Volume 17 Number 19 1989 Nucleic Acids Research

Nucleotide sequence and genetic organization of Hungarian grapevine chrome mosaic nepovirus RNA2

V.Brault, L.Hibrand, T.Candresse, O.Le Gall and J.Dunez

Station de Pathologie Végétale, INRA, BP 131, 33140 Pont de la Maye, France

Received May 24, 1989; Revised and Accepted August 25, 1989 EMBL accession no. X15163

ABSTRACT

The complete nucleotide sequence of hungarian grapevine chrome mosaic nepovirus (GCMV) RNA2 has been determined. The RNA sequence is 4441 nucleotides in length, excluding the poly(A) tail. A polyprotein of 1324 amino acids with a calculated molecular weight of 146 kDa is encoded in a single long open reading frame extending from nucleotides 218 to 4190. This polyprotein is homologous with the protein encoded by the S strain of tomato black ring virus (TBRV) RNA2, the only other nepovirus sequenced so far. Direct sequencing of the viral coat protein and *in vitro* translation of transcripts derived from cDNA sequences demonstrate that, as for comoviruses, the coat protein is located at the carboxy terminus of the polyprotein. A model for the expression of GCMV RNA2 is presented.

INTRODUCTION

As presented in the accompanying paper, hungarian grapevine chrome mosaic virus (GCMV)(1), is a member of the nepovirus group.

We have determined the complete nucleotide sequence of GCMV RNA2. The 4441 nucleotide long sequence encodes a 146 kDa polyprotein which has been compared with the protein encoded by the closely related tomato black ring nepovirus RNA2 (TBRV) (2). The two viral RNAs and the proteins they encode share about 60% sequence homology.

Direct sequencing of GCMV coat protein and *in vitro* translation of transcripts prepared from cloned cDNA sequences have allowed the localization of the coat protein to the carboxy-terminal part of the polyprotein. These results also allow the tentative identification of the coat protein cleavage site.

MATERIALS AND METHODS

Virus and viral RNA purification, cDNA cloning and analysis

These were performed as described in the accompanying paper.

Subcloning cDNA fragments

Restriction fragments generated by the enzymes TaqI, HincII, SphI, SspI, NdeI, XbaI after digestion of the cDNA of clone p112GC covering the 3′ end of RNA2 were subcloned (3−4). Overlapping subclones were constructed by limited digestion of this plasmid with exonuclease III (5) or by gradual deletions in a full-length cDNA as described (6).

Nucleotide sequence determination and analysis

The techniques used to prepare templates, to run the sequencing reactions and to analyze the nucleotides and amino acids sequences are as described in the accompanying paper. Primer extension of a synthetic oligonucleotide on the viral RNA to confirm 5′ terminal nucleotide sequence was performed as described in the accompanying paper.

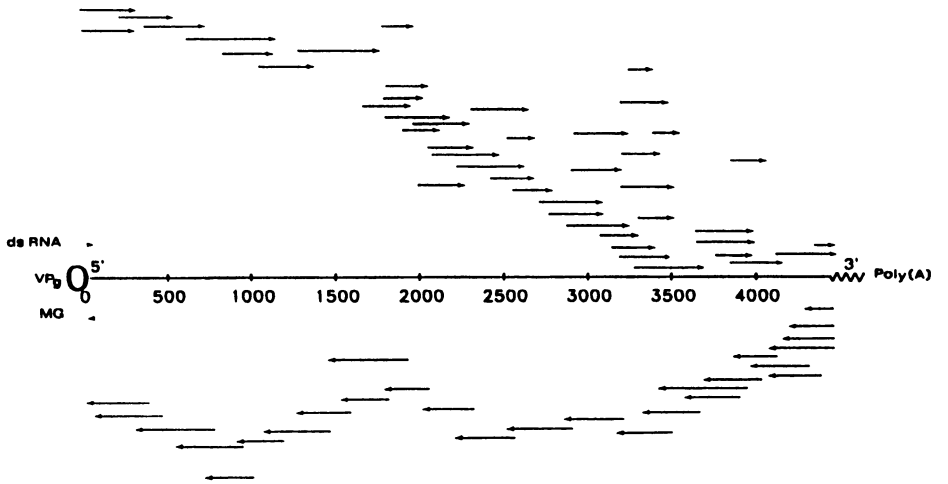


Figure 1: Nucleotide sequencing strategy. The VPg at the 5′ end and the poly(A) tail at the 3′ end of the RNA are indicated. The extent and orientation of the sequences derived from independant cDNA clones are indicated by arrows. The region at the 5′ end of the GCMV RNA2 sequence determined chemically (7) is indicated by **MG. (dsRNA)** is the sequence obtained previously by direct RNA sequencing (2).

Amino-acid sequence determination

The amino-terminal sequence of GCMV coat protein was determined by automated Edman degradation of purified viral particles.

Expression of the coat protein of GCMV in an in vitro translation system

Figure 5 summarizes the procedure for obtaining the GCMV coat protein gene directly linked with GCMV 5′ non-coding sequence. Two constructions (C19 and GC2) linking either of the putative cleavage sites directly in frame with an AUG initiation codon were cloned in plasmid pBS+ (Bluescribe M13+, Stratagene). All recombinant DNA techniques were as described (3). The recombinant plasmids were linearized with HindIII, downstream of the 3′ end of the chimaeric genes, and used as templates for *in vitro* transcription using phage T3 RNA polymerase (BRL) and following the protocol of Promega Biotech.

The transcripts obtained from 250 ng of template were translated in rabbit reticulocyte lysate (Promega), in the presence of ^{35}S-methionine (Amersham, 1200Ci/mmole), according to instructions from the supplier.

