The nucleotide sequence of cowpea mottle virus and its assignment to the genus Carmovirus

X. J. You, J. W. Kim, G. W. Stuart and R. F. Bozarth

1 Department of Life Sciences, Indiana State University, Terre Haute, IN 47809, USA and 2 Department of Microbiology, School of Medicine, Inha University, Inchon 403–751, Korea

The genome of cowpea mottle virus (CPMoV) is a positive ssRNA of 4029 nucleotides with six major open reading frames (ORFs). A non-coding region of 34 nucleotides precedes the first AUG. ORF1 encodes a 25 kDa polypeptide of unknown function and ORF2 encodes a 56 kDa putative RNA replicase. Like other members of carmoviruses, suppression of the amber termination codon of ORF1 would produce a read-through polypeptide of 83 kDa. ORF3 and ORF4 encode two small proteins of 7.8 and 9.8 kDa, respectively. ORF5 encodes the 40 kDa capsid protein. ORF6 is located within ORF5 but has no postulated function. CPMoV RNA is blocked at the 5' end and is not polyadenylated at the 3' end. Comparison of the physicochemical properties, genomic arrangement, and predicted amino acid sequences with those of other viruses justify the assignment of CPMoV to the genus Carmovirus, family Tombusviridae.

Cowpea mottle virus (CPMoV) has been proposed as a member of the Tombusviridae, genus Carmovirus (Kim & Bozarth, 1992). Shoyinka et al. (1978) described the physicochemical properties of the virus and produced antiserum which has been tested against cowpea virus isolates throughout the world. Until recently, this virus was only reported in Nigeria. This 30 nm isometric virus has a sedimentation velocity of S = 122, buoyant density in CsCl of 1.3492 g/cm³, contains a positive-strand RNA genome (molecular mass 1.4 x 10⁶ Da) and a capsid of 180 subunits (39610 Da). The capsid protein gene of 1104 nucleotides is located near the 3' terminus (Kim & Bozarth, 1992). The postulated amino acid sequence of CPMoV capsid protein has 36% and 27% similarity to the S domains of turnip crinkle and carnation mottle virus. Although Kim & Bozarth (1992) reported that it was not serologically related to these carmoviruses, Hari & Bozarth (unpublished) have recently found that CPMoV and bean mild mottle virus (BMMV), an accepted carmovirus, are serologically related in reciprocal tests. Like other members of carmoviruses, CPMoV generates three dsRNAs which are co-terminal at the 3' end, but no subgenomic RNAs were found encapsidated in the virions and no polyadenylation was found at the 3' end of the genomic RNA (Kim & Bozarth, 1992).

In this paper the nucleotide sequence and the genomic organization of CPMoV are reported and compared to the same properties of viruses of the Tombusviridae and related genera. These data support the assignment of CPMoV to the Tombusviridae, genus Carmovirus.

The cowpea strain of CPMoV (ATCC No. PV-955) was propagated in cowpea (Vigna unguiculata cv. California No. 5) and was purified using the procedure of Shoyinka et al. (1978). RNA was extracted from purified virus with phenol–chloroform (1:1), precipitated with ethanol, lyophilized, and then stored at −20 °C. Following the addition of poly(A) tails with poly(A) polymerase, cDNA clones were generated with random primers [including oligo(dT)] and a cDNA synthesis kit (Promega). Additional cDNA clones were generated using synthetic oligonucleotide primers designed to match specific sites within the genomic RNA. The synthetic double-stranded cDNAs were blunt ended by T4 DNA polymerase, then EcoRI adaptors were ligated to the cDNA ends using T4 DNA ligase. A GeneClean II kit (Bio 101) was used to remove the unligated EcoRI adaptors. Following a kinase step, the modified double-stranded cDNAs were ligated to a dephosphorylated pT7/T3 18 U plasmid vector (Pharmacia). The recombinant plasmids were used to transform competent cells (Escherichia coli DH5α) by heat shock according to the manufacturer's protocol (GIBCO BRL). Colonies which were white in the presence of X-Gal were tested for

The nucleotide sequence presented in this paper has been deposited in the GenBank database under the accession number U20976.
Fig. 1. Nucleotide sequence of CPMoV. The amino acids encoded by the six ORFs are shown as p25/83, p56, p7.8, p9, p40 and p28.
Fig. 2. Major ORFs of the CPMoV genome. Only sequences starting with AUG, ending with a termination codon and having 210 or more nucleotides are shown. The readthrough peptide joining ORF1 with ORF2 is indicated by a dotted line.

