

Short communication

Characterisation of a maize-infecting potyvirus from Spain

M. ANGELES ACHON¹, VICENTE MEDINA¹, MICHAEL SHANKS²,
PETER MARKHAM² and GEORGE P. LOMONOSSOFF²

¹ Centre UdL-IRTA, Alcalde Rovira Roure 177, 25006 Lleida, Catalunya, Spain

² Department of Virus Research, John Innes Institute, John Innes Centre, Colney Lane,
Norwich NR4 7 UH, UK

Accepted 17 February 1994

Key words: maize dwarf mosaic virus, nucleotide sequence, potyvirus

Abstract. In order to characterise and classify an unknown maize-infecting potyvirus isolated from fields in northeast Spain, the entire coat protein gene and the C-terminal two-thirds of the large nuclear inclusion protein (Nl) gene were cloned and sequenced. Protein sequencing enabled the cleavage site between the two proteins to be deduced and also revealed that on storage the viral coat protein undergoes a specific degradation in which the N-terminal 39 amino acids are removed. Comparison of the nucleotide sequence of the 3' non-coding region of the viral RNA and the predicted amino acid sequence of the coat protein with the equivalent regions of other members of the potyvirus group revealed that the Spanish virus is closely related to maize dwarf mosaic virus strain A.

Maize crops in northeast Spain are subject to severe losses due to infection with a potyvirus. These losses, estimated by deliberate inoculation of the virus on to maize plants in the field, are found to be between 28–42% of the total crop yield (Achon, unpublished data). Potyviruses which infect maize, sorghum, sugarcane and other poaceous species have been designated as strains of either sugarcane mosaic virus [SCMV; Abbot and Tippet, 1966] or maize dwarf mosaic virus [MDMV; Louie and Knoke, 1975] depending on the host from which they were originally isolated. However, on the basis of serology, some strains of MDMV have been assigned to SCMV [Shepherd, 1965; Snazelle et al., 1971; Pirone, 1972; Jarjees and Uyemoto, 1984]. Recent work, using antibodies specific to the N-terminal portion of the viral coat protein and high performance liquid chromatography (HPLC)-based peptide profiling of tryptic digests of the coat proteins [Shukla et al., 1989; McKern et al., 1991] coupled with host-range studies [Tosic et al., 1990], has suggested that viruses from the USA and Australia, which had been considered isolates of SCMV and MDMV, may actually be four distinct potyviruses.

There is little information currently available on the relationship between European isolates of potyviruses which infect poaceous species and those from the USA and Australia. Biological characterisation of the Spanish virus, which we tentatively designated MDMV-SP, suggested that although it is general similar to isolates of MDMV from the USA, it is distinct from

them in several important respects. For example, it appears to be in sub-division IV rather than I of the potyviruses based on the type of cytoplasmic cylindrical inclusion bodies formed in infected tissue [Edwardson et al., 1984; Lesemann et al., 1992]. It also differs in its host-range as it is able to infect *Bromus rubens* and causes a symptomless in *Sorghum bicolor* cv. TX2786 [Achon et al., 1993]. In order to clarify the nature of this Spanish maize-infecting potyvirus, the region of the viral RNA encoding the viral coat protein and the C-terminal portion of the large nuclear inclusion body (NIb) protein was cloned and its sequence determined.

Virus from an infected maize plant from a field in northeast Spain was propagated by mechanical inoculation on to maize plants (cv. G-4507, Funk's). Virus was purified essentially as described by von Baumgarten et al. [1981] and the virion RNA isolated as described by Wang et al. [1992]. Double-stranded cDNA specific to the 3' terminal region of the viral RNA was synthesised by the RACE procedure of Frohman et al. [1988]. First strand cDNA was synthesised with AMV reverse transcriptase using 5' GACTCGAGTCGACATCGA(T₁₇) 3' as a primer. To amplify the region of the viral RNA expected to encode the carboxy-terminal part of the NIb protein and the entire coat protein the degenerate primer, 5' GGGATCCGA(A,G)ACAGC(T,C,G,A)CT(T,C,G,A)CG(T,C,G,A)AA 3' was synthesised. The primer consists of a *Bam*HI site linked to a sequence corresponding to a region 105–120 nucleotides upstream of the coat protein gene which is partially conserved in MDMV-B, SCMV-SC and Johnson grass mosaic virus (JGMV) [Gough et al., 1987; Frenkel et al., 1991]. The second primer for the PCR reaction was 5' GACTCGAGTCGACATCG 3', a sequence complementary to the extreme 5' end of the first strand cDNA. The PCR reactions were carried out in a Techne PHC-3 thermal cycler using *Taq* polymerase under conditions recommended by the manufacturer. The cycling conditions were as follows: 3 cycles consisting of 1 min at 94°, 1 min at 30° and 3 min at 72°, 30 cycles under the same conditions but with the annealing temperature increased to 57° and a final extension reaction at 72° for 15 min. When analysed by electrophoresis on agarose gels, the major product of this reaction was approximately 1.6 kb in length.

