The nucleotide sequence of a satellite RNA associated with arabis mosaic nepovirus

Yuan Yi Liu,1* C. U. T. Hellen,† J. I. Cooper,2 D. J. Bertioli,1 D. Coates3 and G. Bauer4

1 Department of Plant Sciences, University of Oxford, South Parks Road, Oxford OX1 3RB, 2NERC Institute of Virology, Mansfield Road, Oxford OX1 3SR, 3Genetics Laboratory, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU and 4Max Planck Institut für Biophysikalische Chemie, Nikolaufberg, 3400 Göttingen, F.R.G.

The nucleotide sequence of a satellite RNA (satRNA) associated with a lilac isolate of arabis mosaic virus (ArMV) was determined from cDNA copies. The sequence was 1104 nucleotides in length excluding the poly(A) tail, contained a long open reading frame which encodes a polypeptide of 360 amino acids, with an Mr of 39K. Nucleotide sequence comparisons revealed that the ArMV-associated satRNA shared 83% nucleotide identity with a satRNA from grapevine fanleaf nepovirus, but no extensive sequence homology was observed with other nepoviral satRNAs.

Introduction

Arabis mosaic virus (ArMV) is a nepovirus with many antigenic properties in common with grapevine fanleaf virus (GFLV; Hewitt et al., 1971); however it differs in the breadth of its natural host range and in being carried usually by a distinct species of longidorid nematode. In common with other nepoviruses, ArMV has a bipartite RNA genome each molecule of which is 3'-polyadenylated and terminated at its 5' end by a genome-linked protein, VPg (Mayo et al., 1982; Hellen, 1987). In some instances, natural virus isolates contain a third encapsidated RNA molecule, a satellite RNA (satRNA) sensu Murant & Mayo (1982). Davies & Clark (1983) described an isolate of ArMV (from hops growing in Kent, U.K.) which contained a satRNA that closely resembles the satRNA of tobacco ringspot nepovirus (TobRV; Schneider, 1977; Buzayan et al., 1986) in that it occurs as linear or circular molecules and has the ability to undergo self cleavage. The hop ArMV-associated satRNA has 50% sequence homology with the satRNA of TobRV (Kaper et al., 1988).

When virion RNA from particles of a lilac (Syringa vulgaris) isolate of ArMV was analysed, a third RNA (RNA-3) was detected. This, like the satRNAs of tomato black ring virus (TBRV; Murant et al., 1973 Mayo et al., 1979; Koenig & Fritsch, 1982; Meyer et al., 1984) and GFLV (Pinck et al., 1987; Fuchs et al., 1989), resembled the viral genomic RNAs structurally in being 3'-polyadenylated, and in having mRNA activity and a 5'-linked VPg.

Biological properties of the RNA-3 in lilac ArMV and evidence for its satellite nature will be published elsewhere but here we report its nucleotide sequence and make comparisons with sequences of other satRNAs.

Methods

Virus propagation and nucleic acid preparation. ArMV was propagated in Chenopodium quinoa and purified from systemically infected leaves as described by Harrison & Nixon (1960) except that particles were coacervated with 10% (w/v) polyethylene glycol in the presence of 1% (w/v) sodium chloride. RNA was extracted from virions as described by Massalski & Cooper (1986). Glyoxal-treated satRNA was separated from genomic species by electrophoresis in agarose gels (McMaster & Carmichael, 1977) and was recovered by electroelution using an ISCO model 1750 concentrator as described by Zassenhaus et al. (1982).

Synthesis and cloning of DNA. The RNA-3 separated from the genomic RNAs was used as a template for cDNA synthesis using oligo(dT)12-18 primers and avian myeloblastosis virus (AMV) reverse transcriptase. The reaction conditions for synthesis of cDNA were based on those described by Cann et al. (1983) and Davies et al. (1978). In 50 μl of reaction mixture, 5 μl of satRNA was mixed with 5 μl of reverse transcriptase buffer (× 10) (containing 0.5 M-Tris–HCl pH 8.3, 1.0 M-NaCl, 100 mM-MgCl2, 60 mM-DTT, 5 μl each of the four dNTPs (5 mM), 5 μC[α-32P]dATP (10 μCi/μl, 6000 Ci/mmol), 2 μl oligo(dT)12-18 (1 mg/ml) and 40 units of AMV reverse transcriptase), and incubated at 37°C for 30 min. After extraction with phenol–chloroform, cDNA–RNA hybrids were separated from unincorporated nucleotides...
Sequence analysis. Primers were synthesized by the phosphotriester method with an Applied Biosystems 380 B DNA Synthesizer. With a universal sequencing primer (5' TATAATTGCGCTTATGT 3') or complementary to nucleotides 443 to 461 (3' GAACGAAAGGTTAGGAAGA 5') were used to facilitate extension of the inserts were sequenced using the universal sequencing primer. The terminal sequence was also determined using fluorescent dideoxynucleotide terminators and automated fluorescence detection (Staden, 1982, 1984a, b, 1986; Devereux et al., 1984), available on the Oxford University Computing Service VAX/VMS.