The 5' end labelling experiments indicated that CPMoV RNA is blocked, and most likely capped, similar to the RNA of TMV and carnation mottle virus (CarMV) (Guilley et al., 1985). In all of six experiments the capped control RNA (TMV) and CPMoV RNA remained unlabelled, whereas the uncapped control RNA (MS2) was strongly labelled (data not shown). Further experiments are needed to confirm the nature of the capped structure.

There are six major ORFs in the CPMoV genome (Fig. 2), five of which are found in corresponding positions in other carmoviruses. ORF1 (nt 35–715) encodes a 25 kDa polypeptide of unknown function. ORF2 (nt 776–2263) encodes the 56 kDa putative RNA replicase or part of it. Like many other plant viruses, suppression of the ORF1 amber termination codon would result in a readthrough product of 83 kDa (Carrington & Morris, 1985; Morris & Carrington, 1988). ORF3 (nt 2287–2496) and ORF4 (2411–2677) encode two small proteins of 7.8 and 9.8 kDa, respectively, which appear similar in size and sequence (22–34% similarity) to corresponding ORFs which are located in similar positions in the carmoviruses listed in Table 1. ORF5 is located near the 3' end of the RNA (nt 2674–3777) and encodes the 40 kDa capsid protein subunit. ORF6 (nt 2678–3454) is completely within ORF5, but in a different frame. It is in the same reading frame as ORF4. If the ORF4 termination codon were suppressed, a protein of 38 kDa would be produced similar to the one postulated for maize chlorotic mottle virus (MCMV) (Nutter et al., 1989). However, a comparison of the amino acid sequences of these putative peptides revealed little sequence similarity.

The in vitro translation of CPMoV RNA produced four major polypeptides (Kim & Bozarth, 1992), with estimated molecular masses of about 50, 40, 27 and 11 kDa. The 40 kDa polypeptide reacted with capsid protein antibodies (Kim & Bozarth, 1992) and corresponds to the ORF5 product. The 27 kDa product was the most abundant translation product and was
The protein of approximately 50 kDa could have been encoded by ORF2, the putative replicase gene. However, no polypeptide corresponding to a putative 83 kDa ORF1/ORF2 readthrough product was detected in spite of repeated attempts in which Mg\(^{2+}\) and/or K\(^{+}\) concentrations were varied over a wide range. Furthermore, the use of edeine, which has been reported to promote readthrough of leaky terminators (Kim & Bozarth, 1992), did not result in a detectable p83 product. In this context we note that whereas the readthrough product of CarMV produced a strong band following the ORF1 stop codon and in the same reading frame as ORF2.

ORF1 (pre-readthrough region of p83) had the same general trend in similarities to the viruses compared in Table 1, and we note that MCMV is within the range of the carmoviruses in this comparison. ORF5 (capsid protein) also showed the same trend in similarity with CPMoV, being most similar to the carmoviruses but least similar to MCMV. Kim & Bozarth (1992) discussed the strong similarity of CPMoV to carmoviruses and tombusviruses and observed 39% matches in the S regions of CPMoV and TCV.

Significant sequence similarity has been previously reported among the RNA replicases of carmoviruses, tombusviruses, luteoviruses (Rochon & Tremaine, 1989), necroviruses (Coutts et al., 1991), dianthoviruses (Xiong & Lommel, 1989) and machlomoviruses (Nutter et al., 1989) and MCMV (Nutter et al., 1989). The nucleotide and the derived amino acid sequences of CPMoV were compared to the sequences of tombusviruses and other similar viruses (Table 1). The highest degree of amino acid sequence similarity was observed in the post readthrough part (rtp + ORF2) of the putative RNA replicase gene (Table 1). Similarity was high for all viruses tested, but the strongest similarity was with TCV, MNSV and CCFV (52–54%), which are all carmoviruses. The next best similarity was with CarMV (a carmovirus) and MCMV (a machlomovirus) (45–46%) followed by tomato bushy stunt virus (TBSV), cucumber necrosis virus (CNV) (tombusviruses) and tobacco necrosis virus (TNV) (a necrovirus) (39–40%), and red clover necrotic mosaic virus (RCNMV) and carnation ringspot virus (CRV) (dianthoviruses) (32–33%) in descending order of similarity. Within the carmoviruses (including CPMoV) the conservation was absolute in 157 of 534 positions, including several sets of five to ten contiguous amino acids. Similarity was especially high in the region of the GDD motif, now considered to be a constant in plant virus replicase genes (Goldbach et al., 1991), and was poorest in the last 50 residues of ORF2 (data not shown).