For immunoprecipitation, translation products (400,000 cpm) were mixed with either

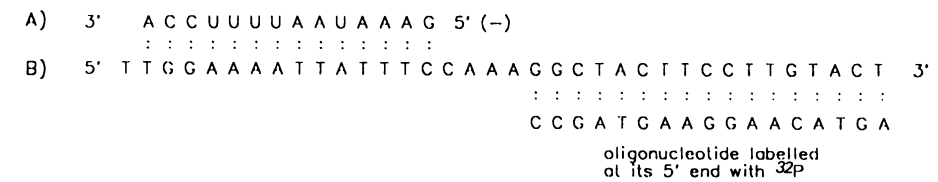


Figure 2: Comparison of the 5′ terminal sequences of RNA2. **A** : 5′ terminal sequences of minus strand RNA obtained by direct dsRNA sequencing (2). **B** : sequence obtained by chemical sequencing of a cDNA to the 5′ end of the viral RNA synthesized using a primer complementary to nucleotide 20 to 36 and shown below.

preimmune or anti-GCMV rabbit antiserum (5μl). The suspensions were shaken overnight at 4°C. *Staphylococus aureus* cells (Pansorbin, Calbiochem) were added and the mixtures further incubated for 2h at room temperature. The suspensions were then centrifuged (5 min, 5000 rpm) and the pellet washed 6 times with 10mM Na-phosphate pH 7.2, 150mM NaCl, 1% Triton X100, 0.5% Na-Deoxycholate and 0.1% SDS. The immunoprecipitates were then denatured by boiling for 5 min in 150mM Tris-HCl pH 6.8, 10% SDS, 25% βmercaptoethanol.

Translation products and immunoprecipitates were analyzed by 12% SDS-polyacrylamide gel electrophoresis followed by autoradiography. Molecular weight markers for gel electrophoresis were purchased from BRL.

RESULTS

Determination of the nucleotide sequence of GCMV RNA2

Using the cDNA cloning conditions of Gubler and Hoffman (7), cDNA clones to RNA2 of GCMV ranging in size from 2.5 kb to 4.5 kb were obtained. A full-length cDNA (FL18) (8) was also used. A restriction map of GCMV RNA2 was obtained by restriction enzyme digestion of the various cDNA clones followed by polyacrylamide or agarose gel electrophoresis.

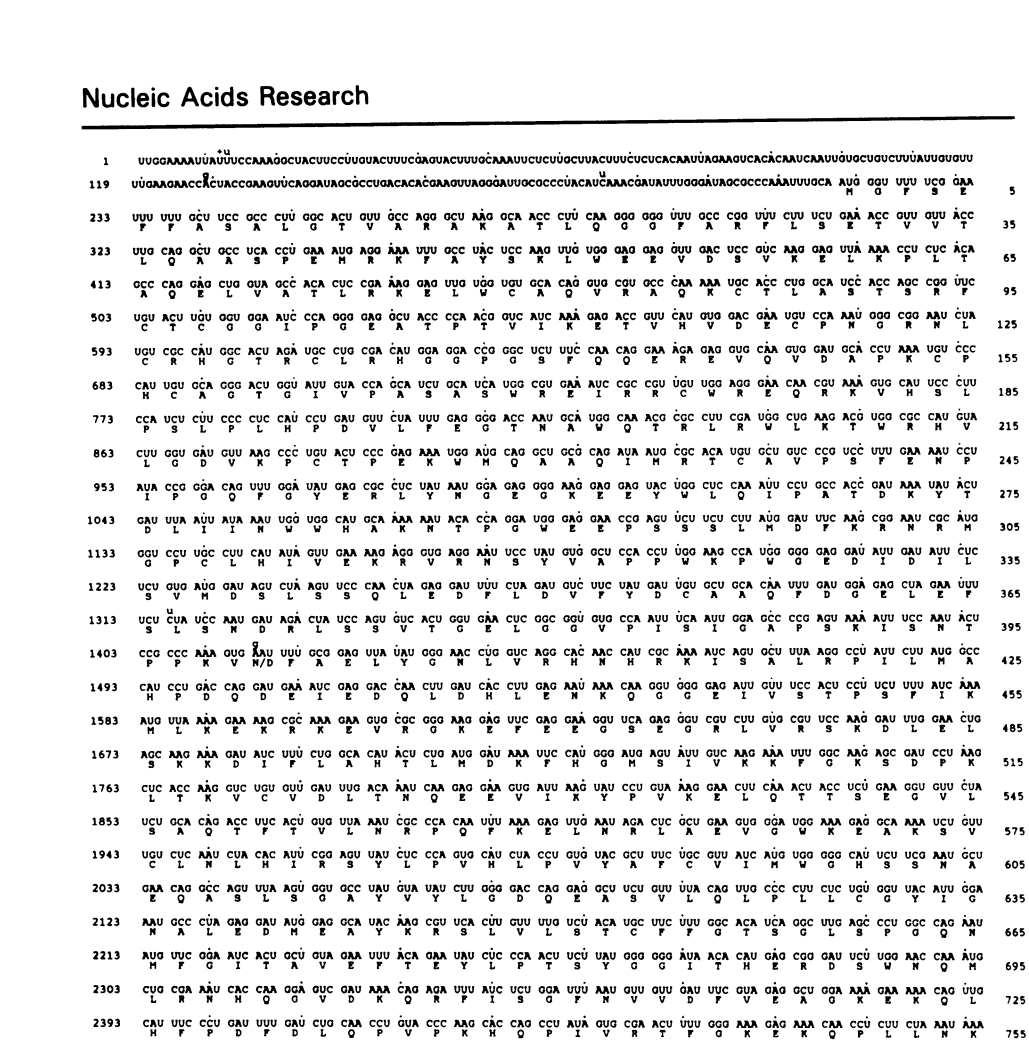
As shown on Figure 1, the nucleotide sequence was completely determined on both strands of the cDNA by supercoiled plasmid sequencing using the Sequenase™ kit (USB). Original clones of the cDNA bank as well as subclones obtained as described in Materials & Methods were used as templates. To ensure that the 5′ end sequence obtained from clone FL18 faithfully represented the end of the viral RNA, primer extension was directly performed on the viral RNA using a synthetic oligonucleotide (AGTACAA-GGAAGTAGCC), complementary to nucleotides 20−36 of the viral RNA. As shown on Figure 2, the sequence obtained is complementary to the 3′ end sequence previously determined for 'minus' strand RNA2 by direct sequencing of double stranded RNAs except for an untemplated U residue at the 5′ end of the 'plus' (virion) strand. This untemplated U is also present in RNA1 of GCMV (see accompanying paper). The VPg is probably linked to this untemplated U residue. The complete nucleotide sequence was finally assembled as described in Materials & Methods.

