

***Plant Gene Register***

# Molecular Cloning and Sequence Analysis of an Andean Potato Mottle Virus cDNA Clone Encoding the Major Coat Protein<sup>1</sup>

Nice Shindo, Paulo S. T. Brioso, Ana C. P. Vicente, Roberto Krengiel, Durvalina B. Felix, Guy Bauw, and Dulce E. de Oliveira\*

*Departamento de Genética, Universidade Federal do Rio de Janeiro, CP 68011, CEP 21944, Rio de Janeiro, Brasil (N.S., R.K., D.E.d.O.); Departamento de Fitopatologia, Universidade Federal Rural do Rio de Janeiro, CP 23851, Rio de Janeiro, Brasil (P.S.T.B.); Departamento de Genética, Fundação Oswaldo Cruz, CP 926, Rio de Janeiro, Brasil (A.C.P.V.); Departamento de Bioquímica, Universidade Federal do Rio de Janeiro, CP 1573, Rio de Janeiro, Brasil (D.B.F., D.E.d.O); and Laboratorium voor Genetica, Rijksuniversiteit Gent, K. L. Ledeganckstraat 35, B-9000 Gent, Belgium (G.B.)*

APMV<sup>2</sup> is an endemic South American virus belonging to the comoviruses, whose host range is mainly restricted to the Solanaceae (2). APMV is a single-stranded positive-sense RNA virus with a bipartite genome, the so called M and B-RNA (3.5 and 5.6 kb), which are separately encapsidated into isometric particles. As for many other comoviruses (6, 9), particles are made up of two CPs, with mol wts of 22,066 and 42,494. Both CPs are encoded by the M-RNA and synthesized as part of a larger polypeptide. Proteolytic cleavage of the polyprotein yields the respective functional proteins (8). Recently a cDNA clone encoding the minor CP of APMV was isolated and showed that the minor CP is localized at the outermost 3' end of the M-RNA (7). Here, we report the cloning and characterization of a larger cDNA from the M-RNA, which encodes, in addition to the minor CP, also the major APMV CP (Table I). Viral RNA purification, cDNA synthesis, and cloning were performed as described earlier (7). cDNA inserts up to 2.3 kb were isolated, assigned by northern hybridization to either APMV genomic RNA, and sequenced according to the method of Sanger et al. (5). The NH<sub>2</sub> terminus of the 42-kD CP was determined by microsequencing the purified CP, after electroblotting onto polybase-coated glass fiber (1).

Nucleotide sequence analysis of clone PS 60 (2309 bp) showed a polyadenylation tail at the 3' end and a long ORF translated into 601 aa (Fig. 1). As previously reported (7), clone PS60 also includes the minor CP at the 3' end (from aa 405 to 601). The homology between both clones within this region is 100%. The NH<sub>2</sub>-terminal sequence of the major

CP, as determined by microsequencing of the purified protein, is indicated on the ORF (starting at position 52). Apparently, the major CP is cleaved NH<sub>2</sub> terminally at a QM bond, in analogy with cowpea mosaic virus, the type member of the comoviruses. The COOH-terminal end of the major CP is localized at position 1212 and represents the proteolytic cleavage site between major and minor APMV CPs. The predicted mol wt (42,494) of the larger CP, derived from the

**Table I.** Characteristics of a cDNA Clone Encoding the 42-kD CP of the APMV

Organism:	Comovirus—Andean potato mottle virus.
Location in Genome:	M-RNA.
Gene Function:	Capsid protein.
Techniques; Methods of Isolation; Subsequent Identification:	cDNA cloning; northern hybridization with APMV RNAs; NH <sub>2</sub> -terminal microsequencing of isolated APMV 42-kD CP; complete dideoxy sequencing of both strands; sequence comparison with related comovirus.
Expression and Regulation:	Protein expression not tested.
Features of Gene Structure:	Putative peptide 387 aa long (mol wt 42,494), presumably processed from a larger polyprotein encoded by the M-RNA.
Codon Usage:	Codons not used include: TTA(L), TCG(S), CGA(R); aa with a high bias to one codon: A(GCA), F(TTT), L(TTG), T(ACA), D(GAT).
(G + C) Content:	41.6%.
Antibodies:	Polyclonal antibodies raised in rabbit against isolated APMV CPs.
GenBank Accession No.:	M83536.

<sup>1</sup> This work was supported by grants from Programa de Apoio ao Desenvolvimento Científico e Tecnológico, Conselho Nacional de Pesquisas, and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Brazil), Algemeen Bestuur voor Ontwikkelings-samenwerking (Belgium), and the Commission of the European Communities.

<sup>2</sup> Abbreviations: APMV, Andean potato mottle virus; CP(s), coat protein(s); ORF, open reading frame; aa, amino acids.

aa sequence, is in good agreement with the estimate of 42,100 previously arrived at by SDS-PAGE (4). Comparison of the aa sequence of the APMV 42-kD CP with the major CP sequence of related comoviruses (3, 9) revealed an overall homology of 36.1% to cowpea mosaic virus and 39.6% to red clover mottle virus, suggesting that they apparently have a common genetic structure.