To determine the nucleotide sequence of the 3' portion of the virus genome, the PCR products were digested with *Bam*HI and *Sal*I and ligated into similarly digested M13mp18 replicative form DNA. The ligation mixture was transformed into *Escherichia coli* strain JM101 and phage harbouring inserts were identified by the *lac* complementation assay. Clones containing part of the viral poly(A) tract were identified by 'T-track' analysis [Sanger et al., 1980]. Two of these clones (M13-MDMV-5 and M13-MDMV-8) were subsequently shown to have inserts of approximately 1.6 kb while the third clone (M13-MDMV-r) had an insert of only 1.1 kb. The inserts from M13-MDMV-5 and M13-MDMV-r were oligo-labelled [Feinberg and Vogelstein, 1983] and shown to hybridise with viral RNA in northern blots (data not shown).

The sequences of all three clones were determined by the 'dideoxy' method using a variety of primers to 'walk' along the DNA. To determine the sequence on the opposite strand, the inserts from all three clones were subcloned into M13mp19 and the process repeated. A total sequence of 2143 nucleotides (excluding the poly(A) tail) could be deduced from the sequence of these clones (Fig. 1). M13-MDMV-5 and -8 had identical 5' termini ending in the sequence of the primer. The non-primer derived portion of these clones starts at nucleotide 596 of the finally derived sequence (Fig. 1). Since this sequence is 446 nucleotides upstream of the coat protein gene rather than the anticipated 105–120, the primer must have fortuitously annealed to a site other than that expected. The two clones differed in sequence from each other at only 5 positions. The third clone, M13-MDMV-r, by contrast, had a poly(A) tract adjacent to a 3' end corresponding to an internal sequence within M13-MDMV-5 and -8 (nucleotide 1102 in Fig. 1) and extended the sequence found in these clones in a 5' direction by 596 nucleotides. Examination of the sequence preceding the 3' end of M13-MDMV-r showed it to be extremely A-rich and it is reasonable to assume that first-strand cDNA synthesis was primed at this position to give rise to an apparent poly(A) tract. The different location of the 5' end of M13-MDMV-r again presumable reflects fortuitous priming by the degenerate primer.

The contiguous sequence shown in Fig. 1 contains a single long open reading frame (ORF) sufficient to code for a polypeptide of 635 amino acids. From the known genome organisation of other potyviruses (for a review, see Reichmann et al. [1992]), it was anticipated that the coat protein gene would be contained in the carboxy-terminal portion of this sequence. To map the cleavage site for the release of the coat protein, a sample of purified virus was subjected to automated Edman degradation using an Applied Biosystems 470A gas phase protein sequencer. An unambiguous sequence of 10 amino acids was obtained which allowed the amino-terminus of the coat protein to be mapped (Fig. 1). The fact that a clear N-terminal sequence could be obtained also demonstrates that the coat protein is not blocked. Assuming that the C-terminus of the coat protein is co-terminal with that of the deduced translation product shown in Fig. 1, the calculated size of the coat protein is 32,682 Da.

When samples of denatured virus were analysed by electrophoresis on polyacrylamide gels run according to Schagger and von Jagow [1987], two forms of the coat protein could be found with estimated sizes of either 32 kDa or 29 kDa (Fig. 2). The proportion of the 29 kDa protein increased during storage of the virus at 4 °C, suggesting that this form was derived from the 32 kDa protein by proteolytic processing. To investigate this possibility further, the two forms were electrophoretically separated, transferred to a ProBlott (Applied Biosystems) membrane and subjected to gas-phase Edman degradation as before. The N-terminal sequence from the 32 kDa protein was identical to that obtained when whole virions were

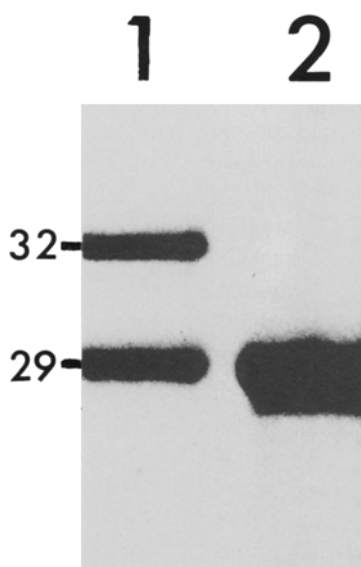


Fig. 2. Coomassie-stained SDS/polyacrylamide gel of the coat protein of the Spanish maize-infecting potyvirus showing the effect of storage. The virus was stored in 50mM sodium phosphate buffer, pH 8.0 at 4° for either 1 week (Lane 1) or 2 weeks (Lane 2) prior to analysis on a 16.5% polyacrylamide gel [Schägger and von Jagow, 1987]. The positions of the 32 and 29kDa forms of the coat proteins are indicated.

