PROTEOLYTIC PROCESSING OF THE PRIMARY TRANSLATION PRODUCTS OF COWPEA MOSAIC VIRUS RNAs

Henk Franssen

Gedrukt door Offsetdrukkerij Kanters B.V., Alblasserdam.

In Memory of HTR

aan de Plak

Veur Spaik

Promoter : Dr. A. van Kammen, hoogleraar in de moleculaire biologie
Co-promoter: Dr. R.W. Goldbach, wetenschappelijk medewerker.

Henk Franssen

PROTEOLYTIC PROCESSING OF THE PRIMARY TRANSLATION PRODUCTS OF COWPEA MOSAIC VIRUS RNAs

Proefschrift
ter verkrijging van de graad van
doctor in de landbouwwetenschappen,
op gezag van de rector magnificus,
dr. C.C. Oosterlee,
in het openbaar te verdedigen
op dinsdag 16 oktober 1984
des namiddags te vier uur in de aula
van de Landbouwhogeschool te Wageningen.

ISN = 154 534-03

CONTENTS

toverd tot een ooglijk geheel, waarvoor mijn dank.
Last but not least wil ik jullie bedanken, alhoewel jullie geheel buiten het CPMV-gebeuren stonden, toch de ups en downs die aan het promotieonderzoek verbonden waren, met mij hebben gedeeld. Aan jullie wil ik dan ook dit proefschrift opdragen.

STELLINGEN

1. De aanwezigheid van een weefselspecifieke enhancer-sequentie in het L-V intron van het muizen κ-immunoglobuline gen verklaart waarom er wel correcte transcriptie van het gecloneerde gen in muizemyeloma-cellen plaatsvindt, maar niet in apeniercellen.

   Gillies, S.D., Morrison, S.L., Di, V.T., and Tonegawa, S., 1983. Cell 33, 717-728.
   Queen, C. and Baltimore, D., 1983. Cell 33, 741-748.

2. Het is erg voorbarig van Morris-Krsinich *et al.* (1983) om op grond van de door hen verkregen resultaten te concluderen dat het eiwit met een molecuul gewicht van 58.000, dat ze vinden bij *in vitro* translatie van RNA-2 van grapevine fanleaf virus, het manteleiwit van dit virus is.

   Morris-Krsinich, B.A.M., Forster, R.L.S., and Mossop, D.W., 1983. Virology 130, 523-526.

3. Bij de berekening van het percentage homologie in de aminozuur-volgorde van de door poliovirus en mond- en klauwzeervirus gecodeerde RNA polymerases is door Robertson *et al.* (1983) een fout gemaakt, waardoor ze uitkomen op een te laag percentage.

   Robertson, B.H., Morgan, D.O., Moore, D.M., Grubman, M.J., Card, J., Fischer, T., Weddell, G., Dowbenko, D., and Yansura, D., 1983. Virology 126, 614-623.

4. De uitspraak van Bresser *et al.* (1983) dat boodschapper RNA uit cellen, die in 12,2 molair natriumjodide geïmmobiliseerd zijn op nitrocellulose nog vertaalbaar is, wordt niet door experimentele gegevens gestaafd.

Bresser, J., Hubbell, R., and Gillespie, D., 1983.
Proc. Nat. Acad. Sci. 80, 6523-6527.

5. De conclusie van Peng en Shih (1984) dat het door B-RNA van cowpea mosaic virus gecodeerde 87K eiwit afkomstig is van het amino-eindstandige gedeelte van het 170K precursoreiwit is niet alleen fout, maar de door hen verkregen resultaten ondersteunen juist de conclusie van Rezelman *et al.* (1980) dat het 87K eiwit gelegen is in het carboxyleindstandige gedeelte van het 170K eiwit.

   Peng, X.X., and Shih, D.S., 1984. J. Biol. Chem. 259, 3197-3201.
   Rezelman, G., Goldbach, R., and Van Kammen, A., 1980. J. Virol. 36, 366-373.

6. De gepubliceerde aminozuurvolgorden voor een eiwit uit *Saccharamyces cerevisiae* dat sequentie homologie vertoont met het door c-ras^{H} proto-oncgen van de mens gecodeerde p21 eiwit, wijzen erop dat er in *Saccharamyces cerevisiae* twee verschillende genen voorkomen die sequentie homologie hebben met het c-ras^{H} gen.

   Gallwitz, D., Donath, C., and Sander, C., 1983. Nature 306, 704-707.
   Defeo-Jones, D., Scolnick, E.M., Koller, R., and Dhar, R., 1983. Nature 306, 707-709.

7. Het verdient aanbeveling om voor onderzoek gericht op de verbetering van de diagnose van lymfomen bij de mens, die geinduceerd worden door leukemie-virus en voor de ontwikkeling van een immunotherapie daartegen, de kat als modelsysteem te gebruiken.

8. De proeven van Schärli en Koch (1984) bewijzen niet dat poliovirusdeeltjes ꝑrotein-kinase activiteit bezitten.

   Schärli, C.E., and Koch, G., 1984. J. Gen. Virol. 65, 129-139.

9. Het bedrijven van politiek vanuit een christelijke levensvisie is innerlijk in tegenspraak met elkaar.

   Jacobus 1:27; 4:4, Het Nieuwe Testament, Willibrord-Vertaling. CDA manifest "Program van uitgangspunten van het christen-democratisch appel".

10. Er dient werk gemaakt te worden van een herbezinning op de inhoud van het begrip "arbeid verrichten".

11. Overleven is erger dan sterven.

12. De televisieprogramma's van de Nederlandse omroepverenigingen zetten aan tot zwart kijken.

Stellingen bij het proefschrift "Proteolytic processing of the primary translation products of cowpea mosaic virus RNAs".

Henk Franssen.

Wageningen, 16 oktober 1984.

CHAPTER I

SCOPE OF THE INVESTIGATION

Cowpea mosaic virus (CPMV) is the type member of a group of plant viruses, the comoviruses, with a genome consisting of two single stranded RNA molecules separately encapsidated in icosahedral particles. A characteristic feature of the two genome RNAs is that they are both polyadenylated at their 3'-terminus and supplied with a small protein at their 5'end. The genetic information encoded in the virus RNAs is expressed by translation of each RNA into large-sized proteins referred to as polyproteins because these primary translation products are subsequently cleaved by specific proteolytic cleavages ("proteolytic processing") into a number of smaller-sized proteins, each with a specific function during virus multiplication. The research reported in this thesis deals with the identification of the proteolytic activities involved in this processing and their specificity.

We have been able to demonstrate that the larger of the two virus RNAs, which contains the information necessary for virus RNA replication, also encodes two different proteolytic activities. One proteolytic activity is responsible for the cleavage of the overlapping polyproteins produced by the smaller of two virus RNAs and releases the two capsid proteins, encoded by this RNA (Chapter III and V), whereas the other proteolytic activity achieves the processing of the polyprotein produced by the larger RNA (Chapter VII). Besides this functional difference the two proteolytic activities recognise peptide bounds between different specific amino acid pairs (Chapter VI and VIII). The results of our studies have led to a detailed model for the processing of the proteins encoded on the two CPMV RNAs.

The striking analogy between the plant comoviruses and the animal picornaviruses, like poliovirus and foot-and-mouth-disease virus, with regard to genome structure, replication, expression strategy and functional organisation of genes has prompted us to study the homology in amino acid sequences between corresponding proteins of the two groups of virus. It was found that some of the non-structural proteins of CPMV and the picornaviruses exhibit significant homology in amino acid se-

quence (Chapter VIII). These results suggest that animal picornaviruses and plant comoviruses have a common ancestor and throw a light on the evolution of RNA viruses.

CHAPTER II

GENERAL INTRODUCTION

1. *Classification of cowpea mosaic virus.*

Based on their type of genome and on their expression mechanism RNA viruses have been divided in four classes (Baltimore, 1971; for an extended description see Luria *et al.*, 1978). One of these classes are the double-stranded RNA viruses which produce upon infection mRNAs by asymmetrical transcription of their double-stranded genome. Examples of this type of viruses are the animal reoviruses and the plant viruses rice dwarf virus and wound tumor virus. A second class of RNA viruses, the retroviruses, encapsidate single-stranded RNA which is copied in the host cell into DNA, that acts as template for the synthesis of viral mRNAs and progeny virus RNA. So far, retroviruses have only been found in animals and not in plants. The third and fourth class of RNA viruses are the positive- and negative-strand RNA viruses respectively. Both classes have single-stranded RNA genomes but the difference is that the genomic RNA of positive-strand viruses can directly act as messenger RNA, whereas the genomic RNA of the negative-strand RNA viruses is not messenger-sense. Particles of negative-strand RNA viruses contain a virus-encoded RNA polymerase which upon infection transcribes the genomic RNA into mRNAs. Such RNA viruses are frequently found among animal viruses, but rarely among plant viruses. Examples are the animal rhabdoviruses (e.g. vesicular stomatitis virus), orthomyxo viruses (e.g. influenza virus) and paramyxoviruses (e.g. sendai virus), and plant rhabdoviruses (e.g. lettuce necrotic yellow virus and sowthistle yellow vein virus). Important groups of positive-strand RNA viruses are the animal picornaviruses, e.g. poliovirus and foot-and-mouth disease virus, and togaviruses while the vast majority of plant viruses also belongs to this class (for a recent, extensive review on RNA plant viruses see Dougerthy and Hiebert, 1984).

The positive-strand RNA viruses can be further divided into two subclasses. One subclass comprises those viruses which generate, in addition to the full-length virus RNA, one or more subgenomic RNAs. Typical

examples of this subclass are the animal togaviruses (e.g. semliki forest virus) and many groups of plant viruses such as the tobamoviruses (e.g. tobacco mosaic virus), tymoviruses (e.g. turnip yellow mosaic virus), bromoviruses (e.g. brome mosaic virus), cucumoviruses (e.g. cucumber mosaic virus) and ilarviruses (e.g. tobacco streak virus).

Viruses of the other subclass of positive-strand RNA viruses, produce only genome-length RNA chains which are translated into large primary translation products. These large proteins are then proteolytically cleaved to generate the functional proteins. Representatives of this subclass are the animal picornaviruses and plant comoviruses. CPMV thus belongs to this second subclass of positive-strand RNA viruses.

Genome structure of plant RNA viruses.

A number of positive strand RNA viruses have their genetic information distributed among two or three single-stranded RNA molecules. With animal viruses this occurs rarely and has so far only been found for the nodaviruses (e.g. nodamura virus and black beetle virus; Friesen and Rueckert, 1981). The two single-stranded RNA molecules of nodaviruses have no base sequences in common (Clewley *et al.*, 1982) and are both required for infectivity (Friesen and Rueckert, 1982). On the other hand, among plant viruses a divided RNA genome is quite common. Moreover, if the genetic information is distributed among two or three positive-strand RNA molecules, these genome segments are almost always separately encapsidated. Ten out of twenty-six different groups of plant viruses have bipartite or tripartite RNA genomes. For plant viruses with a bipartite genome, both RNA molecules are necessary for infectivity and also in case of viruses with a tripartite genome the complete set of RNAs is required for infectivity. Table II.1 shows some examples of mono-, bi- and tripartite RNA plant viruses, and also includes data on the structural features of the virus RNAs and the mechanism used for their translational expression. Three different types of 5'-terminal structures have been found, a m^7Gppp (a cap), a di- or triphosphate and a protein, VPg (= Virus Protein genome-bound),

Table II.1 Structural features of genomic RNAs and mode of expression of single-stranded RNA plantviruses

VIRUS GROUP	TYPE MEMBER VIRUS	TERMINAL STRUCTURES OF GENOMIC RNAs		MODE OF EXPRESSION
		5'end	3'end	
monopartite genome viruses				
Tymoviruses	Turnip yellow mosaic virus(TYMV)	cap	tRNA	generation of subgenomic mRNAs
Sobemoviruses	Southern bean mosaic virus(SBMV)	Vpg	no poly(A)	idem + proteolytic processing of precursor proteins
Tombusviruses	Tomato bushy stunt virus (TBSV)	?	?	?
Tobacco necrosis virus group	Tobacco necrosis virus (TNV)	ppx	?	?
Luteoviruses	Barley yellow dwarf virus (BYDV)	Vpg	no poly(A)	?
Tobamoviruses	Tobacco mosaic virus (TMV)	-	tRNA	generation of subgenomic mRNAs
Potexviruses	Potato virus X (PVX)	cap	no poly(A)	?
Potyviruses	Potato virus Y	Vpg	poly(A)	?
Closteriviruses	Sugar beet yellow virus	?	?	?
bipartite genome viruses				
Tobraviruses	Tobacco rattle virus (TRV)	cap	no poly(A)	generation of subgenomic mRNAs
Comoviruses	Cowpea mosaic virus (CPMV)	Vpg	poly(A)	proteolytic processing of precursor proteins
Nepoviruses	Tobacco ringspot virus (TRSV)	Vpg	poly(A)	idem
Hordeiviruses	Barley stripe mosaic virus(BSMV)	cap	tRNA	
Dianthoviruses	Carnation ringspot virus (CaRV)	?	?	?
Pea enation mosaic virus group	Pea enation mosaic virus (PEMV)	Vpg	?	?
tripartite genome viruses				
Bromoviruses	Brome mosaic virus (BMV)	cap	tRNA	generation of subgenomic mRNAs
Cucumoviruses	Cucumber mosaic virus (CMV)	cap	tRNA	idem
Ilarviruses	Tobacco streak virus (TSV)	?	?	idem
Alfalfamosaic virus group	Alfalfamosaic virus (AMV)	cap	tRNA	idem
Tomato Spotted Wilt virus group	Tomato spotted wilt virus (TSWV)	?	?	?

respectively. RNA segments of divided genome viruses always have the same structures at their 5' ends. The structure of the 3' ends also varies among different viruses. This may be a polyadenylate tail, a tRNA-like structure, which enables these RNAs to be charged with a specific amino acid (Hall *et al.*, 1972; Agranovsky *et al.*, 1981; Loesch-Fries and Hall, 1982), or a 3'-OH end with no additional features.

For comoviruses (Stanley *et al.*, 1980) and nepoviruses (Mayo *et al.*, 1982) it has been demonstrated that VPg is specified by the virus itself, for the other plant viruses with a VPg linked to their 5' end, this has not yet been demonstrated. The role of VPg in virus infectivity and multiplication is not clear. VPg seems not to be essential for the infectivity of the RNAs of CPMV, PEMV and poliovirus (Stanley *et al.*, 1978; Reisman and De Zoeten, 1982; Flanagan *et al.*, 1977). In contrast, the infectivity of nepoviral RNAs is strongly decreased or even abolished upon removal of VPg with proteinase (Mayo *et al.*, 1982). The extent of the decrease of infectivity varies for each nepovirus.

The current hypothesis is that VPg is involved in the initiation of virus RNA synthesis by acting as a primer either directly or after template independant addition of one or two nucleotide residues, or via a specific precursor form. This hypothesis is mainly based on observations with poliovirus RNA (Nomoto *et al.*, 1977). For this virus it has been shown that both minus RNA strands and short nascent positive RNA strands possess VPg molecules (Nomoto *et al.*, 1977; Petterson *et al.*, 1978). The primer dependence of purified poliovirus-specified RNA polymerase (Tuschall *et al.*, 1982) is also consistent with a primer function of VPg. That VPg is required for poliovirus RNA synthesis is, moreover, strongly supported by the finding that anti-VPg serum is able to block poliovirus RNA synthesis *in vitro* (Baron and Baltimore, 1982).

On the other hand, it has been suggested that the genome-linked protein of poliovirus may play a role during the assembly of virus particles but experimental evidence to support such proposal is lacking so far (Wimmer, 1982). The significance of the various specific structures found at the 3'ends of plant viral RNAs has also remained unknown. The

genomic RNAs of bipartite and tripartite genome viruses (Table II. 1) each have a similar structure at their 3'end. This holds for the three genomic RNAs of BMV and CMV, which have a tRNA-like 3'end (Ahlquist *et al.*, 1981) and for the two RNAs of comoviruses and nepoviruses which terminate with a 3' poly(A) tail (El Manna and Bruening, 1973; Mayo *et al.*, 1979). For AMV, which has neither a poly(A) tail nor a tRNA-like structure it has been demonstrated that the nucleotide sequences at the 3'termini of the three genomic RNAs can form the same three-dimensional structure (Koper-Zwarthoff *et al.*, 1979). Since the structures at the 3' ends seem to be conserved in the RNAs of each virus, they may be of functional significance for viral RNA replication. In Table II.1 the positive-strand RNA viruses are further grouped based on the mechanism used for the translational expression of the genome. The mechanism in which one or more subgenomic messenger RNAs are generated for the synthesis of virus-specific proteins is applied by several groups of plant viruses, both mono- and multipartite genome viruses. The RNAs of these viruses all have a cap structure at their 5' terminus while the 3' end is either tRNA-like or without any specific feature. So far, no poly(A) tail has been found at the 3' end of the genome RNAs of viruses which produce subgenomic messenger RNAs. For only one group of plant viruses, the comoviruses, it has been firmly established that it belongs to the subclass of positive-strand RNA viruses in which the virus RNAs are translated into polyproteins which undergo specific proteolytic cleavages. Comoviral RNAs are characterised by a VPg covalently linked to the 5' end and a poly(A) tail at their 3' termini. For many other groups of plant viruses it has remained unclear which mechanism they use for the expression of their genetic information. Some groups may use a combination of mechanisms e.g. both the generation of subgenomic messenger RNAs and specific proteolytic cleavages of precursor proteins. Such a mixed mode of expression has been suggested for sobemoviruses and also for the potyviruses. If in all cases in which the genomic RNAs have a 5' VPg structure, the VPg is a virus-specific protein it may be postulated that proteolytic cleavages play a role at some stage of viral protein synthesis as usually VPg is a small protein ($\sim$ 4K) and should be generated from a precursor protein. If however, VPg is not

a low molecular weight protein, as appears the case for the VPg of pea enation mosaic virus (molecular weight $\sim$ 18,000), this is not necessarily so.

In this thesis we report on the mechanism of expression of CPMV and how that mechanism has been elucidated. In the next sections of this chapter we shall first review some general properties of comoviruses and CPMV in particular as a further introduction to the subject of this thesis.

2.1. *Comoviruses.*

Until now twelve plant viruses have been assigned to the comovirus-group (Bruening, 1978; Stace-Smith, 1981; see table II.2). All members of this group are transmitted by beetles and have a narrow host range. The majority of comoviruses has legumes as their natural hosts with only very few host plant species outside the legumes (RaMV, SqMV and APMV; Stace-Smith, 1981). A low level of seed transmission has been reported for several comoviruses.

For a further description of the biological properties of comoviruses I refer to some reviews, in which these properties are discussed extensively (Bruening, 1978; Van Kammen and De Jager, 1978; Stace-Smith, 1981). Here I like to focus on the molecular biology of the comovirus group. Comoviruses are small icosahedral viruses with a diameter of about 28 nm. Purified preparations of comoviruses usually contain three different centrifugal components, which are referred to as Top-(T) middle- (M) and bottom component (B) with sedimentation coefficients of 54-60 S, 91-100 S and 112-127 S, respectively (Geelen, 1974, Bruening, 1978, 1981). T-, M- and B-components have similar protein capsids but differ in RNA content. M components contain a smaller RNA molecule (M-RNA) and B components a larger RNA molecule (B-RNA), whereas T components are devoid of RNA (Van Kammen, 1972). Both B- and M-components or their RNAs are necessary for infectivity (Van Kammen, 1968; De Jager, 1976), demonstrating that the genetic information for

Table II.2 Members of the comovirus group

Virus	Abbreviation
Andean potato mottle virus	APMV
Bean pod mottle virus	BPMV
Bean rugose mosaic virus	BRMV
Broad bean stain virus	BBSV
Broad bean true mosaic virus	BBTMV
Cowpea mosaic virus	CPMV
Cowpea severe mosaic virus	CPSMV
Pea mild mosaic virus	PMMV
Quail pea mosaic virus	QPMV
Radish mosaic virus	RaMV
Red clover mottle virus	RCMV
Squash mosaic virus	SqMV

virus multiplication is distributed between both genome segments. The molecular weight of B RNAs and M RNAs reported for different comoviruses are in the range of 2.0 - 2.2 x 10^6 and 1.2 - 1.4 x 10^6 respectively. (Reijnders *et al.*, 1974; Bruening, 1978). Top components do not seem to have a specific function in virus infectivity. The amount of T component produced varies for different comoviruses and even for different variants of the same virus, and seems also to be dependent on the conditions of growth of the virus (Van Kammen, 1972; Siler *et al.*, 1976). CPMV is the type member of the group and by far the most studied and best characterised comovirus. In the next section the data on the molecular properties of CPMV will be discussed in more detail.

A common feature of comoviruses is that their protein capsids are constructed of 60 copies of each of two different proteins, a larger one and a smaller one. This has been demonstrated for CPMV (Wu and Bruening, 1971; Geelen *et al.*, 1972), CPSMV (Beier *et al.*, 1981), SqMV (Hiebert and Purcifull, 1981) and RCMV (Oxelfelt, 1976). The molecular

weights of the two coat proteins differ among the various comoviruses and are in the range of 22,000-25,000 (smaller coat protein) and 37,000-44,000 (larger coat protein) respectively (Rottier, 1980; Hiebert and Purcifull, 1981; Beier *et al.*, 1981). The available data indicate that the genetic information for both coat proteins is located on the middle-component RNA (CPMV: Gopo and Frist, 1977; Franssen *et al.*, 1982, SqMV: Hiebert and Purcifull, 1981).

Another characteristic feature of comoviruses is that their genome RNAs have a polyadenylate sequence (poly A tail) at their 3'terminal end and a small protein (VPg) covalently bound to the 5' end. The presence of VPg molecules has been demonstrated for the RNAs of all comoviruses tested sofar, i.e. BBTMV, CPMV, SqMV and CPSMV (Stanley *et al.*, 1978; Daubert *et al.*, 1978; Daubert and Bruening, 1979). The occurrence of a poly(A) tail has been detected in the genome of BPMV, CPMV and RCMV B-RNA (El Manna and Bruening, 1973; Semancik, 1974; Oxelfelt, 1976).

Only for CPMV the expression of the genetic information encoded in the two genome RNAs has been thoroughly studied (see for instance the following chapters of this thesis). The available data indicate that the viral RNAs are translated into large polyproteins, which are subsequently cleaved into functional proteins. The few studies available on SqMV (Hiebert and Purcifull, 1981; Goldbach and Krijt, 1982), CPSMV (Beier *et al.*, 1981; Goldbach and Krijt, 1982; Rezelman, Van der Krol and Goldbach, unpublished results) and BPMV (Gabriel *et al.*, 1982) indicate that these comoviruses apply an expression strategy similar to that of CPMV.

The different comoviruses have genetically considerably diverged which can be concluded from early experiments in which bottom components of one comovirus were combined with middle components of another comovirus. So far such pseudorecombinants have not been found to yield viable virus preparations. The very low nucleotide sequence homology observed between CPMV RNAs and the RNAs of BPMV and RCMV, respectively (Van Kammen, 1972) and between CPMV M RNA and the M RNAs of CPSMV and SqMV (Goldbach and Van der Krol, unpublished results) provide further support for the distant relatedness among comoviruses. The low serolo-

gical relationship observed between CPMV and CPSMV (Goldbach and Krijt, unpublished results) underlines this as well.

2.2. *Cowpea mosaic virus (CPMV)*

The two capsid proteins of CPMV have molecular weights of 37,000 and 23,000 (Wu and Bruening, 1971; Geelen *et al.*, 1972) and are referred to as VP37 and VP23, respectively. The larger capsid protein VP37 is blocked at its N-terminal end by a N-acetylated methionine residue (Bruening, 1981). Partridge *et al.* (1974) have determined the carbohydrate composition of CPMV. From these results it can be calculated that in each protein shell only 6-7% of the capsid proteins are glycosylated. The occurrence of glycoproteins in the viral capsid has been proposed to be linked to seed transmissibility (Partridge *et al.*, 1974). The middle and bottom component RNA of CPMV have molecular weights of 1.37×10^6 and 2.02×10^6 respectively (Reijnders *et al.*, 1974). VPg has been shown to be linked through a phosphodiester bound to the 5' terminal uridilyl residue in both RNAs (Stanley *et al.*, 1978). The amino acid involved in this linkage is probably a serine residue (Zabel *et al.*, 1984). The poly(A) tail at the 3' terminus of B-RNA has been estimated to be about 87 nucleotides in length, whereas the poly(A) tail of M-RNA has an average of about 160 adenylate residues and is more variable in lenght (Ahlquist and Kaesberg, 1979).

Recently, the complete nucleotide sequences of both M and B-RNA have been elucidated (Van Wezenbeek *et al.*, 1983; Lomonossoff and Shanks, 1983). M-RNA has a sequence of 3481 nucleotides and B-RNA of 5889 nucleotides, excluding their poly(A) tails. Zabel *et al.* (1984) have determined that VPg consists of a polypeptide of approximately 30 amino acid residues. Using limited amino acid sequence data they were able to locate the coding region for VPg on B-RNA and to derive the amino acid sequence of VPg.