Primary structure of GCMV RNA2

The assembled GCMV RNA2 nucleotide sequence is 4441 nucleotides in length excluding a variable length (30−50 adenosines) poly(A) tail. The complete nucleotide sequence is shown on Figure 3, along with the amino acid sequence of the long open reading frame found on the RNA. Sequence heterogeneities were observed at three positions in the 5′ non-coding sequence and at three positions (nt 1316, 1415 and 3234) in the coding sequence. It is not known whether those differences reflect errors made during the reverse transcription of the RNA2 or sequence heterogeneities in the RNA populations used. The calculated molecular weight of the RNA is 1.5×10^6, in good agreement with the estimations obtained by denaturing gel electrophoresis.

Open reading frames

The sequence was searched for potential coding regions in all three reading frames of both strands. A single large open reading frame (ORF) has been identified in the 'plus' orientation of the viral RNA beginning at position 218 and extending to an UAG termination codon at position 4190. This open reading frame codes potentially for a protein of 1324 amino acids with a calculated molecular weight of 146 kDa. All other potential ORFs are much



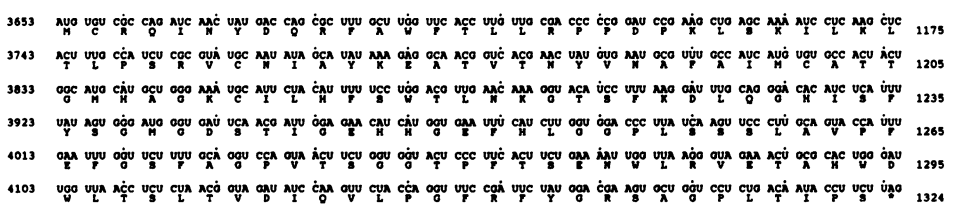


Figure 3: Nucleotide sequence of GCMV RNA2 and deduced amino acid sequence of the large open reading frame. Sequence heterogeneities are shown in low letter case above the sequence. The nucleotide sequence is being deposited at the EMBL data bank.

smaller, the largest one coding for a protein only 122 amino acids long. The AUG codon at position 218 is the first initiation codon found from the 5′ end of the RNA. This codon is in a good context for initiation of translation in plants with G, C and A in positions −3, −2 and −1 respectively and a G in position +4 (9−10). All these results suggest that this codon is indeed the initiation codon used during GCMV RNA2 translation.

The UAG stop codon at position 4190 is followed in phase by two other termination codons. The prediction that GCMV RNA2 encodes one single polyprotein of 146 kDa is confirmed by *in vitro* translation experiments in reticulocyte lysates which show a single translation product of approximately 148 kDa (G. Demangeat, personal communication).

Comparisons between the non-coding regions of GCMV and TBRV-S RNA2

Both the 5′ and 3′ non-coding sequences of GCMV RNA2 are shorter than those of TBRV-S RNA2 whereas the coding sequences have almost the same size, TBRV-S coding for a protein of 1357 amino acids (2). The 3′ non-coding regions are 252 (GCMV) and 304 (TBRV-S) nucleotides long respectively and share a significant homology (73%) (Figure 4A). This high level of homology, in a region which probably contains replication signals may explain, in part, why it is possible to obtain pseudo-recombinants between these two viruses (11).

