Comparative sequence and structure of circular RNAs from two isolates of lucerne transient streak virus

Paul Keese, George Bruening* and Robert H. Symons+

Adelaide University Centre for Gene Technology, Department of Biochemistry, University of Adelaide, Adelaide, SA 5000, Australia

Received 14 June 1983

The sequences of the circular viroid-like RNA 2 (virusoid) from two isolates of lucerne transient streak virus (LTSV) have been determined. They both contain 324 residues and share 98% sequence homology. The proposed secondary structures fit a highly base-paired rod-like model as shown previously for viroids and two other virusoids, velvet tobacco mottle virus (VTMoV) and *Solanum nodiflorum* mottle virus (SNMV). Except for one region of sequence and positional homology (GAUUUU, residues 21–26) LTSV virusoid exhibits limited homology with all presently known virusoids, including those from VTMoV, SNMV and subterranean clover mottle virus.

Lucerne transient streak *Circular RNA* *Virusoid structure* *RNA sequencing*

1. INTRODUCTION

Lucerne transient streak virus (LTSV) [1–3] belongs to a group of novel single-strand RNA plant viruses which also includes velvet tobacco mottle virus (VTMoV), *Solanum nodiflorum* mottle virus (SNMV) and subterranean clover mottle virus (SCMoV) [4–6]. They are allied to the Sobemo-virus group of plant viruses [7] but usually contain two rather than one major encapsidated single-strand RNA species. The larger RNA, RNA 1, is a linear molecule of about 4500 residues while RNA 2 is a circular, covalently closed molecule of 300–400 residues with physical characteristics similar to another group of plant pathogens, the viroids [8,9]. These encapsidated, viroid-like RNAs have been termed virusoids [10]. The sequences of RNA 2 of VTMoV and SNMV have been reported [11].

We report here the sequence and proposed secondary structure of the virusoids of two isolates of LTSV, one from Australia (LTSV-A) [1] and the other from New Zealand (LTSV-N) [2]. In the case of LTSV-N, the RNA 2 is satellite-like since its replication is dependent on the presence of RNA 1 which is capable of autonomous replication [12]. A similar situation presumably exists for LTSV-A and for the isolate reported from Canada [13].

2. MATERIALS AND METHODS

The two isolates of LTSV, LTSV-A and LTSV-N, were kindly provided by Dr R.I.B. Francki (Waite Institute, Glen Osmond SA). Viruses were purified from infected *Chenopodium quinoa* for LTSV-A and from *Nicotiana clevelandii* for LTSV-N. Viral RNAs were extracted essentially as in [3] and fractionated by electrophoresis on 4% polyacrylamide, 7 M urea gels in 90 mM Tris–borate (pH 8.3), 2 mM EDTA. RNA bands were located by staining with toludine blue and RNA 2 eluted from the gel slice by soaking in 0.5 M ammonium acetate, 1% sodium dodecyl sulphate, 1 mM EDTA [14]. ^{32}P-Labelled nucleotides were obtained from Bresa Pty. Ltd. (Adelaide SA).

The two purified virusoids were sequenced by the partial enzymic digestion technique, essentially

+ To whom correspondence should be addressed

* Permanent address: Department of Biochemistry and Biophysics, University of California, Davis, CA 95616, USA

Published by Elsevier Science Publishers B.V.
00145793/83/$3.00

as in [14]. In addition, circular RNA 2 was linearized by heating at 100°C for 2.0 min in 10 mM Tris–HCl (pH 9.0) and converted to double-strand cDNA as in [15]. *Taq*I and *Msp*I restriction enzyme fragments were then ligated into the *Acc*I site of the replicative form of the bacteriophage vector M13mp73 using T4 DNA ligase [16]. *Hae*III restriction enzyme fragments were also cloned as above into the *Sma*I site of the replicative form of M13mp93. Recombinant phage were screened for LTSV RNA 2 inserts by sequencing [17] using a specific M13 primer, TC3AGTCACGACGT (New England Biolabs, Waltham MA).