Thus, we have cloned the APMV ORF that encodes both major and minor CPs. Polymerase chain reaction strategies will be used to synthesize clones that cover each of the CPs separately. We will attempt to engineer CP-mediated protection against APMV. The necessity of one or both CPs in building up resistance will be considered.

#### ACKNOWLEDGMENTS

We thank Raimundo Villarroel for providing oligonucleotides, and Drs. Marc Van den Bulcke and Chris Simoens for interesting discussions.

#### LITERATURE CITED

1. **Bauw G, De Loose M, Inzé D, Van Montagu M, Vandekerckhove J** (1987) Alterations in the phenotype of plant cells studied by NH<sub>2</sub>-terminal amino acid sequence analysis of proteins electroblotted from two dimensional gel-separated total extracts. *Proc Natl Acad Sci USA* **84**: 4806-4810
2. **Fribourg CE, Jones RAC, Koenig R** (1977) Andean potato mottle virus, a new member of cowpea mosaic virus group. *Phytopathology* **67**: 969-974
3. **Myers EW, Miller W** (1988) Description of the alignment method used in program PALIGN. *Cabios* **4**: 11-17
4. **Salazar LF, Harrison BD** (1978) Particle properties and status of Andean potato mottle virus. *J Gen Virol* **39**: 171-175
5. **Sanger F, Nicklen S, Coulson AR** (1977) DNA sequencing with chain terminating inhibitors. *Proc Natl Acad Sci USA* **74**: 5463-5467
6. **Shanks M, Stanley J, Lomonosoff GP** (1986) The primary structure of red clover mottle virus middle component RNA. *Virology* **155**: 697-706
7. **Shindo N, Krenziel R, Brioso PST, Vicente ACP, Weyne M, Oliveira DE, Timmerman B** (1992) Complete nucleotide sequence of the 22 kDa coat protein of Andean potato mottle virus. *Plant Mol Biol* (in press)
8. **Vos P** (1987) Molecular cloning and expression of full-length DNA copies of the genomic RNAs of cowpea mosaic virus. PhD thesis. University of Wageningen, Wageningen, The Netherlands
9. **Wezembeek P, Verver J, Harmsen, J, Vos P, Van Kammen A** (1983) Primary structure and gene organization of the middle-component RNA of cowpea mosaic virus. *EMBO J* **2**: 941-946