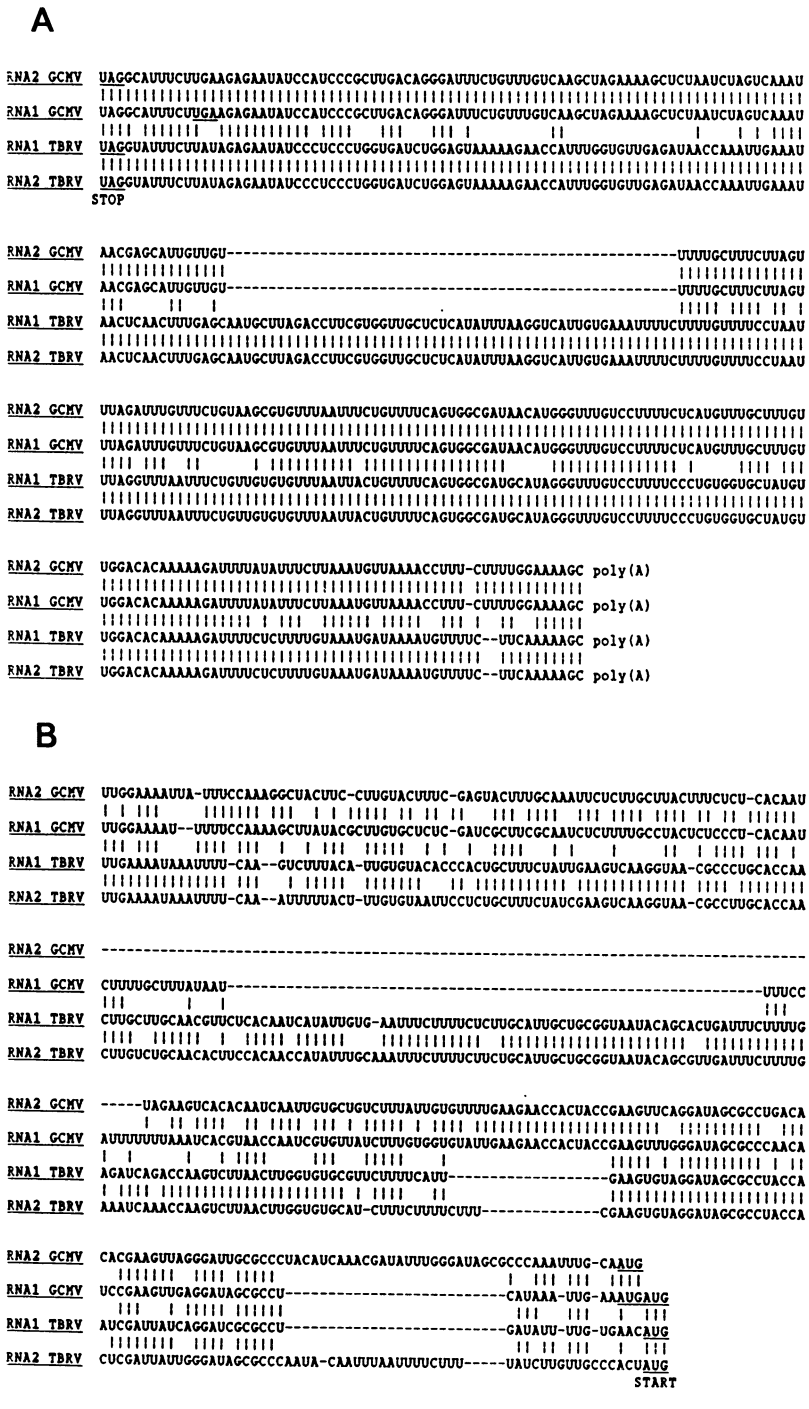
The 5′ non-coding regions are 217 (GCMV) and 300 (TBRV-S) nucleotides long respectively and share only about 50% homology (Figure 4B).

Sequence homologies between the non-coding regions of GCMV RNA1 and RNA2

The 3′ non-coding regions of GCMV RNA1 (accompanying paper) and RNA2 are completely identical but their 5′ non-coding regions, both 217 nucleotides long, are only 68% homologous (Figure 4A−B). These differences in the 5′ non-coding regions might underline differences in translation efficiencies of these two RNAs and thus regulate the respective levels of their protein products in the infected cell.

Localization of the coat protein coding sequence

The NH_2-terminal amino acid sequence of GCMV capsid protein was determined by sequential Edman degradation on purified virions. A sequence, XXXEFAFIHTID (were X represents an unknown amino acid) was obtained, the signal representing only 2% of the input protein. This result indicates that most of the protein in the sample had a blocked NH_2 terminus. The sequence obtained can thus represent the correct end of the capsid protein or, alternatively, the terminal sequence of a minor unblocked contaminant. A search


