Two novel subgenomic RNAs derived from RNA 3 of tomato aspermy cucumovirus

Bu-jun Shi, Shou-wei Ding† and Robert H. Symons

Department of Plant Science, Waite Institute, University of Adelaide, Glen Osmond, SA 5064, Australia

Two abundant subgenomic RNAs, designated RNA 3B and RNA 5, were found to be associated with the V strain of tomato aspermy cucumovirus (V-TAV). Sequence determination showed that the 3′-terminal 323 nucleotides (nt) of RNA 3B was identical to RNA 5, whereas its 5′-terminal 163 nt was a direct repeat (one nt difference) of the 5′-half of RNA 5, and that both RNAs are completely homologous to the 3′-terminal untranslated region of TAV RNA 3. TAV RNAs 3B and 5 were also detected in the infection of a pseudorecombinant virus consisting of TAV RNA 3 and RNAs 1 and 2 from cucumber mosaic virus. Furthermore, only RNA 5, not RNA 3B, was detected in a TAV mutant in which one of the repeats was deleted from RNA 3. These genetic studies clearly show that both RNA species are derived from TAV RNA 3. However, in contrast to TAV RNAs 4 and 4A, which encode coat protein and 2b protein, respectively, RNAs 3B and 5 represent a novel class of subgenomic RNAs from TAV that do not function as mRNAs. Possible functional roles for such a class of viral subgenomic RNAs are discussed.

Introduction

Tomato aspermy cucumovirus (TAV) is an important agent of chrysanthemum disease throughout the world (Kaper & Waterworth, 1981; Shi et al., 1993). TAV contains a single-stranded RNA genome of messenger-sense divided into three species, RNAs 1, 2 and 3 (Murphy et al., 1995). The complete genome sequence of the V strain (V-TAV) (Moriones et al., 1991; Bernal et al., 1991; F. García-Arenal, personal communication), the nucleotide sequences of RNA 3 of the C, P and B strains (O’Reilly et al., 1991, 1994; Salanki et al., 1994), as well as the 3′ termini of the V and N strains (Wilson & Symons, 1981) have been determined. As expected, the genomic RNAs of TAV would express functions similar to those found for the type member of the cucumovirus group, cucumber mosaic virus (CMV) (Palukaitis et al., 1992).

Recently, we demonstrated that RNA 2 of TAV encodes an overlapping gene (2b) in addition to the 2a gene (Shi et al., 1996). Gene 2b is most likely expressed through a subgenomic RNA (sgRNA) derived from RNA 2, which we have called, on a size basis, RNA 4A, because it is equivalent to RNA 4A of CMV (Peden & Symons, 1973; Ding et al., 1994; Shi et al., 1996). The 2b protein of CMV has been demonstrated to be essential for long-distance virus movement and for the expression of systemic symptoms (Ding et al., 1995a).

In this study, we report the isolation and characterization of two non-translated sgRNAs of TAV which are derived from RNA 3 and which we have designated RNA 3B and RNA 5. RNA 3B has two 5′-terminal tandem repeats of 163 nt whereas RNA 5 is exactly one repeat shorter.

Methods

\- Virus and infectious cDNA clones. V-TAV (Habili & Francki, 1974a) was propagated in Nicotiana glutinosa and purified as described by Peden & Symons (1973). RNA 3B and RNA 5 were purified from virion RNAs by PAGE as described by Symons (1978). Infectious full-length cDNA clones of each of the three genomic RNAs of either the Q strain of CMV (Q-CMV) (Ding et al., 1995b) or V-TAV (designated pCassIT1, pCassIT2 and pCassIT3, which correspond to RNA 1, RNA 2 and RNA 3 of V-TAV, respectively; unpublished results) were used and inoculated on N. glutinosa and N. clevelandii, as described by Ding et al. (1995b). Total RNAs were extracted from plants as described by Verwoerd et al. (1989).