3. RESULTS

3.1. *Sequencing procedure*

The approach used successfully for the sequencing of VTMoV RNA 2 and SNMV RNA 2 [11], chrysanthemum stunt viroid (CSV) [14], citrus exocortis viroid (CEV) [15], avocado sunblotch viroid (ASBV) [18], and coconut cadang-cadang viroid (CCCV) [11] was also used here. Purified circular RNA 2 of each of the two isolates of LTSV was partially digested with RNases T_1, U_2 and A and the exposed 5′-ends of the fragments labelled using [γ-^{32}P]ATP and T4 polynucleotide kinase. The 5′-labelled fragments were then fractionated by denaturing polyacrylamide gel electrophoresis, eluted and sequenced using the partial enzymic sequencing procedure [14]. Both RNA 2 isolates were completely sequenced by this method except for some ambiguity in two regions of the molecule (residues 125–130 and 195–210; fig.1,2) which were poorly digested by the single-strand-specific RNases used for sequencing and gave band compression on sequencing gels [14]. It also proved difficult to obtain overlapping fragments correspon-

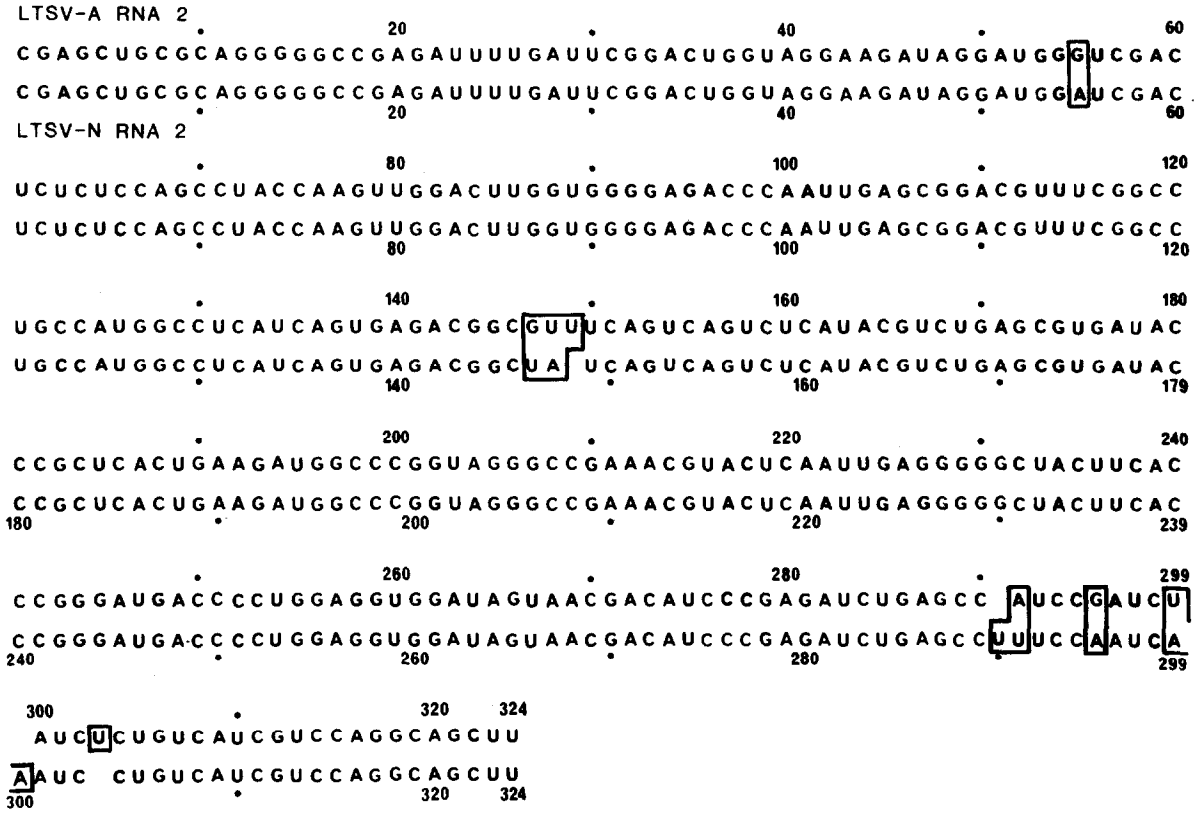

Fig.1. Nucleotide sequences of RNA 2 of LTSV-A and LTSV-N. The circular RNAs are presented in a linear form and are aligned for maximum sequence homology. Residue 1 corresponds to the left-hand end of the secondary structure models of fig.2. The sequence differences between the two RNAs are boxed.