Although both B and M components or their RNAs are necessary for virus multiplication the B component exhibits a partial independence

from the M component. When isolated cowpea mesophyll protoplasts were infected with separate B and M components it appeared that B-RNA is capable to replicate itself (Goldbach *et al.*, 1980). Apparently B-RNA encodes functions involved in viral RNA replication. On the other hand, the replication of M-RNA appeared to be fully dependent on the presence and expression of B-RNA. In protoplasts infected with only B components, all non-structural viral proteins were found but not the capsid proteins, whereas in protoplasts infected with B + M components these proteins were abundantly synthesized. This finding indicated that M-RNA might carry the information for the structural proteins of the virus, and B-RNA (most of) the non-structural proteins. Previous, experiments in which M and B components of different CPMV mutants were combined (Gopo and Frist, 1977; De Jager, 1976, 1978; Bruening, 1977) already suggested that capsid protein synthesis is under control of M-RNA. In the course of our studies we obtained independent and direct evidence of such a distribution of genetic information between M and B-RNA and these data will be presented in chapters III-VII.

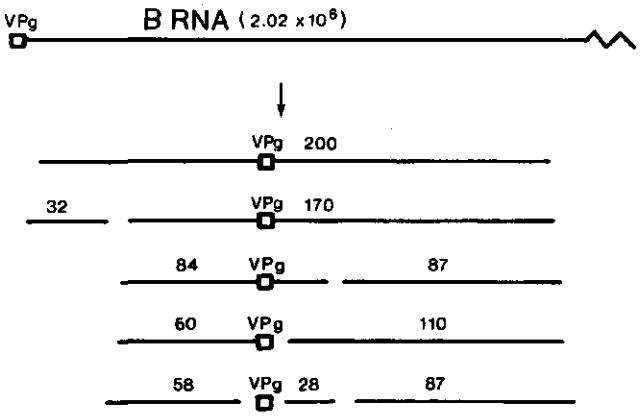
2.3. *Proteins synthesised by CPMV.*

Upon inoculation of cowpea mesophyll protoplasts with CPMV up to 70-90% of the protoplasts can become infected (Hibi *et al.*, 1975; Rottier *et al.*, 1979, 1980). When such infected protoplast suspensions are incubated in the presence of [^{35}S]methionine the synthesis of at least ten virus-specific proteins can be followed. The apparent molecular weights of these proteins as determined by SDS-polyacrylamide gel electrophoresis are 200, 170, 110, 87, 84, 60, 58, 37, 32 and 23K respectively (Rottier *et al.*, 1979, 1980; Goldbach *et al.*, 1980; Rezelman *et al.*, 1980).

Among these virus-specific proteins the 37K and 23K polypeptides represent the capsid proteins VP37 and VP23. The other eight polypeptides are also found in protoplasts infected with B-components alone (Goldbach *et al.*, 1980; Rezelman *et al.*, 1980) and are therefore ap-

parently encoded by B-RNA. The sum of the molecular weights of these polypeptides (∿ 800.000) exceeds the coding capacity of B-RNA, which amounts approximately 230.000 dalton. In order to explain the number of different B-RNA encoded polypeptides the possible precursor-product relationship of these polypeptides was studied both by comparison of their peptide patterns obtained by limited digestion with *Staphylococcus aureus* protease V8, and by serological comparison (Goldbach *et al.*, 1982; Zabel *et al.*, 1982). These studies demonstrated a relationship between the 170K, 110K and 87K polypeptides at the one hand and the 170K, 84K, 60K and 58K polypeptides at the other hand suggesting that the 170K polypeptide is the common precursor to the other polypeptides (Rezelman *et al.*, 1980; Goldbach *et al.*, 1982). Using antiserum against VPg it was shown that the 60K polypeptide is the direct precursor of VPg. The results of these analyses allowed to propose a model for the expression of B-RNA as depicted in Fig. II.1. In this model B-RNA is expressed by translation into a single 200K polyprotein, which is first cleaved into 32K and 170K polypeptides. Subsequently the 170K polypeptide is further cleaved into 60K and 110K polypeptides or, alternatively, into 84K and 87K polypeptides. The 60K polypeptide is cleaved

Fig. II.1.: Model for the expression of CPMV B-RNA.



to release VPg and the 58K polypeptide, the 110K can probably be further cleaved into 28K and 87K polypeptides. In an alternative processing pathway the 84K polypeptide can be cleaved into 58K, VPg and 28K respectively. Thus both processing routes result in the same final cleavage products of respectively 32K, 58K, VPg, 28K and 87K. The model for the expression of B-RNA has been confirmed by *in vitro* translation experiments and by data derived from the complete nucleotide sequence of B-RNA (Lomonossoff and Shanks, 1983).

The 28K polypeptide indicated in Fig. II.1 has so far not been detected *in vivo*. In chapters VII and VIII of this thesis the occurrence and possible function of this polypeptide will be discussed in more detail.

Dorssers *et al.* (1983, 1984) have recently demonstrated that the B-RNA encoded 110K polypeptide is the only virus-specified protein present in highly purified viral RNA replication complexes, which are capable of elongating nascent RNA chains *in vitro* into full-size M and B-RNA molecules. Their findings strongly indicate that the 110K polypeptide may be the viral RNA replicase. In Chapter VIII of this thesis we present data which support this hypothesis. The proteolytic activities involved in the processing of the virus-encoded proteins, and their cleaving specificity are the major subject of this thesis. The results will be discussed in Chapters III, V and VII.

In vitro translation of B and M-RNA.

In both the messenger-dependent rabbit reticulocyte system (Pelham and Stuik, 1977) and in wheat germ extracts (Davies *et al.*, 1977) the CPMV RNAs direct the synthesis of large primary translation products. B-RNA produces a 200K polypeptide and M-RNA two polypeptides with mol. weights of 105K and 95K respectively (Goldbach *et al.*, 1981). The 200K polypeptide specified by B-RNA corresponds to 85% of the coding capacity of B-RNA (Pelham, 1979; Stuik, 1977). The complete nucleotide sequence of B-RNA (5889 nucleotides, excluding the poly(A) tail) reveals the presence of only one large open reading frame starting at position 207 and continuing until a UAG stopcodon at position 5805. This reading

frame corresponds to a primary translation product with a mol.weight of 207.766, which is in good agreement with the experimental value.

The 200K primary translation product obtained by *in vitro* translation of B-RNA is rapidly cleaved into 32K and 170K polypeptides, provided that ATP and 2 mM dithiothreitol are present in the incubation mixture (Pelham, 1979; Goldbach *et al.*, 1981). Further studies using various protease inhibitors (a.o. N-ethylmaleimide, $ZnCl_2$, phenylmethylsulphonylfluoride) reveals that the proteolytic activity involved in this cleavage is a thiol-type protease (Pelham, 1979). In chapter VII our experimental data are presented which indicate that a B-RNA encoded proteolytic activity is engaged in this cleavage step. Pulse labelling experiments indicated that the 32K and 170K polypeptides were derived from the amino terminal and carboxyterminal end of the 200K precursor respectively (Pelham, 1979). By comparison of the peptide patterns obtained from the 170K and 32K polypeptides found in B-infected protoplasts with those produced by *in vitro* translation of B-RNA (Rezelman *et al.*, 1980) showed that the 32K and 170K polypeptides made *in vitro* were identical to those found *in vivo*. In further studies Goldbach and Rezelman (1983) determined the orientation (amino terminally/carboxy terminally) of the 60K and 110K polypeptides within the 170K polypeptide (Fig. II.1) by comparing the proteolytic peptide patterns of the *in vivo* polypeptides with those of 170K polypeptides translated from B-RNA *in vitro* and pulse-labelled at either amino- or carboxy-terminal end. These results on the *in vitro* translation of B-RNA complement the data used for the model of the expression of B-RNA illustrated in Fig. II.1.

The two polypeptides with mol.weights of 105K and 95K translated from M-RNA in *in vitro* protein synthesizing systems, have overlapping amino acid sequences (Pelham, 1979). The larger polypeptide corresponds to about 75% of the coding capacity of M-RNA. Pelham (1979) postulated that the 105K and 95K polypeptides arise by the presence of two initiation sites on M-RNA and that the 105K and 95K polypeptides have identical carboxy-termini. This conclusion was based on ribosome binding experiments which indicated that (a minority of) M-RNA molecules are capable to bind two ribosomes. However, Filipowitz and Haenni

(1979) and Ahlquist *et al.* (1979) later showed that there is no direct correlation between the number of ribosomes bound and the number of translation initiation sites and took away the basis for the hypothesis of Pelham. Our experiments on the processing of the 105K and 95K polypeptides, described in Chapter III of this thesis, independently show that the 105K and 95K polypeptides have the same carboxy terminal end. This was further confirmed when the complet nucleotide sequence of M-RNA was elucidated (Van Wezenbeek *et al.*, 1983). The nucleotide sequence of M-RNA (3481 nucleotides, excluding the poly(A) tail of variable length) contains a single open reading frame starting at position 161 and extending to a UAA stop codon at position 3299, which probably encodes the 105K polypeptide. Two AUG codons at positions 512 and 524 within this reading frame may be used as start codons for the synthesis of the 95K polypeptide.

Pelham (1979) showed that the primary translation products of M-RNA can be cleaved by a proteolytic activity present in the *in vitro* translation products of B-RNA. This proteolytic cleaving generated 41K, 54K and 57K polypeptides but no mature capsid proteins. The protease involved in this cleavage step appeared to be of the thiol type, similar to the proteolytic activity responsible for the cleavage of the B-RNA encoded 200K polypeptide into 32K and 170K polypeptides. In Chapter III of this thesis our studies on the *in vitro* processing of the primary translation products of M-RNA are described. The identification of the protease involved in at least one of the cleavage steps is reported in Chapter V. In Chapter VI we report on the amino acid sequence analyses of M-RNA encoded proteins which allowed the design of a model for the expression of M-RNA.

3. *References.*

- Agranovsky, A., Dolja, V.V., Gorbulev, V.G., Koslov, V.G., Yu, V., and Atabekov, J.G. (1981). Virology, 113, 174-187.
- Ahlquist, P., Dasgupta, R. and Kaesberg, P. (1981). Cell 23, 183-189.
- Ahlquist, P., Dasgupta, R., Shih, D.S., Zimmern, D., and Kaesberg, P. (1979). Nature 281, 277-282.
- Ahlquist, P. and Kaesberg, P. (1979). Nucleic Acids Res. 7, 1195-1204.
- Baltimore, D. (1971). Bacteriol. Rev. 35, 235-241.
- Baron, M.H., and Baltimore, D. (1982) Cell 28, 395-404.
- Beier, H., Issinger, O.G., Deuschle, M. and Mundry, K.W. (1981). J. Gen. Virol. 54, 379-390.
- Bruening, G. (1977). In: (H. Fraenckel-Conrat and R.R. Wagner eds.) "Comprehensive Virology". Vol. 11, pp 55-141, Plenum Press, N.Y.
- Bruening, G. (1978). CMI/AAB Description of Plant Viruses No. 199, pp 6.
- Bruening, G. (1981). In: (Marcus, G. ed.) "The Biochemistry of Plants". Vol. 6, pp 571-631.
- Clewley, J.P., Crump, W.A.L., Avery, R.J., and Moore, N.F. (1982). J. Virol. 44, 767-771.
- Daubert, S.D., Bruening, G. and Najarian, R.C. (1978). Eur. J. Biochem. 92, 45-51.
- Daubert, S. and Bruening, G. (1979). Virology 98, 246-250.
- Davies, J.W., Aalbers, A.M.J., Stuik, E.J. and Van Kammen, A. (1977), FEBS Lett. 77, 265-269.
- De Jager, C.P. (1976). Virology 70, 151-163.
- De Jager, C.P. (1978). Thesis, Agricultural University, Wageningen, The Netherlands.
- Dorssers, L. (1983). Thesis, Agricultural University, Wageningen, The Netherlands.
- Dorssers, L., Van der Krol, S., Van der Meer, J., Van Kammen, A. and Zabel, P. (1984). Proc. Natl. Acad. Sci. USA 81, 1951-1955.
- Dougherty, W.G., and Hiebert, E. (1984). In: J.W. Davies (ed.), CRC Press, in press.

- El Manna, M.M., and Bruening, G. (1973). Virology 56, 198-206.
- Filipowicz, W. and Haenni, A.L. (1979). Proc.Natl.Acad.Sci. USA 76, 3111-3115.
- Flanegan, J.B., Pettersson, R.F., Ambros, V., Heivlett, M.J. and Baltimore, D. (1977). Proc.Natl.Acad.Sci.USA 74, 961-965.
- Franssen, H., Goldbach, R., Broekhuysen, M., Moerman, M. and Van Kammen, A. (1982). J. Virol. 41, 8-17.
- Friesen, P.D., and Rueckert, R.R. (1981). J. Virol. 37, 876-886.
- Friesen, P.D., and Rueckert, R.R. (1982). J. Virol. 42, 986-995.
- Gabriel, L.J., Derrick, K.S., and Shih, D.S. (1982). Virology, 122, 476-480.
- Geelen, J. (1974). Thesis, Agricultural University, Wageningen, The Netherlands.
- Geelen, J.L.M., Van Kammen, A., and Verduin, B.J.M. (1972). Virology 49, 205-213.
- Goldbach, R., and Krijt, J. (1982). J. Virol. 43, 1151-1154.
- Goldbach, R., and Rezelman, G. (1983). J. Virol. 46, 614-619.
- Goldbach, R., Rezelman, G., and Van Kammen, A. (1980). Nature, 286, 297-300.
- Goldbach, R., Rezelman, G., Zabel, P., and Van Kammen, A. (1982). J. Virol. 42, 630-635.
- Goldbach, R.W., Schilthuis, J.G. and Rezelman, G. (1981). Biochem. Biophys. Res. Commun., 99, 89-94.
- Gopo, J.M., and Frist, R.H. (1977). Virology 79, 259-266.
- Hall, T.C., Shih, D.S. and Kaesberg, P. (1972). Biochem. J. 129, 969-976.
- Hibi, T., Rezelman, G., and Van Kammen, A. (1975). Virology 64, 308-318.
- Hiebert, E., and Purcifull, D.E. (1981). Virology 113, 630-635.
- Koper-Zwarthoff, E.C., Brederode, F.Th., Walstra, P., and Bol, J.F. (1979). Nucleic Acids Res. 7, 1887-1900.
- Loesch-Fries, L.S., and Hall, T.C. (1982). Nature 298, 771-773.
- Lomonossoff, G., and Shanks, M. (1983). EMBO J. 2, 2253-2258.
- Luria, S.E., Darnell, J.E., Baltimore, D., and Campbell, A. (1978). In: "General Virology 3rd", Wiley & Sons, New York.

- Mayo, M.A., Barker, H., and Harrison, B.D. (1979). J.Gen.Virol. 59, 149-162.
- Mayo, M.A., Barker, H., and Harrison, B.D. (1982). J.Gen.Virol. 59, 149.
- Nomoto, A., Betien, B., Pozzyti, R. and Wimmer, E. (1977). Nature 268, 208-213.
- Oxelfelt, P. (1976). Virology 74, 73-80.
- Patridge, J.E., Gumpf, D.J., Colbough, P. (1974). Nature 247, 391-392.
- Pelham, H.R.B. (1979). Virology 96, 463-477.
- Pelham, H.R.B., and Stuik, E.J. (1977). Proc. Coll. Nucleic Acids and Protein Synthesis in Plants, CNRS, pp 691-695.
- Reisman, D., and De Zoeten, G.A. (1982). J.Gen.Virol. 62, 187-190.
- Pettersson, R.F., Ambros, V., and Baltimore, D., (1978). J. Virol. 27, 357-365.
- Reynders, L., Aalbers, A.M.J., Van Kammen, A., and Thuring, R.W.J. (1974). Virology 60, 515-521.
- Rezelman, G., Goldbach, R., and Van Kammen, A. (1980). J. Virol. 36, 366-373.
- Rottier, P.J.M., Rezelman, G., and Van Kammen, A. (1979). Virology 92, 299-309.
- Rottier, P.J.M. (1980). Thesis, Agricultural University, Wageningen, The Netherlands.
- Semancik, J.S. (1974). Virology 62, 288-291.
- Siler, D.J., Balcook, J., and Bruening, G. (1976). Virology 71, 560-567.
- Stace-Smith, R. (1981). In: (Kurstak, E. ed.) "Handbook of Plant virus infections and Comparative diagnosis", pp 171-193. Elsevier Press.
- Stanley, J., Goldbach, R., and Van Kammen, A. (1980). Virology 106, 180-182.
- Stanley, J., Rottier, P., Davies, J.W., Zabel, P., and Van Kammen, A. (1978). Nucleic Acids Res. 5, 4505-4522.
- Stuik, E.J. (1979). Thesis, Agricultural University, Wageningen, The Netherlands.

- Thongmeearkom, P., and Goodman, R.M. (1978). Virology 85, 75-83.
- Tuschall, D.M., Hiebert, E. and Flanegan, J.B. (1982). J. Virol. 44, 209-216.
- Van Kammen, A. (1968). Virology 34, 312-318.
- Van Kammen, A. (1972). Annual Review of Phytopathology 10, 125-150.
- Van Kammen, A., and De Jager, C.P. (1978). In: CMI/AAB "Description of Plantviruses", No. 197.
- Van Wezenbeek, P., Verver, J., Harmsen, J., Vos, P., and Van Kammen, A. (1983). EMBO J. 2, 941-946.
- Wimmer, E. (1982). Cell 28, 199-201.
- Wu, G., and Bruening, G. (1971). Virology 46, 596-612.
- Zabel, P., Moerman, M., Lomonossoff, G., Shanks, M., and Bayreuther, K. (1984). EMBO J. in press.
- Zabel, P., Moerman, M., Van Straaten, F., Goldbach, R., and Van Kammen, A. (1982). J. Virol. 41, 1083-1088.

CHAPTER III

EXPRESSION OF MIDDLE-COMPONENT RNA OF COWPEA MOSAIC VIRUS: IN VITRO GENERATION OF A PRECURSOR TO BOTH CAPSID PROTEINS BY A BOTTOM-COMPONENT RNA-ENCODED PROTEASE FROM INFECTED CELLS

Henk Franssen, Rob Goldbach, Martien Broekhuijsen, Marja Moerman and Albert van Kammen.

J. Virol. 41, 8-17 (1982).

JOURNAL OF VIROLOGY, Jan. 1982, p. 8–17
0022-538X/82/010008-10$02.00/0

Vol. 41 No. 1

Expression of Middle-Component RNA of Cowpea Mosaic Virus: In Vitro Generation of a Precursor to Both Capsid Proteins by a Bottom-Component RNA-Encoded Protease from Infected Cells

HENK FRANSSEN, ROB GOLDBACH,* MARTIEN BROEKHUIJSEN, MARJA MOERMAN, AND ALBERT VAN KAMMEN

Department of Molecular Biology, Agricultural University, 6703 BC Wageningen, The Netherlands

Received 15 June 1981/Accepted 28 August 1981

The expression of the middle-component (M) RNA of cowpea mosaic virus was studied by means of in vitro translation. In both the wheat germ extract and the rabbit reticulocyte lysate, M RNA was translated into two overlapping polypeptides of 95 and 105 kilodaltons. Incubation of these polypeptides with 30,000 × *g* supernatant fractions from cowpea mesophyll protoplasts inoculated with complete virus or with separate bottom (B) components alone resulted in extensive processing, yielding polypeptides of 60, 58, 48, and 47 kilodaltons. Similar proteolytic activity was found associated with the in vitro translation products from the bottom-component RNA, demonstrating that the protease present in infected cells is encoded by B RNA. Using antisera raised against the separate capsid proteins VP23 and VP37, it was shown that the 60-kilodalton cleavage product is the precursor to both capsid proteins. Cleavage of nascent 95- and 105-kilodalton polypeptides by the in vivo protease demonstrated that this capsid protein precursor is located C terminally within both polypeptides and that the synthesis of these two overlapping polypeptides is the result of two initiation sites on middle-component RNA. In addition, a second virus-induced proteolytic activity, capable of releasing VP23 from the 95- and 105-kilodalton polypeptides, was detected in leaves of infected plants, but not in infected mesophyll protoplasts. A model for the expression of the middle-component RNA is presented.

In cowpea mesophyll protoplasts infected with cowpea mosaic virus (CPMV), at least nine virus-coded polypeptides have been detected, with sizes of 170, 110, 87, 84, 60, 37, 32, 23, and 4 kilodaltons (K) (9, 19). The 37K and 23K polypeptides represent the capsid proteins (denoted as VP37 and VP23) and the 4K polypeptide represents the genome-linked protein VPg, attached to the 5′ ends of both bottom (B)- and middle (M)-component RNA (4, 27). Upon infection of protoplasts with purified B components, the seven noncapsid polypeptides are still synthesized, but VP37 and VP23 are lacking (9, 19, 26). On the other hand, inoculation with the M component does not result in detectable synthesis of any viral protein. These findings have been explained by the independent replication of B RNA (molecular weight, 2.02×10^6 [18]), a property not associated with M RNA (molecular weight, 1.37×10^6 [18]) (9). Therefore, direct in vivo studies on the expression of B RNA have been possible, and a model has been proposed in which this RNA is translated into a 200K polyprotein which is processed by three different proteolytic cleavages to produce the 110K, 87K, 84K, 60K, and 32K polypeptides (19). Among these B RNA-encoded polypeptides the 60K polypeptide represents the direct precursor to VPg (P. Zabel, M. Moerman, F. van Straaten, R. Goldbach, and A. van Kammen, manuscript in preparation). As M RNA is not replicated in the absence of B RNA, the expression in vivo of this RNA cannot be studied directly. VP23 and VP37 are not formed in B-infected protoplasts and are therefore probably coded for by M RNA. Larger M RNA-coded (precursor-) polypeptides have not been detected in vivo (9, 19). Therefore, we have studied the expression of M RNA by in vitro translation, using two different cell-free systems, the wheat germ system and the rabbit reticulocyte lysate. In both systems M RNA is translated into two large polypeptides of approximately 95K and 105K with overlapping amino acid sequences (10, 16). Since B RNA is translated properly in these systems, resulting in a 200K polypeptide which is cleaved into 170K and 32K polypeptides also found in CPMV-infected cells (19), the M RNA-encoded 95K or

105K polypeptide, or both, may well represent proper primary translation products, which remain undetectable in vivo. Previously, it has been shown by Pelham (16) that either the 170K or the 32K polypeptide translated from B RNA in reticulocyte lysate possesses proteolytic activity which accomplishes cleavage of these two M RNA-encoded in vitro products. Here we show that this B RNA-encoded protease is actually produced in CPMV-infected cells and that it cleaves the M RNA-coded 105K and 95K primary translation products in vitro, generating a 60K precursor to both capsid proteins VP23 and VP37. Besides, a second virus-specific proteolytic activity, present in leaves of infected plants but not detectable in infected protoplasts, is able to cleave only VP23 from the M RNA-coded in vitro products. A model for the expression of M RNA is discussed.

MATERIALS AND METHODS

Virus and RNA. CPMV was grown in cowpea plants (*Vigna unguiculata* L., "California Blackeye"), and B and M components were purified and separated as previously described (9, 12). CPMV RNAs were extracted from separated components as described by Davies et al. (6). Quality and purity of B and M RNA preparations were tested by electrophoresis in 1% agarose gels (8).

Translation in wheat germ extracts. Wheat germ (General Mills Inc., Vallejo, Calif.) was extracted as described by Davies et al. (5). RNA (0.5 μg) was added to a 15-μl reaction mixture containing 7.5 μl of wheat germ extract, 20 mM *N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid (HEPES)-KOH, pH 7.5, 2.9 mM magnesium acetate, 90 mM potassium acetate, 0.4 mM spermidine-hydrochloride, 2.5 mM ATP, 0.375 mM GTP, 10 mM creatine phosphate, 10 μg of creatine kinase per ml, 2 mM dithiothreitol, 5 μg of human placental RNase inhibitor (2, 21) per ml, 25 μmol of each amino acid (except methionine), and 2 to 5 μCi of [^{35}S]methionine (1,000 to 1,100 Ci/mmol, Radiochemical Centre, Amersham, England). Incubation was for 1 h at 30°C.

Translation in reticulocyte lysates. Translation in an mRNA-dependent rabbit reticulocyte lysate (a generous gift of H. R. B. Pelham and R. J. Jackson, Department of Biochemistry, University of Cambridge, Cambridge, England) was performed as previously described (10, 14, 16). [^{35}S]methionine (2 to 5 μCi per 15-μl reaction mixture) was used as radioactive amino acid, and incubation was for 1 h at 30°C.

Preparation of subcellular fractions from uninfected and CPMV-infected cells. Extracts were prepared from both uninfected and CPMV-infected cowpea leaves and cowpea mesophyll protoplasts. Extracts from leaves were prepared by the method of Zabel et al. (30, 31) for the preparation of solubilized viral replicase. Portions of 10 to 12 g of freshly harvested primary leaves of 13-day-old cowpea plants (uninfected or infected on day 9) were homogenized at 0°C in 35 ml of homogenization buffer (HB buffer), containing Tris-acetate (pH 7.4), 10 mM potassium acetate, 1 mM EDTA, 10 mM dithioerythritol, and 0.5 mM phenylmethylsulfonyl fluoride. The homogenate was filtered and centrifuged at 1,000 × *g* for 15 min at 4°C. The supernatant was adjusted to 20% (vol/vol) glycerol and centrifuged at 30,000 × *g* for 30 min at 4°C. The 30,000 × *g* pellet was extracted by resuspension in TGKEDP buffer, containing 50 mM Tris-acetate (pH 8.2), 25% (vol/vol) glycerol, 50 mM potassium acetate, 1 mM EDTA, 10 mM dithioerythritol and 0.5 M phenylmethylsulfonyl fluoride (1 ml for each gram of leaf tissue used), and was centrifuged at 30,000 × *g* for 60 min at 4°C. The resulting 30,000 × *g* pellet extract was used for the in vitro processing experiments and has been stored at −80°C for at least several months without any loss of proteolytic activity. Extracts from protoplasts were prepared as follows. Mesophyll protoplasts were prepared and infected with either complete virus (B + M) or purified B component as previously described (19, 20). Forty hours after infection, portions of 2.5 × 10^6 protoplasts were collected by centrifugation (2 min, 600 × *g*) and disrupted by homogenization for 5 min in a small Thomas tissue homogenizer with 0.5 ml of HB buffer containing 10% sucrose. Intact cells were removed by centrifugation for 15 min at 1,000 × *g* and 4°C. The homogenate was then centrifuged for 30 min at 30,000 × *g* and 4°C to give the 30,000 supernatant, to be used in the in vitro processing experiments. Such preparations have been stored at −80°C for at least several months without any loss of proteolytic activity.