A
RNA2 GCMV UAGGCAUUUCUUGAAGAGAAUAUCCAUCCCGCUUGACAGGGAUUUCUGUUUGUCAAGCUAGAAAAGCUCUAAUCUAGUCAAAU
RNA1 GCMV UAGGCAUUUCUUGAAGAGAAUAUCCAUCCCGCUUGACAGGGAUUUCUGUUUGUCAAGCUAGAAAAGCUCUAAUCUAGUCAAAU
RNA1 TBRV UAGGUAUUUCUUAUAGAGAAUAUCCCUCCCUGGUGAUCUGGAGUAAAAAGAACCAUUUGGUGUUGAGAUAACCAAAUUGAAAU
RNA2 TBRV UAGGUAUUUCUUAUAGAGAAUAUCCCUCCCUGGUGAUCUGGAGUAAAAAGAACCAUUUGGUGUUGAGAUAACCAAAUUGAAAU
STOP
RNA2 GCMV AACGAGCAUUGUUGU----------UUUUGCUUUCUUAGU
RNA1 GCMV AACGAGCAUUGUUGU----------UUUUGCUUUCUUAGU
RNA1 TBRV AACUCAACUUUGAGCAAUGCUUAGACCUUCGUGGUUGCUCUCAUAUUUAAGGUCAUUGUGAAAUUUUCUUUUGUUUUCCUAAU
RNA2 TBRV AACUCAACUUUGAGCAAUGCUUAGACCUUCGUGGUUGCUCUCAUAUUUAAGGUCAUUGUGAAAUUUUCUUUUGUUUUCCUAAU
RNA2 GCMV UUAGAUUUGUUUCUGUAAGCGUGUUUAAUUUCUGUUUUCAGUGGCGAUAACAUGGGUUUGUCCUUUUCUCAUGUUUGCUUUGU
RNA1 GCMV UUAGAUUUGUUUCUGUAAGCGUGUUUAAUUUCUGUUUUCAGUGGCGAUAACAUGGGUUUGUCCUUUUCUCAUGUUUGCUUUGU
RNA1 TBRV UUAGGUUUAAUUUCUGUUGUGUGUUUAAUUACUGUUUUCAGUGGCGAUGCAUAGGGUUUGUCCUUUUCCCUGUGGUGCUAUGU
RNA2 TBRV UUAGGUUUAAUUUCUGUUGUGUGUUUAAUUACUGUUUUCAGUGGCGAUGCAUAGGGUUUGUCCUUUUCCCUGUGGUGCUAUGU
RNA2 GCMV UGGACACAAAAAGAUUUUAUAUUUCUUAAAUGUUAAAACCUUU-CUUUUGGAAAAGC poly(A)
RNA1 GCMV UGGACACAAAAAGAUUUUAUAUUUCUUAAAUGUUAAAACCUUU-CUUUUGGAAAAGC poly(A)
RNA1 TBRV UGGACACAAAAAGAUUUUCUCUUUUGUAAAUGAUAAAAUGUUUUC--UUCAAAAAGC poly(A)
RNA2 TBRV UGGACACAAAAAGAUUUUCUCUUUUGUAAAUGAUAAAAUGUUUUC--UUCAAAAAGC poly(A)
B
RNA2 GCMV UUGGAAAAUUA-UUUCCAAAGGCUACUUC-CUUGUACUUUC-GAGUACUUUGCAAAUUCUCUUGCUUACUUUCUCU-CACAAU
RNA1 GCMV UUGGAAAAU--UUUUCCAAAAGCUUAUACGCUUGUGCUCUC-GAUCGCUUCGCAAUCUCUUUUGCCUACUCUCCCU-CACAAU
RNA1 TBRV UUGAAAAUAAAUUUU-CAA--GUCUUUACA-UUGUGUACACCCACUGCUUUCUAUUGAAGUCAAGGUAA-CGCCCUGCACCAA
RNA2 TBRV UUGAAAAUAAAUUUU-CAA--AUUUUUACU-UUGUGUAAUUCCCUCUGCUUUCUAUCGAAGUCAAGGUAA-CGCCUUGCACCAA
RNA2 GCMV ----------
RNA1 GCMV CUUUUGCUUUAUAAU----------UUUCC
RNA1 TBRV CUUGCUUGCAACGUUCUCACAAUCAUAUUGUG-AAUUUCUUUUCUCUUGCAUUGCUGCGGUAAUACAGCACUGAUUUCUUUUG
RNA2 TBRV CUUGUCUGCAACACUUCCACAACCAUAUUUGCAAAUUUCUUUUCUUCUGCAUUGCUGCGGUAAUACAGCGUUGAUUUCUUUUG
RNA2 GCMV -----UAGAAGUCACACAAUCAAUUGUGCUGUCUUUAUUGUGUUUUGAAGAACCACUACCGAAGUUCAGGAUAGCGCCUGACA
RNA1 GCMV AUUUUUUUAAAUCACGUAACCAAUCGUGUUAUCUUUGUGGUGUAUUGAAGAACCACUACCGAAGUUUGGGAUAGCGCCCAACA
RNA1 TBRV AGAUCAGACCAAGUCUUAACUUGGUGUGCGUUCUUUUCAUU----------GAAGUGUAGGAUAGCGCCUACCA
RNA2 TBRV AAAUCAAACCAAGUCUUAACUUGGUGUGCAU-CUUUCUUUUCUUU----------CGAAGUGUAGGAUAGCGCCUACCA
RNA2 GCMV CACGAAGUUAGGGAUUGCGCCCUACAUCAAACGAUAUUUGGGAUAGCGCCCAAAUUUG-CAAUG
RNA1 GCMV UCCGAAGUUGAGGAUAGCGCCU----------CAUAAA-UUG-AAAUGAUG
RNA1 TBRV AUCGAUUAUCAGGAUCGCGCCU----------GAUAUU-UUG-UGAACAUG
RNA2 TBRV CUCGAUUAUUGGGAUAGCGCCCAAUA-CAAUUUAAUUUUCUUU-----UAUCUUGUUGCCCACUAUG
START

of the sequence of the predicted translation product of GCMV RNA2 revealed the sequence AGGEFAFIHTID at position 811−822. A protein containing this sequence and ending at the carboxy terminus of the polyprotein (position 1324) would be 514 amino acids long and would have a molecular weight of 56.7 kDa, as compared to a molecular weight of 52 kDa estimated by SDS-polyacrylamide gel electrophoresis for the coat protein isolated from virions. This carboxy-terminus position for the coat protein would also be in agreement with the results of Meyer *et al.* (2) who positioned the coat protein of TBRV-S in the same region on the basis of amino acid composition. However, the dipeptide cleavage site yielding the AGGEFAFIHTID NH_2-end is an arginine/alanine (R/A), a site not so far observed in the processing of picorna-like RNA viruses polyproteins (12). Search for known plant virus protease cleavage sites in that region revealed a glutamine/alanine (Q/A) dipeptide at position 855−856, which would give a 52 kDa product in better agreement with the estimated molecular weight of the capsid protein.

These conflicting results prompted us to try to identify the location of the coat protein and of its cleavage site in the RNA2-encoded polyprotein. As described in Materials and Methods, two constructions were obtained, allowing the expression, under the control of GCMV 5′ non-coding region, of proteins having (except for an additional methionine) either of the two putative NH_2 termini we had determined (Figure 5). These constructions were transcribed *in vitro* and the transcripts were further translated in a rabbit reticulocyte lysate *in vitro* translation system as described in Materials and Methods. The ^{35}S labelled translation products were analyzed by SDS-polyacrylamide gel electrophoresis followed by autoradiography before (Figure 6, tracks 1 and 4), or after (Tracks 2 and 5) immunoprecipitation with anti-GCMV rabbit immunoglobulins. As can be seen, the polypeptides obtained upon *in vitro* translation of both constructions are immunoprecipitated by anti-virions immunoglobulins and the protein corresponding to the R/A cleavage site pinpointed by direct protein sequencing (track 1) has the same electrophoretic mobility as coat protein extracted from the virions (noted CP). Taken together, these results demonstrate that GCMV coat protein corresponds to the carboxy-terminal part of the RNA2-encoded polyprotein and that the cleavage site freeing it from the polyprotein is probably the R/A dipeptide located at position 810−811. We cannot completely rule out the possibility of the primary (viral) cleavage being located just upstream of the R/A site and followed by a secondary (host cell) serine protease cleavage at the R/A site.

