cDNA cloning and sequence analysis of the 3'-terminal region of zucchini yellow mosaic virus RNA

Rebecca Grumet* and Guowei Fang

Department of Horticulture, Michigan State University, East Lansing, Michigan 48824, U.S.A.

The 3' half of the RNA of the cucurbit potyvirus zucchini yellow mosaic virus (ZYMV) was genetically cloned and the cDNA sequence of a portion of the putative RNA polymerase gene, the complete coat protein gene and the 3' untranslated region was determined. The coat protein and putative polymerase genes are both part of a continuous open reading frame as is the case for other potyviruses whose genomes are expressed as polyproteins. The Q/S protease cleavage site for the N terminus of the coat protein was deduced by alignment of the coat protein and polymerase genes with other potyviral sequences. The resulting protein has 279 amino acids and a calculated Mr of 31 214. The predicted amino acid sequence indicates a ZYMV-unique N-terminal region and potyvirus-characteristic central and C-terminal regions. These data also verify that ZYMV is distinct from the cucurbit potyvirus watermelon mosaic virus 2.

Cucurbit crops are subject to severe losses due to a relatively new and highly aggressive member of the potyvirus group, zucchini yellow mosaic virus (ZYMV). This virus was first described in Italy and France in 1981 (Lisa et al., 1981), since when it has spread throughout the world (Lisa & Lecoq, 1984). It is capable of outcompeting other cucurbit potyviruses in establishing infections in rub inoculation (Davis & Mizuki, 1987), aphid acquisition (Lecoq & Pitrat, 1985) and field studies (Alderz, 1987). Several new strains have been identified including one capable of overcoming resistance to the dominant Zym gene in muskmelon (Lisa & Lecoq, 1984).

In the past 3 years there have been several demonstrations of genetically engineered plant virus resistance utilizing viral capsid genes as a novel source of resistance genes (reviews: Baulcombe, 1989; Grumet, 1990). As a first step towards determining whether the coat protein of ZYMV can be used to engineer resistance genetically, we cloned approximately half of the ZYMV genome (5-5 kb) and sequenced part of the putative RNA polymerase gene, the complete coat protein gene and the 3' untranslated region.

The Connecticut strain of ZYMV was isolated from infected zucchini (Cucurbita pepo cv. Black Jack) plants 3 to 5 weeks after rub inoculation using the papaya ringspot virus purification method 1 (Parcifull et al., 1984b) with the following changes. Prior to the first PEG (polyethylene glycol 8000) precipitation the virions were solubilized with 1% (v/v) Triton X-100 for 20 min. After the second PEG precipitation the resuspended virus was layered onto a step gradient of 0 to 1·2 m-cæsium sulphate in PE (0·1 m-potassium phosphate pH 7, 10 mM-EDTA) buffer with 30% sucrose and centrifuged in a Beckman SW28 rotor for 4 h at 25 000 r.p.m. The virus band was diluted with 3 vol. of PE, reprecipitated with 5% PEG and 0·3 M-NaCl, and resuspended in a minimal volume of PE with 1 mM-sodium azide.

The yield of virions was 1 to 5 mg/100 g fresh weight with an absorbance ratio (260/280 nm) of 1·2 to 1·35; the preparations were infective. SDS-PAGE gave a single Coomassie blue-stained protein band with an estimated Mr of 32700 (calculated from three preparations, data not shown).

Viral RNA was prepared using ammonium carbonate denaturation and linear (7·5 to 30%) sucrose gradients based on the protocols of Yeh & Gonsalves (1985) and Hellmann et al. (1980), respectively. The viral RNA was predominantly full length and was located in the first 4 ml from the bottom. cDNA cloning was based on the method of Gubler & Hoffman (1983) using oligo(dT) primers and Amersham supplies. After second strand synthesis, procedures were as described by Maniatis et al. (1982). The cDNA was extracted in phenol:chloroform:isoamyl alcohol (25:24:1), ethanol-precipitated, blunt end-ligated into phosphatase-treated (bacterial alkaline phosphatase), SmaI-cut Bluescript KS+ plasmid (Strategene) and used to transform competent DHz Escherichia coli cells. cDNA clones of 0·3 to 5·5 kb were obtained.