Processing of in vitro translation products by subcellular fractions from CPMV-infected cells. One volume of in vitro translation products from CPMV M RNA (as obtained after 1 h of translation at 30°C in either wheat germ extract or reticulocyte lysate) was mixed with an equal volume of extract from either CPMV-infected or uninfected cells and incubated for 1 h at 30°C. As a control, a 1:1 (vol/vol) mixture of M RNA in vitro products and TGKEDP buffer (see above) was incubated under the same conditions. Processing was followed by electrophoresis of samples in polyacrylamide gels.

SDS-polyacrylamide slab gel electrophoresis. Samples of radiolabeled proteins were mixed with one-third volume of a fourfold concentrated sample buffer (4 × SB: 40 mM Tris-hydrochloride (pH 8.0), 4 mM EDTA, 40% [vol/vol] glycerol, 8% [wt/vol] sodium dodecyl sulfate [SDS], 20% [vol/vol] β-mercaptoethanol, and 0.004% [wt/vol] bromophenol blue). After heating for 3 min at 100°C, samples were analyzed in polyacrylamide gels containing 12.5% acrylamide (with 0.09% bisacrylamide) as previously described (13), using spacers of 4% acrylamide with 0.10% bisacrylamide. Gels were dried and autoradiographed with Kodak Royal X-Omat film.

Antisera. Antisera against the electrophoretically separated capsid proteins VP23 and VP37 were prepared by directly immunizing rabbits with polyacrylamide containing denatured antigen, in principle as described by Tjian et al. (29) and Schiff and Grandgenett (22) but with some modifications. CPMV particles (1.5 mg) were diluted in 1 × SB (see above), heated for 3 min at 100°C, and electrophoresed in a 15% polyacrylamide gel (containing 0.10% bisacrylamide). After electrophoresis, capsid proteins were visualized by staining with Coomassie brilliant blue. The separate

VP37 and VP23 bands were excised from the gel and washed successively with 25% isopropanol and 10% methanol. Washed gel slices were then chopped up, lyophilized, and ground in a mortar. The resulting powder was mixed with approximately 1 to 2 ml of 10 mM sodium phosphate (pH 7.2)–0.9% NaCl and emulsified with an equal volume of Freund complete adjuvant (final volume, 2 to 4 ml). The emulsion was injected subcutaneously in the neck region of New Zealand White rabbits; injection was repeated twice, at 4-week intervals and using Freund incomplete adjuvant. Specificity of the anti-VP23 and anti-VP37 sera was tested as described later in the text.

Immunoprecipitation. For immunoprecipitation 5 to 10 μl of protein sample was adjusted to PBSTDS (10 mM sodium phosphate, pH 7.2, 0.9% NaCl, 1% Triton X-100, 0.5% sodium deoxycholate, and 0.1% SDS) in a total volume of 100 μl and was incubated for 16 h at 4°C in the presence of 5 μl of preimmune serum, anti-VP23 serum, or anti-VP37 serum. Finally, 25 μl of a 10% (wt/vol) suspension of *Staphylococcus aureus* cells (IgGsorb from the Enzyme Center Inc., Boston, Mass.) in PBSTDS (containing 10 mg of bovine serum albumin per ml) was added, and incubation at 4°C was continued for 1 h. Samples were then centrifuged through a sucrose cushion, consisting of a 0.5-ml layer

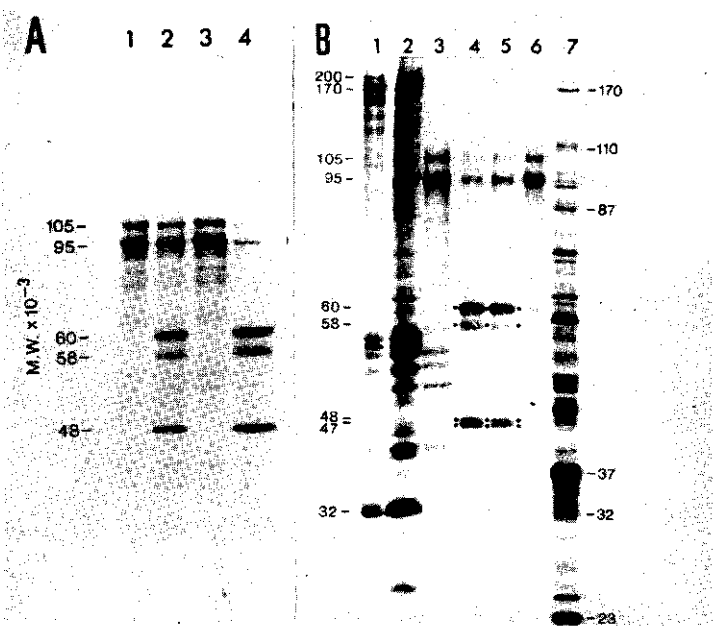


FIG. 1. Detection of a B RNA-encoded protease activity in extract of CPMV-infected cowpea mesophyll protoplasts. Panel A: M RNA was translated in rabbit reticulocyte lysate, and portions (4 μl) of the products obtained were added to an equal volume of HB buffer (lane 1) or 30,000 × *g* supernatant fraction from protoplasts which were inoculated with either complete virus (lane 2) or B components alone (lane 4) or were uninfected (lane 3). Incubation was continued for 1 h at 30°C, and samples were analyzed in a 12.5% polyacrylamide gel. Panel B: B RNA was translated for 1 h in reticulocyte lysate either in the presence of [^{35}S]methionine or unlabeled methionine. The products formed were added to [^{35}S]methionine-labeled in vitro translation products from M RNA, and incubation was continued for 1 h at 30°C. Lane 1, labeled in vitro translation products from B RNA; lane 3, labeled in vitro translation products from M RNA; lane 2, same as lane 3 but mixed with an equal volume of labeled in vitro translation products from B RNA; lane 4, same as lane 3 but mixed with an equal volume of 30,000 × *g* supernatant from B-inoculated protoplasts; lanes 5 and 6, same as lane 3 but mixed with 1 (lane 6) or 10 (lane 5) volumes of unlabeled in vitro translation products from B RNA. Lane 7 contains [^{35}S]methionine-labeled polypeptides from CPMV-infected protoplasts. Samples were analyzed on a 12.5% polyacrylamide gel. Numbers indicated at the left side of the gel refer to the molecular weights ($\times 10^{-3}$) of the in vitro polypeptides; numbers indicated at the right side of the gel refer to the molecular weights ($\times 10^{-3}$) of the viral polypeptides visible in lane 7. The endogenous activity (no RNA added) of the reticulocyte lysate used was undetectable (data not shown).

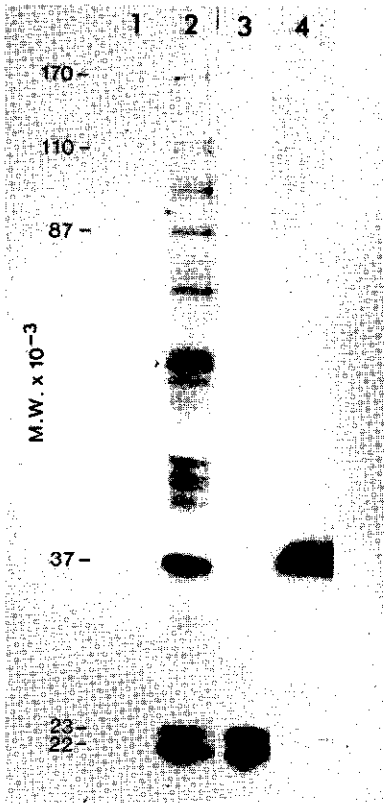


FIG. 2. Characterization of the antisera against the separate capsid proteins VP23 and VP37. Unlabeled proteins from CPMV-infected protoplasts were separated in a 12.5% SDS-polyacrylamide gel and blotted onto nitrocellulose filters (0.45-μm pore size) by the method of Bowen et al. (3). The nitrocellulose filters were incubated with preimmune serum (lane 1), anti-VP23 serum (lane 3), or anti-VP37 serum (lane 4), and immunocomplexes were visualized by using ^{125}I-labeled protein A from *S. aureus* as described in detail elsewhere (P. Zabel, M. Moerman, F. van Straaten, R. Goldbach, and A. van Kammen, manuscript in preparation). To test the fidelity of the protein-blotting technique ^{35}S-labeled proteins from CPMV-infected protoplasts were included in the original gel (lane 2). The small capsid protein VP23 occurs in two size classes (23K and 22K) (20).

of 1 M sucrose and a 0.25-ml layer of 0.5 M sucrose, both in PBSTDS, at 17,000 × *g* for 30 min. The precipitate was washed three times in PBSTDS, dissolved in 1 × SB, heated for 3 min at 100°C, and analyzed in polyacrylamide gels.

RESULTS

B RNA-coded protease active in CPMV-infected cells. Pelham (16) has shown that one of the in vitro translation products from B RNA (i.e., either the 170K or the 32K polypeptide) is able to cleave the M RNA-encoded 95K and 105K polypeptides. To verify whether this B RNA-encoded protease occurs in vivo, we prepared 30,000 × *g* supernatant fractions from both infected and uninfected protoplasts. M RNA in vitro products (as obtained after 1 h of translation) were mixed with these protoplast extracts and incubated at 30°C for another hour. The results shown in Fig. 1A indicate that indeed a proteolytic activity was present in cells infected with complete virus (lane 2) or with B component alone (lane 4), but not in uninfected protoplasts (lane 3). This B component-induced activity was capable of cleaving both 95K and 105K polypeptides into polypeptides of 60K, 58K, and 48K. To test whether this activity was identical to the proteolytic activity of B RNA in vitro products previously described (16), the 95K and 105K polypeptides were mixed with B RNA in vitro products. As in vitro translation of isolated B RNA yielded, along with the 200K, 170K, and 32K polypeptides, a large number of minor products (Fig. 1B, lanes 1 and 2) (16, 17) which obscured the emergence of new polypeptide bands in acrylamide gels, B RNA was translated in an unlabeled reaction. The unlabeled products were mixed with labeled M RNA products, and the mixture was incubated for 1 h at 30°C. Significant processing occurred only upon addition of 10 volumes of unlabeled B RNA products to 1 volume of labeled M RNA products (cf. lanes 5 and 6 in Fig. 1B), resulting in the same cleavage products (60K, 58K, 48K, and sometimes 47K) as obtained with extract of B-infected protoplasts (cf. lanes 4 and 5 in Fig. 1B). This experiment proves that the protease present in infected cells is indeed coded for by B RNA. Additional support for this conclusion comes from the fact that the in vivo and in vitro protease activities were both sensitive to the same inhibitors: $ZnCl_2$, *N*-ethylmaleimide, and chymostatin (results not shown).

Characterization of the cleavage products. As the capsid proteins VP23 and VP37 are absent in protoplasts inoculated with only B components, suggesting that they are coded for by M RNA (9), the in vitro cleavage products were analyzed by using antisera raised against the separate capsid proteins. To test the specificity of the antisera used in these experiments, the anti-VP23 and anti-VP37 sera were incubated with proteins from virus-infected cells which were separated in a 12.5% SDS-polyacrylamide gel and subsequently blotted onto a nitrocellulose

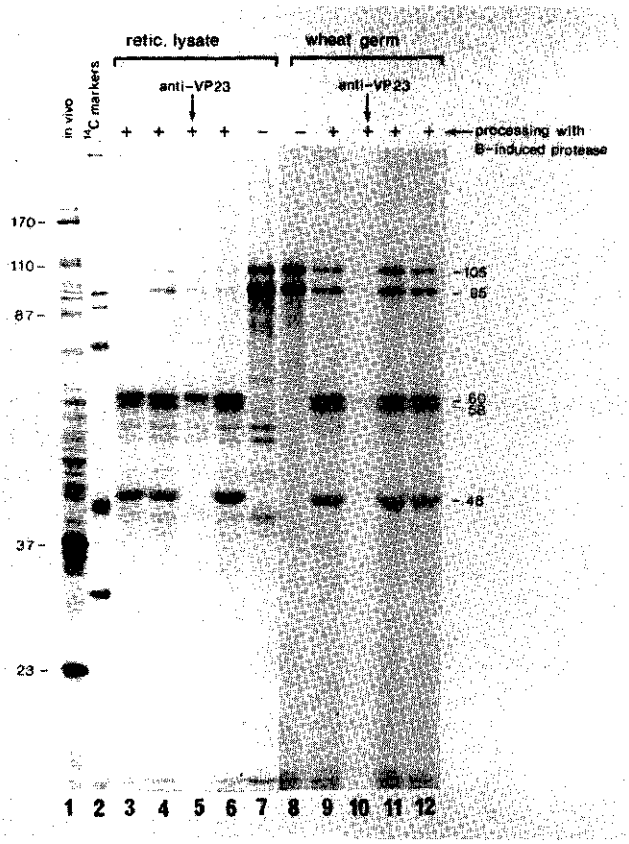


FIG. 3. In vitro processing and immunoprecipitation of translational products from CPMV M RNA. M RNA was translated for 1 h in either reticulocyte lysate (lanes 3–7) or wheat germ extract (lanes 8–12). Portions (2 μl) of the translation products were mixed with an equal volume of either HB buffer (lanes 7 and 8), 30,000 × *g* supernatant fraction from B-inoculated protoplasts (lanes 5, 6, 9, and 10), or unfractionated homogenate of M + B-inoculated protoplasts (lanes 3, 4, 11, and 12). Incubations 4 and 11 included unfractionated CPMV RNA (250 μg/ml). All mixtures were incubated for 1 h at 30°C, and samples 5 and 10 were immunoprecipitated with anti-VP23 serum. Molecular weight markers included ^{35}S-labeled polypeptides from CPMV-infected protoplasts (lane 1) and (in lane 2) ^{14}C-methylated myosin (210,000), phosphorylase *b* (100,000 and 92,500), bovine serum albumin (68,000), ovalbumin (46,000), and carbonic anhydrase (30,000).

filter (as described by Bowen et al. [3]). Of these polypeptides only VP23 reacted with anti-VP23 serum and only VP37 reacted with anti-VP37 serum, without any cross-reaction (Fig. 2). Of the processed M RNA in vitro products the only polypeptide precipitating with anti-VP23 serum was the 60K polypeptide, except for some minor polypeptides, which were also present in untreated M RNA product (Fig. 3, lanes 5 and 10). The 58K and 48K cleavage products did not react with anti-VP23 serum. The 60K polypeptide was also the only cleavage product precipitated by anti-VP37 serum (Fig. 4, lane 1). The results shown in Fig. 3 and 4 demonstrate conclusively that M RNA encodes both capsid proteins. In view of its size, it is attractive to propose the 60K polypeptide as the common precursor to VP23 and VP37. Since the 95K and

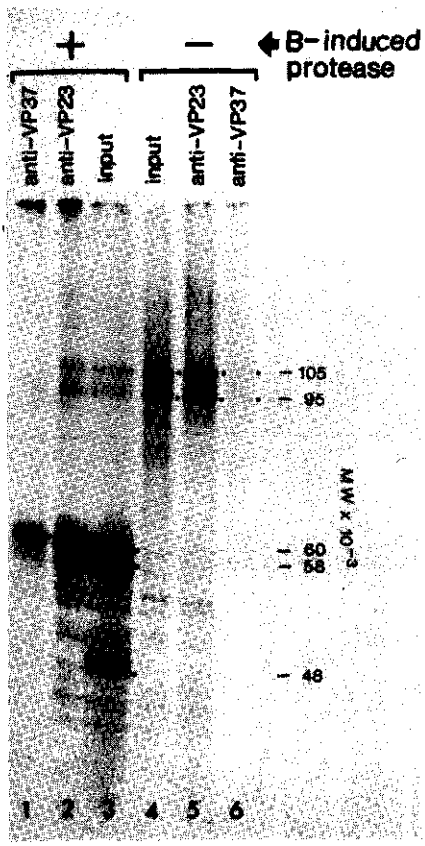


FIG. 4. Immunoprecipitation of CPMV M RNA translational products. M RNA was translated in reticulocyte lysate, and the products obtained were incubated for 1 h at 30°C with either HB buffer (lanes 4–6) or 30,000 × *g* supernatant fraction from B-inoculated protoplasts (lanes 1–3). Samples 1 and 6 were immunoprecipitated with anti-VP37 serum; samples 2 and 5, with anti-VP23 serum. Analysis was on a 12.5% polyacrylamide gel.

105K polypeptides were both efficiently precipitated by at least the anti-VP23 serum (Fig. 3, lane 5, and Fig. 4, lanes 2 and 5), the 60K polypeptide must originate from both these primary products. The size difference between the 95K and 105K polypeptides should therefore correlate with the difference in size between the cleavage products of 58K and 48K. A reason that anti-VP37 serum reacted only weakly with 95K and 105K may be the considerably lower titer of this antiserum compared with that of the anti-VP23 serum used, and possibly the internal location of the VP37-sequence within both 95K and 105K (see below).

Mapping of the cleavage products. To map the cleavage products more precisely within 95K and 105K, M RNA was translated in a wheat germ extract, and samples were taken at intervals of 10 min and processed by addition of 30,000 × *g* supernatant of B-inoculated protoplasts. Full-length (95K and 105K) products were detectable only after 30 min (Fig. 5). Processing of the 20-min sample, exclusively containing incomplete primary products, revealed that the 48K and 58K sequences were synthesized first (and simultaneously) and are therefore located N terminally within both the 95K and 105K polypeptides, respectively (Fig. 5). The 60K capsid protein precursor clearly appeared last, when full-size 95K and 105K polypeptides were produced, indicating that its sequence is located C terminally within both 105K and 95K polypeptides. This finding implies that the full-size 95K and 105K polypeptides differ in their N termini and confirms that there are two initiation sites on M RNA in vitro. Whether two different initiation sites on M RNA are actually used in vivo or are just the result of artificial events in the cell-free systems remains to be answered, however.

Attempts to cleave the 60K polypeptide into mature capsid proteins. The finding that a B RNA-coded protease present in vivo is capable of specifically releasing a 60K capsid protein precursor from M RNA in vitro products strongly suggests that the in vitro studies reflect the situation within the infected cell. The question arises then of why the 60K precursor was not further processed into the mature capsid proteins. A possibility might be that this final cleavage is generated by a second protease which is membrane bound and therefore absent from the 30,000 × *g* supernatant fraction of CPMV-infected protoplasts. Unfractionated homogenate of protoplasts infected with complete virus (M + B) displayed only the same proteolytic activity as was found in the 30,000 × *g* supernatant fraction, however (Fig. 3, lanes 3 and 12). A second possibility is that the 60K polypeptides have to be assembled into a procapsid structure and that the presence of unpackaged viral RNA is necessary for maturation of this procapsid. Addition of excess viral RNA to a mixture of M RNA in vitro product and unfractionated homogenate of CPMV-infected protoplasts did not show any effect (Fig. 3, lanes 4 and 11). On the other hand, when M RNA products were incubated with a membrane protein fraction (the so-called 30,000 × *g* pellet extract; for details, see Materials and Methods) prepared from intact leaves of infected plants, 93K, 79K, and 23K

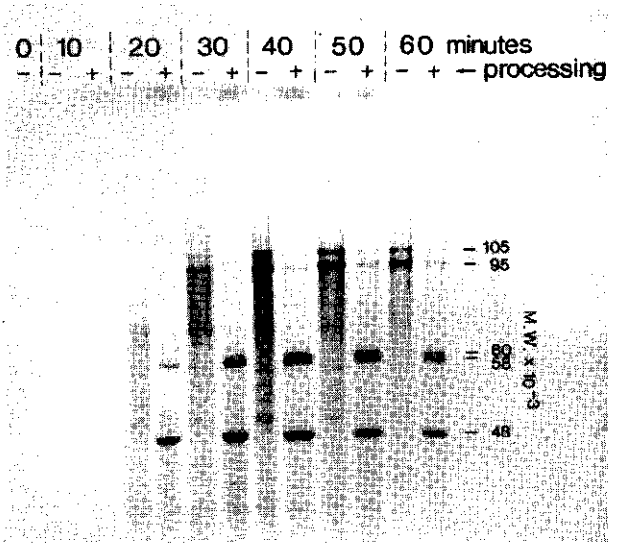


FIG. 5. Time course of appearance of in vitro translation products. CPMV M RNA was translated in wheat germ extract. Samples were removed from the translation mixture at the times indicated and were either treated with an equal volume of 30,000 × *g* supernatant fraction from B-inoculated protoplasts (+) or left untreated (−); they were analyzed on a 12.5% polyacrylamide gel.

polypeptides were generated, in addition to the 60K, 58K, and 48K products (Fig. 6, lanes 4 and 5). Extracts from uninfected leaves did not contain this proteolytic activity (Fig. 6, lane 2). The 23K cleavage product comigrated with the small capsid protein VP23 (cf. lanes 3–5 in Fig. 6) and precipitated with anti-VP23 serum (result not shown). Since VP37 was not produced by this leaf extract, the 95K and 105K polypeptides were apparently cleaved into a polypeptide of 79K and VP23 and into a polypeptide of 93K and VP23, respectively, by a proteolytic activity recognizing the cleavage site between the VP23 and VP37 sequences. This activity seemed to be independent from the activity present in infected protoplasts, as incubation of the 95K and 105K polypeptides in a 1:1 (vol/vol) mixture of leaf and protoplast extracts did not result in further cleavage of the 60K polypeptide (data not shown). The release of VP23 and not of VP37 from 95K and 105K polypeptides enables us to map both capsid proteins within the 60K precursor, as summarized in the cleavage model of Fig. 7.

DISCUSSION

We have previously found (9, 19) that in cowpea protoplasts B RNA is replicated independently, i.e., in the absence of M RNA, allowing in vivo studies on the expression of this RNA. On the other hand, M RNA is not replicated independently from B RNA. Therefore, a direct study of the expression of M RNA in vivo has been impossible. The results presented in this paper show that the translation strategy and coding function of M RNA can be studied by supplying extracts of CPMV-infected plant cells to the M RNA-encoded 95K and 105K polypeptides synthesized in cell-free systems. It has been demonstrated now that a B RNA-coded protease present in both infected leaves and infected mesophyll protoplasts is capable of cleaving the M RNA in vitro translation products, generating a 60K precursor to both capsid proteins VP23 and VP37. Complementation studies using particles from different CPMV isolates have suggested that at least one of the capsid proteins is coded for by M RNA (11, 28). We have now definitely established that M RNA encodes both VP23 and VP37. The cleavage map shown in Fig. 7 summarizes the results described in this paper. The model proposes that M RNA is translated in vitro into two overlapping polypeptides of 95K and 105K which are cleaved into pieces of 60K, 58K, 48K, and 47K by a protease translated from B RNA in vivo. Of

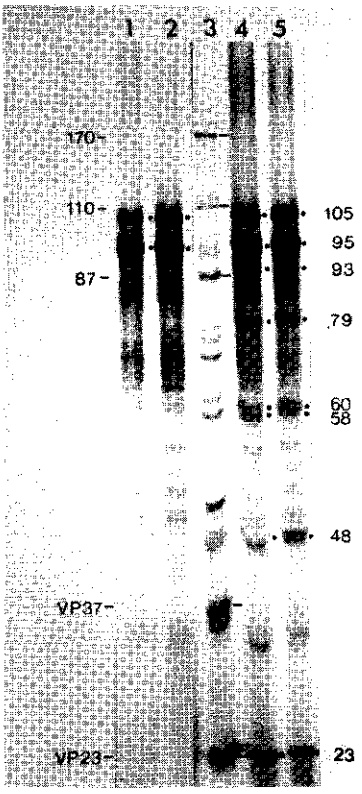


FIG. 6. In vitro processing of CPMV M RNA translational products using extracts from CPMV-infected cowpea leaves. M RNA was translated in wheat germ extract, and portions (4 μl) of the products obtained were mixed with an equal volume of either TGKEDP buffer (sample 1) or 30,000 × *g* pellet extract (for details, see Materials and Methods) from CPMV-infected (samples 4 and 5) or uninfected (sample 2) cowpea leaves. They were incubated for 1 h at 30°C. Samples 4 and 5 were treated with different leaf extracts. Products were analyzed in a 12.5% polyacrylamide gel. Lane 3 contains [^{35}S]methionine-labeled polypeptides from CPMV-infected protoplasts. The numbers indicated at the right side of the gel refer to the molecular weights ($\times 10^{-3}$) of the (processed) in vitro products.

these cleavage products the 60K polypeptide represents the direct precursor to VP23 and VP37, whereas the function of the other products is as yet unknown. Although a second virus-specific protease activity has been detected in extracts from infected leaves, capable of generating cleavage between the VP23 and VP37 sequence within 95K and 105K, the conditions necessary for final processing of the 60K precursor are not yet understood and are the subject of further investigation. Neither mixing of leaf and protoplast extract nor addition of excess viral RNA resulted in cleavage of the 60K polypeptide. A possibility is that the 60K polypeptides should be arranged in a specific procapsid structure to allow final cleavage and that this assembly does not occur in vitro. A similar phenomenon has been described for the assembly and maturation of picornaviruses (for review, see 23). After the viral RNA has been packaged, the final maturation step in poliovirus (and encephalomyocarditis [EMC] virus) particles is cleavage of VP0 (EMC virus: ε) into the mature capsid proteins VP2 and VP4 (EMC virus: ε → β + δ).

