codon while the four possible polypeptides from the minus strand terminate at different stop codons. UAG is not found as the first stop codon after a start codon in either strand. Potential polypeptide products range in length from 4 to 63 residues for the plus strand and 2 to 60 residues for the minus strand.

DISCUSSION

The sequence of 247 residues of ASBV is appreciably shorter than the PSTV sequence of 359 residues (7) and the CSV sequence of 356 residues (8). However, the sequence of the smallest RNA component of cadang cadang viroid (28,29) is even shorter with only 243 residues (Haseloff, Mohamed and Symons, unpublished).

In view of the 69% sequence homology of our isolate of CSV with PSTV

(8) and of the properties of a European isolate of CSV (3) and of a viroid isolated from Columnea erythrophae (31), we considered it likely that there exists a group of viroids, including PSTV, which share common sequences and possible secondary structures and which may be derived from a common ancestral viroid (8). Citrus exocortis viroid (CEV) must be added to this group in view of its extensive sequence homology with PSTV and CSV (Visvader and Symons, unpublished). ASBV shows a much lower (18%) sequence homology with PSTV and CSV (Fig. 2) which indicates that it should be allocated to a separate viroid group.

In spite of the limited and scattered regions of sequence homology between ASBV and PSTV and CSV, there are intriguing relationships between ASBV and the other two viroids. For example, the central unpaired loopout sequence, GAAACC, in the ASBV secondary structure can be aligned vertically with the corresponding sequence in PSTV and CSV (Fig. 2) so that a few sequences at the right hand and left hand ends of the ASBV structure occupy similar positions with the homologous sequences in PSTV and CSV (Fig. 2). Further, this common GAAACC sequence is part of a larger, central, highly conserved region in the proposed secondary structure of PSTV, (7), CSV (8), CEV (Visvader and Symons, unpublished) and the four RNA components of cadang cadang viroid (Haseloff, Mohamed and Symons, unpublished). It is conceivable that ASBV may have arisen from an ancestral viroid similar to PSTV and CSV by elimination of about 90 residues from the left hand end of the molecule (Fig. 2) together with substantial sequence divergence in the rest of the molecule relative to PSTV and CSV. The complete sequences of all known viroids are obviously of importance in extending this discussion on the evolutionary origins of viroids, as will be sequences and structures of the encapsidated viroid-like RNAs (32-34).

In a recent discussion on the origins of viroids, Diener (36) considered that they may be escaped introns. One testable corollary of his hypothetical model was that complementary sequences of all viroids (or the sequences of viroids) should have in common a sequence complementary to the 5'-end of the mammalian small nuclear U1 RNA. Although such a sequence is present in PSTV (residues 257 to 279) (36) and CSV (residues 256 to 278)(8), no such sequence is present in the plus or minus strand of ASBV.

The approach used here for sequencing ASBV was the same as that used successfully for sequencing CSV (8). A large number of 5'-labelled fragments were rapidly obtained from partial RNase digests of only a few μ g of viroid while the sequencing of these fragments by the partial enzymic cleavage

method is rapid and reliable. One problem with this overall approach is that some regions of the viroid molecule are always found in partial RNase digests and are repeatedly sequenced while other regions are poorly represented. In the case of ASBV, a poorly represented region was the left hand end of the molecule which is consistent with its proposed open secondary structure in Fig. 2. Unambiguous overlaps in these areas were difficult to obtain so that the sequencing of a large number of fragments had to be carried out. That the sequence at the left hand end of the molecule was correct was confirmed by the dideoxynucleotide chain termination technique for sequencing RNA (20,35) using a specific DNA primer (corresponding to residues 101 to 153) isolated from a recombinant DNA clone of ASBV.

ACKNOWLEDGMENTS

The author thanks Miss J.L. Rosey for assistance, J. Haseloff for helpful advice on RNA sequencing and Dr. P.J. Murphy for a cloned DNA fragment of ASBV. This work was supported by the Australian Research Grants Committee and the Rural Credits Development Fund of the Reserve Bank of Australia.

REFERENCES

1. Horne, W.T. and Parker, E.R. (1931) *Phytopathology* **21**, 235-238.
2. Wallace, J.M. and Drake, R.J. (1962) *Phytopathology* **52**, 237-241.
3. Da Graca, J.V. and Van Vuuren, S.P. (1981) *Plant Disease* **65**, 46-47.
4. Da Graca, J.V. and Van Vuuren, S.P. (1980) *Plant Disease* **64**, 475.
5. Palukaitis, P., Hatta, T., Alexander, D. McE. and Symons, R.H. (1979) *Virology* **99**, 145-151.
6. Allen, R.N., Palukaitis, P. and Symons, R.H. (1981) *Aust. Plant Pathology* **10**, 31-32.
7. Gross, H.J., Domdey, H., Lossow, C., Jank, P., Raba, M. and Alberty, H. (1978) *Nature* **273**, 203-208.
8. Haseloff, J. and Symons, R.H. (1981) *Nucleic Acids Res.* **9**, 2741-2752.
9. Palukaitis, P. and Symons, R.H. (1980) *J. Gen. Virol.* **46**, 477-489.
10. Palukaitis, P., Rakowski, A.G., Alexander, D. McE. and Symons, R.H. (1981) *Ann. Appl. Biol.* **98**, 439-449.
11. Donis-Keller, H. (1980) *Nucleic Acids Res.* **8**, 3133-3142.
12. 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.
13. 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.
14. Johnson, R.A. and Walseth, T.F. (1979) *Advances in Cyclic Nucleotide Research*, Brooker, G., Greengard, P. and Robison, G.A., Eds., Vol. 10, pp. 135-167, Raven Press, New York.
15. Maxam, A.M. and Gilbert, W. (1980) *Methods in Enzymology* **65**, 499-560.
16. Symons, R.H. (1977) *Nucleic Acids Res.* **4**, 4347-4355.
17. Symons, R.H. (1974) *Methods in Enzymology* **29**, 102-115.

18. Donis-Keller, H., Maxam, A.M. and Gilbert, W. (1977) *Nucleic Acids Res.* **4**, 2527-2538.
19. Krupp, G. and Gross, H.J. (1979) *Nucleic Acids Res.* **6**, 3481-3490.
20. Symons, R.H. (1979) *Nucleic Acids Res.* **7**, 825-837.
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. Hall, T.C., Wepprich, R.K., Davies, J.W., Weathers, L.G. and Semancik, J.S. (1974) *Virology* **61**, 486-492.
24. Davies, J.W., Kaesberg, P. and Diener, T.O. (1974) *Virology* **61**, 281-286.
25. Semancik, J.S., Conejero, V. and Gerhart, J. (1977) *Virology* **80**, 218-221.
26. Kozak, M. (1979) *Nature* **280**, 82-85.
27. Matthews, R.E.F. (1978) *Nature* **276**, 850.
28. Randles, J.W. (1975) *Phytopathology* **65**, 163-167.
29. Randles, J.W. and Hatta, T. (1979) *Virology* **96**, 47-53.
30. Gross, H.J., Domdey, H. and Sanger, H.L. (1977) *Nucleic Acids Res.* **4**, 2021-2028.
31. Owens, R.A., Smith, D.R. and Diener, T.O. (1978) *Virology* **89**, 388-394.
32. Randles, J.W., Davies, C., Hatta, T., Gould, A.R. and Francki, R.I.B. (1981) *Virology* **108**, 111-122.
33. Gould, A.R. (1981) *Virology* **108**, 123-133.
34. Gould, A.R. and Hatta, T. (1981) *Virology* **109**, 137-147.
35. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467.
36. Diener, T.O. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 5014-5015.