Nucleotide sequence of tomato aspermy virus RNA 2

Enrique Moriones, Marilyn J. Roossinck† and Fernando García-Arenal*

Departamento de Patología Vegetal, ETSI Agrónomos, Ciudad Universitaria, 28040 Madrid, Spain

RNA 2 of the V strain of tomato aspermy virus (TAV) consists of 3074 nucleotides and contains one open reading frame of 2487 nucleotides. Thus, it resembles RNA 2 of cucumber mosaic virus (CMV) strains Q and Fny (62% identical to both), brome mosaic virus (42% identical) and cowpea chlorotic mottle virus (40% identical). In comparisons between amino acid sequences, three different regions of similarity could be distinguished. These were the central part (amino acids 224 to 757 for V-TAV), which was most similar among the four viruses, and the N and C ends; sequences conserved among RNA polymerase species were found in the C half of the central part. Hydrophobicity patterns, and distributions of acidic and basic amino acids in the proteins encoded by V-TAV RNA 2, Q-CMV RNA 2 and Fny-CMV RNA 2 were very similar except at the extreme ends of the molecules. Structures that have been reported to act as regulatory signals for minus- and plus-strand synthesis were found in the 5' and 3' non-coding regions of the RNA.

Introduction

Tomato aspermy virus (TAV) (cucumovirus group) resembles cucumber mosaic virus (CMV) in having a tripartite, plus-sense RNA genome, but differs from it in physicochemical properties (Habili & Francki, 1974a, b), in host range (Marrou et al., 1975) and serologically (Habili & Francki, 1974a; Devergne & Cardin, 1975). However TAV is able to support the replication of the satellite RNAs of CMV (CMV satRNA), although TAV differs from CMV in its properties as a helper (Mossop & Francki, 1979; our unpublished results); and pseudorecombinants can be formed between TAV and CMV RNAs (Habili & Francki, 1974c). Sequence data for TAV RNAs are limited to the 190 5'-terminal nucleotides (Wilson & Symons, 1981). Here we report the complete nucleotide sequence of the genomic RNA 2 of TAV, and compare it to that of CMV RNA 2.

Methods

Virus. The V strain of TAV (V-TAV), from Chrysanthemum sp. (Habili & Francki, 1974a), was obtained from R. I. B. Francki (Waite Agricultural Research Institute, Adelaide, Australia). The virus was propagated in, and purified from, Nicotiana clevelandii Gray; virions were purified, and RNAs were extracted as described by Habili & Francki (1974a).

† Present address: Department of Plant Pathology, Cornell University, Ithaca, New York 14853, U.S.A.
Fig. 1. Nucleotide sequence of V-TAV RNA 2 and encoded amino acid sequence for the long ORF. The initiation and stop codons are underlined.
clones in M13. The complete sequence of both the viral encapsidated plus strand and the minus strand was determined; each nucleotide was determined an average of four times.

**Genetic organization**

The genomic RNA 2 of TAV was 3074 nucleotides (nt) long (Fig. 1) and had an overall base composition of 31% U, 26-3% A, 21-6% C and 21-1% G. Computer analysis revealed one long open reading frame (ORF) in the plus strand, starting at the first AUG (positions 88 to 90) codon and extending to position 2574. Its putative translation product (P2a), 828 amino acid residues long, was also found for the RNA 2 of Q- and Fny-CMV (Rizzo & Palukaitis, 1988). Nucleotide sequence identities over the whole RNA 2 molecule with other tripartite genome viruses were, with the cucumoviruses Q-CMV (Rezaian et al., 1984) 62% and Fny-CMV (Rizzo & Palukaitis, 1988) 59%, and with the bromoviruses brome mosaic virus (BMV) (Ahquist et al., 1984) 42% and cowpea chlorotic mottle virus (CCMV) (Allison et al., 1989) 40%.

**Non-coding regions and regulatory signals**

TAV RNA 2 extracted from virions could only be labelled at the 5' end with [γ-32P]ATP and polynucleotide kinase after treatment with NaIO4, aniline and alkaline phosphatase (Dasgupta et al., 1976). This indicates that a cap site is present at the 5' end.