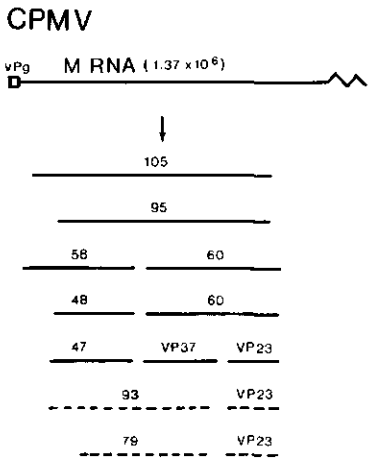


FIG. 7. Model for the expression of CPMV M RNA. In cell-free protein-synthesizing systems M RNA is translated into two overlapping polypeptides of 95 and 105K, which differ in their N termini as a result of two active initiation sites on M RNA in vitro. Both polypeptides are cleaved by a B RNA-encoded protease present in infected cells, resulting in polypeptides of 60, 58, 48, and 47K. The 60K polypeptide represents a precursor to both capsid proteins VP23 and VP37. As this precursor is generated by a viral protease from infected cells, it is proposed that the model presented reflects the translation strategy of M RNA in vivo. The dotted lines represent the products generated by a second virus-induced protease detectable in infected leaves but not in infected protoplasts. The genome-linked protein VPg at the 5'-end of both CPMV RNAs (4, 27) has been indicated with a box, and the polyadenylate tail at the 3' end (1, 7) has been indicated with a zigzag line. It is not known whether translation of CPMV RNA is preceded by removal of VPg.

In reticulocyte lysates polioviral (and EMC viral) RNA is translated into a large polyprotein which is then processed extensively (15, 24, 25). One of the products is VPO (EMC virus: ε), which is not processed further, however, indicating that final assembly of virions also has been impeded.

The experiments presented in this paper confirm the observation of Pelham (16) that in vitro M RNA contains two active initiation sites. The overlapping 95K and 105K polypeptides are synthesized in two different cell-free systems (Fig. 3) (10), but it is not known whether they are both produced in vivo. The production of polypeptides with completely overlapping amino acid sequences does not appear useful and might be the result of the artificial conditions in cell-free extracts. In view of the faithful translation of B RNA in both in vitro translation systems (10, 19) and the generation of a specific 60K capsid protein precursor upon cleavage by an in vivo B RNA-coded protease, at least one of the M RNA in vitro products probably represents a correct primary translation product even if it has not been detected in vivo. The use of in vitro protein-synthesizing systems, in combination with well-defined subcellular fractions from infected cells, may provide the means for studying the remaining coding functions of M RNA and the mechanism of CPMV assembly.

ACKNOWLEDGMENTS

We thank Geertje Rezelman and Robert Pels Rijcken for help in some of the experiments, Peter Blackburn for the gift of human placental RNase inhibitor, Hugh Pelham and Richard Jackson for the gift of rabbit reticulocyte lysate, Pim Zabel for suggesting the immunization procedure and critical reading of the text, and Annelies Bruins for typing the manuscript.

This work was supported by the Netherlands Foundation for Chemical Research, with financial aid from the Netherlands Organization for the Advancement of Pure Research.

LITERATURE CITED

1. **Ahlquist, P., and P. Kaesberg.** 1979. Determination of the length distribution of poly(A) at the 3' terminus of the virion RNAs of EMC virus, poliovirus, rhinovirus, RAV-61 and CPMV and of mouse globin mRNA. Nucleic Acids Res. 7:1195–1204.
2. **Blackburn, P.** 1979. Ribonuclease inhibitor from human placenta: rapid purification and assay. J. Biol. Chem. 254:12484–12487.
3. **Bowen, B., J. Steinberg, U. K. Laemmli, and H. Weintraub.** 1980. The detection of DNA-binding proteins by protein blotting. Nucleic Acids Res. 8:1–20.
4. **Daubert, S. D., G. Bruening and R. C. Najarian.** 1978. Protein bound to the genome RNAs of cowpea mosaic virus. Eur. J. Biochem. 92:45–51.
5. **Davies, J. W., A. M. J. Aalbers, E. J. Stnik, and A. van Kammen.** 1977. Translation of cowpea mosaic virus RNA in a cell-free extract from wheat germ. FEBS Lett. 77:265–269.
6. **Davies, J. W., J. W. G. Verver, R. W. Goldbach, and A. van Kammen.** 1978. Efficient reverse transcription of cowpea mosaic virus RNAs. Nucleic Acids Res. 5:4643–4661.
7. **El Manna, M. M., and G. Bruening.** 1973. Polyadenylate sequences in the ribonucleic acids of cowpea mosaic virus. Virology 56:198–206.
8. **Goldbach, R. W., P. Borst, J. E. Bollen-de Boer, and E. F. J. van Bruggen.** 1978. The organization of ribosomal RNA genes in the mitochondrial DNA of *Tetrahymena pyriformis* strain ST. Biochim. Biophys. Acta 521:169–186.
9. **Goldbach, R., G. Rezelman, and A. van Kammen.** 1980. Independent replication and expression of B-component RNA of cowpea mosaic virus. Nature (London) 286:297–300.
10. **Goldbach, R. W., J. G. Schilthuis, and G. Rezelman.** 1981. Comparison of *in vivo* and *in vitro* translation of cowpea mosaic virus RNAs. Biochem. Biophys. Res. Commun. 99:89–94.
11. **Gopo, J. M., and R. H. Frist.** 1977. Location of the gene specifying the smaller protein of the cowpea mosaic virus capsid. Virology 79:259–266.
12. **Klootwijk, J., I. Klein, P. Zabel, and A. van Kammen.** 1977. Cowpea mosaic virus RNAs have neither m^7GpppN . . . . nor mono-, di-, or triphosphates at their 5'-ends. Cell 11:73–82.
13. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of the bacteriophage T_4. Nature (London) 227:680–685.
14. **Pelham, H. R. B.** 1978. Leaky UAG termination codon in tobacco mosaic virus RNA. Nature (London) 272:469–471.
15. **Pelham, H. R. B.** 1978. Translation of encephalomyocarditis virus RNA *in vitro* yields an active proteolytic processing enzyme. Eur. J. Biochem. 85:457–462.
16. **Pelham, H. R. B.** 1979. Synthesis and proteolytic processing of cowpea mosaic virus proteins in reticulocyte lysates. Virology 96:463–477.
17. **Pelham, H. R. B.** 1979. Translation of fragmented viral RNA *in vitro*. FEBS Lett. 100:195–199.
18. **Reijnders, L., A. M. J. Aalbers, A. van Kammen, and R. W. J. Thuring.** 1974. Molecular weights of plant viral RNAs determined by gel electrophoresis under denaturing conditions. Virology 60:515–521.
19. **Rezelman, G., R. Goldbach, and A. van Kammen.** 1980. Expression of bottom component RNA of cowpea mosaic virus in cowpea protoplasts. J. Virol. 36:366–373.
20. **Rottier, P. J. M., G. Rezelman, and A. van Kammen.** 1979. The inhibition of cowpea mosaic virus replication by actinomycin D. Virology 92:299–309.
21. **Scheele, G., and P. Blackburn.** 1979. Role of mammalian RNase inhibitor in cell-free protein synthesis. Proc. Natl. Acad. Sci. U.S.A. 76:4898–4902.
22. **Schiff, R. D., and D. P. Grandgenett.** 1980. Virus-coded origin of a 32,000-dalton protein from avian retrovirus cores: structural relatedness of p32 and the β polypeptide of the avian retrovirus DNA polymerase. J. Virol. 28:279–291.
23. **Scraba, D. G.** 1979. The picornavirion: structure and assembly, p. 1–23. In R. Perez-Bercoff (ed.), The molecular biology of picornaviruses. Plenum Press, New York.
24. **Shih, D. S., C. T. Shih, O. Kew, M. Pallansch, R. Rueckert, and P. Kaesberg.** 1978. Cell-free synthesis and processing of the proteins of poliovirus. Proc. Natl. Acad. Sci. U.S.A. 75:5807–5811.
25. **Shih, D. S., C. T. Shih, D. Zimmern, R. R. Rueckert, and P. Kaesberg.** 1979. Translation of encephalomyocarditis virus RNA in reticulocyte lysates: kinetic analysis of the formation of virion proteins and a protein required for processing. J. Virol. 30:472–480.
26. **Stanley, J., R. Goldbach, and A. van Kammen.** 1980. The genome-linked protein of cowpea mosaic virus is coded by RNA from the bottom component. Virology 106:180–182.
27. **Stanley, J., P. Rottier, J. W. Davies, P. Zabel, and A. van Kammen.** 1978. A protein linked to the 5' termini of both RNA components of the cowpea mosaic virus genome. Nucleic Acids Res. 5:4505–4522.
28. **Thongmeearkom, P., and R. M. Goodman.** 1978. Complementation and pseudorecombination between ribonucleic

acids from two natural isolates of cowpea mosaic virus (severe subgroup). Virology 85:75–83.
29. **Tjian, R., D. Stinchcomb, and R. Losick.** 1974. Antibody directed against *Bacillus subtilis* σ factor purified by sodium dodecyl sulfate slab gel electrophoresis. J. Biol. Chem. 250:8824–8828.
30. **Zabel, P., I. Jongen-Neven, and A. van Kammen.** 1976. In vitro replication of cowpea mosaic virus RNA. II. Solubilization of membrane-bound replicase and the partial purification of the solubilized enzyme. J. Virol. 17:679–685.
31. **Zabel, P., I. Jongen-Neven, and A. van Kammen.** 1979. In vitro replication of cowpea mosaic virus RNA. III. Template recognition by cowpea mosaic virus RNA replicase. J. Virol. 29:21–33.

CHAPTER IV

LIMITS TO THE INDEPENDENCE OF BOTTOM COMPONENT RNA OF COWPEA MOSAIC VIRUS

G. Rezelman[1], H.J. Franssen[1], R.W. Goldbach[1], T.S. Ie[2] and A. van Kammen[1].

1. Department of Molecular Biology, Agricultural University, De Dreijen 11, 6703 BC Wageningen, The Netherlands.
2. Department of Virology, Agricultural University, Binnenhaven 11, 6709 PD Wageningen, The Netherlands.

J.gen.Virol. 60, 335-342, (1982).

J. gen. Virol. (1982), **60**, 335–342. *Printed in Great Britain*

Key words: *CPMV/bottom component/transport/electron microscopy*

Limits to the Independence of Bottom Component RNA of Cowpea Mosaic Virus

By G. REZELMAN,[1] H. J. FRANSSEN,[1] R. W. GOLDBACH,[1]* T. S. IE[2] AND A. VAN KAMMEN[1]

[1]*Department of Molecular Biology, Agricultural University, De Dreijen 11, 6703 BC Wageningen and* [2]*Department of Virology, Agricultural University, Binnenhaven 11, 6709 PD Wageningen, The Netherlands*

(*Accepted 19 January 1982*)

SUMMARY

Electron microscopic analyses have revealed that the bottom (B) component of cowpea mosaic virus alone induces cytopathic structures in cowpea mesophyll protoplasts, similar to those induced by the complete virus [i.e. B plus middle (M) components]. This indicates that the development of such structures is not linked to accumulation of virus particles but to virus RNA replication and expression. When purified B component was inoculated to primary cowpea leaves, symptoms were not produced and B component RNA was incapable of spreading to surrounding cells. The results are discussed in terms of limits to the independence of B component RNA and of possible functions for M component RNA-encoded proteins.

INTRODUCTION

Cowpea mosaic virus (CPMV) is a plant virus with an RNA genome distributed between two nucleoprotein particles, the bottom (B) and middle (M) component (Van Kammen, 1972). Both components, or their RNAs (denoted B RNA and M RNA, with mol. wt. $2{\cdot}02 \times 10^6$ and $1{\cdot}37 \times 10^6$ respectively; Reijnders *et al.*, 1974) are necessary for virus multiplication (Bruening, 1977; Van Kammen, 1968). We have shown previously that B component RNA is able to replicate independently (i.e. in the absence of M component), a property not associated with M component RNA (Goldbach *et al.*, 1980). Direct studies on the expression of B component RNA *in vivo* have therefore been possible (Rezelman *et al.*, 1980; Goldbach *et al.*, 1982).

CPMV infection is accompanied by the appearance of characteristic cytopathic structures (Assink *et al.*, 1973; De Zoeten *et al.*, 1974). These structures consist of amorphous electron-dense material and a large number of vesicles. We have now investigated whether the B component alone is able to induce these cytopathic structures in protoplasts. Furthermore, as the replication and expression of B component RNA has mainly been studied in isolated mesophyll protoplasts, we have now followed its fate in the intact host. We have studied whether B RNA is capable of spreading to surrounding cells in the absence of M RNA. This property has been reported for the large RNA (RNA-1) of another two-component virus, tobacco rattle virus (TRV) (Sänger & Brandenburg, 1961; Cadman, 1962; Lister, 1968, 1969).

The results presented in this paper demonstrate that the independence of the B component is limited.

METHODS

Virus purification and separation of B and M components. CPMV was propagated in

0022-1317/82/0000-4916 $02.00

cowpea plants (*Vigna unguiculata* L. 'California Blackeye') and purified as described previously (Van Kammen, 1967). B and M components were separated in a linear 15 to 30% sucrose gradient by zonal centrifugation (Beckman Ti 15 rotor, 16 h, 23000 rev/min at 10 °C). This procedure was repeated twice to give an M component free of B, and a B component contaminated with less than 0·2% of M, as determined by the local lesion infectivity test (De Jager, 1976).

Protoplast isolation and inoculation. Cowpea mesophyll protoplasts were isolated and inoculated with CPMV components [5 μg of M components, 5 μg of B components, or 10 μg of a 1:1 (w/w) mixture of both components per 5×10^5 protoplasts per ml inoculum] or left untreated as described in detail previously (Hibi *et al.*, 1975; Rottier *et al.*, 1979). Inoculated protoplasts were incubated in culture medium at 25 °C as described by Rottier *et al.* (1979). In some experiments protoplasts were reinoculated with CPMV components 24 h after their preparation. For that purpose, protoplasts incubated for 23 h in culture medium were sedimented and resuspended in a solution of 0·5% (w/v) cellulase in 0·6 M-mannitol pH 5·6. After 1 h incubation at 25 °C the protoplasts were washed twice with 0·6 M-mannitol and resuspended in 0·6 M-mannitol, 0·01 M-potassium citrate pH 5·2, containing 0·5 or 1·0 μg poly-L-ornithine (PLO) (Pilot Chemicals, New England Nuclear) per ml. Concurrently, a solution of unfractionated virus (5 μg/ml) or purified M component (2·5 μg/ml) was made in the same buffer (also containing 0·5 or 1·0 μg/ml PLO). Both the virus solution and the protoplast suspension were kept for 5 min at room temperature. The protoplasts were then sedimented and resuspended in the virus solution. After 15 min the protoplasts were washed three times in 0·6 M-mannitol containing 10 mM-$CaCl_2$, and were finally resuspended and incubated in culture medium.

Electron microscopy. Samples of protoplasts were collected by centrifugation (2 min, 600 *g*) 24 h or 40 h after inoculation and incubated in 2% glutaraldehyde, 0·1 M-sodium phosphate pH 7, 0·6 M-mannitol for 1 h at 4 °C. After five successive washes with 0·1 M-sodium phosphate pH 7, they were fixed in 1% osmium tetroxide in 0·1 M-sodium phosphate pH 7 for 1 h at 4 °C, washed in double-distilled water and in 0·14 M-veronal acetate pH 5 (containing 0·577 g sodium barbiturate and 0·38 g sodium acetate per 100 ml) successively, and stained for 1 h in 2% uranyl acetate in 0·14 M-veronal acetate pH 5. Finally, the fixed protoplasts were washed in veronal acetate (0·14 M, pH 5), dehydrated in ethanol and acetone, and embedded in a prepolymerized mixture of methacrylate and divinyl benzene by the method of Kushida (1961). Polymerization took place in gelatin capsules at 50 °C for 48 h. Ultrathin sections made with an LKB Ultrotome III ultramicrotome were stained with uranyl acetate and lead citrate by the method of Reynolds (1963), and examined in a Siemens Elmiskop 101 electron microscope, operated at an accelerating voltage of 80 kV.

Assay for spreading of CPMV components in cowpea leaves. Primary leaves of 8- or 9-day-old cowpea plants were inoculated with B component (5 μg in 100 μl 0·01 M-sodium phosphate pH 7 per leaf) or with a mixture of M + B components (5 μg of each in 100 μl of the same buffer). After 24 or 48 h protoplasts were isolated from the inoculated leaves. Protoplasts prepared from B-inoculated leaves were divided into three aliquots at a concentration of 5×10^5 protoplasts per ml. One portion was inoculated with M component (5 μg per 5×10^5 protoplasts), and, to verify whether the protoplasts were infectable with CPMV, a second portion was inoculated with a mixture of B and M components (5 μg of each per 5×10^5 protoplasts). A third portion was left untreated. Protoplasts isolated from M + B-infected leaves were left untreated to measure the spreading of the complete virus. To determine the percentage of cells containing virus particles, samples of protoplasts were stained with fluorescent antibodies against CPMV (Hibi *et al.*, 1975) 0, 25 and 42 h after inoculation.

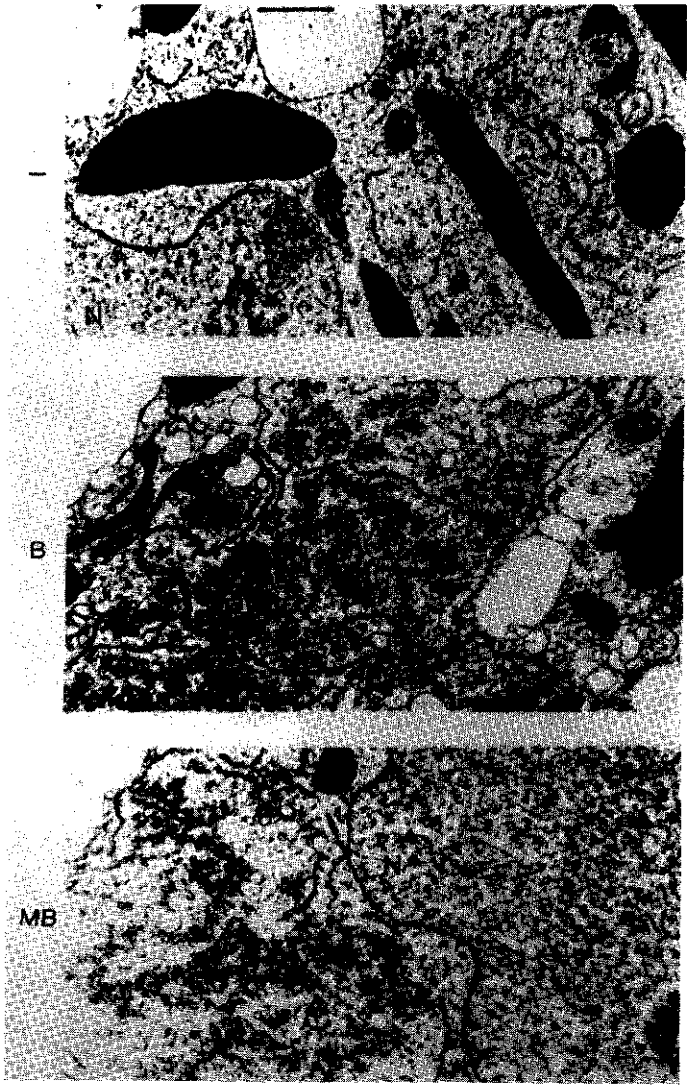


Fig. 1. Electron micrographs of cowpea mesophyll protoplasts: (*a*) uninoculated; (*b*) inoculated with B component; (*c*) inoculated with a 1:1 (w/w) mixture of B + M components. Portions of protoplasts were mounted for electron microscopy 24 h after inoculation. Arrows indicate the characteristic cytopathological structures (i.e. electron-dense material); N, nucleus. Bar marker represents 1 μm.

RESULTS

Electron microscopy

Cowpea mesophyll protoplasts, inoculated with B, M, or B + M components, were analysed by electron microscopy 24 h and 40 h after inoculation. Cytopathic structures were

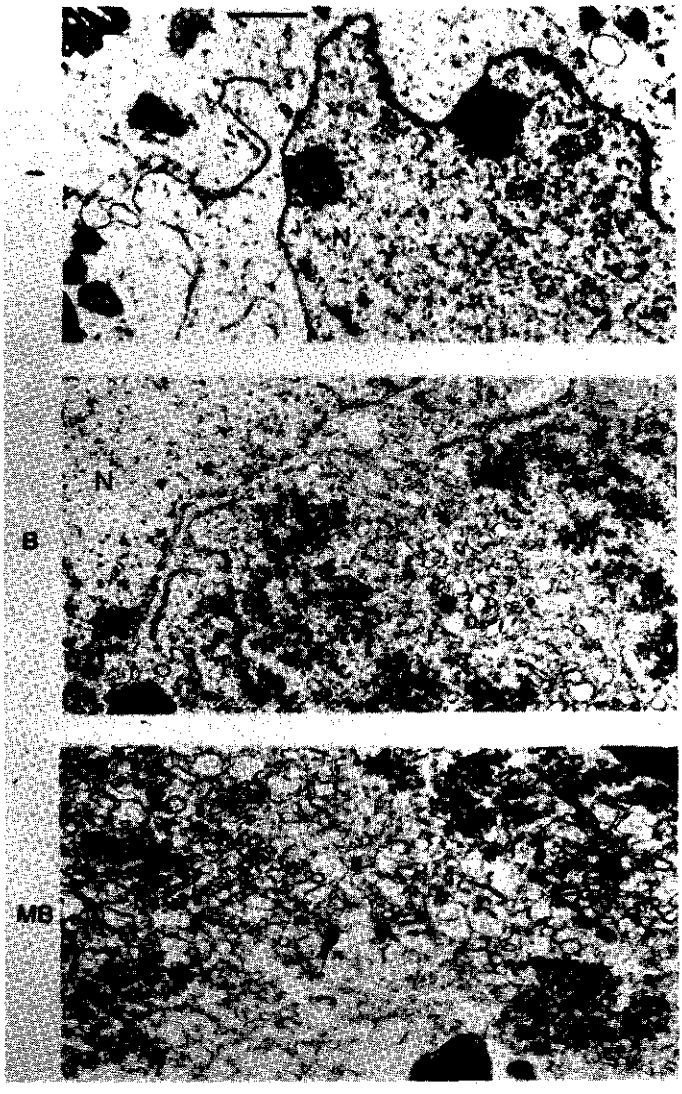


Fig. 2. Electron micrographs of cowpea mesophyll protoplasts: (*a*) uninoculated; (*b*) inoculated with B component; (*c*) inoculated with a 1:1 (w/w) mixture of B + M components. Portions of protoplasts were mounted for electron microscopy 40 h after inoculation. Arrows indicate the characteristic cytopathological structures (i.e. vesicles); N, nucleus. Bar marker represents 1 μm.

found in protoplasts inoculated with complete virus and in protoplasts inoculated with only B component, but not in protoplasts inoculated with M component (Fig. 1, 2). The morphology of the cytopathic structures in the inoculated protoplasts changed during the time after inoculation. Twenty-four h after inoculation (Fig. 1) these structures contained amorphous

Table 1. *Induction of cytopathic structures in cowpea protoplasts by components of CPMV*

Inoculum	Percentage of cells containing	
	c.p.s.*	CPMV particles†
–	0	0
B	71	4
M	0	0
B + M	68	67

* c.p.s., Cytopathic structures.
† As determined by staining with fluorescent anti-CPMV serum.

electron-dense material clustered within membrane structures. Forty h after inoculation (Fig. 2) the amount of electron-dense material diminished, but the number of vesicles surrounding the remaining electron-dense material increased. Such structures were absent in both healthy protoplasts (Fig. 1*a*, 2*a*) and in M component-inoculated protoplasts (data not shown). The results indicate that B component alone is capable of inducing the cytological alterations typical of CPMV infection. The data presented in Table 1 show that in the case of inoculation with B + M particles the percentage of protoplasts containing cytopathic structures (68%) and the percentage of infected protoplasts as determined by fluorescence (67%) were in good agreement. This correlation strongly suggests that all virus-containing cells contain cytopathic structures. In the case of B-inoculated protoplasts, the percentage of cells containing cytopathic structures was as high (71%) as for the B + M-inoculated cells, but the percentage of fluorescent cells was only 4%. Production of capsid proteins by these cells must have resulted from minor (0·2% or less) contamination of the B component preparation with M components. Therefore, it appears that B component alone is capable of inducing cytopathic structures in protoplasts with the same efficiency as the complete virus.

Is B RNA capable of spreading to surrounding cells?