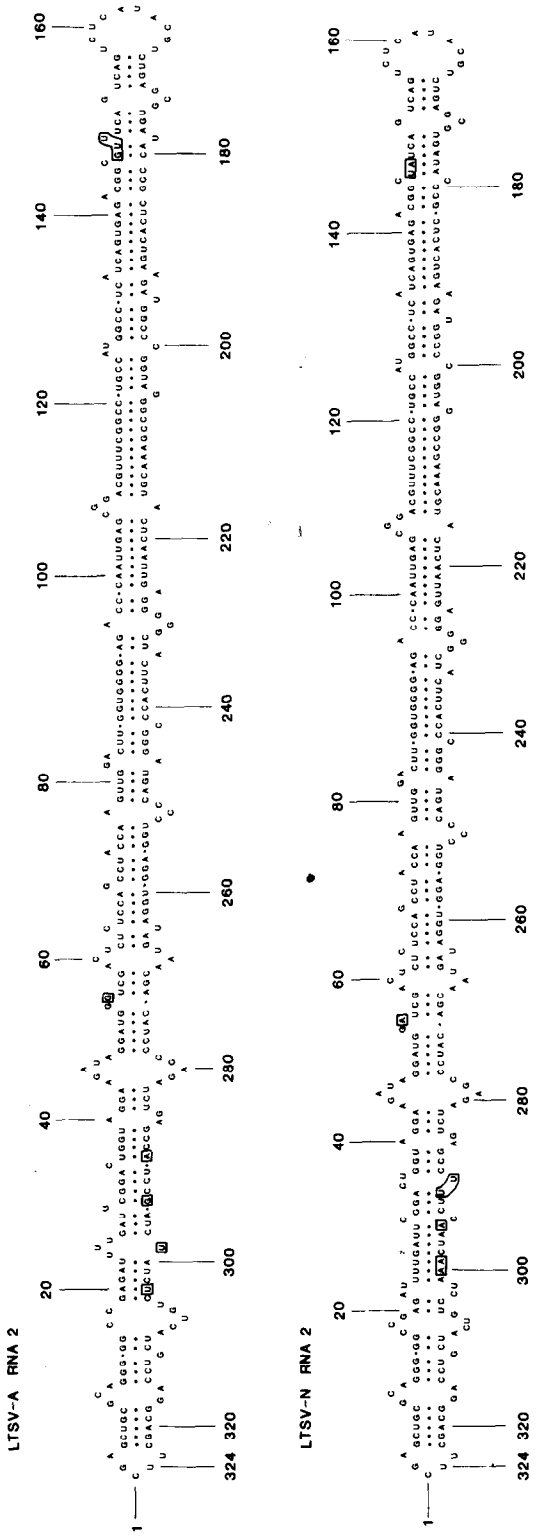

Fig.2. Predicted secondary structures of RNA 2 of LTSV-A and LTSV-N. The sequence differences between the two RNAs are boxed.

ding to the right-hand loop in the proposed secondary structure (fig.2) due to the marked sensitivity of this region to nuclease digestion. These difficulties were resolved successfully by cloning double-strand cDNA fragments into the replicative form of the bacteriophage vectors M13mp73 and M13mp93 and sequencing the cloned inserts by the dideoxynucleotide chain termination method. RNA 2 of one isolate, LTSV-A, was sequenced completely by both methods.

3.2. *Sequence and proposed secondary structure of RNA 2 from LTSV-A and LTSV-N*

The sequences of both isolates of LTSV RNA 2 are given in fig.1. Although the RNAs are covalently closed circular molecules, the sequences are presented in linear form for convenience and ease of comparison. Both species contain 324 residues and share extensive sequence homology with only 8 residues different. About half of these differences are in one small region of the molecule between residues 290 and 303 (fig.1).