\- Sequence determination. Purified RNA 3B and RNA 5 were either 5′ end-labelled using [γ-32P]ATP and T4 polynucleotide kinase,

Author for correspondence: Bob Symons.
Fax +61 8 8303 7102. e-mail bsymons@waite.adelaide.edu.au

† Present address: Institute of Molecular Agrobiology, National University of Singapore, 59A, The Fleming, Science Park Drive, Singapore 118240.

0001-4329 © 1997 SGM
following dephosphorylation, or 3’ end-labelled with [5’-32P]pCp (England et al., 1980) and then sequenced by partial enzymatic digestion as described by Haseloff & Symons (1981) and Forster et al. (1990).

The rapid amplification of cDNA ends (RACE) technique (Frohman et al., 1988) was used for synthesis of cDNAs to RNA 3B and RNA 5. Purified RNA 3B and RNA 5 were polyadenylated with Escherichia coli poly(A) polymerase and then reverse transcribed with an oligo(dT) primer [5’ GACTCGAGTGCACATCGA(T)17 3’]. The resulting cDNAs were 3’ dA-tailed using terminal deoxynucleotidyl transferase, then PCR-amplified with the same oligo(dT) primer. The amplified product was cloned into the SmaI site of pBluescript SK(+) (Stratagene) and sequenced by partial enzymatic digestion following dephosphorylation, or 3’ end-labelled with [5’-32P]pCp (England et al., 1980) and then sequenced by partial enzymatic digestion as described by Haseloff & Symons (1981) and Forster et al. (1990).

Direct dideoxynucleotide RNA sequencing and primer extension, in which no ddNTPs were added (Fichot & Girard, 1990), were carried out using virion RNAs, total V-TAV-infected plant RNAs, RNA 3B and RNA 5 using an oligonucleotide primer complementary to a known sequence at the 5’-terminal region of RNA 3B and RNA 5 obtained as above.

Pseudorecombinants between V-TAV and Q-CMV. Infectious cDNA clones of RNA 3 of V-TAV and Q-CMV were exchanged between infectious cDNA 1 and 2 clones of V-TAV and Q-CMV to make two different pseudorecombinant plasmid mixtures, C1C2T3 and T1T2C3, in which numbers 1, 2 and 3 correspond to genomic RNAs 1, 2 and 3 of each virus and C and T correspond to Q-CMV and V-TAV, respectively. The mixtures were then inoculated directly onto two common host plants, N. glutinosa and N. clevelandii.

Deletion cDNA clone of V-TAV RNA 3. Deletion of one copy of the 163 nt tandem repeat in V-TAV RNA 3 was constructed by sequentially digesting pCassIT3 with specific restriction enzymes, followed by religation and confirmation by dideoxynucleotide sequencing. This deletion construct was designated pCassIT3·163.

Results

RNA 3B and RNA 5 are both associated with V-TAV

RNAs 1, 2 and 3 (genomic RNAs) and RNA 4 (a subgenomic RNA of RNA 3 and the mRNA for coat protein) of V-TAV have been reported (Habili & Francki, 1974; Moriones et al., 1991; Bernal et al., 1991; F. García-Arenal, personal communication). We have recently shown that TAV RNA 2 also has a subgenic RNA, RNA 4A (702 nt), which most likely functions as mRNA for a newly discovered overlapping gene (2b) whose translation product has been detected in infected plants (Shi et al., 1996).

As shown by Northern blot hybridization analysis (Fig. 1b), two additional smaller RNA species, designated RNA 3B and RNA 5 (see below), were found to be encapsidated in virions of V-TAV (lane 5). Both hybridized to an RNA probe (probe T) complementary to the 3’-terminal 100 residues of V-TAV RNA 3 (Fig. 1a), indicating that these two RNA species are of a TAV origin and that they most likely contain conserved 3’-terminal sequences. These two RNAs accumulated in TAV-infected plants (lane 4) to levels comparable to the known viral RNAs (lane 5). An RNA of similar size to TAV RNA 3B was detected in total plant RNAs (TR) of WAIL-CMV (a subgroup I strain) infected plants (Fig. 1b, lane 3). However, it was not detected in virion RNAs (VR) of WAIL-CMV (data not shown), in VR of Q-CMV (a subgroup II strain) (data not shown) or in TR of Q-CMV infected plants (Fig. 1b, lane 2) using an RNA probe (probe C in Fig. 1a) complementary to the 3’-terminal 495 residues of Q-CMV RNA 3. RNA 5 was detected in both VR of Q-CMV (data not shown) and TR of Q-CMV infected plants (Fig. 1b, lane 2) using probe C. In addition, TAV RNAs 3B and 5 were detected at early stages of virus infection (6 days after inoculation) in leaves inoculated with infectious cDNA clones of V-TAV (data not shown), indicating that these two RNAs may not be defective interfering (DI) RNAs, as the appearance of DI RNAs usually requires several passages at high multiplicity of infection (Hillman et al., 1987). Therefore,
Subgenomic RNAs derived from TAV RNA 3

