Sequence and Antigenic Activity of the Region 93 to 113 of the Coat Protein of Strain U2 of Tobacco Mosaic Virus

(Accepted 3 September 1980)

SUMMARY

The amino acid sequence of the coat protein of the U2 strain of tobacco mosaic virus (TMV) has been re-examined and completed. Four incorrect residue allocations in the published U2 sequence were identified. These were located in the region corresponding to residues 96 to 105. In addition, the identity of residues 106 to 112 was also determined. This latter region is of particular interest in antigenic studies, since the homologous region 108 to 112 in TMV (common strain) corresponds to an antigenic determinant of the depolymerized coat protein. Tryptic peptide 6 of strain U2 (residues 93 to 113) also contains an antigenic determinant, as shown by its ability to inhibit the reaction between U2 protein and specific antibodies.

The U2 strain of tobacco mosaic virus (TMV) belongs to a group of strains that produce field infections of tobacco in Europe and the U.S.A. and differ markedly from TMV (common strain) in chemical and biological properties (Singer et al., 1951; Siegel & Wildman, 1954; Knight et al., 1962; Bald et al., 1974; Wetter & Bernard, 1977). Strain U2 can be distinguished from TMV by its much lower electrophoretic mobility at pH 7 to 9 (Ginoza & Atkinson, 1955; Van Regenmortel, 1972) and is easily identified by serological intra-gel cross-absorption tests (Van Regenmortel, 1967; Wetter & Bernard, 1977).

The amino acid sequence of the coat protein of strain U2 has been determined (Wittmann, 1965; Rentschler, 1967), with the exception of residues 106 to 112. This region is of particular interest for immunochemical studies, since the region 108 to 112 in TMV has been shown to correspond to an antigenic determinant of the depolymerized coat protein (Benjamini, 1977; Milton & Van Regenmortel, 1979).

In the course of studies aimed at elucidating the complete antigenic structure of TMV (Milton & Van Regenmortel, 1979; Milton et al., 1980) we utilized two approaches: the first was to determine the antigenic activity of small peptides corresponding to discrete regions of the protein sequence; the second was to compare the antigenic activity of corresponding regions from different strains (Van Regenmortel, 1975). Studies of the antigenic structure of myoglobin (Atassi, 1977; Hurrell et al., 1977) and lysozyme (Atassi & Lee, 1978; Ibrahimi et al., 1979) have shown that both approaches are essential to account fully for the antigenic properties of globular proteins. To utilize strain U2 in immunochemical studies, it was necessary, therefore, to complete the sequence of its coat protein.

Strain U2 was obtained in 1965 from Dr A. C. Knight. The virus and coat protein were purified as described (Van Regenmortel, 1967).

Trypsin (treated with L-1-tosylamido-2-phenylethylchloromethyl ketone) and carboxy-peptidases A and B were obtained from Worthington Biochemicals (Freehold, N.J., U.S.A.). The staphylococcal protease was a gift from Dr G. R. Drapeau (University of Montreal, Canada). DABITC (4-N,N-dimethylaminoazobenzene-4'-isothiocyanate) was purchased from Fluka (Buchs, Switzerland) and all chemicals used for the Edman degradation were analytical grade products (Pierce, Rockford, Ill., U.S.A.).
Table 1. Amino acid composition of peptides of U2 protein*

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>Tryptic peptide 6 (93–113)</th>
<th>Peptide SP-1 (93–95)</th>
<th>Peptide SP-2 (96–106)</th>
<th>Peptide SP-3 (107–113)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I† II‡ III§</td>
<td>I II III</td>
<td>I II III</td>
<td>I II III</td>
</tr>
<tr>
<td>Asp</td>
<td>4.3 4 4</td>
<td>3.1 3 3</td>
<td>1.0 1 1</td>
<td>1.0 1 1</td>
</tr>
<tr>
<td>Thr</td>
<td>3.0 3 3</td>
<td>1.8 2 2</td>
<td>0.9 1 1</td>
<td>0.9 1 1</td>
</tr>
<tr>
<td>Glu</td>
<td>4.0 4 4</td>
<td>2.1 2 2</td>
<td>1.1 1 1</td>
<td>1.1 1 1</td>
</tr>
<tr>
<td>Pro</td>
<td>1.9 2 2</td>
<td>1.8 2 1</td>
<td>0 0 1</td>
<td>0 0 1</td>
</tr>
<tr>
<td>Ala</td>
<td>2.2 2 2</td>
<td>1.0 1 2</td>
<td>0.9 1 1</td>
<td>0.9 1 1</td>
</tr>
<tr>
<td>Val</td>
<td>1.6 2 2</td>
<td>0.8 1 1</td>
<td>0.5 1 1</td>
<td>0.5 1 1</td>
</tr>
<tr>
<td>Ile</td>
<td>1.4 3 3</td>
<td>0.8 2 2</td>
<td>0 0 0</td>
<td>0 0 0</td>
</tr>
<tr>
<td>Arg</td>
<td>1.0 1 1</td>
<td>1 1 1</td>
<td>0.8 1 1</td>
<td>0.8 1 1</td>
</tr>
</tbody>
</table>