Since coat protein genes of potyviruses are located at the extreme 3' end of the viral RNA (Dougherty &
Carrington, 1988), the approx. 1.55 kb clone 187 was estimated to be sufficiently long for sequencing of the coat protein gene. Nested deletions were made in both directions according to the Stratengene protocol based on Yanisch-Perron et al. (1985); eight overlapping clones were chosen for each strand. Dideoxynucleotide sequencing was performed according to Sanger (1981) using [35S]dCTP and single-stranded DNA was prepared according to Vieira & Messing (1987). Sequence analyses were performed using the DNASIS program (Hitachi Software Engineering) and sequence data from the following potyviruses: watermelon mosaic virus 2 (WMV-2; Yu et al., 1989), tobacco etch virus (TEV; Allison et al., 1985b, 1986), tobacco vein mottling virus (TVMV; Domier et al., 1986), potato virus Y (PVY; Shukla et al., 1986; van der Vlugt et al., 1989), plum pox virus (PPV; Ravelonandro et al., 1986; Maiss et al., 1989) and Johnson grass mosaic virus (JGMV; Gough et al., 1987).

The cDNA sequence including a portion of the putative polymerase gene, the full coat protein gene and the 3' untranslated region [excluding the poly(A) tail] is shown in Fig. 1 along with the predicted amino acid sequence. The 211 nucleotide 3' untranslated region is within the size range of other potyviral untranslated regions. However, the possible yeast polyadenylation signal, TATGT, has also been noted for TVMV (Domier et al., 1986; Dougherty & Carrington, 1988), the approx. 1.55 kb clone 187 was nested according to Sanger (1981) using [35S]dCTP and single-stranded DNA was prepared according to Vieira & Messing (1987). Sequence analyses were performed using the DNASIS program (Hitachi Software Engineering) and sequence data from the following potyviruses: watermelon mosaic virus 2 (WMV-2; Yu et al., 1989), tobacco etch virus (TEV; Allison et al., 1985b, 1986), tobacco vein mottling virus (TVMV; Domier et al., 1986), potato virus Y (PVY; Shukla et al., 1986; van der Vlugt et al., 1989), plum pox virus (PPV; Ravelonandro et al., 1986; Maiss et al., 1989) and Johnson grass mosaic virus (JGMV; Gough et al., 1987).

Since potyviral genomes are expressed as polyproteins, the full coat protein gene and the 3' untranslated region [excluding the poly(A) tail] is shown in Fig. 1 along with the predicted amino acid sequence. The 211 nucleotide 3' untranslated region is within the size range of other potyviral untranslated regions. However, the possible yeast polyadenylation signal, TATGT, has also been noted for TVMV (Domier et al., 1986; Dougherty & Carrington, 1988), the approx. 1.55 kb clone 187 was nested according to Sanger (1981) using [35S]dCTP and single-stranded DNA was prepared according to Vieira & Messing (1987). Sequence analyses were performed using the DNASIS program (Hitachi Software Engineering) and sequence data from the following potyviruses: watermelon mosaic virus 2 (WMV-2; Yu et al., 1989), tobacco etch virus (TEV; Allison et al., 1985b, 1986), tobacco vein mottling virus (TVMV; Domier et al., 1986), potato virus Y (PVY; Shukla et al., 1986; van der Vlugt et al., 1989), plum pox virus (PPV; Ravelonandro et al., 1986; Maiss et al., 1989) and Johnson grass mosaic virus (JGMV; Gough et al., 1987).

The cDNA sequence including a portion of the putative polymerase gene, the full coat protein gene and the 3' untranslated region [excluding the poly(A) tail] is shown in Fig. 1 along with the predicted amino acid sequence. The 211 nucleotide 3' untranslated region is within the size range of other potyviral untranslated regions. However, the possible yeast polyadenylation signal, TATGT, has also been noted for TVMV (Domier et al., 1986; Dougherty & Carrington, 1988), the approx. 1.55 kb clone 187 was nested according to Sanger (1981) using [35S]dCTP and single-stranded DNA was prepared according to Vieira & Messing (1987). Sequence analyses were performed using the DNASIS program (Hitachi Software Engineering) and sequence data from the following potyviruses: watermelon mosaic virus 2 (WMV-2; Yu et al., 1989), tobacco etch virus (TEV; Allison et al., 1985b, 1986), tobacco vein mottling virus (TVMV; Domier et al., 1986), potato virus Y (PVY; Shukla et al., 1986; van der Vlugt et al., 1989), plum pox virus (PPV; Ravelonandro et al., 1986; Maiss et al., 1989) and Johnson grass mosaic virus (JGMV; Gough et al., 1987).