Fig. 2. Complete nucleotide sequences of RNA 3B and RNA 5 and their alignment with the 3′-terminal sequence of V-TAV RNA 3 (kindly provided by F. García-Arenal) or the sequence deduced from our own cDNA clones (RNA 3#). The differences between the two RNA 3 sequences are indicated. Residue numbers in parentheses refer to the equivalent position in V-TAV RNA 3. Positions of identity are indicated by dots. A dash represents one missing nucleotide. The repeat sequences at the 5′-end of either RNA 3B or RNA 5 are underlined (first repeat with single underline and second repeat with double underlines). The two single nucleotide differences between the two repeats are indicated by an asterisk above. The lengths of RNA 3, RNA 3B and RNA 5 are given at the 3′-end. The full-length sequence of RNA 3# has yet to be determined.

it is most likely that RNAs 3B and 5 are natural subgenomic RNAs of TAV.

RNAs 3B and 5 are a novel class of subgenomic RNAs derived from RNA 3

To determine the nature and the molecular structure of RNAs 3B and 5, the complete sequences of both RNAs were determined after purification from virions by a combination of three sequencing strategies described in Methods, and are presented in Fig. 2. As the 5′-end-labelling of both RNAs required dephosphorylation but not de-capping, it is likely that both RNAs 3B and 5 are 5′-phosphorylated and not capped. RNA 3B is 486 nt long and RNA 5 is 323 nt long, which is identical to the 3′-terminal 323 nt of RNA 3B (Fig. 2). The 5′-terminal sequence (326 nt) of RNA 3B consists of two tandem repeats of 163 nt that differ only by a single nt (marked by an asterisk in Fig. 2) while RNA 5 contains only one repeat. Importantly, both RNAs encode no open reading frame (ORF) of 21 codons or longer.

The nucleotide sequences of RNAs 3B and 5 were further compared to those of V-TAV RNAs 1 and 2 (Moriones et al., 1991; Bernal et al., 1991) as well as to the unpublished sequence of RNA 3 of V-TAV (F. García-Arenal, personal communication). RNA 3B and RNA 5 showed 99.4% and 99.7% identity to the 3′-terminal 486 and 323 of RNA 3 of V-TAV, respectively. The difference between either RNA 3B or RNA 5 and the RNA 3 sequences include two substitutions (G1973 of RNA 3 to U1973 of RNA 3B; U1973 of RNA 3 to G1973 of RNA 3B) and one insertion (U1469 of RNA 3B or U1976 of RNA 5 between nucleotides 2367 and 2371 of RNA 3). These came up consistently by all three sequencing methods used. However, RNA 3B is identical to the 3′-terminal 486 nt, and RNA 5 to the 3′-terminal 323 nt, of V-TAV RNA 3 determined from our own four RNA 3 cDNA clones (B.-J. Shi and others, unpublished results). As RNA 3 contains a 3′-untranslated region of 502 nt which encodes no ORF of 21 codons or longer (F. García-Arenal, personal communication), we conclude that both RNAs 3B and 5 are derived completely from the 3′-untranslated region of RNA 3 and are unlikely to function as mRNAs. The lack of 5′-capping of RNAs 3B and 5 further supports a non-mRNA role for these RNAs.