1 CAACAGAGCATGTTGATTTCCAACGTCGCAATGTTGTTGCACAGGCACAAATGGAGGTT  
 Q Q S M L Y F Q R R N V V A Q A Q M E Y  
 61 GCGCAGTTTGGCGGTYGCTGGATGATACCAAGAGTCTCAAGTTGCAAAGTCTGTTGAAT  
A O F G R I L L D D T K S L K L Q S L L N  
 121 AGCAGAATTGCAGCAGGTAGATTTTCAATACCCATGACCGCAGTCAAAGGTACTGTTGTT  
 S R I A A G R F S I P M T A V K G T V V  
 181 TTCGATGGTTTGTGGCATCACTTATTGGCACAACCTTGAGGGGAGCACCTATGTTTAGG  
 F D G L L A S L I G T T L R G A P M F R  
 241 CATACTTACAGGCAGACCAAAGCTGCGTTTTATTTTACCATAAATGTTCCAATATCA  
 H T Y R Q S T K L R F I F T I N V P I S  
 301 ACTGGTATAGGACTTATGGTAGGGTACAACAGTGTACATCTGATAAACACCTGACCAAC  
 T G I G L M V G Y N S V T S D K E L T N  
 361 GAGTATACAATCTCATCTGAAGAAAGCGTTGTGTGGAATCTGCGTGTGAGGGAGTGCTC  
 E Y T I S S E S V V W N P A C Q G V L  
 421 GAGTTTTAGTCACTAGTCTAATCCATGTGGGATGACTGGTCAATGATTATTTTAGGCAA  
 E F S V S P N P C G M Y W S Y D Y F R Q  
 481 ACAGGATCAGCCTTTCCATTTGTGTAATTTACCATGGAGTGAACACCCACAACCTGAC  
 T G S R L S I C V I S P W S A T P T T D  
 541 TGTGCTGTCGATGGCAAATCCATGTGGATGATGAGCAGATGACTATGTCAATTTTAAAT  
 C A V A W Q I H V D D E Q M T M S I P N  
 601 CCAACGCAAGCTCAGCAGTTTTGGCTGTGAAAAGATGGATGGGTAATCTTATTTTAA  
 P T Q A P A V L P V K R W M G N L I F K  
 661 CAGGGAGCACAGGACAGTGAAGAAAATGCCTTTGGCAATGGAGCAGCAGTGGGAGAT  
 Q G A Q E Q V K K M P L A I G A A V G D  
 721 GATAAACAGCTGTCTAGACAATGCCAATCTTTGGCAGCAATGTGGAATACCAGATT  
 D K T A V M T M P N S L A A M W N Y Q I  
 781 GGTACATTTAATTTGAATTTTACAAAATGAGTAGTCCCTTCATCAAAGGAACGCTTTTG  
 G T F N F E F T K L S S P F I K G T L L  
 841 GCTTTTATAGCAATGGATCAGGACGTGTCTTATAGTTTGAAGAGTTGCAAATTTTCCA  
 A F I A M D Q D V S Y S L E E L Q N F P  
 901 AACAAAGATTGTCAATTTGATGAAAAGATGGAAGGCATATGTATCTTTTGGTGAAGAG  
 N K I V Q P F D E K D G R A Y V S P G E E  
 961 CACTTTGCACAAGCTGGTCCACTCAGGTTTTCAGGAGCTGTAACATCAGCCAAAAGGGGT  
 H F A Q A W S T Q V S G A V T S A K R G  
 1021 TGTCTTATTGTGTTGTTAGTAAAGATTGCATAGCTTCTACCATTGTGGTGATTTT  
 C P Y L Y V V S K D C I A S T I C G D F  
 1081 CAAGTGGGAGTTAAGTTGCTATCTATAGAAAATTTACCATGTGGTTACAATCCCTGGA  
 Q V G V K L L S I E N Y S P C G Y N P G  
 1141 CTTGTTGAGCTTCTACAATAGTGCAGAATACTGCAGGTTCAAACCTACATCTTTGTTG  
 L V V A S T I V Q N T A G S N S T S L L  
 1201 GCGTGGCCTCAGTTTGTAGTCCATGATAAATGTTTGGAGTGAAGTTTGTGCATTAGAT  
 A W P Q<sup>1</sup> F C S P C I N V W S E F C A L D  
 1261 ATTCTCTGTTGGACACAACCTAAGGTTAATTTTGGCCAATATCTCTGGATCTTGTGAAT  
 I P V V D T T K V N F A Q Y S L D L V N  
 1321 CCAACAGTTTCTGCAATGCCTCTGGGCGTAATGGAGGTTTGTCTTATACCTTCTCCC  
 P T V S A N A S G R N W R F V L I P S P  
 1381 ATGGTGTATTACTCAAACCTCAGACTGAAAAGAGGAAAGTTGCATTTTAAAGCTTAAA  
 M V Y L L Q T S D W K R G K L H F K L K  
 1441 ATACTGGGAAATCCAATGTTAAACGATCTGAATGGAGTAGCACAAGCAGGATAGATGTG  
 I L G K S N V K R S E W S S T S R I D V  
 1501 AGAAGAGCACCTGGTACAGAGTATTTAAATGCTATCACTGTTTCTACTGCTGAGCCACAT  
 R R A P G T E Y L N A I T V F T A E P H  
 1561 GCAGATGAGATAAATTTGAAAATAGAAAATTTGTTGGCCAAACAACGGATTGAGATGTGG  
 A D E I N F E I E I C G P N N G F E M W  
 1621 AATGCTGATTTTGGAAACCAATGCTTGGATGGCAAATGTTGTTATGGAATCCTGAT  
 N A D F G N Q L S W M A N V V I G N P D  
 1681 CAAGCGGTATACATCAATGGTATGTTAGGCCAGGAGAAAATTTGAGGTTGAGGAAAT  
 Q A G I H Q V Y V R P G E N F E V A G N  
 1741 AGGATGGTTCAACCTTAGCGCTTCTGGGAGGATGGTACTGGTATGCTTCCAATACTA  
 R M V Q P L A L S G E D G T G M L P I L  
 1801 AAGTAGCCAATAAATTTGATTTGTGCGTGTCTCTCTGAGAAACGCTCTGGTGTGCATT  
 K -  
 1861 CACCACCTAGGAGCTAGGACTCTGGGTTTTAATGCAAATGTTTTAATTTTGTATTATAA  
 1921 ATGGTTTCTTTAATTTAAGTAGTCGTAATCGTTTATCTGGAATTTTAAACAAGTTTTTAC  
 1981 GTTACTGAGCTTCTGCCCTTGTAAAGAACTGTGTAAATTTGTAGTTTATAAACTTAG  
 2041 TTTGTATTCTGGATTTGTGCGTGTCTCTGAGAAACGCTCTGGTGTGCATTACCAC  
 2101 CTAGGAGGTAGGACTCTGGGTTTCAATGCAAACCTGTTTTAATTTTGTACTTAAATAGT  
 2161 TTGCTTTAGTTCAAGTAATATGATCGTTTATTGGAATTCATAAGTTCTTGGCTTCT  
 2221 GAGCTTCTGCCCTTGTAAAGAACTTGTGTAATTTGTAGTTTACAACTTTGGTTTTCGA  
 2281 TTTGTGTTTTAAAAA

Figure 1. Nucleotide sequence of the clone PS60. The deduced aa sequence of the 42-kD APMV CP is in bold. Arrowheads indicate both cleavages sites, QM (NH<sub>2</sub>) and QF (COOH). The 42-kD CP NH<sub>2</sub>-terminal sequence as determined by microsequencing is underlined.