In protoplasts inoculated with B component alone, B RNA is replicated and expressed to the same extent as it is in protoplasts inoculated with B + M components (Goldbach *et al.*, 1980; Rezelman *et al.*, 1980). To investigate whether B RNA is also able to act independently from M RNA in the host plant, we have followed the fate of this RNA upon infection of primary cowpea leaves with B component. Since transfer of B RNA from cell to cell is necessary for detectable expression of B component RNA in leaves, we have determined the number of leaf cells containing B RNA upon inoculation with B component. For that purpose, protoplasts from B-inoculated leaves were prepared at various times after inoculation, and these protoplasts were subsequently inoculated with excess M particles. The rationale was that B RNA-containing cells which become inoculated with M particles should begin production of B and M particles, which can easily be detected by staining these cells with fluorescent antibodies against CPMV. Table 2 shows that the complete virus (M + B) rapidly spreads throughout the leaf, reaching up to 40% of the cells within 48 h. In contrast, the number of fluorescent protoplasts from B-inoculated leaves hardly increased (from 1 to 4%, Table 2 and Fig. 3) within the first 48 h after inoculation of leaves. Since addition of excess M particles to these cells did not further increase the number of fluorescent cells (Table 2 and Fig. 3) we conclude that the number of cells containing only B component RNA was negligible. The reliability of the assay used in this experiment was verified by inoculating one portion of protoplasts from B-inoculated leaves with a 1:1 (w/w) mixture of B + M components. Approximately 40 to 50% (Table 2) of these protoplasts produced virus particles, demonstrating that they were accessible to infection by CPMV particles.

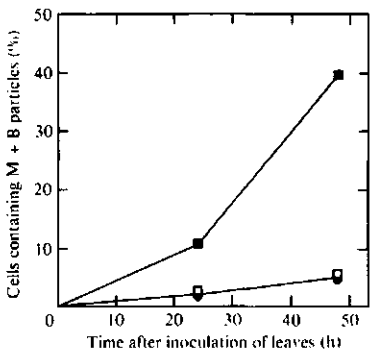


Fig. 3. Transport of CPMV components throughout the cowpea leaf. Leaves were inoculated with a 1 : 1 (w/w) mixture of B + M components, or with B component alone. Protoplasts were prepared from these leaves at the times indicated. ■, B + M-inoculated leaves; ●, B-inoculated leaves. One portion of protoplasts prepared from B-inoculated leaves was inoculated with excess M component (□). After 42 h incubation of the protoplasts, percentages of infected cells were determined by staining with fluorescent anti-CPMV serum.

Table 2. *Spreading of CPMV components throughout the cowpea leaf*

Leaf inoculum	Preparation of protoplasts (h post-inoculation)*	Protoplast inoculum	CPMV-infected cells (%)† 0 h‡	25 h‡	42 h‡
B	24	–	0	0·6	1·4
		M		0·7	1·8
		B + M		48·4	53·0
B	48	–	4·4	3·0	3·8
		M		3·5	4·4
		B + M		41·7	37·0
B + M	2	–			<1·0
B + M	24	–	5·5	8·2	10·6
B + M	48	–	39·9	35·5	39·7

* Protoplasts were isolated from leaves at the times indicated.
† As determined by staining with fluorescent anti-CPMV serum.
‡ Time of incubation of protoplasts.

Table 3. *Reinoculation of cowpea protoplasts with CPMV components*

Inoculum*	Reinoculum†	PLO (μg/ml)‡	CPMV-infected cells (%)§
–	B + M	–	1·5
–	B + M	0·5	32
–	B + M	1·0	35
B	–	–	6
B	B + M	0·5	12
B	B + M	1·0	20
B	M	0·5	16
B	M	1·0	20

* Inoculated immediately after isolation of protoplasts.
† Reinoculated 24 h after isolation of protoplasts.
‡ Concentration of poly-L-ornithine used for reinoculation.
§ As determined by staining with fluorescent anti-CPMV serum 67 h after isolation of protoplasts.

Independence of CPMV B RNA

To exclude the possibility that, in protoplasts isolated from B-inoculated leaves, B RNA was not able to support the replication and expression of newly entered M RNA, the following control experiment was performed. Protoplasts inoculated with B component were first incubated for 23 h under standard conditions. They were then treated with cellulase (to remove the newly formed cell walls) and reinoculated with M component. Forty-four h later (i.e. 67 h after the first inoculation) they were stained with fluorescent anti-CPMV to determine the number of CPMV-producing cells. Whereas PLO is not necessary for efficient infection of freshly prepared protoplasts (Hibi *et al.*, 1975; Rottier *et al.*, 1979), its presence appeared to be a prerequisite for infection of protoplasts aged for 24 h. In the absence of PLO, only 1·5% of such protoplasts were infected upon inoculation with CPMV (B + M), whereas in the presence of PLO (0·5 or 1·0 μg/ml) this percentage was approx. 30 to 35% (Table 3). Using 1·0 μg/ml PLO, reinoculation of protoplasts with M component, 24 h after they had been inoculated with B component, resulted in 20% CPMV-producing cells, compared to only 6% if these protoplasts had not been reinoculated with M component (Table 3). As the same percentage (20%) was reached after reinoculation with complete virus (Table 3) this result shows that at least in a major part of cells, which have been first inoculated with B component and 24 h later with M component, B RNA can support the replication and expression of M RNA. The low number of protoplasts from B-inoculated leaves which fluoresce after inoculation with M component (Table 2) therefore reflects the inability of B RNA to spread independently to surrounding cells.

DISCUSSION

It has been previously shown that the B component RNA of CPMV is capable of self replication (Goldbach *et al.*, 1980) whereas M RNA is not. In studies on the expression of B RNA in mesophyll protoplasts, at least seven B RNA-coded proteins with sizes of 170K, 110K, 87K, 84K, 60K, 32K and 4K could be detected (Rezelman *et al.*, 1980; Stanley *et al.*, 1980). Although functions could not be assigned to any of these polypeptides, at least one of them should be involved in virus RNA replication (Goldbach *et al.*, 1980) and another should represent the protease responsible for the *in vitro* cleavage of the M RNA-coded primary translation products (Pelham, 1979; Goldbach *et al.*, 1981; Franssen *et al.*, 1982). As it has recently been shown that M RNA encodes both capsid proteins (Franssen *et al.*, 1982), B RNA should exist as an unencapsidated molecule in B component-inoculated protoplasts. In this paper we have shown that this unencapsidated B RNA and its products, rather than the accumulation of virus particles, are responsible for the induction of the characteristic cytopathic structures in CPMV-infected cells. Our results are consistent with the finding of Hibi *et al.* (1975) that RNA replication is associated with the appearance of the cytopathic structures. Furthermore, the electron microscopic analyses presented here show that 24 h after infection these structures mainly consist of electron-dense material as described previously (Assink *et al.*, 1973; Hibi *et al.*, 1975) but that with advancing time (i.e. 40 h after infection) the infected cell develops a large number of vesicles. Although B RNA appears to replicate independently in isolated protoplasts, our results demonstrate that non-packaged B RNA is unable to spread throughout leaves (Table 2 and Fig. 3). This indicates that M RNA encodes one or more proteins essential for the transport of the virus RNA molecules. In this context, it is worth mentioning that apart from the capsid proteins, two other polypeptides (58K and 48K) are translated from M RNA *in vitro* (Franssen *et al.*, 1982). For TRV, a rod-shaped two-component virus, it has been shown (Lister, 1968, 1969) that the RNA from the long particle (which does not carry the information for the virus coat protein) can replicate by itself and this replication is accompanied by spreading of the virus RNA through the infected plant and by the production of symptoms. Although B RNA of CPMV is unable to cause local lesions in cowpea leaves, this does not mean *a priori* that B RNA is incapable

of spreading as expression of B RNA in the leaf may not have been associated with visible symptoms. We have now demonstrated that the lack of symptom development is related to the inability of B RNA to spread. The question as to whether encapsidation is solely essential for transport of CPMV RNAs or whether another M RNA-encoded function is involved in this process needs further investigation.

We thank R. Brouwer for help in initial experiments and H. Lohuis for technical assistance. This work was supported by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for the Advancement of Pure Research (ZWO).

REFERENCES

ASSINK, A. M., SWAANS, H. & VAN KAMMEN, A. (1973). The localization of virus-specific double-stranded RNA of cowpea mosaic virus in subcellular fractions of infected Vigna leaves. *Virology* **53**, 384–391.

BRUENING, G. (1977). Plant covirus systems: two component systems. In *Comprehensive Virology*, vol. 11, pp. 55–141. Edited by H. Fraenkel-Conrat & R. R. Wagner. New York: Plenum Press.

CADMAN, C. H. (1962). Evidence for association of tobacco rattle virus nucleic acid with a cell component. *Nature, London* **193**, 49.

DE JAGER, C. P. (1976). Genetic analysis of cowpea mosaic virus mutants by supplementation and reassortment tests. *Virology* **70**, 151–163.

DE ZOETEN, G. A., ASSINK, A. M. & VAN KAMMEN, A. (1974). Association of cowpea mosaic virus induced double-stranded RNA with a cytopathic structure in infected cells. *Virology* **59**, 341–355.

FRANSSEN, H., GOLDBACH, R., BROEKHUIJSEN, M., MOERMAN, M. & VAN KAMMEN, A. (1982). Expression of middle component RNA of cowpea mosaic virus: *in vitro* generation of a precursor to both capsid proteins by a bottom component RNA-encoded protease from infected cells. *Journal of Virology* **41**, 8–17.

GOLDBACH, R. W., REZELMAN, G. & VAN KAMMEN, A. (1980). Independent replication and expression of B-component RNA of cowpea mosaic virus. *Nature, London* **286**, 297–300.

GOLDBACH, R. W., SCHILTHUIS, J. G. & REZELMAN, G. (1981). Comparison of *in vivo* and *in vitro* translation of cowpea mosaic virus RNAs. *Biochemical and Biophysical Research Communications* **99**, 89–95.

GOLDBACH, R., REZELMAN, G., ZABEL, P. & VAN KAMMEN, A. (1982). Expression of the bottom component RNA of cowpea mosaic virus: evidence that the 60 kilodalton VPg-precursor is cleaved into single VPg and a 58 kilodalton polypeptide. *Journal of Virology* (in press).

HIBI, T., REZELMAN, G. & VAN KAMMEN, A. (1975). Infection of cowpea mesophyll protoplasts with cowpea mosaic virus. *Virology* **64**, 308–315.

KUSHIDA, H. (1961). A new embedding method for ultrathin sectioning using a methacrylate resin with three dimensional polymer structure. *Journal of Electron Microscopy* **10**, 194–197.

LISTER, R. M. (1968). Functional relationship between virus-specific products of infection by viruses of the tobacco rattle type. *Journal of General Virology* **2**, 43–58.

LISTER, R. M. (1969). Tobacco rattle NETU, virus in relation to functional heterogeneity of plant viruses. *Federation Proceedings* **28**, 1875–1889.

PELHAM, H. R. B. (1979). Synthesis and proteolytic processing of cowpea mosaic virus proteins in reticulocyte lysates. *Virology* **96**, 463–477.

REUNDERS, L., AALBERS, A. M. J., VAN KAMMEN, A. & THURING, R. W. J. (1974). Molecular weights of plant viral RNAs determined by gel electrophoresis under denaturing conditions. *Virology* **60**, 515–521.

REYNOLDS, E. S. (1963). The use of lead citrate at high pH as an electron opaque stain in electron microscopy. *Journal of Cell Biology* **17**, 208–212.

REZELMAN, G., GOLDBACH, R. W. & VAN KAMMEN, A. (1980). Expression of bottom component RNA of cowpea mosaic virus in cowpea protoplasts. *Journal of Virology* **36**, 366–373.

ROTTIER, P. J. M., REZELMAN, G. & VAN KAMMEN, A. (1979). The inhibition of cowpea mosaic virus replication by actinomycin D. *Virology* **92**, 299–309.

SÄNGER, H. L. & BRANDENBURG, E. (1961). Über die Gewinnung von Infektiosem aus 'Wintertyp'-Pflanzen des Tabak-Rattle-Virus durch Phenolextraktion. *Naturwissenschaften* **48**, 391.

STANLEY, J., GOLDBACH, R. W. & VAN KAMMEN, A. (1980). The genome-linked protein of cowpea mosaic virus is coded by RNA from the bottom component. *Virology* **106**, 180–182.

VAN KAMMEN, A. (1967). Purification and properties of the components of cowpea mosaic virus. *Virology* **31**, 635–642.

VAN KAMMEN, A. (1968). The relationship between the components of cowpea mosaic virus. I. Two ribonucleoprotein particles necessary for the infectivity of CPMV. *Virology* **34**, 312–318.

VAN KAMMEN, A. (1972). Plant viruses with a divided genome. *Annual Review of Phytopathology* **10**, 125–150.

(Received 22 September 1981)

CHAPTER V

EVIDENCE THAT THE 32,000-DALTON PROTEIN ENCODED BY THE BOTTOM-COMPONENT RNA OF COWPEA MOSAIC VIRUS IS A PROTEOLYTIC PROCESSING ENZYME

Henk Franssen, Marja Moerman, Geertje Rezelman and Rob Goldbach.

J. Virol. 50, 183-190 (1984).

JOURNAL OF VIROLOGY, Apr. 1984, p. 183–190
0022-538X/84/040183-08$02.00/0

Vol. 50, No. 1

Evidence That the 32,000-Dalton Protein Encoded by Bottom-Component RNA of Cowpea Mosaic Virus is a Proteolytic Processing Enzyme

HENK FRANSSEN, MARJA MOERMAN, GEERTJE REZELMAN, AND ROB GOLDBACH*

Department of Molecular Biology, Agricultural University, 6703 BC Wageningen, The Netherlands

Received 23 August 1983/Accepted 12 December 1983

Translation of middle-component RNA of cowpea mosaic virus in vitro produced two polypeptides of 95 and 105 kilodaltons (95K and 105K, respectively) with overlapping amino acid sequences, which were specifically cleaved by a protease encoded by the bottom-component RNA. The proteolytic cleavage was studied by the addition of antibodies raised against various bottom-component RNA-encoded proteins to extracts prepared from bottom-component RNA-inoculated cowpea protoplasts. Since antiserum to the 32K polypeptide efficiently inhibited the proteolytic activity of such extracts, although antiserum to VPg or to the 170K polypeptide did not, evidence was obtained which indicates that the 32K polypeptide represents the protease involved. Fractionation of proteolytically active extract by glycerol gradient centrifugation demonstrated that 32K polypeptides do not exist as free proteins but are aggregated to the bottom-component RNA-encoded 170K, 84K, 60K, or 58K polypeptides. Maximal proteolytic activity was observed for 32K polypeptides associated with 170K polypeptides, suggesting that the activity was unstable and confined to newly synthesized molecules.

Among plant viruses, cowpea mosaic virus (CPMV), the type member of the comoviruses, has become one of the most extensively studied viruses. The genome of this virus consists of two separately encapsidated plus-stranded RNA molecules with molecular weights of 2.01×10^6 (bottom-component [B] RNA) and 1.22×10^6 (middle-component [M] RNA), respectively (18, 26, 27). Both RNAs possess a genome-linked protein, denoted as VPg, at their 5'-terminus and are polyadenylated (1–3, 22, 23). Expression of the viral RNAs does not involve the generation of subgenomic mRNAs but occurs by means of proteolytic cleavages of long, primary translation products. Thus, it has been shown that B-RNA is translated into a 200-kilodalton (200K) polyprotein which is first cleaved to give a 32K and a 170K polypeptide (8, 16). The 170K polypeptide is then further cleaved to give either 60K and 110K or 84K and 87K polypeptides (20). Subsequent cleavage of the 60K polypeptide provides VPg and a 58K polypeptide (7, 28). M-RNA is translated into two polypeptides of 95K and 105K with overlapping amino acid sequences (4, 16). These polypeptides are proteolytically cleaved at the same position to give polypeptides of 60K, 58K, and 48K (4, 16). The 60K polypeptide is derived from the COOH-terminal half of both these polypeptides and represents the precursor to both capsid proteins VP37 and VP23, whereas the 58K and 48K cleavage products are derived from the NH_2-terminal parts of the longer (105K) and smaller (95K) polypeptides, respectively. The function of the overlapping 58K and 48K polypeptides is unknown as yet, but they may be involved in the spreading of viral RNA throughout the leaf (19). The protease responsible for the cleavage of the M-RNA-encoded primary translation products is obviously coded for by the B-RNA, since both in vitro translation products obtained from this RNA (i.e., the 170K and 32K polypeptides [8, 16]) and extract from cowpea mesophyll protoplasts inoculated with purified B-RNA (4) exhibit this activity. Until now, however, it has not been possible to elucidate which B-RNA-encoded polypeptide actually represents the protease involved. This report presents evidence that the proteolytic activity resides in the 32K polypeptide.

MATERIALS AND METHODS

Virus and RNA. CPMV was propagated in cowpea plants (*Vigna unguiculata* L. "California Blackeye") as described previously (12, 25). Separate B- and M-components were obtained by three cycles of centrifugation in a linear 15 to 30% (wt/vol) zonal sucrose gradient (16 h, 23,000 rpm at 10°C; Beckman Ti 15 rotor) as previously described (6, 20). M-RNA was isolated as follows: purified M-components were disrupted by adding an equal volume of 4% (wt/vol) Sarkosyl NL97–2% (wt/vol) sodium tri-isopropylnaphthalene sulfonate–0.02 M Tris-hydrochloride (pH 7.4)–0.2 M NaCl–0.004 M EDTA and by heating for 5 min at 60°C. The RNA was purified by three extractions with phenol (saturated at pH 8) and precipitated with two volumes of ethanol at −20°C. The precipitate was dried and dissolved in water and then subjected to a linear 15 to 30% (wt/vol) sucrose gradient centrifugation (16 h, 22,500 rpm at 20°C; Beckman SW27 rotor). Fractions containing intact RNA were pooled and precipitated with two volumes of ethanol at −20°C. The RNA precipitate obtained was washed three times with absolute ethanol, dried, and dissolved in water at a concentration of 1 mg/ml.

Incubation of protoplasts, labeling of proteins, and subcellular fractionation. Cowpea mesophyll protoplasts were prepared, inoculated, and incubated as described previously (11, 20). When labeled proteins were required, protoplast suspensions (5 ml, 5×10^5 cells per ml) were supplied with portions (150 μCi) of [^{35}S]methionine (1,100 Ci/mmol; New England Nuclear Corp.) at 18 and 25 h after inoculation. Forty-four hours after inoculation, protoplasts were collected by centrifugation (2 min, 600 × *g*), suspended in 0.5 ml of

* Corresponding author.

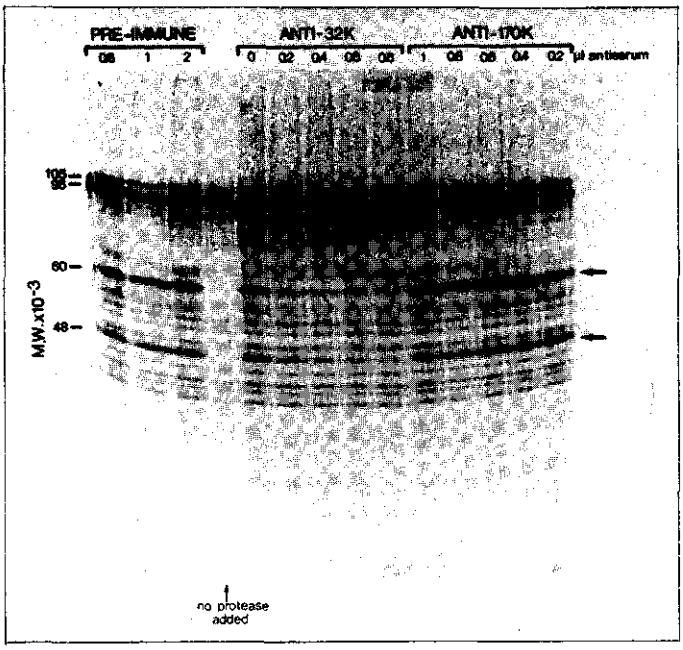


FIG. 1. Inhibition of the protease activity by antiserum raised against the 32K polypeptide. Increasing volumes of anti-32K, anti-170K, or preimmune sera were added to 2-μl portions of the S30 fraction from B-component-inoculated protoplasts, in a total volume of 20 μl of TKED buffer. After incubation for 1.5 h at 0°C, 2 μl of [^{35}S]methionine-labeled in vitro translation products of M-RNA was added, and incubation was continued for 1.5 h at 30°C. Proteolytic cleavage of the 105K and 95K polypeptides was determined by electrophoresis in a 12.5% polyacrylamide gel. The arrows indicate the 60K and 48K cleavage proteins from the 95K polypeptide (the input amounts of the 105K polypeptide were too low to allow detection of the 58K cleavage product). Molecular weights (M.W. $\times 10^{-3}$) are indicated to the left of the figure.

TKEDP buffer (50 mM Tris-acetate [pH 7.4], 10 mM potassium acetate, 1 mM EDTA, 10 mM dithioerythritol, 1 mM phenylmethylsulfonyl fluoride) containing 10% (wt/vol) sucrose, and disrupted by homogenization for 2.5 min at 0°C in a small Thomas tissue homogenizer. The homogenate thus obtained was centrifuged for 30 min at 30,000 × *g* and 4°C to give the 30,000 × *g* supernatant (S30) fraction, which contained (most of) the virus-specific proteolytic activity.

SDS-polyacrylamide slab gel electrophoresis. Portions of radiolabeled proteins were mixed with one-third volume of a fourfold-concentrated sample buffer (40 mM Tris-hydrochloride [pH 8.0], 4 mM EDTA, 40% [vol/vol] glycerol, 8% [wt/vol] sodium dodecyl sulfate [SDS], 20% [vol/vol] β-mercaptoethanol, 0.004% [wt/vol] bromophenol blue) and heated for 3 min at 100°C. The samples were then electrophoresed at 150 V in an SDS-polyacrylamide gel containing 12.5% acrylamide (with 0.09% bisacrylamide) by using a stacking gel of 4% acrylamide and 0.10% bisacrylamide, as previously described (13, 20). After electrophoresis, gels were dried either with or without preceding staining and autoradiographed with Kodak Royal X-omat X-ray film.

Antisera and immunoprecipitation. Antiserum against the electrophoretically separated virus-coded 170K polypeptide was prepared by directly immunizing a New Zealand white rabbit with polyacrylamide containing the denatured antigen, in principle as described by Tijan et al. (24) and Schiff and Grandgenett (21) but with some modifications as described elsewhere (P. Zabel and F. van Straaten, manuscript in preparation). Antiserum against the electrophoretically separated 32K polypeptide was prepared by immunizing rabbits with 32K polypeptide eluted with buffer (12.5 mM Tris-hydrochloride [pH 6.8], 0.1 mM EDTA) from gel slices of a nonfixed polyacrylamide gel. Antiserum against the genome-linked protein VPg was prepared as described previously (28). Specificity of the various antisera raised against CPMV-encoded proteins was tested by immunoprecipitation of proteins from the S30 fraction of radiolabeled B-component-inoculated protoplasts in a buffer (TKE-TDS) containing 50 mM Tris-acetate (pH 7.4), 10 mM potassium acetate, 1 mM EDTA, 1% Triton X-100, 0.5% sodium deoxycholate, and 0.1% SDS. Immunoprecipitation was performed in TKE-TDS buffer as described previously (4), except as stated otherwise. None of the antiserum preparations used showed reactivity against host proteins.

Glycerol gradient centrifugation. Typically, 80 μl of the S30 fraction obtained from B-component-inoculated proto-

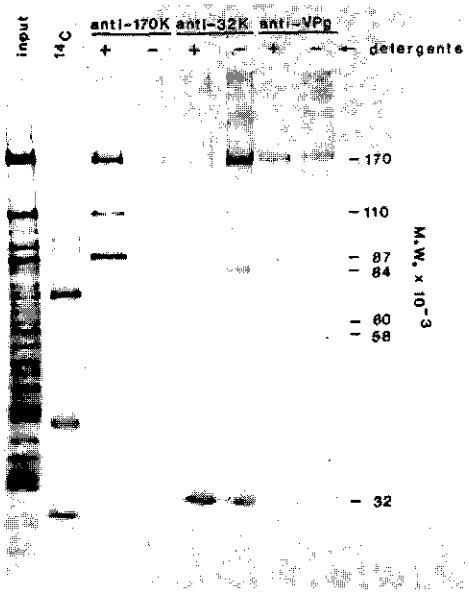


FIG. 2. Autoradiogram of a polyacrylamide gel with immunoprecipitated CPMV polypeptides. Portions of the S30 fraction (2 μl) from [^{35}S]methionine-labeled B-component-inoculated protoplasts were incubated with 2 μl of anti-32K serum, 5 μl of anti-170K serum, or 5 μl of anti-VPg serum, in a final volume of 20 μl of TKE buffer (without detergents, −) or TKE-TDS buffer (with detergents, +). The two lanes on the left contain the [^{35}S]methionine-labeled polypeptides from B-RNA-inoculated protoplasts without immunoprecipitation (input) and ^{14}C-methylated protein markers (^{14}C) (myosin [M.W., 210,000]; phosphorylase *b* [M.W., 100,000 and 92,500]; bovine serum albumin [M.W., 68,000]; ovalbumin [M.W., 46,000]; and carbonic anhydrase [M.W., 30,000]), respectively. The numbers to the right of the figure refer to the molecular weights of the B-RNA-encoded polypeptides.

plasts was layered on linear 15 to 30% (vol/vol) glycerol gradients in TKED buffer (i.e., TKEDP buffer minus phenylmethylsulfonyl fluoride). Centrifugation was carried out in a Beckman SW41 or SW50 rotor under conditions described in the figure legends. After the run, gradients were fractionated in 500-μl portions which were assayed for both protease activity and occurrence of viral proteins.