The 5', pyrimidine-rich (U+C=66-6%) leader sequence extended 87 nt up to the beginning of the P2a ORF. The leader was similar in length to leader sequences of Q- and Fny-CMV RNA 2, and very similar (91 to 93%) to them in the 5' extreme 44 nt. In these 44 nt (Fig. 2) a sequence corresponding in part (9/11) to the B box of the internal control regions of tRNA promoters (ICR2) was found for TAV as reported for BMV, CMV (Marsh et al., 1989) and CCMV (Allison et al., 1989). Mutational analysis of ICR2 in BMV has shown it to act as a promoter for plus-strand synthesis (Pogue et al., 1990), a role that could be extended to both CMV and TAV. In CMV, the ICR2 would be part of a sequence able to form a hairpin structure; this sequence has been proposed to interact with complementary sequences in CMV sat-RNAs (Fig. 2) (Rezaian et al., 1985), and it was suggested that this interaction might have a role in the regulation of the replication of viral and satRNAs. The presence of CMV satRNAs is associated with a general depression in the synthesis of CMV RNA, in particular of RNAs 1 and 2 more than RNAs 3 and 4 (Kaper & Tousignant, 1977; Mossop & Francki, 1979). However CMV and TAV differ as helpers of CMV satRNA in that the presence of CMV satRNA is not associated with a differential depression of TAV RNAs 1 and 2, nor indeed with a general depression of TAV RNA synthesis (Harrison et al., 1987; Mossop & Francki, 1979; our unpublished results). The interaction of CMV satRNA with the ICR2 motif-including sequence would be the same for TAV RNA 2 and for CMV RNA 2 (Fig. 2) and thus, this putative interaction does not explain the effect of CMV satRNA on the replication of CMV, for which other, unidentified, regulatory signal(s) must be responsible.

The 3' non-coding region of TAV RNA 2 was 500 nt long, considerably longer than for Q- and Fny-CMV RNA 2. This was due to the presence of a stop codon at position 2572. The sequence of the 3' extreme 190 nt corresponds to what has been reported for V-TAV RNA 1 (Wilson & Symons, 1981) except for the change G→U at nt 2737. It has been shown that the 3' extreme 149 nt of TAV RNA can be folded into a tRNA-like structure (Joshi & Haenni, 1986) like those of CMV and BMV RNAs (Joshi et al., 1983; Rietveld et al., 1983) in spite of relatively low sequence similarity (67% TAV/Q-CMV, 46% TAV/Fny-CMV, 62% Q-CMV/Fny-CMV). These structural similarities suggest that the 3' end of TAV RNA 2 has a role in the regulation of minus-strand synthesis as shown for the 3' end of BMV RNA (Bujarski et al., 1986; Dreher & Hall, 1988; Dreher et al., 1984). The sequence homologies for the 3' non-coding regions upstream of the tRNA-like structure in TAV, Q-CMV and Fny-CMV RNAs were similar among the RNA species of the three viruses (62% TAV/Q-CMV, 60% TAV/Fny-CMV, 64% Q-CMV/Fny-CMV).

**The TAV RNA 2 translation product**

*In vitro* translation, in the rabbit reticulocyte lysate cell-
Table 1. Similarities between V-TAV RNA 2-encoded protein and Q-CMV, Fny-CMV, BMV and CCMV RNA 2-encoded proteins

<table>
<thead>
<tr>
<th>Region</th>
<th>Amino terminus</th>
<th>Central</th>
<th>Carboxy terminus</th>
</tr>
</thead>
<tbody>
<tr>
<td>V-TAV</td>
<td>1–223</td>
<td>224–757</td>
<td>758–828</td>
</tr>
<tr>
<td>Q-CMV</td>
<td>1–215</td>
<td>216–749</td>
<td>750–839</td>
</tr>
<tr>
<td>Fny-CMV</td>
<td>1–218</td>
<td>219–753</td>
<td>754–857</td>
</tr>
<tr>
<td>BMV</td>
<td>1–182</td>
<td>183–724</td>
<td>725–822</td>
</tr>
<tr>
<td>CCMV</td>
<td>1–176</td>
<td>177–713</td>
<td>714–808</td>
</tr>
<tr>
<td>Similarity (%)</td>
<td>Q-CMV 37</td>
<td>71</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>Fny-CMV 35</td>
<td>71</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>BMV 21</td>
<td>43</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>CCMV 19</td>
<td>46</td>
<td>8</td>
</tr>
</tbody>
</table>