Secondary structure models for the two RNAs were constructed as in [19] and are given in fig.2. Both RNAs form extensively base-paired rod-like structures which are similar to those described for viroids [9] and for the 2 circular virusoids of VTMoV and SNMV [11]. The structures are consistent with the known sites of high sensitivity to RNases under the conditions of high salt concentration and low temperature used to generate specific RNA fragments for direct enzymatic sequencing [14]. Thus, the terminal single-strand hairpin loops and two internal regions (residues 43–48 and 227–231 in fig.2) were especially susceptible.

The properties of the proposed secondary structures of the virusoids of LTSV are summarized in table 1 and are compared with those of the published structures of RNA 2 from VTMoV and SNMV [14], and of 5 viroids, potato spindle tuber viroid (PSTV) [20], ASBV [18], CSV [14], CEV [15] and CCCV [10]. All RNAs, except ASBV, contain a similar proportion of G:C basepairs while the percentage of residues base-paired varies in the narrow range from 66–73%. Circular RNA molecules of random sequence and the same size as viroids were calculated to contain about 55% of their residues base-paired [21]. The thermodynamic properties of the proposed models for

Table 1

Properties of proposed secondary structures for two isolates of LTSV RNA 2 and other virusoids and viroids

RNA	No. of residues	No. basepairs			G:C base-pairs (% of total)	Residues base-paired (%)	ΔG[a] (kJ/mol) at 25°C in 1 M NaCl	ΔG/No. of residues	Ref.
		A:U	G:C	G:U					
LTSV-A RNA 2	324	40	65	11	56	72	−440	−1.36	Here
LTSV-N RNA 2	324	42	61	12	53	71	−445	−1.37	Here
VTMoV RNA 2	365	39	72	13	58	68	−350	−0.96	[11]
SNMV RNA 2	377	41	76	20	55	73	−455	−1.21	[11]
ASBV	247	43	28	12	34	67	−280	−1.13	[18]
PSTV	359	37	73	16	58	70	−610	−1.70	[20]
CSV	356	44	64	16	52	70	−540	−1.52	[14]
CEV	371	34	72	18	58	67	−590	−1.59	[15]
CCCV	246	20	55	6	68	66	−320	−1.30	[10]

[a] Parameters for calculation provided by Dr D. Riesner et al. (personal communication)

the LTSV RNAs 2 were calculated using values kindly provided by Dr D. Riesner et al. (personal communication) and are compared with those calculated for the other RNAs. The value of −440 kJ/mol for LTSV-A RNA 2 (table 1) is consistent with its thermal denaturation properties in which LTSV-A RNA 2 gave a T_m of 70°C in 0.15 M NaCl, 0.015 M sodium citrate (pH 7) [3]. Under the same conditions, VTMoV RNA 2 and SNMV RNA 2 gave T_m values of 57°C and 64°C, respectively [5].

3.3. *Possible polypeptide translation products from LTSV RNA 2 species and their complements*

Although eukaryotic ribosomes do not interact with circular RNAs [22], sub-genomic linear fragments derived from either the infectious plus strand or its complement could act as mRNAs. Thus, 7 potential polypeptides are encoded by RNA 2 and its complement for each of the isolates of LTSV (fig.3). All possible translation products are <75 amino acids long; the gene coding for the coat protein (about 300 amino acids [2,6]) must therefore reside in the RNA 1 species. Despite the considerable sequence homology between the 2 virusoid isolates of LTSV, only 2 of these polypeptides, of 6 and 29 amino acids (fig.3), are shared between the isolates. This limited conservation of possible translation products between the highly