Since potyviral genomes are expressed as polyproteins, the gene products begin and end at protease cleavage sites (review: Dougherty & Carrington, 1988). The proposed N-terminal amino acid based on coat protein size, possible protease cut sites [Q/G, Q/S or Q/A (Domier et al., 1986; Dougherty & Carrington, 1988)] and the alignment of the putative polymerase gene is underlined and marked with a slash. This results in a protein 279 amino acids long, with a predicted Mr of 31214. Consistent with expression of the potyviral genome as a polyprotein, a single open reading frame continues for the length of clone no. 187, at least 170 amino acids upstream of the putative cut site for the N terminus of the coat protein.

It was not possible directly to determine the N-terminal amino acid of the ZYMV coat protein by chemical sequencing, presumably because the protein was N-terminally blocked. This problem has been observed for several other potyviral coat proteins (Shukla et al., 1988; Allison et al., 1985a; Domier et al., 1986), particularly those beginning with S rather than G or A (Shukla et al., 1988). The predicted Mr (approx. 31200) obtained from the proposed Q/S site is consistent with the value

**Fig. 1.** The cDNA sequence and predicted amino acid sequence of the terminal 3' 1546 nucleotides of ZYMV. The proposed polymerase–coat protein protease cut site is underlined and marked with a slash. The beginning of the conserved trypsin-resistant core protein is marked with a backslash. Possible polyadenylation signals are underlined.
The predicted amino acid sequence of the ZYMV coat protein was compared with the published amino acid sequence data, examination of the polymerase gene also aided in assigning the protease cleavage site. Sequence data from the adjoining putative polymerase gene were aligned with available sequence data from other potyviral polymerase genes beginning at the highly conserved amino acid span, GDD, which is found in virus-encoded polymerases regardless of their plant, animal or bacterial origin (Kamer & Argos, 1984). This alignment also resulted in assignment of the Q/S dipeptide as the ZYMV polymerase-coat protein junction (data not shown).

The central part of the protein was calculated to begin at the trypsin-resistant core as marked in Fig. 1.

Table 1. Percentage amino acid sequence homology between ZYMV coat protein and other potyvirus coat proteins

<table>
<thead>
<tr>
<th>Virus*</th>
<th>Overall</th>
<th>N terminus</th>
<th>Central and C terminus†</th>
</tr>
</thead>
<tbody>
<tr>
<td>WMV-2</td>
<td>67.1</td>
<td>11.6</td>
<td>76.9</td>
</tr>
<tr>
<td>PYV-D</td>
<td>59.2</td>
<td>7.0</td>
<td>67.2</td>
</tr>
<tr>
<td>TEV-HAT</td>
<td>58.8</td>
<td>4.7</td>
<td>67.7</td>
</tr>
<tr>
<td>JGMV</td>
<td>55.0</td>
<td>7.6</td>
<td>65.9</td>
</tr>
<tr>
<td>TYMV</td>
<td>50.5</td>
<td>2.3</td>
<td>56.8</td>
</tr>
<tr>
<td>PPV</td>
<td>47.5</td>
<td>6.5</td>
<td>61.1</td>
</tr>
</tbody>
</table>

* Abbreviations and sources of sequence data are listed in text.
† The trypsin-resistant core as marked in Fig. 1.

We thank Dr R. Davis for generously supplying the inoculum, Dr K. Zagula for advice with the virus purification protocol and Sue Hammar for technical assistance. This work was in part supported by the Michigan Agriculture Experiment Station, and by grants to R. Grumet from the Michigan State Biotechnology Research Center and the Michigan Research Excellence Fund. This is Michigan State University Agriculture Experiment Station Journal Article no. 13171.

References


DOMIER, L. L., FRANKLIN, K. M., SHAHABUDDIN, M., HELLMANN, G. M., OVERTMEYER, J. H., HIREMATH, S. T., SIW, M. F. E.,


(Received 2 January 1990; Accepted 23 March 1990)