In vitro translation of M-RNA. CPMV M-RNA was translated in an mRNA-dependent rabbit reticulocyte lysate (a generous gift of R. J. Jackson, Department of Biochemistry, University of Cambridge, Cambridge, England) under conditions as previously described (8, 16, 17). [^{35}S]methionine (10 μCi per 10 μl of reaction mixture) was used as the radioactive amino acid, and incubation was for 1 h at 30°C.

Viral protease assay. For detection of the viral protease, 2 to 20 μl of samples to be analyzed was mixed with 2 μl of [^{35}S]methionine-labeled in vitro translation products from CPMV M-RNA (as obtained after 1 h of translation) in a total volume of 22 μl of TKED buffer and incubated at 30°C for 1.5 h. To determine proteolytic activity, 5-μl portions of each sample were electrophoresed in a 12.5% polyacrylamide gel.

RESULTS

Inhibition of the protease activity by anti-32K serum. Since translation products obtained from B-RNA after 1 h of in vitro translation were proteolytically active (4, 16), either the 32K polypeptide or the 170K polypeptide must be responsible for this activity. If the proteolytic activity resides in the 170K polypeptide, then it might be expected that one of the known, final cleavage products of the 170K polypeptide represents the mature protease. To discriminate between these possibilities, antisera raised against the 32K and 170K polypeptides were prepared and tested for their ability to inhibit the proteolytic activity. For this purpose, increasing amounts of anti-32K, anti-170K, or preimmune sera were added to 2-μl portions of the S30 fraction from B-component-inoculated protoplasts in a total volume of 20 μl of TKED buffer and incubated for 1.5 h at 0°C to allow binding of the immunoglobulins to the viral proteins. Possible inhibition of the proteolytic activity in these mixtures was then determined by addition of 2 μl of [^{35}S]methionine-labeled (95K and 105K) in vitro translation products from M-RNA and by analysis for the lack of appearance of the 60K, 58K, and 48K cleavage products. A significant decrease in proteolytic activity was obtained with anti-32K serum, even at lower amounts (Fig. 1). On the other hand, preimmune serum, tapped from the anti-32K rabbit before immunization, and anti-170K serum did not show any inhibitory effect (Fig. 1). These results suggest that the proteolytic activity resides in the 32K polypeptide and not in the 170K polypeptide. Some complications which may be important should be regarded, however. First, although the titer of the anti-170K serum used was reasonable under normal immunological conditions (i.e., in TKE-TDS buffer), reacting with both the 170K polypeptide and the 110K and 87K polypeptides derived from this polypeptide, this antiserum showed a reduced reactivity under conditions where its possible inhibitory effect on the proteolytic cleavage was tested (in TKE buffer, i.e., in the absence of any detergent [Fig. 2]). Moreover, in the absence of detergents, the anti-32K serum apparently bound to complexes of 32K polypeptides with the viral 58K, 60K, 84K, and 170K polypeptides (Fig. 2). Therefore, the possibility that one of these four other B-RNA-encoded polypeptides represented the protease, the activity of which was inhibited by sterical hindrance of immunoglobulin G (IgG) molecules bound to associated 32K polypeptides, had to be considered. The following observations, however, provide evidence against this idea. First, since the 87K and 110K polypeptides were not detectable in complexes with the 32K polypeptide (Fig. 2), the conclusion can be drawn that at least these polypeptides were not responsible for the proteolytic cleavage. Second, antiserum raised against VPg did not interfere with the proteolytic cleavage (data not shown). This antiserum was capable of binding the VPg-containing 170K, 60K, and, to a lesser extent, 84K polypeptides, both in the presence and absence of detergents (Fig. 2). The results obtained with all three antisera are therefore consistent and indicate that neither the 170K polypeptide nor any of its cleavage products represent the protease, but instead, point out that the 32K polypeptide bears the activity in question.

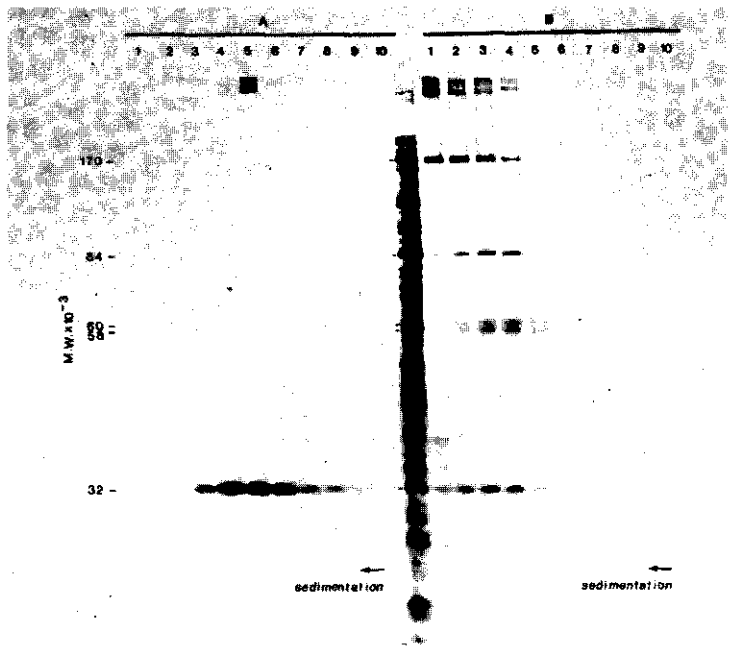


FIG. 3. Sedimentation of [^{35}S]methionine-labeled CPMV polypeptides in a glycerol gradient. Eighty microliters of an S30 fraction of radiolabeled B-component-inoculated protoplasts was layered on a 10 to 30% (vol/vol) linear glycerol gradient either in TKE-TDS buffer (A) or in TKE buffer (B). Centrifugation was carried out in a Beckman SW50 rotor for 17 h at 48,000 rpm and 4°C. After the run, 500-μl fractions were collected and immunoprecipitated with anti-32K serum in TKE-TDS buffer (A) or in TKE buffer without detergents (B). Samples were analyzed in a 12.5% polyacrylamide gel. Lanes 1 correspond to the bottom fractions of each gradient. The central lane between (A) and (B) contains unfractionated [^{35}S]methionine-labeled proteins from B-component-inoculated protoplasts. Numbers to the left of the figure refer to the molecular weights of some of the B-RNA-encoded proteins.

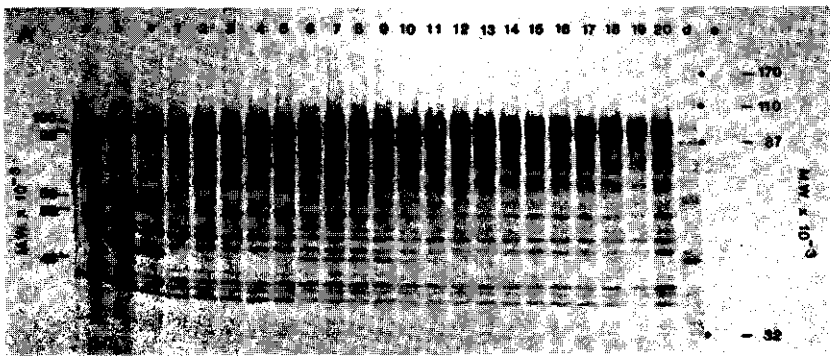


FIG. 4. Glycerol gradient fractionation and protease activity of B-RNA-encoded proteins. Eighty microliters of an S30 fraction of [^{35}S]methionine-labeled, B-component-inoculated protoplasts was layered on a 15 to 30% (vol/vol) linear glycerol gradient (in TKED buffer) and centrifuged for 42 h in a Beckman SW41 rotor at 38,000 rpm and 4°C. After centrifugation, the gradient was fractionated into 20 portions of 500 μl which were tested for protease activity (A) and the presence of viral polypeptides (B and C). (A) From each fraction, 20-μl portions were taken and tested for protease activity on M-RNA-encoded polypeptides, as described in the text. Proteolytic cleavage was analyzed in a 12.5% polyacrylamide gel (lanes 1 to 20). Lane 1 corresponds to the bottom fraction of the gradient. Lane a contains 2 μl of [^{35}S]methionine-labeled in vitro translation products of M-RNA not further treated. Lane b contains the same polypeptides which were incubated for 1.5 h at 30°C with 2 μl of S30 fraction from B-component-inoculated protoplasts. Lane c is as lane b, but the S30 fraction was preincubated on ice for 42 h after addition of glycerol to a final concentration of 20% (vol/vol). Lanes d and e contain [^{35}S]methionine-labeled proteins of B-component-inoculated (lane d) and noninoculated (lane e) protoplasts. Numbers indicated to the left of the gel refer to the molecular weights

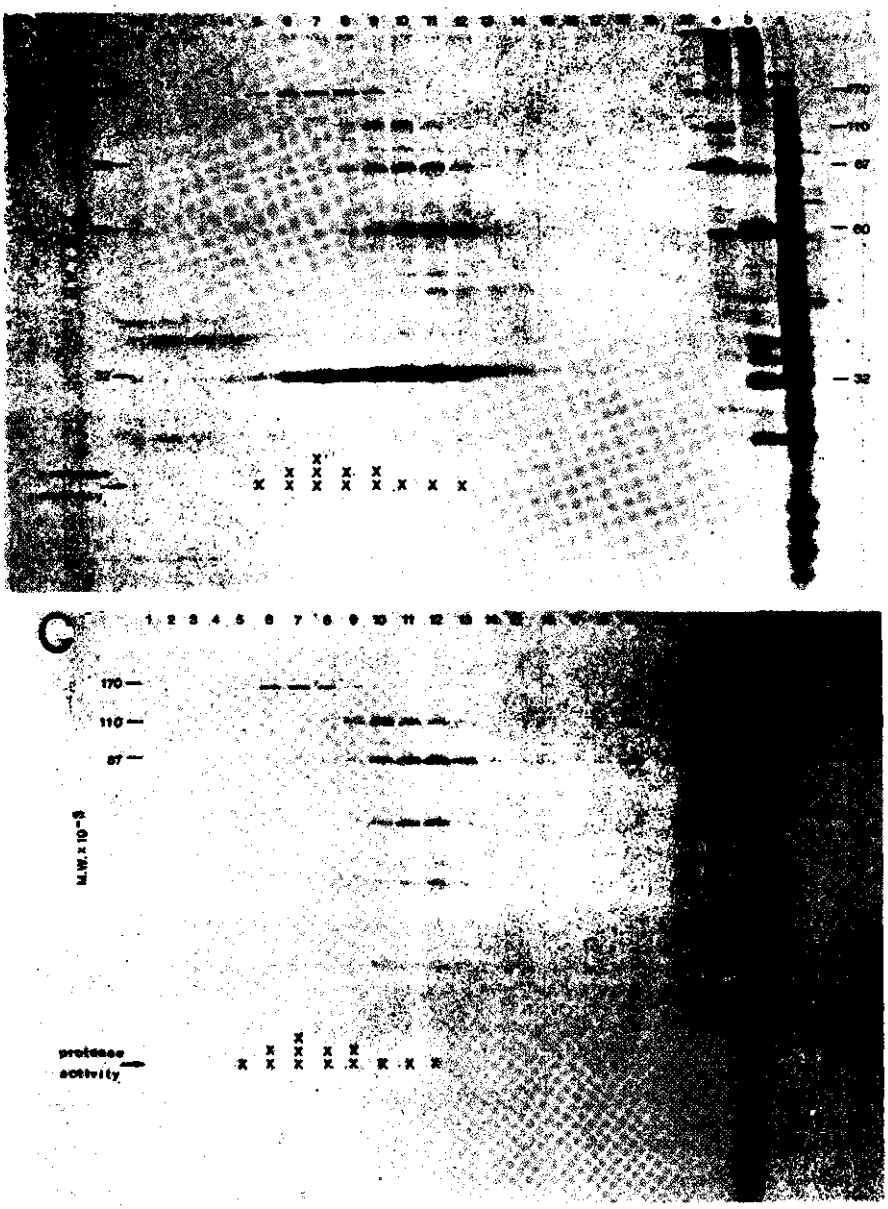


FIG. 4—*Continued*

of the in vitro translation products of M-RNA and the polypeptides generated by the proteolytic activity. Numbers to the right of the gel refer to the molecular weights of some of the B-RNA-encoded polypeptides. (B and C) Equal amounts (150 μl) of the glycerol gradient fractions were incubated either with a mixture of anti-32K and anti-VPg sera (B) or with anti-170K serum (C) in TKE-TDS buffer, and the immunoprecipitates were analyzed in 12.5% polyacrylamide gels. Each gel includes lanes containing the unfractionated [^{35}S]methionine-labeled S30 fraction of B-component-infected protoplasts (input, lane a) and immunoprecipitates of unfractionated S30 fraction obtained with anti-VPg and anti-32K sera (lane b) or with anti-170K serum (lane c) in TKE-TDS buffer. Numbers indicated at both sides of the gels refer to the molecular weights of the viral polypeptides. The number of crosses below the gel indicates the amount of proteolytic activity in the fractions, as deduced from (A).

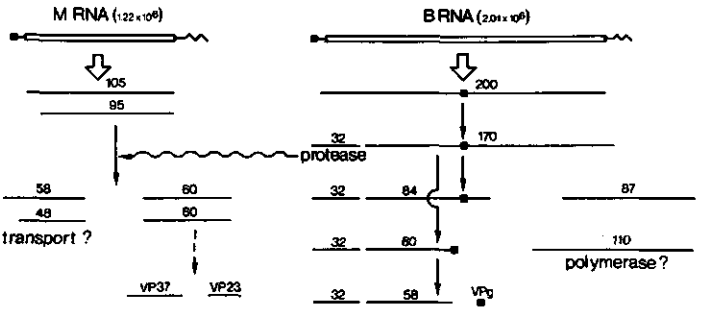


FIG. 5. Model for the proteolytic processing and complexation of the CPMV RNA-encoded polypeptides. After the primary cleavage of the B-RNA-encoded, 200K (200) polyprotein, 32K (32) and 170K (170) polypeptides remain tightly bound in a complex in which the 32K polypeptide is responsible for the cleavage of the (95K [95] and 105K [105]) primary translation products from M-RNA. Further cleavage of the 170K polypeptide follows two different pathways (20), resulting in the release of either the 87K (87) polypeptide or the 110K (110) polypeptide from the complex. The 60K (60) polypeptide is the direct precursor to VPg (28), and the 110K polypeptide has recently been detected in the viral replication complex and might represent an RNA-dependent RNA polymerase (L. Dorssers, S. van der Krol, J. van der Meer, A. van Kammen, and P. Zabel, submitted for publication). VPg is supposed to be involved in viral RNA replication by acting as a primer (7, 28) and might be directly linked to nascent RNA chains. Open reading frames in both RNAs are indicated with double-lined bars, and the polyadenylated tails are indicated with a zigzag line. VPg is indicated with a black box. 58, 48, and 84 indicate polypeptides of 58K, 48K, and 84K, respectively.

Glycerol gradient centrifugation of the proteolytic activity. Complexation of the 32K polypeptide to other B-RNA-encoded polypeptides was confirmed and further investigated by glycerol gradient centrifugation of the S30 fraction from protoplasts inoculated with B-RNAs and labeled with [^{35}S]methionine. In a 5-ml linear 10 to 30% (vol/vol) glycerol gradient made up in TKE-TDS buffer and run for 17 h at 4°C and 48,000 rpm in a Beckman SW50 rotor, 32K polypeptides sedimented as free proteins to approximately the middle of the gradient (lanes 4 through 6, Fig. 3A). When a similar gradient was made up in TKE buffer only (i.e., buffer without detergents), the 32K polypeptides sedimented faster (lanes 1 through 5 from the bottom, Fig. 3B). Under these conditions, the 32K polypeptides were apparently associated with the cosedimenting 58K, 60K, 84K, and 170K polypeptides, which, indeed, were all coprecipitated by anti-32K serum. Since in this gradient a slight separation was obtained between the 170K polypeptides on the one hand (peak fractions 1 through 3 [Fig. 3, lanes 1 through 3]) and the 84K, 60K and 58K polypeptides on the other hand (peak fractions 2 through 5 [Fig. 3, lanes 2 through 5]) a longer and less steep glycerol gradient volume of 11 ml (15 to 30% [vol/vol] glycerol, Beckman SW41 rotor) was run with the purpose of optimizing this separation and for assaying the separated complexes for protease activity. Again, no detergents were present in the gradient, but for obtaining a complete picture of the position of the viral proteins throughout the gradient, and not only of those proteins associated with 32K polypeptides, immunoprecipitations with anti-32K, anti-VPg, and anti-170K sera were carried out in buffer containing detergents (TKE-TDS). In this gradient, the protease activity sedimented into lanes 5 through 12 (Fig. 4A) from the bottom, with the main activity confined to lanes 6 through 9 (Fig. 4A). Screening of the gradient fractions with a mixture of anti-VPg and anti-32K sera (Fig. 4B) or with anti-170K serum (Fig. 4C) indicated that the protease-containing fractions all contained 32K polypeptide (Fig. 4B). The 84K and 60K polypeptides (Fig. 4B) and the 110K and 87K polypeptides (Fig. 4C) were mainly present in lanes 9 through 12 (Fig. 4B and C) and appeared not to be responsible for the (main) protease activity. The occurrence of the 170K polypeptide in lanes 6 through 9 (Fig. 4B and C), however, corresponded strikingly to the peak of the protease activity in the gradient (cf. A, B, and C in Fig. 4), although this polypeptide could not be detected in lanes 11 and 12, which contained minor proteolytic activity. On the other hand, the broad sedimentation profile of the 32K polypeptide which must be caused by separate association to either 170K polypeptides (lanes 5 through 9) or to the smaller 84K, 60K, and 58K polypeptides (lanes 9 through 13), respectively, did not nicely fit with the protease activity profile. Lanes 6 through 9 (with most of the activity) and lanes 10 through 12 (with significantly less activity) contained approximately equal amounts of 32K polypeptide (Fig. 4B). Therefore, it seems that not all 32K polypeptide molecules possessed proteolytic activity. A possible explanation for this finding will be discussed below.

DISCUSSION

As was first shown by Pelham (16), in vitro translation products from B-RNA possess a proteolytic activity which is capable of cleaving the 95K and 105K primary translation products from M-RNA. Later experiments (4) demonstrated that one of the cleavage products represents a specific precursor to both capsid proteins VP37 and VP23. Since upon in vitro translation of B-RNA only the 32K and 170K polypeptides are produced (8, 16), it can be deduced that either the 32K or the 170K polypeptide possesses the proteolytic activity involved. The protease inhibition studies presented in this report provide evidence that the 32K polypeptide represents the protease in question. The proteolytic cleavage of the M-RNA-encoded 95K and 105K polypeptides by extract from B-component-inoculated protoplasts is efficiently inhibited by antiserum raised against the

32K polypeptide, but not by anti-170K, anti-VPg, or preimmune sera (Fig. 1). Further experiments, in which the reactivity of these antisera was analyzed under conditions where this inhibition was tested, demonstrated, however, that care should be taken from directly drawing conclusions from such experiments. Anti-32K immunoglobulins appeared to bind to 32K polypeptides associated in complexes with 58K, 60K, 84K, and 170K polypeptides. The binding of heterologous viral polypeptides, apparently present in protein aggregates, has also been reported for poliovirus with antiserum directed against the polioviral protease (10). Additional experiments, in which proteolytic extract from B-component-inoculated cells was fractionated in glycerol gradients, demonstrated that the protease activity can only be correlated with 32K polypeptides or uncleaved 170K polypeptides. Therefore, two possibilities still remain: the protease activity resides either in the 170K precursor polypeptide or in the 32K polypeptide. For several reasons the first possibility seems to be unlikely. (i) If the 170K polypeptide is proteolytically active, then one might expect that, in analogy with, for instance, the animal picornaviruses (9, 10, 14, 15), one of the cleavage products from this precursor represents the mature protease. There is no evidence for such a mature protease, however. The experiments shown in Fig. 2 and 4 indicate that none of the 110K, 87K, 84K, or 60K polypeptides possess significant proteolytic activity. (ii) Anti-VPg immunoglobulins which are capable of reacting with the 170K (and 60K) polypeptides under nondenaturing conditions (Fig. 2) do not interfere with the proteolytic cleavage. Therefore, it is reasonable to propose that the 32K polypeptide represents the protease. Indeed, antiserum raised against this polypeptide efficiently inhibits the proteolytic activity (Fig. 1). Although other viral polypeptides were associated with the 32K polypeptides which might have been blocked in their function by the binding of IgG molecules to the 32K polypeptide, it should be noted that the antibody-antigen complexes were not removed from the protease reaction mixture. The possibility that the 32K polypeptide is an activator of a cellular enzyme can readily be excluded by the observation that the proteolytic activity also resides in the in vitro product from B-RNA (16). A puzzling observation is that not all 32K polypeptides appear to be active molecules. Only 32K polypeptides associated with 170K polypeptides seem to be involved in the proteolytic cleavage of the M-RNA-encoded primary translation products (cf. Fig. 4A and B). A possible explanation for this phenomenon is that the 32K polypeptides connected to 170K polypeptides represent newly synthesized molecules which are freshly cleaved from the 200K primary translation product. The 32K polypeptides associated with the 84K, 60K, or 58K polypeptides may represent significantly older molecules, which may have lost much of their activity. Indeed, in vitro cleavage of the 200K primary translation product into the 32K and 170K polypeptides occurs as soon as the 200K polypeptide chain has been completed, whereas further cleavage of the 170K polypeptide into the 110K, 87K, 84K, and 60K polypeptides takes, at least in vitro, a significantly longer period of time, with the first cleavage products only visible after 6 to 8 h (H. Franssen, unpublished data).

The information available suggests the following pathway of the B-RNA-encoded polypeptides (Fig. 5): soon after its synthesis, the 200K primary translation product is cleaved into the 32K and 170K polypeptides, which remain associated after cleavage. The 32K polypeptide is the protease involved in the cleavage of the primary translation products of M-RNA, and its activity decreases over the course of time. Further cleavage of the 170K polypeptide occurs at a slower rate, giving rise to complexes consisting of the 32K polypeptide, with decreased activity, and 84K, 60K, or 58K polypeptides. The 110K and 87K polypeptides, which are cleaved from the COOH-terminal part of the 170K polypeptide (5), do not remain in the complexes but are released (Fig. 2). It should be mentioned, however, that the viral protein complexes described here have been found in the S30 fractions, whereas considerable amounts of the 58K and 60K polypeptides have been found in the membrane fraction (i.e., 30,000 × *g* pellet fraction) of infected cells (7, 28). Therefore, to obtain a complete picture of the viral protein complexes in infected cells, additional studies on this membrane fraction are necessary. The 32K polypeptide is probably not the only protease involved in the proteolytic processing of the viral polyproteins. According to the mapping of the coding sequences of VP37 and VP23 on the M-RNA sequence (27), the cleavage site of the 32K polypeptide should be the dipeptide sequence glutamine-methionine. The cleavage site used to release VP37 and VP23 from their 60K precursor (4), however, is a glutamine-glycine sequence (27). Therefore, at least a second (plant or viral?) protease should be involved in the generation of the M-RNA-encoded polypeptides. The cleavage sites used for the processing of the B-RNA-encoded polyprotein remain to be elucidated, and information about the nature of the protease charged with the cleavage of this protein is also not available.

ACKNOWLEDGMENTS

We wish to thank Bep Kuiper for technical assistance, Richard Jackson for the gift of rabbit reticulocyte lysates, Ab van Kammen for encouraging discussions and critical reading of the text, Piet Madern for printing the photographs, and Marie-José van Neerven for typing the manuscript.

This work was supported by the Netherlands Foundation for Chemical Research, with financial aid from the Netherlands Organization for the Advancement of Pure Research.

LITERATURE CITED

1. **Ahlquist, P., and P. Kaesberg.** 1979. Determination of the length distribution of poly(A) at the 3'-terminus of the virion RNAs of EMC virus, poliovirus, rhinovirus, RAV-61 and CPMV and of mouse globin mRNA. Nucleic Acids Res. **7**:1195–1204.
2. **Daubert, S. D., G. Bruening, and R. C. Najarian.** 1978. Protein bound to the genome RNAs of cowpea mosaic virus. Eur. J. Biochem. **92**:45–51.
3. **El Manna, M. M., and G. Bruening.** 1973. Polyadenylate sequences in the ribonucleic acids of cowpea mosaic virus. Virology **56**:198–206.
4. **Franssen, H., R. Goldbach, M. Broekhuijsen, M. Moerman, and A. van Kammen.** 1982. Expression of middle-component RNA of cowpea mosaic virus: in vitro generation of a precursor to both capsid proteins by a bottom-component RNA-encoded protease from infected cells. J Virol. **41**:8–17.
5. **Goldbach, R., and G. Rezelman.** 1983. Orientation of the cleavage map of the 200-kilodalton polypeptide encoded by the bottom-component RNA of cowpea mosaic virus. J. Virol. **46**:614–619.
6. **Goldbach, R., G. Rezelman, and A. van Kammen.** 1980. Independent replication and expression of B-component RNA of cowpea mosaic virus. Nature (London) **286**:297–300.
7. **Goldbach, R., G. Rezelman, P. Zabel, and A. van Kammen.** 1982. Expression of the bottom-component RNA of cowpea mosaic virus: evidence that the 60-kilodalton VPg-precursor is cleaved into single VPg and a 58-kilodalton polypeptide. J. Virol. **42**:630–635.
8. **Goldbach, R. W., J. G. Schilthuis, and G. Rezelman.** 1981. Comparison of in vivo and in vitro translation of cowpea mosaic virus RNAs. Biochem. Biophys. Res. Commun. **99**:89–94.