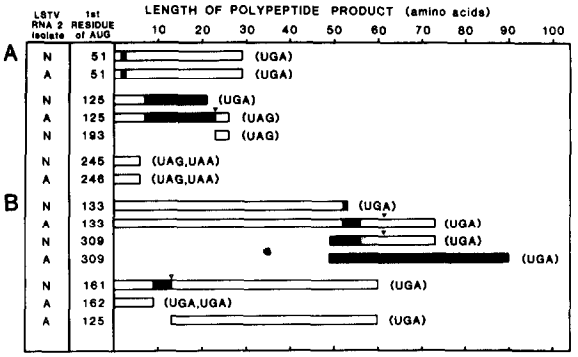

Fig.3. Possible polypeptide products of RNA 2 of LTSV-A and LTSV-N in (A) and their putative complementary RNAs (B) are shown in schematic form. Each possible translation product is given with the residue number of the first residue of the AUG initiation codon and termination codon(s) in parentheses. For the complementary sequences, the same residue numbers are retained and therefore run in the 3′-to-5′ direction. The clear areas represent regions of amino acid sequence homology and the black areas of non-homology for each group of polypeptides. Inverted triangles indicate sites of internal methionine residues.

conserved RNA 2 species of the two LTSV isolates suggests that they may lack functional mRNA activity in vivo. This property is characteristic of viroids [23] and of the satellite RNA of tobacco ringspot virus [24].

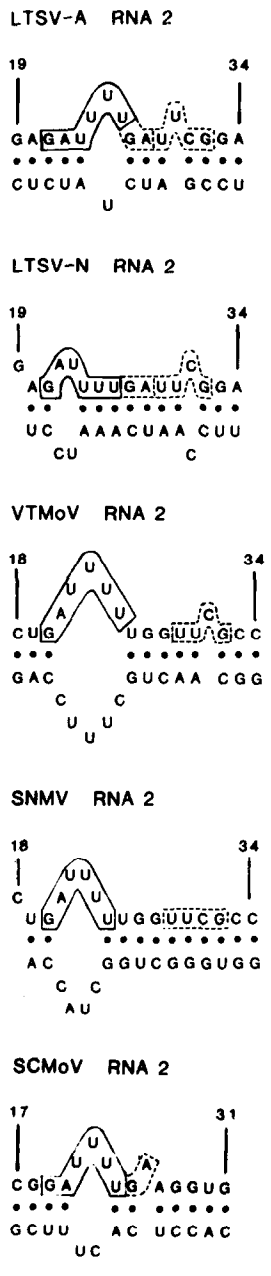

Fig.4. Portion of the predicted native structures of RNA 2 of LTSV-A and LTSV-N that share sequence homology with RNA 2 of VTMoV and SNMV [11] and SCMoV [26]. Solid boxes indicate sequence homology in all species. Broken boxes indicate sequence homology between at least two species. Residues are numbered from the left-hand end of the secondary structure model for each RNA. VTMoV RNA 2 and SNMV RNA 2 share >90% sequence homology [11] and the viruses are closely related serologically [4].

4. DISCUSSION

In overall structure, RNA 2 of LTSV resembles both the viroids and the virusoids of VTMoV and SNMV in consisting of small single-strand covalently closed circular RNA molecules with extensive base-paired regions interspersed with single-strand regions. Although the predicted free energy of LTSV RNA 2 is similar to that of VTMoV RNA 2 and SNMV RNA 2 (table 1), the smaller size of LTSV RNA 2, its higher T_m and greater resistance to ribonuclease (not shown) indicates that it is more stable than the other two virusoids. LTSV RNA 2 may show, therefore, hydrodynamic properties more closely related to viroids such as PSTV and CCCV which demonstrate greater stiffness in solution than the RNA 2 species of VTMoV and SNMV [25].

The virusoids of VTMoV and SNMV share with all viroid sequences published a common sequence of GAAAC which is positioned in a central single-strand loop in each molecule (residues 86–90) [11]. However, this sequence in LTSV RNA 2 occurs in a highly base-paired region (residues 210–214, fig.2) which is positionally distinct. Furthermore, although LTSV RNA 2 exhibits only limited sequence homology when compared individually with the virusoids of VTMoV and SNMV [11] and of SCMoV [26], there is one small region of sequence, structural and positional homology shared by the four known virusoids (fig.4). This involves a GAUUUU string (fig.4 (■)) between residues 19–26 where residue 1 corresponds to the left-hand end of their respective secondary structure models. The broken boxes indicate sequence homology between the virusoids of LTSV and either one or two of the virusoids of VTMoV, SNMV and SCMoV. Whether or not this region of sequence homology reflects a functional relationship between the known virusoids remains to be determined since LTSV RNA 1 can support not only its own virusoid but also those of SNMV [27] and SCMoV (unpublished).