Comparison of the amino acid sequence of the polyproteins coded by GCMV and TBRV-S RNA2

The results of *in vitro* translation experiments (G. Demangeat, personal communication) show that the GCMV RNA2 polyprotein is cleaved by an RNA1-encoded protease to yield the mature coat protein and an 84 kDa protein which is further cleaved into two products of approximately 46 and 48 kDa. This mechanism for the expression of the coat protein seems to be general in nepoviruses since similar results have been reported for GFLV, TobRV and TBRV (13−14 and C. Fritsch personal communication). So far, a search of the sequence of TBRV and GCMV proteins has not yielded any known viral protease sites that could account for the 84 kDa protein cleavage.

No significant homologies have been observed between GCMV RNA2 polyprotein and

Figure 4: Nucleotide sequence homologies between GCMV and TBRV-S RNA1 and RNA2 non coding regions. **A** : 3′ ends homologies . **B** : 5′ ends homologies .

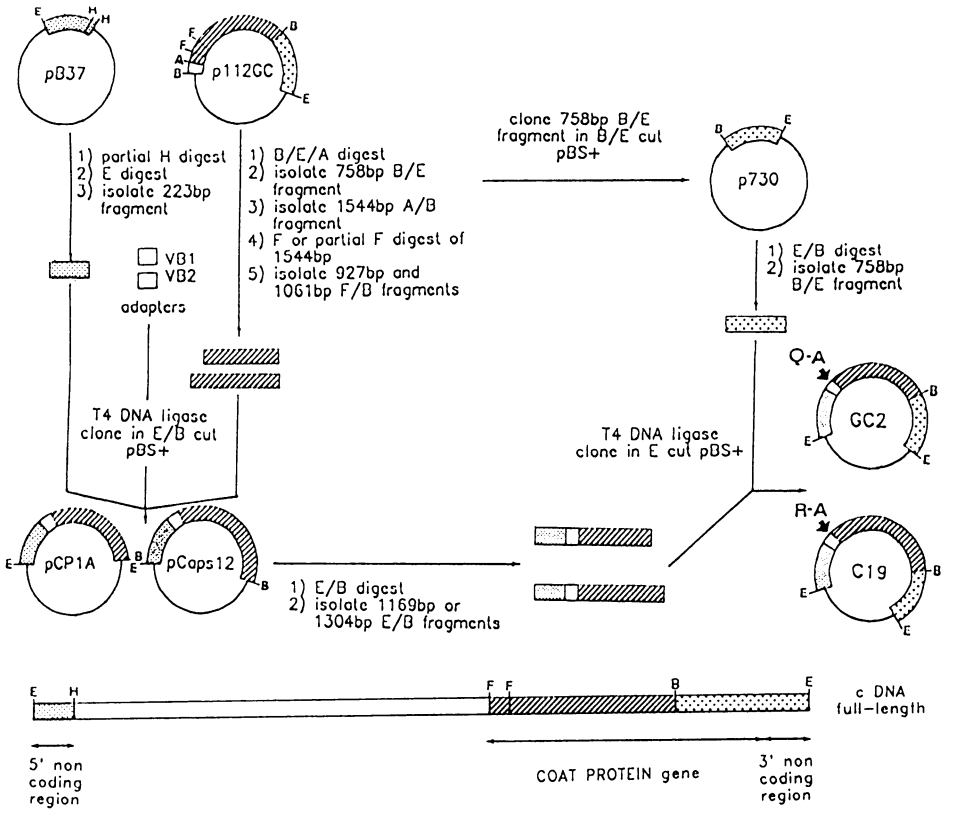


Figure 5: Construction of plasmids GC2 and C19 containing chimaeric GCMV coat protein genes. GCMV RNA2 cDNA is denoted by solid line. From plasmid pB37, a 223bp fragment containing 207 nucleotides of the 5′ non-coding region of GCMV RNA2 and 16 nucleotides from the polylinker was isolated following EcoRI and partial HaeII digestion. From plasmid p112GC, a 758bp BamHI/EcoRI fragment (from nucleotide 3715 to nucleotide 4441 plus a A_{16} tail and 16 nucleotides derived from the polylinker) containing the carboxy terminal end of the coat protein gene and the 3′ non coding region of GCMV RNA2 was isolated and subcloned in pBS+. Also from plasmid p112GC, a BamHI/BamHI fragment of 1544bp (from nucleotide 2170 to nucleotide 3714) was isolated and submitted to partial FokI digestion. Two fragments were isolated in this way. An 927bp fragment (nucleotides 3714 to 2787 and containing the amino end of the coat protein cleaved at the Q/A cleavage site) and a 1061bp fragment (nucleotides 3714 to 2653 and containing the amino end of the coat protein starting from the R/A cleavage site). Plasmids pCaps12 and pCP1A were obtained by four point ligations in EcoRI/BamHI digested pBS+ of the 223 bp EcoRI/HaeII fragment from plasmid pB37, one of the the FokI/BamHI fragments derived from plasmid p112GC, and synthetic oligonucleotides adapters VB1 or VB2 restoring the end of the 3′ non-coding region and the beginning of the two possible coat protein coding sequences. Oligonucleotide **VB1** (CCAAGTCGACAATGGCGGTGCGCGGGTTCAGCTGTTACCGCCCACCACT) was used to link the 1061bp and 223bp fragments and oligonucleotide **VB2** (CCAAGTCGACAATGGCTGGGCGCGGGTTCAGCTG-TTACCGACCCATAC) to link the 927bp and 223bp fragments. The final constructions, plasmids GC2 and C19 were respectively obtained by ligating the BamHI/EcoRI fragment from plasmid p730 to the BamHI/EcoRI fragments from plasmids pCaps12 or pCP1A and cloning in EcoRI digested pBS+. **F**, **B**, **A**, **H**, and **E** : restriction enzymes FokI, BamHI, AccI, HaeII and EcoRI, respectively. At the bottom of the figure, the fragments used to make the hybrid genes, are located on the full-length cDNA.