9. **Gorbalenya, A. E., Y. V. Svitkin, Y. A. Kazachkov, and V. I. Agol.** 1979. Encephalomyocarditis virus-specific polypeptide p22 is involved in the processing of the viral precursor polypeptides. FEBS Lett. **108**:1–5.
10. **Hanecak, R., B. L. Semler, C. W. Anderson, and E. Wimmer.** 1982. Proteolytic processing of poliovirus polypeptides: antibodies to polypeptide P3-7c inhibit cleavage at glutamine-glycine pairs. Proc. Natl. Acad. Sci. U.S.A. **79**:3973–3977.
11. **Hibi, T., G. Rezelman, and A. van Kammen.** 1975. Infection of cowpea mesophyll protoplasts with cowpea mosaic virus. Virology **64**:308–318.
12. **Klootwijk, J., I. Klein, P. Zabel, and A. van Kammen.** 1977. Cowpea mosaic virus RNAs have neither m^7GpppN . . . nor mono-, di-, or triphosphates at their 5'-ends. Cell **11**:73–82.
13. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
14. **Palmenberg, A. C., M. A. Pallansch, and R. R. Rueckert.** 1979. Protease required for processing picornaviral coat protein resides in the viral replicase gene. J. Virol. **32**:770–778.
15. **Palmenberg, A. C., and R. R. Rueckert.** 1982. Evidence for intramolecular self-cleavage of picornaviral replicase precursors. J. Virol. **41**:244–249.
16. **Pelham, H. R. B.** 1979. Synthesis and proteolytic processing of cowpea mosaic virus proteins in reticulocyte lysates. Virology **96**:463–477.
17. **Pelham, H. R. B., and R. J. Jackson.** 1976. An efficient mRNA-dependent translation system from reticulocyte lysates. Eur. J. Biochem. **67**:247–256.
18. **Reijnders, L., A. M. J. Aalbers, A. van Kammen, and R. W. J. Thuring.** 1974. Molecular weights of plant viral RNAs determined by gel electrophoresis under denaturing conditions. Virology **60**:515–521.
19. **Rezelman, G., H. J. Franssen, R. W. Goldbach, T. S. Ie, and A. van Kammen.** 1982. Limits to the independence of bottom component RNA of cowpea mosaic virus. J. Gen. Virol. **60**:335–342.
20. **Rezelman, G., R. Goldbach, and A. Van Kammen.** 1980. Expression of bottom component RNA of cowpea mosaic virus in cowpea protoplasts. J. Virol. **36**:366–373.
21. **Schiff, R. D., and D. P. Grandgenett.** 1980. Virus-coded origin of a 32,000-dalton protein from avian retrovirus cores: structural relatedness of p32 and the β polypeptide of the avian retrovirus DNA polymerase. J. Virol. **28**:279–291.
22. **Stanley, J., R. Goldbach, and A. van Kammen.** 1980. The genome-linked protein of cowpea mosaic virus is coded by RNA from the bottom component. Virology **106**:180–182.
23. **Stanley, J., P. Rottier, J. W. Davies, P. Zabel, and A. van Kammen.** 1978. A protein linked to the 5' termini of both RNA components of the cowpea mosaic virus genome. Nucleic Acids Res. **5**:4505–4522.
24. **Tijan, R., D. Stinchcomp, and R. Losick.** 1974. Antibody directed against Bacillus subtilis σ factor purified by sodium dodecyl sulfate slab gel electrophoresis. J. Biol. Chem. **250**:8824–8828.
25. **Van Kammen, A.** 1967. Purification and properties of the components of cowpea mosaic virus. Virology **31**:633–642.
26. **Van Kammen, A.** 1972. Plant viruses with a divided genome. Annu. Rev. Phytopathol. **10**:125–150.
27. **Van Wezenbeek, P., J. Verver, J. Harmsen, P. Vos, and A. van Kammen.** 1983. Primary structure and gene organization of the middle component RNA of cowpea mosaic virus. EMBO J. **2**:941–946.
28. **Zabel, P., M. Moerman, F. van Straaten, R. Goldbach, and A. van Kammen.** 1982. Antibodies against the genome-linked protein VPg of cowpea mosaic virus recognize a 60,000-dalton precursor polypeptide. J. Virol. **41**:1083–1088.

CHAPTER VI

MAPPING OF THE CODING REGIONS FOR THE CAPSID PROTEINS OF COWPEA MOSAIC VIRUS ON THE NUCLEOTIDE SEQUENCE OF MIDDLE COMPONENT RNA

Henk Franssen[1], Wim Roelofsen[2], Ab van Kammen[1] and Rob Goldbach[1].

1. Department of Molecular Biology, Agricultural University, De Dreijen 11, 6703 BC Wageningen, The Netherlands.
2. Department of Microbiology, Agricultural University, Hesselink van Suchtelenweg 4, 6703 CT Wageningen, The Netherlands.

ABSTRACT

The amino-terminal and carboxy-terminal amino acid sequences of the two capsid proteins VP37 and VP23 of cowpea mosaic virus have been analysed. The results allow the mapping of the coding regions for VP37 and VP23 on the middle-component (M) RNA using the nucleotide sequence data of Van Wezenbeek *et al.* (EMBO J., 2, 941-946, 1983). These regions are contiguous and indicate that the cleavage sites used to release VP37 and VP23 from the primary translation products of M RNA are a glutamine-methionine and a glutamine-glycine dipeptide sequence, respectively.

INTRODUCTION

Cowpea mosaic virus (CPMV), typemember of the comoviruses, has a bipartite positive-stranded RNA genome, divided between two components, the bottom (B) and middle (M) component (for a review see Bruening, 1977, 1981). Both B and M RNA are supplied with a protein, called VPg, at their 5' end (Daubert *et al.*, 1978; Stanley *et al.*, 1978) and a poly(A) track at their 3' end (El Manna and Bruening, 1973; Ahlquist and Kaesberg, 1979). Both B and M particles have identical capsids made up of 60 copies of two different proteins, denoted VP37 (mol. weight 37,000) and VP23 (mol. weight 23,000). Genetic (Wood, 1972; Gopo and Frist, 1977) as well as biochemical (Franssen *et al.*, 1982) data have indicated that the coat proteins are encoded by the M RNA. *In vitro* translation studies have revealed that M RNA is translated into two overlapping 105K and 95K polyproteins (Pelham 1979; Goldbach *et al.*, 1981) which both contain the sequences of the capsid proteins (Franssen *et al.*, 1982). These polyproteins are cleaved into products with sizes of 48K, 58K and 60K (Franssen *et al.*, 1982) by the recently identified 32K protease encoded by B-RNA (Franssen *et al.*, 1984).

The 60K cleavage product is derived from the carboxy-terminal half of the 105K and the 95K polypeptides and represents the direct precursor of the capsid proteins, VP37 and VP23 (Franssen *et al.*, 1982). Since the complete nucleotide sequence of M RNA has recently been elucidated (Van Wezenbeek *et al.*, 1983) it has now become possible to map the capsid proteins exactly on the M RNA. For this purpose the amino acid composition of VP37 and VP23 was determined and the amino-terminal and carboxy-terminal amino acid sequences of these proteins were analysed. The information obtained reveals the presence of two different proteolytic cleavage sites in the primary translation products of M RNA. Parts of the results described here, have briefly been mentioned in a previous paper (Van Wezenbeek *et al.*, 1983).

MATERIALS AND METHODS

Virus purification and top component isolation.

CPMV was grown in *Vigna unguiculata* L. (California Blackeye) and purified by butanol-chloroform extraction (Steere, 1956). Virus was precipitated with polyethyleenglycol/NaCl (Van Kammen, 1967), resuspended in sodium phosphate buffer (pH 7.2) and layered on a 50% sucrose cushion in the same buffer. The virus was pelleted by centrifugation in a Beckman Ti60 rotor for 16 h. at 30,000 rpm and 4°C. Top (T) components were separated from B and M components by zonal centrifugation in a linear 15 to 30% sucrose gradient (Beckman Ti15 rotor, 16 h., 23,000 rpm at 10°C). T components thus obtained were dialysed against distilled water and lyophilised.

Carboxymethylation of purified virus.

Virus particles were disintegrated in guanidium chloride-LiCl according to Wu (1970). The precipitated nucleic acid was removed by low-speed centrifugation and the dissociated capsid proteins present in the supernatant were then subjected to carboxymethylation with iodoacetamid as described by Wu (1970) and Geelen (1974).

Separation of carboxymethylated capsid proteins.

Thirty mg of lyophylised, carboxymethylated proteins were dissolved in 6 M urea and applied to a Sephadex G-200 column (100 x 1.5 cm). Proteins were separated by elution with 5 M urea (flow rate of 2 ml/h). Fractions containing proteins were located by measuring the optical density at 280 nm. The protein content of these fractions was analysed in 15% SDS polyacrylamide gels (Laemmli, 1970; Franssen *et al.*, 1982) and visualised by silverstaining (Morrissey, 1981). Fractions containing respec-

tively VP37 and VP23 were separately pooled and dialysed extensively against distilled water, which resulted in precipitation of the proteins. The precipitated protein was collected by centrifugation, lyophilised and stored at -80°C.

Determination of amino acid composition.

Carboxymethylated VP37 and VP23 were hydrolysed with 6N HCl (constant boiling) in evacuated sealed tubes for 20-24 h at 110°C. Amino acid analyses were performed on a Kontron Liquimat 111 amino acid analyser as described by Vereijken *et al.* (1982).

Determination of amino-terminal amino acid sequences.

Samples of protein were dissolved in 70% formic acid and subjected to automatic Edman degradation in a Beckman spinning cup model 890C protein sequenator using a program adapted from Brauer *et al.* (1975) with the modifications according to Hunkapiller and Hood (1978). The amino-terminal residues were removed in heptafluorobutyric acid (HFBA) and the phenylthiohydantoin derivatives of amino acids were analysed by high-performance liquid chromatography (Lichro Sorb. Si60-10 Merck) as described by Frank and Strubert (1973).

Isolation of an amino-terminal peptide from VP37.

Since the amino-terminal methionine residue in VP37 is acetylated (Bruening, 1981) this protein can not directly be subjected to Edman degradation. In an effort to characterise the amino-terminal sequence of VP37 an amino-terminal oligopeptide from this protein was isolated. Five to ten mg of purified carboxymethylated VP37 was digested with 50 µg of chymotrypsin for 20 h. at 25°C. The digest was then adjusted to pH 3 with acetic acid and applied to a Dowex 50W-X8 column (5 x 1 cm)

equilibrated with 0.1 M acetic acid (Narita *et al.*, 1975). The eluate containing the blocked amino-terminal oligopeptide was lyophilised and the residue was resuspended in 6N HCl. Complete hydrolysation was achieved by incubating at 110°C for 24 h. The amino acid composition of the amino-terminal chymotryptic peptide was then determined as described above.

Digestion with carboxypeptidase Y.

Samples of protein were dissolved in 0.1 M ammonium acetate (pH 5.5), 0.01% SDS and incubated with 50 µg/ml carboxypeptidase Y, according to Hayashi (1977). Carboxypeptidase Y (Worthington) freshly dissolved in distilled water to a concentration of 1 mg/ml was used. In each experiment the protein-enzyme ratio was approximately 50:1 (w/w). The mixture was incubated at 25°C and samples were taken after 0.5, 1 and 2 h of incubation. Undigested peptide chains were precipitated in 2% sulphosalicylic acid at 4°C and removed by centrifugation. The supernatant fractions containing the released amino acids were diluted with lithium citrate to a final concentration of 0.12 M, corresponding to a pH of approximately 1.9, and applied to a Biotronik LC 6000 E amino acid analyser (Durrum - DC6a cationion exchange resin, packaged in lithium citrate buffer pH 2.8). Norleucine was included as internal standard. Elution of the amino acids was performed in a discontinuous lithium citrate buffer system as described by Hamilton (1963) and Benson (1973).

RESULTS

Purification and separation of carboxymethylated VP37 and VP23.

CPMV capsid proteins were properly separated after carboxymethylation and gelfiltration on a Sephadex G200 column in 5M urea (Fig. 1, panel A). Upon SDS-polyacrylamide gelelectrophoresis of the purified carboxymethylated proteins it appeared that the carboxymethylated VP37 ran significantly slower than the untreated protein. This might indicate a relatively high content of cysteine residues in VP37. The column fractions with carboxymethylated VP23 contained a mixture of at least two forms of this protein, designated VP23 (intact protein) and VP20, respectively (Fig. 1, panel B) the relative amounts of which depended on the age of the virus preparations used. It has been shown that this heterogeneity of the small capsid protein is due to the loss of a specific carboxyterminal sequence. (Geelen *et al.*, 1972; Kridl and Bruening, 1983).

Amino-terminal amino acid sequences of VP23 and VP37.

Since the amino-terminal of VP37 is blocked and the size heterogeneity of VP23 only occurs at the carboxy-terminal of this polypeptide (Geelen *et al.*, 1972; Kridl and Bruening, 1983) the amino-terminal amino acid sequence of VP23 could be determined by Edman degradation applied to purified top components. Such Edman degradation resulted indeed in a single unambiguous amino acid sequence. Inspection of the nucleotide sequence of the open reading frame in M RNA (Van Wezenbeek *et al.*, 1983) revealed that this amino acid sequence matched the sequence from nucleotide 2660 downstream (Fig. 2, sequences underlined ——). Assuming that the coding region of VP23 extends from this position to the end of the open reading frame at nucleotide 3299, the size of this protein would be 23.6K. This in agreement with the experimental determined size for this protein and supports the mapping of the amino-terminal amino acid of VP23 at nucleotide 2660 of M RNA. *In vitro*

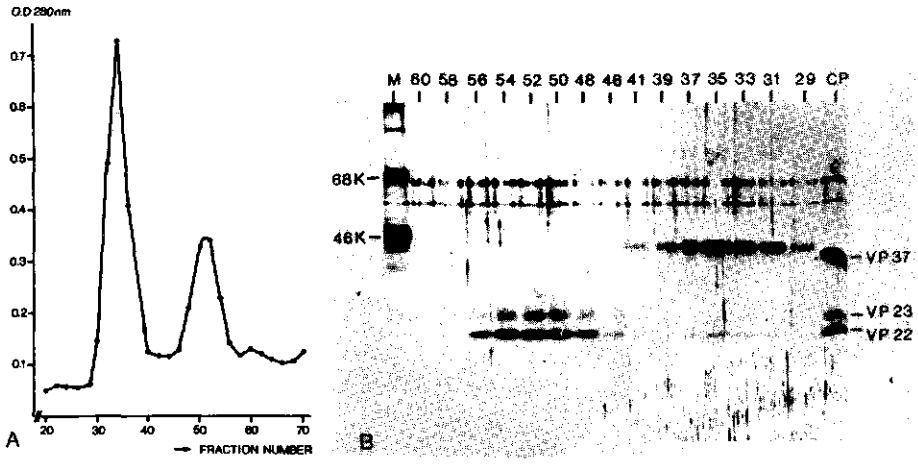


Fig. 1A: Elution profile of carboxymethylated capsid proteins from CPMV on a Sephadex G-200 column (100 x 1.5 cm). Elution was performed in 5 M urea. Profiles were obtained by measuring the optical density at 280 nm.

Fig. 1B: Gel electrophoretic analysis of the protein content of the Sephadex G-200 column fractions. The numbers at the top of the gel correspond to the numbers of the fractions in panel A. The marker lane (M) contains bovine serum albumine (68K) and ovalbumine (46K). The lane indicated with CP contains unfractionated, non-carboxymethylated capsid proteins. Protein bands were visualised by silverstaining. The background bands in the gel around 55K and 68K are probably due to the staining method (Morrissey, 1981; Dorssers, 1983; Tasheva and Dessev, 1983; Ochf, 1983; Dorssers *et al.*, 1984).

translation studies have indicated that within the overlapping primary translation products of M RNA the VP37 sequence precedes the sequence of VP23 (Franssen *et al.*, 1982). As the amino-terminal end of VP37 is known to be an aminoacetylated methionine residue (Bruening, 1981) and with the amino-terminal end of VP23 now being mapped, it has become possible to make a bid for the AUG codon within the open reading frame of M RNA, which might be coding the amino-terminal methionine of VP37. In the M RNA sequence there are only two AUG codons at positions which

allow a size for VP37 approximating its apparent molecular weight. These triplets are found at nucleotide positions 1537 and 1649 and would result in VP37 molecules with sizes of 41.2K and 37.2K respectively. In order to discriminate between these two possibilities VP37 was digested with chymotrypsin. The blocked amino-terminal peptide was isolated by chromatography on a Dowex 50W-X8 column and analysed for its amino acid composition. Inspection of the M RNA sequence (Van Wezenbeek *et al.*, 1983) learns that if the methionine residue corresponding to nucleotide position 1537 represents the amino-terminal amino acid of VP37, the amino-terminal blocked chymotryptic peptide is expected to contain methionine, glutamic acid, glutamine, asparagine, leucine and phenylalanine in approximately equal amounts. On the other hand, if the methionine of position 1649 represents the amino-terminal residue this peptide should contain methionine, alanine, glycine, aspartic acid, valine, leucine, glutamic acid and tyrosine in a 1:1:2:2:1:2:1:1 ratio, respectively. The experimentally determined amino acid composition (Table I) reveals that the blocked chymotryptic peptide mainly consists

Table I. Amino acid composition of the chymotryptic peptide from the amino-terminal end of VP37.

Amino acid	nmol/100 µl*
glutamic acid/glutamine	3.8
aspartic acid/asparagine	2.7
glycine	2.3
leucine	2.2
phenylalanine	1.9

* Portions of 100 µl of the hydrolysed chymotryptic peptide dissolved in 0.12 M lithiumcitrate pH 1.8-2.0, were used for determining the amino acid composition. For details see Materials and Methods.

of glutamic acid, aspartic acid, glycine, leucine and phenylalanine residues, in a ratio of roughly 2:1:1:1:1. Since the detected glutamic acid and aspartic acid residues may have arisen from glutamine and asparagine respectively by the acidic hydrolysis procedure, this result indicates that if the methionine of position 1537 represents the first amino acid of VP37 five out of the six expected amino acid residues are recovered (Fig. 2, sequences underlined ----). On the other hand, if the methionine residue of position 1649 would represent the amino-terminal amino acid of VP37 only four out of eight expected amino acids are found. In addition the relative amounts in which the amino acids are found clearly disagree with the latter possibility. Therefore it is reasonable to conclude that the coding region of VP37 starts with the AUG triplet of position 1537.

Amino acid composition of purified VP37 and VP23.

To verify the mapping of both VP37 and VP23 on M RNA the amino acid composition of both proteins was determined and compared with the composition of these proteins derived from the nucleotide sequence of the open reading frame in M RNA (Table II). The good agreement between the experimental values (Table II, lane A) and the theoretical values (lane B) confirm the mapping of VP37 and VP23 with their amino termini on position 1537 and 2660, respectively.

Analyses of the carboxy-terminal amino acid sequences of VP23 and VP37.

To obtain additional evidence that the cistron for VP23 continues as far as the stopcodon at the end of the open reading frame of M RNA, the carboxy-terminal amino acids of VP23 were determined by limited digestion with carboxypeptidase Y.

VP23 is known to be heterogeneous in size due to the loss of a carboxy-terminal peptide (Wu and Bruening, 1972; Geelen *et al.*, 1972; Siler *et al.*, 1976; Kridl and Bruening, 1983). This nibbling at the

Table II: Amino acid composition of the capsid proteins of CPMV: a comparison of the amino acid composition determined experimentally (lane A) and derived from the nucleotide sequence of M RNA (lane B).

	VP37		VP23 + VP20□	
amino acid	A	B	A	B
asp + asn	10.5*	10.2	12.0	11.7
thr	8.4	8.6	6.7	7.0
ser	8.8	8.8	8.3	7.0
glu + gln	8.4	7.0	8.6	7.5
pro	4.6	3.7	8.3	8.9
gly	8.9	8.0	7.6	5.6
ala	6.7	6.4	7.4	8.4
val	6.9	7.0	7.1	8.0
met	3.0	2.9	2.7	3.3
ile	5.1	5.9	4.7	5.6
leu	8.2	8.3	5.5	4.7
tyr	2.8	2.7	2.5	2.4
phe	5.9	5.6	5.9	6.6
his	1.4	1.3	1.2	1.4
lys	3.7	4.0	2.2	2.4
arg	4.7	4.3	4.5	5.6
trp	N.D.	N.D.	N.D.	N.D.
cys	2.0†	2.9	4.7†	1.4

* presented as mol percentages.

† detected as carboxymethylated cysteine residue

□ the small capsid protein preparation was a mixture of VP23 and VP20 in roughly 1:1 ratio.

carboxy-terminal end mainly results in a shortened polypeptide of approximately 20K, in this paper referred to as VP20. Due to this heterogeneity, digestion of VP23 with carboxypeptidase Y is expected to produce a complex mixture of amino acids. The protein preparation used for carboxypeptidase Y digestion contained the two forms of the small capsid protein in approximately 1:1 ratio, as determined by SDS-polyacrylamide gelelectrophoresis (results not shown).

Table III shows that carboxypeptidase Y digestion of this preparation mainly released two amino acid species: alanine and leucine in approximately equal amounts. Alanine matches the triplets at the very end of the reading frame of M RNA, where two adjoining codons for alanine are present (Fig. 2, sequences underlined ....). Therefore, these alanine residues probably originates from the carboxy terminal end of VP23. Upstream in the VP23 coding region the first leucine codons are found at nucleotide positions 3225 and 3228. Since the

Table III. Amino acids released from the carboxy termini of the CPMV capsid proteins upon incubation with carboxypeptidase Y.

Amino acid	VP23/VP20	VP37
alanine	4*	1.5
leucine	4	1.2
valine	2.5	-
methionine	1	-
phenylalanine	1	-
glycine	1.5	1.8
isoleucine	-	0.4
glutamic acid/glutamine	-	0.8

* The numbers indicate nmols per 100 µl reaction mixture.

digestion mixture contained alanine and leucine in about 1:1 ratio it is probable that the leucine originated from the carboxy terminal of VP20 molecules ending with the leucine-residues encoded by codons at positions 3225 or 3228 (Fig. 2, sequences underlined ·-·-·-). Since carboxypeptidase Y is not capable of releasing proline at carboxyl ends in a polyproline sequence (Hayashi, 1977) further release of residues upstream of these two leucine residues is not expected. The small amounts of valine, glycine and methionine detected among the released amino acids (Table III, lane VP23) may be due to a contamination with one or more polypeptides with a chain length between those of VP20 and VP23.

Carboxypeptidase Y digestion of VP37 revealed the presence of glycine, alanine, leucine, glutamic acid/glutamine and isoleucine residues at the carboxyterminus of this capsid protein. (Table III, lane VP37). Since these amino acids can all be mapped on the M RNA nucleotide sequence in a short region preceding the coding region of VP23 (Fig. 2, sequences underlined ··-···-), this finding confirms that the coding regions VP37 and VP23 are contiguous.

Fig. 2: Mapping of the carboxy-terminal and amino-terminal ends of VP37 and VP23 on the amino acid sequence derived from the open reading frame in CPMV-M RNA. The underlined sequence (——) of the first eleven amino acids at the amino-terminal end of VP23 was determined by Edman degradation.
The amino acids underlined by the broken lines (----) were detected in the hydrolysate of the blocked chymotryptic oligopeptide from the amino-terminal end of VP37.
The amino acids released from the carboxy-terminal end of VP37 by limited digestion with carboxypeptidase Y are underlined with ._._._.
The amino acids released from the carboxy-terminal ends of VP23 and VP20 by limited digestion with carboxypeptidase Y are underlined by ..... and ·-·-·; respectively.
The nucleotide sequence of M-RNA was taken from Van Wezenbeek *et al.*, (1983). For details see the text.