RNAs 3B and 5 are derived from RNA 3 as shown by genetic studies

The following two sets of genetic studies further support the conclusion that both RNAs 3B and 5 are derived from RNA 3. Using infectious cDNA clones of the three genomic RNAs of V-TAV (B.-J. Shi and others, unpublished results) and Q-CMV (Ding et al., 1995a), two pseudorecombinant viruses, T1T2C3 and C1C2T3, were reconstituted by exchanging RNA 3 cDNA clones between the two viruses. Both pseudorecombinants were viable, as systemic virus infections were established in N. glutinosa plants, in agreement with earlier work using purified RNAs to prepare pseudorecombinants (Habili & Francki, 1974b). Total RNAs were extracted from systemically infected leaves and analysed by Northern blot hybridization using probes T and C specific for the 3′-conserved untranslated regions of V-TAV and Q-CMV, respectively (Fig. 1a). RNA 3B was detected in C1C2T3 infections (Fig. 3a, lane 2) and not in T1T2C3 infections (lane 3), demonstrating that TAV RNA 3 is the genetic source of RNA 3B.

In contrast to RNA 3B, RNA 5 of V-TAV was detected in infections of both pseudorecombinants (Fig. 3b, lanes 2 and 3); The specificity of probe T for the 3′-conserved untranslated region of V-TAV was demonstrated by the absence of hybridization with any of the six known RNAs of Q-CMV RNAs (lane 6). Hence, this indicates that V-TAV RNA 5 can be
Fig. 3. Northern blot analysis of total RNAs extracted from *N. glutinosa* inoculated with different inocula. Total RNAs extracted from *N. glutinosa* inoculated with a TAV mutant [T1T2T3(\[163\]) (lanes 1 and 8), C1C2T3 (lanes 2 and 9), T1T2C3 (lanes 3 and 10), WAII-CMV (lanes 4 and 11), V-TAV (lanes 5 and 12), Q-CMV (lanes 6 and 13) or sterile water (lanes 7 and 14) were electrophoresed and hybridized with probe T (lanes 1–7) or probe C (lanes 8–14). Probes T and C are shown in Fig. 1(a). The positions of viral RNAs 1, 2, 3, 4, 4A, 3B and 5 are indicated.

**Discussion**

Our sequence analysis showed that RNA 3B is 486 nt in length and contained two 5'-terminal tandem repeats of 163 nt, whereas RNA 5 is 323 nt in length and contained only one repeat, and that both RNAs are identical in sequence to the 3'-terminal 486 and 323 nt of RNA 3, respectively, and have no ability to encode any proteins. Interestingly, such tandem repeats are absent in strains B, C and P of TAV sequenced to date (O'Reilly *et al*., 1991, 1994; Salanki *et al*., 1994). Genetic studies demonstrated that these two novel sgRNAs, RNAs 3B and 5, are derived from TAV RNA 3. RNA 5 of V-TAV was also detected in infections of the pseudorecombinant (T1T2C3) suggesting that RNA 5 may be a mixed population derived from RNA 3 and RNAs 1 and/or 2 as found in CMV (Blanchard *et al*., 1996).

RNA 3B and RNA 5 have the same 5'-sequence, suggesting that they might have the same mechanism for generation. Using a radioactive probe complementary to the whole of RNA 3 except for the region common to RNA 3B, we detected only full-length RNAs 3 and 4 and could not detect any other fragments derived from RNA 3, even after long exposure (data not shown), indicating that RNA 3B and RNA 5 are not the result of an endonucleolytic cleavage of RNA 3 at a specific site. Hence, these subgenomic RNAs are presumably transcribed from the (-) RNA 3 sequence (Miller *et al*., 1985).