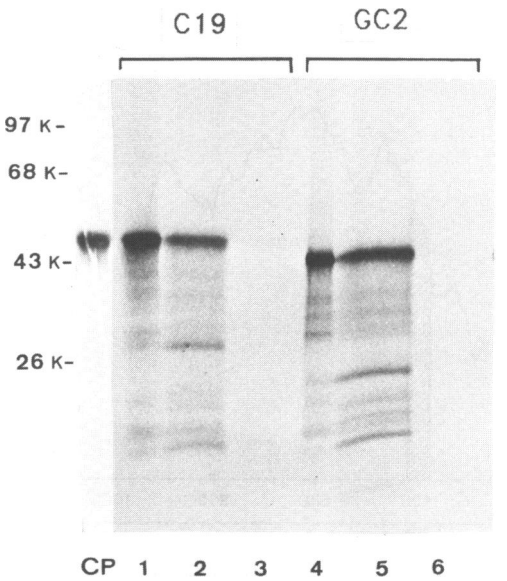


Figure 6: *In vitro* translation of *in vitro* transcripts and immunoprecipitation of translational products. Plasmids GC2 and C19 restricted with HindIII were used as template for phage T3 RNA polymerase. Transcripts were translated in reticulocyte lysate with ^{35}S Methionine (lanes 1 and 4). Samples were immunoprecipitated with a GCMV-specific antiserum (lanes 2 and 5) or a non-immune serum (lanes 3 and 6). Authentic capsid (CP) was stained with Coomassie brillant blue.

the polyprotein of CPMV M RNA or any of the other viral proteins that were assayed, except for TBRV-S RNA2-encoded polyprotein. These two proteins are 1324 and 1357 amino acids long respectively and share an overall 60% homology. The two proteins are strictly colinear except for one 28-amino acid gap in the sequence of the GCMV polyprotein corresponding to amino acids 388−416 of the TBRV-S polyprotein. A curve plotting the percentage of homology between the two proteins is presented on Figure 7. The position of the R/A capsid cleavage site and the estimated position of the putative cleavage site inside the 84 kDa protein are also presented on this figure. It can be observed that the central portion of the polyprotein is more highly conserved than either the amino or carboxy termini. The division into three domains with different levels of homology fits roughly the cleavage of the polyprotein into three mature products. It has been reported (2) that the central, most conserved region (590−800) of TBRV protein has reduced local homologies with proteins involved in viral cell-to-cell movement (30K protein of tobacco mosaic virus and the 48/58K protein of cowpea mosaic virus). However, these homologies are too low to allow definitive assignment of such a function to this region of the polyproteins of GCMV or TBRV.

The coat proteins, at the C-terminus of the polyprotein, occupy a domain of high global hydrophobicity (not shown) and of reduced homology (54% average), which could explain the hydrophobic properties of these proteins and the very low serological relations that exist between GCMV and TBRV-S.

So far, no function has been attributed to the less conserved protein (43% homology average) located in the NH_2-terminal domain of the polyprotein.

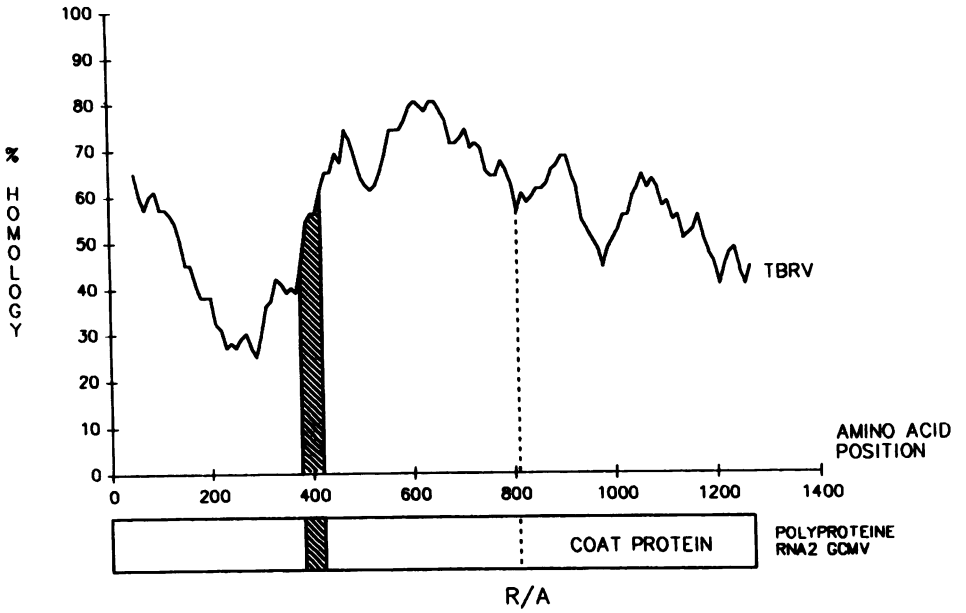


Figure 7: Amino acid homologies between the polyproteins encoded by GCMV RNA2 and TBRV-S RNA2. Percent homology are calculated in a window of 100 residues. On the GCMV polyprotein at the bottom of the figure are located the putative coat protein and its cleavage site R/A determined by chemical sequencing of the NH_2-end. The hatched aera represents a sequence of amino acids where cleavage of the 84 kDa protein should occur to yield 46 kDa and 48 kDa proteins.