analysed (Fig. 1). The N-terminus of the 29 kDa product could also be unambiguously identified (Fig. 1) and starts at residue 40 of the coat protein. This result implies that the 29 kDa protein arises from the 32 kDa protein by the loss of precisely 39 amino acids from the N-terminus. The effect of this loss reduces the calculated size of the viral coat protein to 28,468 Da, in good agreement with size reduction as observed on polyacrylamide gels. The reduction in size of potyviral coat proteins has been reported previously [Hiebert et al., 1984; Lain et al., 1988; Shukla et al., 1988] and the lability of the N-terminal region is presumably due to the fact that it is exposed on the surface of the potyvirus particle [Allison et al., 1985; Shukla et al., 1988].

To clarify the taxonomic position of MDMV-SP, the nucleotide sequence of the 3' non-coding region was compared with those from a number of monocot-infecting potyviruses. By far the greatest identity (94.9%) was

found with the 3' non-coding region of maize dwarf mosaic virus strain A (MDMV-A; J. Jilka and J. M. Clark, unpublished data) from the USA. By contrast the identity with the 3' non-coding region of sugarcane mosaic virus (SCMV-SC) and MDMV strain B [Frenkel et al., 1991] was approximately 50%. Similarly, when the amino acid sequences of the coat proteins were compared, that of MDMV-SP showed an identity of 96.2% with MDMV-A as opposed to approximately 70% when compared with SCMV-SC and MdMV-B. The homology between MDMV-SP and MDMV-A is in the range of variability found between viruses which have been designated as strains of JGMV [Shukla et al., 1992] and is similar to the variability found between strains of other potyviruses [Shukla and Ward, 1988; Frenkel et al., 1989]. The high degree of identity between the sequences of the N-terminal 70 amino acids of the coat proteins of the Spanish virus and MDMV-A (63 out of 70 residues identical) is also similar to the level of identity found in this region between strains of JGMV. This is particularly striking since this region is known to be extremely variable between strains of SCMV [Frenkel et al., 1991; Xiao et al., 1993]. From these homology studies we conclude that MDMV-SP is, indeed, a strain of MDMV and is closely related to MDMV-A.

To our knowledge this is the first example of a European isolate of a maize-infecting potyvirus being characterised at the molecular level. The work also illustrates the usefulness of PCR-based cloning methods in the field of virus taxonomy and epidemiology. The biological significance, if any, of the proteolytic cleavage of the coat protein which occurs on storage of virus is unclear though its occurrence may explain the variation in estimates of the size of the coat protein of MDMV-A [von Baumgarten and Ford, 1981; Jensen et al., 1986].

Acknowledgments

We are indebted to Dr. J.M. Clark of the University of Illinois, USA for providing the nucleotide sequence of MDMV-A prior to publication. We thank Pat Barker (Institute of Animal Physiology and Genetics Research, Babraham, Cambridge) for help with the protein sequence analysis. We acknowledge the contribution of UPC-OLIVETTI and the British Council/Acciones Integradas programme in supporting this work.

References

- Abbott EV and Tippet RL (1966) Strains of sugarcane mosaic virus. US Department of Agriculture Research Services Technical Bulletin 1340: 1-25
- Achon MA, Shanks M, Medina V, Markham PG and Lomonosoff GP (1993) Characterisation of a Spanish maize dwarf mosaic virus isolate. International Congress of Virology, Glasgow, Abstract P76-12