The promoter regions for sgRNA synthesis have been investigated with other members of the *Bromoviridae*: brome mosaic virus (BMV) (French & Ahlquist, 1987, 1988; Marsh *et al*., 1988; Smirnyagina *et al*., 1994), cowpea chlorotic mottle virus (Allison *et al*., 1989; Pacha & Ahlquist, 1992), alfalfa mosaic virus (AlMV) (Van der Kuyl *et al*., 1990, 1991; Van der Vossen *et al*., 1995) and CMV (Boccard & Baulcombe, 1993). Most of these promoter regions have been shown to lie predominantly upstream of the transcription initiation site (e.g., of BMV and CMV sgRNA 4).
As RNAs 3B and 5 have the same initiation sequence, they may have the same promoter-like sequences. However, RNA 3 does not have a third copy of the 163 nt and thereby the sequence upstream of RNA 3B on RNA 3 is different from that upstream of RNA 5 on RNA 3, suggesting that the organization of the RNA 3 promoters for sgRNA 3B and RNA 5 might be, like the beet necrotic yellow vein virus RNA 3 sgRNA promoter (Balmori et al., 1993), situated downstream, and within, each repeat of RNAs 3B and 5 on RNA 3.

The mechanism of generation of sequence repeats remains unresolved. TAV RNA 3, as well as RNA 3B, contains two tandem repeats of 163 nt, suggesting an association with replication slippage (Hancock et al., 1995). Although internal in these tandem repeats of RNA 3, a single nucleotide difference in two repeats, A1966 and G2139 (Fig. 2), could arise by a copy error of an RNA polymerase replicating a single RNA species (Keese et al., 1988). However, due to the sequence from residue C2534 to A2821 of RNA 2 (Moriones et al., 1991), or from U3131 to A3161 of V-TAV RNA 1 (Bernal et al., 1991) corresponding to the region around A1966 in the first repeat of V-TAV RNA 3, the two tandem repeats of RNA 3 could also have arisen by RNA polymerase jumping (Keese & Symons, 1985, 1987) from the primary template RNA 3 to the nascent complementary RNA 1 or RNA 2. Such a replicate-mediated copy choice-type model has been proposed to explain the generation of DI RNAs (Pogany et al., 1995).

Direct repeats have been investigated in other plant RNA viruses (Ding et al., 1989). In ALMV, a tandem repeat at the 5' untranslated region of RNA 3 affected accumulation of P3 protein (Van der Vossen et al., 1993) and may be involved with ribosome binding (Pinck et al., 1981). In ononis yellow mosaic tymovirus, the repeat of the 5' untranslated region has been proposed to form a stem-loop structure to modulate initiation of translation of the overlapping tymoviral genes (Ding et al., 1989).

In our study, a TAV mutant in which one of the repeats was precisely removed from RNA 3 did not apparently affect the symptoms and host range compared to wild-type TAV (data not shown), indicating that the repeat may not be of major biological significance. In addition, the TAV mutant was stable in the presence of the wild-type TAV (data not shown), suggesting that the repeat did not interact with the other viral RNAs. However, it is feasible that the repeat sequence may provide some selective advantage under specific host and/or environmental conditions.

Sequence analysis showed that neither RNA 3B nor RNA 5 contain any ORFs longer than 63 nt. It is unlikely that these two RNAs encode any protein although they have been clearly shown to be genuine sgRNAs. A similar sgRNA with no coding ability has been found in barley yellow dwarf luteovirus (Kelly et al., 1994). Pseudorecombinant experiments showed that the level of accumulation of RNA 3B is higher than that of RNA 4 in C1C2T3 (Fig. 3 b, lane 2) while it is lower than that of RNA 4 in wild-type V-TAV (Fig. 3 b, lane 5). Combined with the fact that RNAs 3B and 5 are abundantly transcribed during infection, this leads us to speculate that both RNAs might have a regulatory role in viral RNA replication.

The authors wish to thank Dr F. García-Arenal for kindly providing the full-length RNA 3 sequence of V-TAV, Wanxiang Li for technical help, Neil Shirley and Jing Li for oligonucleotide synthesis and David Warrillow for providing an oligo(dT) primer. This work was supported by the Australian Research Council Special Research Centre for Basic and Applied Plant Molecular Biology. B.-J.S. is the recipient of a University of Adelaide Overseas Postgraduate Research Scholarship.

References


Received 8 August 1996; Accepted 13 November 1996