The protein of approximately 50 kDa could have been encoded by ORF2, the putative replicase gene. However, no polypeptide corresponding to a putative 83 kDa ORF1/ORF2 readthrough product was detected in spite of repeated attempts in which Mg\(^{2+}\) and/or K\(^{+}\) concentrations were varied over a wide range. Furthermore, the use of edeine, which has been reported to promote readthrough of leaky terminators (Kim & Bozarth, 1992), did not result in a detectable p83 product. In this context we note that whereas the readthrough product of CarMV produced a strong band in translation studies (Carrington & Morris, 1985), similar experiments with turnip crinkle virus (TCV) (Dougherty & Kaesberg, 1981) and BMMV (Hari et al., 1988) produced very weak bands relative to the other translation products. Others merely postulated the presence of a readthrough product on the basis of similar genomic organization for cardamine chlorotic fleck virus (CCFV) (Skotnicki et al., 1993) and melon necrotic spot virus (MNSV) (Riviere & Rochon, 1990). While there is strong evidence that the ORF1/ORF2 readthrough protein is a characteristic of carmoviruses, it has yet to be demonstrated in three of the six carmoviruses studied (including CPMoV).

The nucleotide and the derived amino acid sequences of CPMoV were compared to the sequences of tombusviruses and other similar viruses (Table 1). The highest degree of amino acid sequence similarity was observed in the post readthrough part (rtp + ORF2) of the putative RNA replicase gene (Table 1). Similarity was high for all viruses tested, but the strongest similarity was with TCV, MNSV and CCFV (52–54%), which are all carmoviruses. The next best similarity was with CarMV (a carmovirus) and MCMV (a machlomovirus) (45–46%) followed by tomato bushy stunt virus (TBSV), cucumber necrosis virus (CNV) (tombusviruses) and tobacco necrosis virus (TNV) (a necrovirus) (39–40%), and red clover necrotic mosaic virus (RCNMV) and carnation ringspot virus (CRV) (dianthoviruses) (32–33%) in descending order of similarity. Within the carmoviruses (including CPMoV) the conservation was absolute in 157 of 534 positions, including several sets of five to ten contiguous amino acids. Similarity was especially high in the region of the GDD motif, now considered to be a constant in plant virus replicase genes (Goldbach et al., 1991), and was poorest in the last 50 residues of ORF2 (data not shown).

ORF1 (pre-readthrough region of p83) had the same general trend in similarities to the viruses compared in Table 1, and we note that MCMV is within the range of the carmoviruses in this comparison. ORF5 (capsid protein) also showed the same trend in similarity with CPMoV, being most similar to the carmoviruses but least similar to MCMV. Kim & Bozarth (1992) discussed the strong similarity of CPMoV to carmoviruses and tombusviruses and observed 39% matches in the S regions of CPMoV and TCV.

Significant sequence similarity has been previously reported among the RNA replicases of carmoviruses, tombusviruses, luteoviruses (Rochon & Tremaine, 1989), necroviruses (Coutts et al., 1991), dianthoviruses (Xiong & Lommel, 1989) and machlomoviruses (Nutter et al., 1989). The nucleotide and the derived amino acid sequences of the putative RNA replicase of positive-strand RNA viruses revealed that RNA replicase is the only viral protein containing motifs conserved throughout this class of viruses (Koonin, 1991). CPMoV has the overall physicochemical properties (Shoyinka et al., 1978) and the genomic attributes of carmoviruses (Morris & Carrington, 1988), which are now placed in the family Tombusviridae (Mayo & Martelli, 1993). The 5' end of CPMoV has two ORFs which are similar in size and sequence to the two ORFs found in similar positions in other carmoviruses. Like other carmoviruses, suppression of the termination codon of the first ORF would result in a readthrough
product encompassing both ORFs (Riviere & Rochon, 1990; Hacker et al., 1992; Skotnicki et al., 1993). The central portion of the CPMoV genome has ORFs of approximately the same size as those which have been shown to be associated with virus movement (Hacker et al., 1992). The 3' region of the genome codes for the capsid protein which is similar in properties to other carmoviruses and tombusviruses (Kim & Bozarth, 1992), and a serological relationship has been found with the capsid protein of BMMV, a carmovirus (V. Hart & R. F. Bozarth, unpublished). Like CarMV (Guilley et al., 1985) it has a blocked 5' end which may be capped and is not polyadenylated at the 3' end. On the basis of the above, CPMoV should be accepted as a legitimate member of the genus Carmovirus.

This project was supported by a research grant from Indiana Academy of Science to X. J. You (1993) and a Faculty Research Grant to R. F. Bozarth from Indiana State University. We thank V. Hart for permission to report his unpublished experiments.

References


(Received 22 February 1995; Accepted 7 July 1995)