- Allison RF, Dougherty WG, Parks TD, Willis L, Johnston RE, Kelly M and Armstrong FB (1985) Biochemical analysis of the capsid protein gene and capsid protein of tobacco etch virus: N-terminal amino acids are located on the virion's surface. *Virology* 147: 309–316
- Edwardson JR, Christie RG and Ko, NJ (1984) Potyvirus cylindrical inclusions – Subdivision-IV. *Phytopathology* 74: 1111–1114
- Feinberg AP and Vogelstein B (1983) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Analytical Biochemistry* 132: 6–13
- Frenkel MJ, Ward CW and Shukla DD (1989). The use of 3' non-coding nucleotide sequences in the taxonomy of potyviruses: Application to watermelon mosaic virus 2 and soybean mosaic virus-N. *Journal of General Virology* 70: 2775–2783
- Frenkel MJ, Jilka JM, McKern NM, Strike PM, Clark JM, Shukla DD and Ward CW (1991). Unexpected sequence diversity in the amino-terminal ends of the coat proteins of strains of sugarcane mosaic virus. *Journal of General Virology* 72: 237–242
- Frohman MA, Dush MK and Martin GR (1988) Rapid production of full-length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer. *Proceedings of the National Academy of Sciences USA* 85: 8998–9002
- Gough KH, Azad AA, Hanna PJ and Shukla DD (1987) Nucleotide sequence of the capsid and nuclear inclusion protein genes form the Johnsongrass strain of sugarcane mosaic virus RNA. *Journal of General Virology* 68: 297–304.
- Hiebert E, Tremaine JH and Ronald WP (1984) The effect of limited proteolysis on amino acid composition of five potyviruses and on the serological reaction and peptide map of the tobacco etch virus capsid protein. *Phytopathology* 74: 411–416
- Jarjees MM and Uyemoto JK (1984) Serological relatedness of strains of maize dwarf mosaic and sugarcane mosaic viruses as determined by microprecipitin and enzyme-linked immunosorbent assays. *Annals of Applied Biology* 104: 497–501
- Jensen SG, Long-Davidson B and Seip L (1986) Size variation among proteins induced by sugarcane mosaic viruses in plant tissue. *Phytopathology* 76: 528–532
- Lain S, Riechmann JL, Mendez E and Garcia JA (1988) Nucleotide sequence of the 3' terminal region of plum pox potyvirus RNA. *Virus Research* 10: 325–342
- Lesemann DE, Shukla DD, Tosic M, & Huth W (1992) Differentiation of the four viruses of the sugarcane mosaic virus subgroup based on cytopathology, pp 353–361 in "Potyvirus Taxonomy", Barnett OW (ed) *Archives of Virology Supplement* 5: Springer, Wien/New York
- Louie R and Knoke JK (1975) Strains of maize dwarf mosaic virus. *Plant Disease Reporter* 59: 518–522
- McKern NM, Shukla DD, Toler RW, Jensen SG, Tosic M, Ford RE, Leon O and Ward CW (1991) Confirmation that the sugarcane mosaic virus subgroup consists of four distinct potyviruses by using peptide profiles of coat proteins. *Phytopathology* 81: 1025–1029
- Pirone, TP (1982) Sugarcane mosaic virus. CMI/AAB Descriptions of Plant Viruses No. 88
- Riechmann JL, Lain S and Garcia JA (1992) Highlights and prospects of potyvirus molecular biology. *Journal of General Virology* 73: 1–16.
- Sanger F, Coulson AR, Barell BG, Smith AJH and Roe BA (1980) Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. *Journal of Molecular Biology* 143: 161–178
- Schägger H and von Jagow G (1987) Tricine-sodium dodecyl sulphate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Analytical Biochemistry* 166: 368–379.
- Shepherd RJ (1965) Properties of a mosaic virus of corn and Johnsongrass and its relation to sugarcane mosaic virus. *Phytopathology* 55: 1250–1256.
- Shukla DD and Ward CW (1988) Amino acid sequence homology of coat proteins as a basis for identification and classification of the potyvirus group. *Journal of General Virology* 69: 2703–2710.
- Shukla DD, Strike PM, Tracy SL, Gough KH and Ward CW (1988) The N and C termini of

- the coat proteins of potyviruses are surface-located and the N terminus contains the major virus-specific epitopes. *Journal of General Virology* 69: 1497–1508
- Shukla DD, Tosić M, Jilka J, Ford RE, Toler RW and Langham, MAC (1989) Taxonomy of potyviruses infecting maize, sorghum, and sugarcane in Australia and the United States as determined by reactivities of polyclonal antibodies directed towards virus-specific N-termini of the coat proteins. *Phytopathology* 78: 223–229
- Shukla DD, Frenkel ML, McKern NM, Ward CW, Jilka J and Ford RE (1992). Current status of the sugarcane mosaic virus subgroup, pp 363–373 in “Potyvirus Taxonomy”, Barnett OW (ed) *Archives of Virology Supplement 5*: Springer, Wien/New York
- Snazelle TE, Bancroft JB and Ullstrup AJ (1971) Purification and serology of maize dwarf mosaic and sugarcane mosaic viruses. *Phytopathology* 61: 1059–1063
- Tosić M, Ford RE, Shukla DD and Jilka J (1990) Differentiation of sugarcane, maize dwarf, Johnson grass, and sorghum mosaic viruses based on reactions of oat and some sorghum cultivars. *Plant Disease* 74: 544–552
- Von Baumgarten G and Ford RE (1981) Purification and partial characterization of maize dwarf mosaic virus strain A. *Phytopathology* 71: 36–41
- Wang D, Hayes IM and Maule AJ (1992) Procedures for the efficient purification of pea seed-borne mosaic virus and its genomic RNA. *Journal of Virological Methods* 36: 223–230
- Xiao XW, Frenkel MJ, Teakle DS, Ward CW and Shukla DD (1993) Sequence diversity in the surface-exposed N-terminal region of the coat proteins of seven strains of sugarcane mosaic virus correlates with their host-range. *Archives of Virology* 132: 399–408