* The total amino acid composition of tryptic peptide 6 agrees with the data of Rentschler (1967). However, several discrepancies were found in the composition of peptides SP-2 and SP-3.
† Composition determined from this work.
‡ Integral values of column I.
§ Composition expected from the data of Rentschler (1967). For peptide SP-2, the position 106 (unidentified by Rentschler) was taken as glutamic acid on the basis of the SP enzyme specificity.
|| The sequences Ile-Ile and Ile-Val are hydrolysed at very low yield under the conditions used (5-7 M-HCl, 20 h at 108 °C). Tsugita et al. (1960) reported that only 50% yield of Ile is obtained from the sequence Ile-Ile.
Residue number

<table>
<thead>
<tr>
<th>Residue number</th>
<th>93</th>
<th>95</th>
<th>97</th>
<th>99</th>
<th>101</th>
<th>103</th>
<th>105</th>
<th>107</th>
<th>109</th>
<th>111</th>
<th>113</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequence of U2 protein*</td>
<td>I</td>
<td>I</td>
<td>E</td>
<td>V</td>
<td>N</td>
<td>Q</td>
<td>P</td>
<td>A</td>
<td>P</td>
<td>N</td>
<td>T</td>
</tr>
<tr>
<td>SP-1</td>
<td>I</td>
<td>I</td>
<td>Z</td>
<td>V</td>
<td>B</td>
<td>B</td>
<td>A</td>
<td>Z</td>
<td>P</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>Sequence of TMV protein‡</td>
<td>I</td>
<td>I</td>
<td>E</td>
<td>V</td>
<td>E</td>
<td>N</td>
<td>Q</td>
<td>A</td>
<td>N</td>
<td>P</td>
<td>T</td>
</tr>
</tbody>
</table>

* The peptides isolated are indicated by double-headed arrows. Residues identified by Edman degradation using the DABITC reagent (→) and by digestion with carboxypeptidase (←) are indicated below the sequence.
† The sequence of residues given in parentheses (106 to 112) was unknown. Circled residues indicate wrong allocations in the published sequence of Rentschler (1967).
‡ The sequence of the equivalent region of TMV protein is given for comparison purposes.

due to losses of this hydrophobic peptide during the extraction procedures of the different steps of the Edman degradation. However, when these 11 residues were compared with the published sequence of U2 protein (Rentschler, 1967), discrepancies were found at positions 100, 101 and 103 (see Table 2).

After SP digestion of peptide T6, three peptides were isolated by the fingerprinting technique. The amino acid compositions of the peptides SP-1 (positions 93 to 95), SP-2 (96 to 106) and SP-3 (107 to 113) are given in Table 1. Compared with the composition expected from the data of Rentschler (1967), discrepancies in the amounts of proline and alanine are found in peptides SP-2 and SP-3. The sequence of peptide SP-3 was determined by the method of Chang et al. (1978) (Table 2). Digestion with carboxypeptidases A and B confirmed that residue 109 is asparagine and residue 112 glutamine. Peptide SP-2 was sequenced in the region 96 to 104. Digestion with carboxypeptidase A (digestion at pH 5.5, followed by digestion at pH 8) identified the sequence 104 to 106 as Thr-Thr-Glu. All these results made it possible to align unambiguously the sequence of peptide T6 and thus to complete the sequence of U2 protein (Table 2). The discrepancies between our results and the published sequence of Rentschler (1967) concern positions 100, 101, 103 and 105.

The antigenic activity of tryptic peptide 6 of strain U2 was demonstrated by its ability to inhibit the complement fixation reaction between U2 protein and its specific antibodies. Using the procedure described previously (Milton & Van Regenmortel, 1979) it was found that 10 nmol of peptide T6 caused 30% inhibition of complement fixation. Since the coat protein conformation of strains U2 and TMV is very similar, it seems reasonable to assume that in peptide T6 of strain U2, it is also the region 108 to 112 that possesses antigenic activity (Benjamini, 1977). Further work with synthetic peptides is in progress to verify this assumption.

Institut de Biologie Moléculaire
et Cellulaire du CNRS
15 rue Descartes
67084 Strasbourg Cedex, France

D. ALTSCHUH
J. REINBOLT
M. H. V. VAN REGENMORTEL*

REFERENCES

SHORT COMMUNICATIONS


(Received 2 July 1980)