free system, of V-TAV RNAs 1 and 2 yielded a product of the expected size for the putative protein encoded by V-TAV RNA 2 (A. Fraile, unpublished results). The initiation codon for this putative translation product was not in the optimal sequence context for plant mRNA (Lütcke et al., 1987) according to the ribosome-scanning model (Kozak, 1986), as is also the case for CMV RNA 2. Notably, the putatively important G at position +4 was substituted for a U. The effect, if any, of this suboptimal context in the regulation of the translation of TAV RNA 2 remains to be investigated, but highly expressed plant viral mRNAs [i.e. for the coat protein of tobamoviruses (Lehto & Dawson, 1990; Solis & Garcia-Arenal, 1990)] also have a U at this position. For the ORF of the 2a protein, codon usage deviated from random, and codons other than those for Leu, Ile and Val mostly had a U or an A at the third position.

TAV P2a was 11 and 29 amino acids shorter than those of Q-CMV and Fny-CMV RNA 2 (Q-CMV P2a, Fny-CMV P2a), respectively, due to differences at the C ends of the molecule; its sequence (Fig. 1) showed an overall similarity of 57% with both of them. Similarity was 37% with the product of BMV RNA 2 (BMV P2a) and of CCMV RNA 2 (CCMV P2a), and no significant similarity was found with the product of alfalfa mosaic virus RNA 2 (AIMV P2a) (Cornelissen et al., 1983). Matrix analyses of amino acid sequences showed homology to be unevenly distributed in the molecule, being maximal for a central part between amino acids 224 and 757 (759 to 2358 nt) of TAV P2a (Table 1). In the C half of this conserved central region of P2a four consensus motifs of polymerases were found, including that considered to be the nucleus of the polymerase activity (KLIFSGDDSLG, nt 1909 to 1941) (Argos, 1988; Kamer & Argos, 1984). Divergence was highest for the C-terminal part of the molecule for which no significant homologies were found. Regardless of the different degree of similarity shown for the three different regions (Table 1) in TAV, Q-CMV and Fny-CMV P2a, hydrophobicity patterns for all three P2as were extremely similar except for the 75 N-terminal and for the 45 (TAV) to 80 (Fny-CMV) C-terminal residues (not shown), indicating conservative changes for variable positions. Also, the distribution of acidic and basic amino acids is the same for TAV, Q-CMV and Fny-CMV P2as. As shown for Q-CMV (Rezaian et al., 1984) acidic amino acids predominate at the N terminus, basic amino acids predominate at the C terminus, and neither type predominates in the central conserved part.

In addition to the presence of consensus motifs for polymerases, a role for CMV and BMV P2a in viral RNA replication is substantiated by the ability of CMV and BMV RNAs 1 and 2 to replicate in protoplasts (Kiberstis et al., 1981; Nitta et al., 1988) and by the presence of CMV and BMV P2a in the RNA polymerase complexes (Hayes & Buck, 1990; Horikoshi et al., 1988; Quadt et al., 1988), and a similar function could be attributed to TAV P2a. The high degree of similarity found for TAV, Q-CMV and Fny-CMV P2a agreed with reported similarities of these viruses in properties related to RNA replication such as (i) TAV may be a helper of CMV satRNAs (Gould et al., 1978; Mossop & Francki, 1979) and differences between V-TAV, Q-CMV and Fny-CMV as helpers of CMV satRNAs cannot be associated with sequence features in 2a proteins; (ii) TAV RNAs 1 and 2 are able to replicate with CMV RNA 3, and CMV RNAs 1 and 2 are able to replicate with TAV RNA 3 in stable pseudorecombinants (Habili & Francki, 1974c; Mossop & Francki, 1979; our unpublished results). On the other hand, the heterologous exchange of CMV and TAV RNAs 1 and 2 is not permitted. For BMV and CCMV it has been shown that the less conserved N- and C-terminal regions of the P2as are involved in virus-specific interactions needed for RNA replication (Allison et al., 1989). A similar phenomenon may explain why viable CMV–TAV pseudorecombinants exchanging RNAs 1 and 2 have not been obtained.

We want to thank R. J. B. Francki for providing V-TAV, and Aurora Fraile for helpful discussions. This work was in part supported by grant PA 86-0353, Comisión Interministerial de Ciencia y Tecnología, Spain. E. Moriones was the recipient of a Formación de Personal Investigador fellowship, Ministerio de Educación y Ciencia, Spain. M. J. Roossinck was supported by grant CCB-84-017 from the US-Spain Joint Committee for Scientific and Technological Cooperation.

References