Results and Discussion

The complete nucleotide sequence of RNA-3 from ArMV (lilac) comprises 1104 nucleotides and is presented in Fig. 1. The clone A3/3 almost fully represented the satRNA and contained a run of 30 A residues (not shown in Fig. 1) indicating that the RNA sequence at its 3' end had been copied. The sequence at the 5' end, obtained when RNA-3 from virions was used as a template with the reverse transcription kit, terminated with two indecipherable nucleotides. These were identified as U and A using fluorescent dideoxynucleotide terminators; neither was present in clone A3/3.

The positive-sense RNA-3 sequence contains one long open reading frame whereas the negative-sense strand contains only short ones; the longest corresponds to 88 amino acids, a protein of approximately 9K. The first AUG, at positions 2 to 4, corresponds to an initiation codon for a putative peptide containing eight amino acids. In this, as in many other respects, the ArMV satRNA sequence closely resembles that described by Fuchs et al. (1989); in the GFLV RNA-3, the corresponding peptide has 22 amino acids. The longest ORF of ArMV RNA-3 has three in-frame methionine codons at positions 15 to 17, 126 to 128 and 216 to 218 in the 5' end region. Proteins initiated at each of these methionines would have Mr values of 38849, 34673 and 31310 respectively. When ArMV RNA-3 was translated in rabbit reticulocyte lysates (Hellen, 1987), three products were obtained with Mr values of 41K, 37K and 36K. The 41K product was the most abundant, the 37K protein was present in lesser amounts and the 36K product was barely detectable. By contrast, Fuchs et al. (1989), using wheatgerm extracts and GFLV RNA-3, obtained only one translation product although the first three methionines fall in identical positions within the longest ORF of this sequence. We have not yet excluded the possibility that the three translation products obtained from ArMV RNA-3 result from premature termination after initiation at the first AUG or the possibility that another initiation codon is occasionally selected. Furthermore, although it is conceivable that the satRNA of ArMV is a mixture of molecules differing in the positions of their stop codons, we did not find sequence heterogeneity between the clones A3/3 and A3/2.

The 3' end of the RNA-3 from ArMV has a non-coding region of seven nucleotides and is thereby shorter than the corresponding regions in the satRNAs of GFLV (74; Fuchs et al., 1989) or TBRV (74; Meyer et al., 1986). The region does not contain the polyadenylation signal AAUAAA implicated by Benoist et al. (1980) for animal mRNAs and viral RNAs or that proposed by Messing et al. (1983) for plant mRNAs i.e. G/AAUAA1-3. The 5' end of the ArMV RNA-3 sequence has a non-coding leader of 14 nucleotides preceding the first AUG of the longest ORF. The 5' terminal 11 nucleotides (UAUGAAAAAUU) fits the consensus sequence at the 5' end of nepovirus RNAs noted by Fuchs et al. (1989);UG/UGAAAAAUU/AU/AU/A).

The satRNA of lilac ArMV has a base composition of 22.7% A, 25.7% C, 28.0% G and 23.8% U. The distribution of bases in the third codon position of the longest ORF is 22.8% A, 22.5% C, 25.8% G and 28.5% U. These values indicate a lesser preference for U than
Fig. 1. Alignments of the nucleotide sequences of ArMV (lilac) RNA-3 (upper) and GFLV RNA-3 (lower). Bold and underlined nucleotides are sites recognized by HpaII; the single HindIII cleavage site is underlined. Spaces introduced to optimize the alignments are marked with hyphens (-); ▽ and * indicate the start and stop codons, respectively, of the longest ORF of ArMV RNA-3.
many plant viral RNAs (Dasgupta & Kaesberg, 1982; van Wezenbeek et al., 1983; Meyer et al., 1986).

Nucleotide sequence comparisons showed no significant homology between the ArMV satRNA and those of TBRV (Meyer et al., 1984), TobRV (Buzayan et al., 1986), ArMV hop isolate (Kaper et al., 1988), or cucumber mosaic virus (CARNA5; Collmer et al., 1983).

It was surprising to discover that the ArMV satRNA shares 83% of its nucleotide sequence with that of the satRNA of GFLV (Fuchs et al., 1989); to make the best possible fit in the alignments between the nucleotide sequences of these two satRNAs (Fig. 1) only four breaks are required. With regard to the fact that these two satRNA sequences are so similar, two of these breaks (in regions A, at positions 338 and 339 and positions 360 to 366) could be treated as self-correcting ‘frameshift’ mutations resulting in a loss of three amino acids to the coding potential of ArMV satRNA. The possible ‘frameshift’ area A* (at positions 108 to 118) in the alignments between the deduced amino acid sequences of these two satRNAs (Fig. 2) shows this effect, with only two amino acid identities maintained, and a three amino acid break. Following this ‘frameshift’ area, there is a region of low amino acid sequence homology corresponding to amino acid positions 120 to 143. Although this region can be considered in-frame (nucleotide sequence homology of 73%), the amino acid sequence differs greatly (with 43% homology) and most of the differences are at the first and second base positions. The other two breaks (Fig. 1, in B and C) may be compound ‘frameshift’ alterations which completely change the amino acid sequence at the C terminus of the predicted proteins (Fig. 2, B* and C*).

This work was done while D. J. Bertioli was in receipt of a NERC Studentship and in part under a contract (PECD/7/8/106) with the Department of the Environment which has authorized publication. We are also grateful to Susan Jackson for text capture.

References


Nucleotide sequence of a nepovirus satRNA


(Received 14 November 1989; Accepted 6 February 1990)