DISCUSSION

We have determined the nucleotide sequence of hungarian grapevine chrome mosaic nepovirus RNA2. As previously observed for the closely related tomato black ring virus, GCMV RNA2 harbors one single large open reading frame encoding a polyprotein which is processed by an RNA1-encoded protease. A similar expression mechanism is used by cowpea mosaic virus (CPMV), the type member of the comovirus group (15). The homologies between nepoviruses and comoviruses even extend further since our results demonstrate that the coat protein of GCMV is located at the carboxy-terminal end of the RNA2 polyprotein, as is also the case for the coat proteins of CPMV. However, the RNA2 polyprotein of nepoviruses has the capacity to encode three proteins in contrast to only two proteins for the comoviruses since the protein located at the NH_2-terminal end of the nepovirus polyprotein has no counterpart in the comovirus genome. No significant homologies have been observed between this putative protein and any accession in the PIR database and, at the moment, we do not have any clues as to what might be its function *in vivo*. One hypothesis is that this protein could be required for nematode transmission of the virus since this characteristic has been found to be associated with RNA2 in another nepovirus, raspberry ringspot virus (16).

The arginine/alanine dipeptide that we have tentatively assigned as the cleavage site liberating the capsid protein from its precursor is quite unusual and had not previously been observed in other viruses belonging to the 'picorna-like' superfamily of viruses (12). This might be explained by the difference in structure of GCMV protease as discussed

in the accompanying paper. In this respect, it would be interesting to determine precisely the coat protein cleavage site in the case of TBRV because this virus appears to have a slightly larger coat protein (56 kDa as estimated by SDS-polyacrylamide gel electrophoresis) and since the R/A dipeptide is not conserved between the two viruses.

Results in our laboratory have shown that, during cross protection experiments in *Chenopodium quinoa*, the two viruses seem to replicate independently and that, in particular, the severe, superinfecting TBRV is not affected in its replication by the presence of the mild, cross-protecting GCMV (17). This situation is in contrast with other cross protection systems in which infection of the plant by and replication of the severe strain are drastically reduced by the presence of the protecting strain. We are currently engineering the GCMV coat protein expressing constructs for transformation of tobacco plants. It will be of interest to see if, as was observed for other models (18−19−20), a protection against TBRV can be obtained in this way and if the behaviour of TBRV in these transgenic plants is similar to its behaviour in classically cross-protected plants.

ACKNOWLEDGEMENTS

We would like to thank C. Fritsch and members of her group for communicating results before publication and Kathryn Mayo-Candresse for help with the English of the manuscript. V.B. and H.L. are supported by grants from the French Ministry of Science Research and of University Education.

REFERENCES

1. Martelli, G.O. and Quacquarelli, A. (1972) in Murant, A.F. and Harrisson, B.D. (Eds), C.M.I./A.A.B. Descriptions of plant viruses, Old Woking, England, # 103.
2. Meyer, M., Hemmer, O., Mayo, M.A., and Fritsch, C. (1986) J. Gen. Virol. 67, 1257−1271.
3. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning : A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
4. Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. (1987) Current protocols in molecular biology, Greene publishing associates and Wiley-interscience, New-York.
5. Henikoff, S. (1984) Gene 28, 351−359.
6. Lin, H.C., Lei, S.P., and Wilcox, G. (1985) Anal. Biochem. 147, 114−119.
7. Gubler, U. and Hoffman, B.J. (1983) Gene 25, 263−269.
8. Le Gall, O., Candresse, T., Brault, V., Bretout, C., Hibrand, L., and Dunez, J. (1988) Gene 73 (1), 67−75.
9. Lütcke, H.A., Chow, K.C., Mickel, F.S., Moss, K.A., Kern, H.F., and Sceele, G.A. (1987) EMBO J. 6, 43−48.
10. Kozak, M. (1987) J. Mol. Biol. 196, 947−950.
11. Doz, B., Delbos, R., and Dunez, J. (1982) Les colloques de l'INRA 11, 29−44.
12. Wellink, J., and Van Kammen, A., (1988) Arch. Virol. 98, 1−26.
13. Morris-Krsinich, B.A.M., Forster, R.L.S., and Mossop, D.W. (1983) Virology 130, 523−526.
14. Forster, R.L.S., and Morris-Krsinich, B.A.M. (1985) Virology 144, 516−519.
15. Franssen, H., Goldbach, R., Broekhuisen, M., Moerman, M., and Van Kammen, A. (1982) J. Gen. Virol. 41, 8−17.
16. Harrison, B.D., Murant, A.F., Mayo, M.A., and Roberts, I.M. (1974) J. Gen. Virol. 22, 233−247.
17. Bretout, C. (1987) Contribution à l'étude du phénomène de la prémunition entre deux souches du virus des anneaux noirs de la tomate (TBRV) sur *Chenopodium quinoa* Wild. Ph.D. thesis, Université de Bordeaux II.
18. Powell-Abel, P., Nelson, R.S., De, B., Hoffmann, H., Rogers, S.G., Fraley, R.T. and Beachy, R.N. (1986) Science 232, 738−743.
19. Tumer, N.E., Clark, W.G., Tabor, G.J., Hironaka, C.M., Fraley, R.T. and Shah, D.M. (1986) Nucl. Acid. Res. 8, 3325−3342.
20. Van Dun, C.M.P., John, F., Bol and Lous Vloten-doting (1987) Virology 159, 299−305.

This article, submitted on disc, has been automatically converted into this typeset format by the publisher.