ACKNOWLEDGEMENTS

The authors thank Dr R.I.B. Francki for the supply of virus isolates and for glasshouse facilities, J. Haseloff for helpful discussions, Jenny Rosey and Sharon Freund for assistance, and the Australian Research Grants Scheme for financial support.

REFERENCES

[1] Blackstock, J.McK. (1978) Aust. J. Agric. Res. 29, 291–304.
[2] Forster, R.L.S. and Jones, A.T. (1979) Ann. Appl. Biol. 93, 181–189.
[3] Tien-Po, Davies, C., Hatta, T. and Francki, R.I.B. (1981) FEBS Lett. 132, 353–356.
[4] Randles, J.W., Davies, C., Hatta, T., Gould, A.R. and Francki, R.I.B. (1981) Virology 108, 111–122.
[5] Gould, A.R. and Hatta, T. (1981) Virology 109, 137–147.
[6] Francki, R.I.B., Randles, J.W., Hatta, T., Davies, C., Chu, P.W.G. and McClean, G.D. (1983) Plant Pathol. 32, 47–59.
[7] Matthews, R.E.F. (1982) Intervirology 17, 1–199.
[8] Diener, T.O. (1979) Viroids and Viroid Diseases, Wiley, New York.
[9] Gross, H.J. and Riesner, D. (1980) Angew. Chem. Int. Ed. Engl. 19, 231–243.
[10] Haseloff, J., Mohamed, N.A. and Symons, R.H. (1982) Nature 299, 316–321.
[11] Haseloff, J. and Symons, R.H. (1982) Nucleic Acids Res. 10, 3681–3691.
[12] Jones, A.T., Mayo, M.A. and Duncan, G.H. (1983) J. Gen. Virol. 64, in press.
[13] Paliwal, Y.C. (1982) Phytopathology 7, 989.
[14] Haseloff, J. and Symons, R.H. (1981) Nucleic Acids Res. 9, 2741–2752.
[15] Visvader, J.E., Gould, A.R., Bruening, G.E. and Symons, R.H. (1982) FEBS Lett. 137, 288–292.
[16] Messing, J., Crea, R. and Seeburg, P.H. (1981) Nucleic Acids Res. 9, 309–321.
[17] Sanger, F., Coulson, A.R., Barrell, B.G., Smith, A.J.H. and Roe, B.A. (1980) J. Mol. Biol. 143, 161–178.
[18] Symons, R.H. (1981) Nucleic Acids Res. 9, 6527–6537.
[19] Tinoco, I., Uhlenbeck, O.C. and Levine, M.D. (1971) Nature 230, 362–367.
[20] Gross, H.J., Domdey, H., Lossow, C., Jank, P., Raba, M. and Alberty, H. (1978) Nature 272, 203–208.
[21] Riesner, D., Henco, K., Rokohil, U., Klotz, G., Kleinschmidt, A.K., Domdey, H., Jank, P., Gross, H.J. and Sänger, H.J. (1979) J. Mol. Biol. 133, 85–115.
[22] Kozak, M. (1979) Nature 280, 82–85.
[23] Sänger, H.L. (1982) in: Nucleic Acids and Proteins in Plants, vol.2 (Parthier, B. and Boulter, D. eds) pp.368–454, Springer, Berlin, New York.
[24] Owens, R.A. and Schneider, I.R. (1977) Virology 80, 222–224.
[25] Riesner, D., Kaper, J.M. and Randles, J.W. (1982) Nucleic Acids Res. 10, 5587–5598.
[26] Haseloff, J. and Symons, R.H. (1983) submitted.
[27] Jones, A.T. and Mayo, M.A. (1983) J. Gen. Virol., in press.