VP 37

1505 GAU CCA CGA AAU GGG AAU GUG GCU UUU CCA CAA AUG GAG CAA AAC UUG UUU GCC CUU UCU
asp pro arg asn gly asn val ala phe pro gln met glu gln asn leu phe ala leu ser

1565 UUG GAU GAU ACA AGC UCA GUU CGU GGU UCU UUG CUU GAC ACA AAA UUC GCA CAA ACU CGA
leu asp asp thr ser ser val arg gly ser leu leu asp thr lys phe ala gln thr arg

1625 GUU UUG UUG UCC AAG GCU AUG GCU GGU GGU GAU GUG UUA UUG GAU GAG UAU CUC UAU GAU
val leu leu ser lys ala met ala gly gly asp val leu leu asp glu tyr leu tyr asp

1685 GUG GUC AAU GGA CAA GAU UUU AGA GCU ACU GUC GCU UUU UUG CGC ACC CAU GUU AUA ACA
val val asn gly gln asp phe arg ala thr val ala phe leu arg thr his val ile thr

1745 GGC AAA AUA AAG GUG ACA GCU ACC ACC AAC AUU UCU GAC AAC UCG GGU UGU UGU UUG AUG
gly lys ile lys val thr ala thr thr asn ile ser asp asn ser gly cys cys leu met

1805 UUG GCC AUA AAU AGU GGU GUG AGG GGU AAG UAU AGU ACU GAU GUU UAU ACU AUC UGC UCU
leu ala ile asn ser gly val arg gly lys tyr ser thr asp val tyr thr ile cys ser

1865 CAA GAC UCC AUG ACG UGG AAC CCA GGG UGC AAA AAG AAC UUC UCG UUC ACA UUU AAU CCA
gln asp ser met thr trp asn pro gly cys lys lys asn phe ser phe thr phe asn pro

1925 AAC CCU UGU GGG GAU UCU UGG UCU GCU GAG AUG AUA AGU CGA AGC AGA GUU AGG AUG ACA
asn pro cys gly asp ser trp ser ala glu met ile ser arg ser arg val arg met thr

1985 GUU AUU UGU GUU UCG GGA UGG ACC UUA UCU CCU ACC ACA GAU GUG AUU GCC AAG CUA GAC
val ile cys val ser gly trp thr leu ser pro thr thr asp val ile ala lys leu asp

2045 UGG UCA AUU GUC AAU GAG AAA UGU GAG CCC ACC AUU UAC CAC UUG GCU GAU UGU CAG AAU
trp ser ile val asn glu lys cys glu pro thr ile tyr his leu ala asp cys gln asn

2105 UGG UUA CCC CUU AAU CGU UGG AUG GGA AAA UUG ACU UUU CCC CAG GGU GUG ACA AGU GAG
trp leu pro leu asn arg trp met gly lys leu thr phe pro gly gly val thr ser glu

2165 GUU CGA AGG AUG CCU CUU UCU AUA GGA GGC GGU GCU GGU GCG ACU CAA GCU UUC UUG GCC
val arg arg met pro leu ser ile gly gly gly ala gly ala thr gln ala phe leu ala

2225 AAU AUG CCC AAU UCA UGG AUA UCA AUG UGG AGA UAU UUU AGA GGU GAA CUU CAC UUU GAA
asn met pro asn ser trp ile ser met trp arg tyr phe arg gly glu leu his phe glu

2285 GUU ACU AAA AUG AGC UCU CCA UAU AUU AAA GCC ACU GUU ACA UUU CUC AUA GCU UUU GGU
val thr lys met ser ser pro tyr ile lys ala thr val thr phe leu ile ala phe gly

2345 AAU CUU AGU GAU GCC UUU GGU UUU UAU GAG AGU UUU CCU CAU AGA AUU GUU CAA UUU GCU
asn leu ser asp ala phe gly phe tyr glu ser phe pro his arg ile val gln phe ala

2405 GAG GUU GAG GAA AAA UGU ACU UUG GUU UUC UCC CAA CAA GAG UUU GUC ACU GCU UGG UCA
glu val glu glu lys cys thr leu val phe ser gln gln glu phe val thr ala trp ser

2465 ACA CAA GUA AAC CCC AGA ACC ACA CUU GAA GCA GAU GGU UGU CCC UAC CUA UAU GCA AUU
thr gln val asn pro arg thr thr leu glu ala asp gly cys pro tyr leu tyr ala ile

2525 AUU CAU GAU AGU ACA ACA GGU ACA AUC UCC GGA GAU UUU AAU CUU GGG GUC AAG CUU GUU
ile his asp ser thr thr gly thr ile ser gly asp phe asn leu gly val lys leu val

2585 GGC AUU AAG GAU UUU UGU GGU AUA GGU UCU AAU CCG GGU AUU GAU GGU UCC CGC UUG CUU
gly ile lys asp phe cys gly ile gly ser asn pro gly ile asp gly ser arg leu leu

VP23 →
2645 GGA GCU AUA GCA CAA GGA CCU GUU UGU GCU GAA GCC UCA GAU GUG UAU AGC CCA UGU AUG
gly ala ile ala gln gly pro val cys ala glu ala ser asp val tyr ser pro cys met

2705 AUA GCU AGC ACU CCU CCU GCU CCA UUU UCA GAC GUU ACA GCA GUA ACU UUU GAC UUA AUC
ile ala ser thr pro pro ala pro phe ser asp val thr ala val thr phe asp leu ile

2765 AAC GGC AAA AUA ACU CCU GUU GGU GAU GAC AAU UGG AAU ACG CAC AUU UAU AAU CCU CCA
asn gly lys ile thr pro val gly asp asp asn trp asn thr his ile tyr asn pro pro

2825 AUU AUG AAU GUC UUG CGU ACU GCU GCU UGG AAA UCU GGA ACU AUU CAU GUU CAA CUU AAU
ile met asn val leu arg thr ala ala trp lys ser gly thr ile his val gln leu asn

2885 GUU AGG GGU GCU GGU GUC AAA AGA GCA GAU UGG GAU GGU CAA GUC UUU GUU UAC CUG CGC
val arg gly ala gly val lys arg ala asp trp asp gly gln val phe val tyr leu arg

2945 CAG UCC AUG AAC CCU GAA AGU UAU GAU GCG CGG ACA UUU GUG AUC UCA CAA CCU GGU UCU
gln ser met asn pro glu ser tyr asp ala arg thr phe val ile ser gln pro gly ser

3005 GCC AUG UUG AAC UUC UCU UUU GAU AUC AUA GGG CCG AAU AGC GGA UUU GAA UUU GCC GAA
ala met leu asn phe ser phe asp ile ile gly pro asn ser gly phe glu phe ala glu

3065 AGC CCA UGG GCC AAU CAG ACC ACC UGG UAU CUU GAA UGU GUU GCU ACC AAU CCC AGA CAA
ser pro trp ala asn gln thr thr trp tyr leu gly cys val ala thr asn pro arg gln

3125 AUA CAG CAA UUU GAG GUC AAC AUG CGC UUC GAU CCU AAU UUC AGG GUU GCC GGC AAU AUC
ile gln gln phe glu val asn met arg phe asp pro asn phe arg val ala gly asn ile

VP20 ←
3185 CUG AUG CCC CCA UUU CCA CUG UCA ACG GAA ACU CCA CCG UUA UUA AAG UUU AGG UUU CGG
leu met pro pro phe pro leu ser thr glu thr pro pro leu leu lys phe arg phe arg

VP23 ←
3245 GAU AUU GAA CGC UCC AAG CGU AGU GUU AUG GUU GGA CAC ACU GCU ACU GCU GCU UAA
asp ile glu arg ser lys arg ser val met val gly his thr ala thr ala ala

3002 CUCUGGUUUCAUUAAAUUUUCUUUAGUUUGAAUUUACUGUUAUUUGGUGUGCAUUUCUAUGUUUGGUGAGCGGUUUUCU
GUGCUCAGAGUGUGUUUAUUUUAUGUAAUUUAAUUUCUUUGUGAGCUCCUGUUUAGCAGGUCGUCCCUUCAGCAAGGAC
ACAAAAAGAUUUUAAUUUUAUU-poly(A)

DISCUSSION

Previous translation studies have demonstrated that the capsid proteins of CPMV are produced by processing of a 60K precursor protein derived from the C-terminal halves of the primary translation products from M RNA (Franssen *et al.*, 1982). The amino acid analyses described in this paper allow a more precise mapping of the coding region of VP37 and VP23 in the 3'-terminal part of the M RNA sequence.

The coding region of VP23 starts on the M RNA sequence at nucleotide 2660 (Fig. 2) and continues probably to the end of the open reading frame in M RNA (stopcodon at position 3299). Furthermore, the results presented here, provide evidence that the sequence of the smaller form of the small capsid protein, denoted VP20, ends with a leucine residue which probably corresponds with the triplet at nucleotide position 3228 to 3230. This is in agreement with the previous finding of Niblett and Semancik (1969) that the carboxy terminal residue of VP20 is a leucine. Based on the tentative mapping of both VP23 and VP20 on the nucleotide sequence of M RNA the difference between VP20 and VP23 would be a carboxy-terminal peptide of 23 amino acids and a molecular weight of 3,157. This fits properly with the difference in apparant molecular weight of these forms as determined by gel electrophoresis (Geelen, 1974; Rottier *et al.*, 1980). The release of minor amounts of glycine, valine and methionine from VP23 upon carboxypeptidase Y digestion (Table III) suggests the possible occurrence of one or more additional degradation products of VP23 ending near or at the triplet for glycine at position 3278 of the M RNA sequence. Such products would only be 6-9 amino acids shorter than full-length VP23 and are probably not detected as a separate band upon SDS-polyacrylamide gelelectrophoresis (Fig. 1B). The occurrence of some such forms of the small capsid protein has been detected previously by Rottier (1980).

The amino acid composition of the blocked amino-terminal peptide of VP37 (Table I) strongly suggests that the coding region of VP37 starts with the AUG codon at position 1537. This mapping is further supported by the good agreement of the experimentally determined amino acid com-

position of VP37 with the predicted amino acid composition of this protein assuming that its sequence starts at the methionine codon at position 1537. N-acetylated amino acids are often followed by a hydrophilic acidic amino acid residue (Driessen, 1983). This condition is indeed fulfilled at the methionine residue of position 1537, but not at the methionine residue of position 1641.

Besides, evidence was obtained that the coding regions of VP37 and VP23 are contiguous since all amino acids released from the carboxy terminus of VP37 by carboxypeptidase Y are found in the amino acid sequence within the primary translation product of M RNA preceding the amino-terminal glycine residue of VP23 (Table III). The mapping of both capsid proteins as described in this paper implies molecular weights of 41,216 and 23,681 for VP37 and VP23 respectively. The mapping allows furthermore for the identification of two different proteolytic cleavage sites in the 105K and 95K primary translation products from M RNA. Since in the amino acid sequence of these primary translation products, the amino-terminal methionine of VP37 on position 1537 is preceded by a glutamine residue (Fig. 2) it appears that a cleavage of glutaminyl-methionine dipeptide sequence is utilized to generate the 60K capsid protein precursor (Franssen *et al.*, 1982). This site is recognised by the 32K protease encoded by B RNA (Franssen *et al.*, 1984). At the other hand, since the coding regions for VP37 and VP23 are contiguous a glutaminyl-glycine bond should be cleaved to release VP37 and VP23 from their common precursor (Fig. 2). The protease responsible for this cleavage has not yet been identified.

ACKNOWLEDGEMENT

We thank Piet Madern for printing the photographs and Marie-José van Neerven for typing the manuscript.

REFERENCES

1. Ahlquist, P., and Kaesberg, P. (1979). Nucleic Acids Res. 7, 1195-1204.
2. Benson, J.R. (1973). In: Simmons, J.L., and Ewing, G.W. (eds.). Applications of the newer techniques of analysis. p. 223. Plenum Publishing Corp. N.Y.
3. Brauer, A.W., Marqalies, M.N., and Haber, E. (1975). Biochemistry 14, 3029-3035.
4. Bruening, G. (1977). In: Fraenkel-Conrat, H. and Wagner, R.R. (eds.). Comprehensive Virology, vol. 11, p. 55-141. Plenum Publishing Corp., N.Y.
5. Bruening, G. (1981). In: Marcus, E. (ed.). The Biochemistry of plants, vol. 6, p. 571-631, Academic Press.
6. Daubert, S.D., Bruening, G., and Najarian, R.C. (1978). Eur.J. Biochem. 92, 45-51.
7. Dorssers, L. (1983). Thesis, Agricultural University, Wageningen, The Netherlands.
8. Dorssers, L., Van der Krol, S., Van der Meer, J., Van Kammen, A., and Zabel, P. (1984). Proc. Natl. Acad. Sci. U.S.A., in press.
9. Driessen, H. (1983). Thesis, University of Nijmegen, Nijmegen, The Netherlands.
10. El Manna, M., and Bruening, G. (1973). Virology 56, 198-206.
11. Frank, G. and Strubert, W. (1973). Chromatographia 6, 522-524.
12. Franssen, H., Goldbach, R., Broekhuysen, M., Moerman, M., and Van Kammen, A. (1982). J. Virol. 41, 8-17.
13. Franssen, H., Moerman, M., Rezelman, G., and Goldbach, R. (1984). J. Virol., in press.
14. Geelen, J., Van Kammen, A. and Verduin, B.J.M. (1972). Virology 49, 205-213.
15. Geelen, J. (1974). Thesis, Agricultural University, Wageningen, The Netherlands.
16. Goldbach, R.W., Schilthuis, J.G., and Rezelman, G. (1981). Biochem. Biophys. Res. Commun. 99, 89-94.

17. Gopo, J.M. and Frist, R.H. (1977). Virology 79, 259-266.
18. Hamilton, P. (1963). Anal. Chem. 35, 2055.
19. Hayashi, R. (1977). In: Methods in Enzymology 47, 84-103.
20. Hunkapiller, M.W., and Hood, L.E. (1978). Biochemistry 17, 2124-2133.
21. Kridl, J.C. and Bruening, G. (1983). Virology 129, 369-380.
22. Laemmli, U.K. (1970). Nature (London) 227, 680-685.
23. Morrissey, J.H. (1981). Anal. Biochem. 117, 307-310.
24. Narita, K., Matsuo, H., and Nakajima, T. (1975). In: Needleman (ed.). Protein sequence determination, p. 30-104. Springer-Verlag, Berlin.
25. Niblett, C.L., and Semancik, J.S. (1969). Virology 38, 685-693.
26. Ochf, D. (1983). Anal. Biochem. 135, 470-474.
27. Pelham, H.R.B. (1979). Virology 96, 463-477.
28. Rottier, P.J.M., Rezelman, G., and Van Kammen, A. (1980). J. Gen. Virol. 51, 359-371.
29. Siler, D.J., Babcook, J., and Bruening, G. (1976). Virology 71, 560-567.
30. Stanley, J., Rottier, P., Davies, J.W., Zabel, P. and Van Kammen, A. (1978). Nucleic Acids Res. 5, 4505-4522.
31. Steere, R.L. (1956). Phytopathology 46, 60-69.
32. Tasheva, B. and Dessev, G. (1983). Anal. Biochem. 129, 98-102.
33. Van Kammen, A. (1967). Virology 31, 633-642.
34. Van Wezenbeek, P., Verver, J., Harmsen, J., Vos, P., and Van Kammen, A. (1983). EMBO J. 2, 941-946.
35. Vereijken, J.M., Schwander, E.H., Soeter, N.M., and Beintema, J.J; (1982). Eur. J. Biochem. 123, 283-289.
36. Wood, H.A. (1972). Virology 49, 592-598.
37. Wu, G.J. (1970). Thesis, University of California, Davis, U.S.A.
38. Wu, G.J. and Bruening, G. (1971). Virology 46, 596-612.

CHAPTER VII

TRANSLATION OF BOTTOM COMPONENT RNA OF COWPEA MOSAIC VIRUS IN RETICULOCYTE LYSATE: FAITHFUL PROTEOLYTIC PROCESSING OF THE PRIMARY TRANSLATION PRODUCT

Henk Franssen, Rob Goldbach and Ab van Kammen

Virus Research 1, 39-49 (1984).

Virus Research, 1 (1984) 39–49
Elsevier

VRR 00102

Translation of bottom component RNA of cowpea mosaic virus in reticulocyte lysate: faithful proteolytic processing of the primary translation product

Henk Franssen, Rob Goldbach and Ab van Kammen

Department of Molecular Biology, Agricultural University, De Dreijen 11, 6703 BC Wageningen, The Netherlands

(Accepted 23 November 1983)

Summary

The expression of bottom component (B) RNA of cowpea mosaic virus (CPMV) was studied by in vitro translation, using the rabbit reticulocyte lysate system. Translation of B-RNA produced within 1 h a 200 kilodalton (K) primary translation product, which was immediately cleaved into 32K and 170K polypeptides. This primary cleavage also occurred on nascent polypeptide chains longer than 110K. Upon prolonged incubation of the translation mixture (18 h) the 170K polypeptide was further cleaved into 110K, 87K, 84K and 60K polypeptides, corresponding to those found in infected cowpea cells. In addition a fifth cleavage product of approximately 28K was detected, not found in vivo thus far. The faithful in vitro processing observed indicates that the proteolytic enzyme involved is coded for by the B-RNA itself. A model is proposed in which the proteolytic activity involved in the processing of the 200K primary translation product is located in the 28K polypeptide.

CPMV, reticulocyte lysate, in vitro translation, proteolysis

Introduction

The extensive studies on its proteins (see e.g. Rezelman et al., 1980; Goldbach et al., 1982) and the recent elucidation of the complete nucleotide sequence of its

genome (van Wezenbeek et al., 1983; Lomonossoff and Shanks, 1983) make cowpea mosaic virus (CPMV) one of the best characterised plant viruses. The genome of this virus consists of two separately encapsidated RNA molecules of plus-strand polarity, which are 3481 (M-RNA) and 5889 nucleotides (B-RNA) long, respectively, each containing one, long open reading frame (van Wezenbeek et al., 1983; Lomonossoff and Shanks, 1983). Both RNAs are polyadenylated at their 3′ terminus (El Manna and Bruening, 1973; Ahlquist and Kaesberg, 1979) and carry a small protein, called VPg, at their 5′ terminus (Stanley et al., 1978; Daubert et al., 1978). Characterisation of the polypeptides produced by CPMV in cowpea mesophyll protoplasts (Rezelman et al., 1980; Goldbach et al., 1982; Goldbach and Rezelman, 1983) as well as those translated from its RNA in vitro (Pelham, 1979; Goldbach et al., 1981; Franssen et al., 1982) has revealed that this virus follows a translation strategy which involves proteolytic processing of large, primary translation products. Thus, it has been demonstrated that in vivo B-RNA is translated into a 200K polyprotein, which undergoes several successive cleavages. The primary cleavage generates 32K and 170K polypeptides. Next, the 170K polypeptide is cleaved into 60K and 110K polypeptides, or, alternatively, into 84K and 87K polypeptides (secondary cleavages). The 60K product represents the direct precursor to VPg and a 58K polypeptide, which both arise from the 60K polypeptide by a single further cleavage (Zabel et al., 1982; Goldbach et al., 1982). Furthermore, since two pairs of overlapping proteins arise after the alternative, secondary cleavages, it has been proposed that the 110K and 84K polypeptides may undergo additional cleavage to give the 87K and 60K polypeptides, respectively, together with a polypeptide of approximately 24K (Rezelman et al., 1980). Until now the existence of such a 24K polypeptide has remained a matter of speculation.

In several aspects CPMV resembles the animal picornaviruses, like poliovirus and encephalomyocarditis virus (EMC) (for reviews see Rueckert et al., 1979; Koch et al., 1981; Putnak and Phillips, 1981). These viruses also have a positive-stranded RNA genome, supplied with both a VPg molecule and a polyadenylate tail. Moreover, they use the same translation strategy, i.e. proteolytic processing of a large 'polyprotein'. In rabbit reticulocyte lysates translation of EMC RNA has been reported to be followed by almost complete processing of the primary translation product to give the proteins also found in vivo (Pelham, 1978; Shih et al., 1979; Palmenberg, 1982). In contrast to these results, so far translation of CPMV B-RNA in reticulocyte lysates has been reported to be followed by the primary cleavage only, giving the 32K and 170K polypeptides (Pelham, 1979; Goldbach et al., 1981). In the present paper we show, however, that in this system, after prolonged incubation, also further cleavage of the 170K polypeptide can be achieved, resulting in the production of most of the B-RNA-encoded polypeptides found in infected cells. The experiments described here confirm and extend the translation map of B-RNA as based on our previous in vivo studies, and demonstrate that the reticulocyte lysate may be particularly useful for studying the proteolytic pathways followed and the protease(s) involved.

Materials and Methods

Virus and RNA

CPMV was grown in *Vigna unguiculata* (L.) 'California Blackeye' and B and M components separated and purified as described before (Klootwijk et al., 1977; van Kammen, 1967). CPMV RNA molecules were extracted from separated components as described by Davies et al. (1978) and fragmented molecules removed by sucrose gradient centrifugation (Franssen et al., 1984).

In vitro translation of CPMV RNA

B-RNA was translated in a messenger-dependent lysate of rabbit reticulocytes (a generous gift of Dr. R.J. Jackson, Department of Biochemistry, University of Cambridge, Cambridge, U.K.) in general under the conditions as previously described (Pelham and Jackson, 1976; Pelham, 1979), using a dithiothreitol concentration of 2 mM. [^{35}S]Methionine (New England Nuclear, specific activity approximately 1200 Ci/mmol) was used as radioactive amino acid (10 μCi per 10 μl reaction mixture) and incubation was at 30°C for the times indicated in the text. The endogenous activity of the lysate used (no RNA added) was undetectable.

SDS–polyacrylamide gel electrophoresis

Labelled protein samples were made up in sample buffer (SB) (Rezelman et al., 1980), heated for 3 min at 100°C, and electrophoresed in SDS–polyacrylamide gels using the buffer system of Laemmli (1970). Gels contained 12.5% or 20% acrylamide with 0.09% or 0.07% bisacrylamide, respectively, using a stacking gel of 4% acrylamide (with 0.07% bisacrylamide). After electrophoresis (at 150 V) gels were dried either with or without prior staining, and autoradiographed.

Immunoprecipitations

Antiserum directed against the electrophoretically separated 32K polypeptide was prepared as described previously (Franssen et al., 1984). Immunoprecipitations of protein samples were performed in buffer containing 10 mM sodium phosphate, pH 7.2, 0.9% NaCl, 1% Triton X-100, 0.5% sodium deoxycholate and 0.1% SDS, as described before (Franssen et al., 1982).

Peptide mapping

Bands of radiolabeled polypeptides, localized by autoradiography, were cut out from dried, unstained gels. The gel pieces were swollen in buffer (120 mM Tris-HCl (pH 6.8)/1 mM EDTA/0.4% SDS) containing 150 μg of *Staphylococcus aureus* V8 protease (Miles Laboratories, Inc., Elkhart, Ind., U.S.A.) per ml, and incubated for 1 h at 30°C. The gel pieces were then transferred into slots of a 20% SDS–polyacrylamide gel, covered with a layer of 2 × SB (see above) and electrophoresed at 150 V. After the run the gel was fluorographed using Enhance (New England Nuclear).

Results

The primary cleavage of the 200K translation product

In the reticulocyte lysate B-RNA is translated into a 200K polypeptide, which is then, within 1 h of incubation, cleaved into a 32K and a 170K polypeptide (Pelham, 1979; Rezelman et al., 1980). Since the 32K polypeptide is derived from the amino-terminal part of the 200K polypeptide (Goldbach and Rezelman, 1983), first

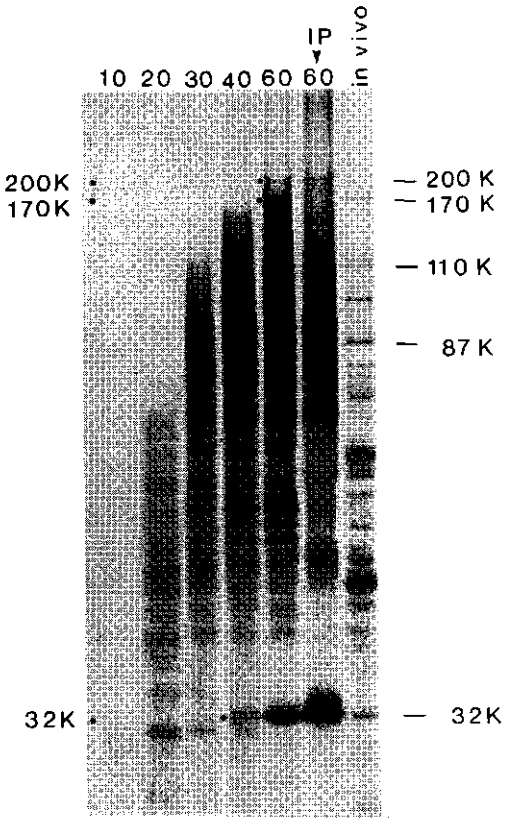


Fig. 1. Time course of appearance of the B-RNA-encoded 32K and 170K polypeptides. B-RNA was translated in a rabbit reticulocyte lysate at 30°C. Samples were taken at 10, 20, 30, 40 and 60 min after start of protein synthesis and analysed in a 12.5% SDS–polyacrylamide gel. An additional sample taken at 60 min was immunoprecipitated, using antiserum raised against the 32K polypeptide, and the immunoprecipitate obtained included in the gel (lane 60, IP). The right lane (in vivo) contains the soluble [^{35}S]methionine-labeled polypeptides from B-component-inoculated cowpea protoplasts, containing the B-RNA-encoded 170K, 110K and 87K polypeptides.

it was determined whether this cleavage occurs already on short, nascent polypeptide chains, or only on polypeptide chains which have been completed. For this purpose samples were taken from a translation mixture at various times after addition of B-RNA. After 30 min of translation only incomplete polypeptide chains were synthesized, with maximal lengths of approximately 110K (Fig. 1). After 40 min these chains were further elongated to approximately 150K. Neither full-length 200K or 170K polypeptides were synthesized yet, but clearly a band of 32K polypeptide appeared (Fig. 1). Full-length 200K product was only detectable after 60 min of translation and, as follows from the amounts of 170K and 32K polypeptides present then, most of the chains had already undergone primary cleavage (Fig. 1). Part of the 60 min sample was immunoprecipitated using anti-32K serum. It is clear that this antiserum efficiently precipitated both the 200K and 32K polypeptide, but not the 170K polypeptide (Fig. 1, lane indicated with IP). Most of the polypeptide chains after 1 h of translation still shorter than 170K were also precipitated by anti-32K serum. Such short products are thought to be 'early quitters', due to hidden breaks or secondary structure in the RNA. B-RNA is known to be a difficult messenger for in vitro translation (Pelham, 1979; Rezelman et al., 1980). The results of Fig. 1 indicate that the primary cleavage can take place on nascent 200K chains. Such cleavage only occurs if the nascent chains have a size exceeding 110K, however.

Further cleavage of the 170K polypeptide

Upon prolonged incubation of the B-RNA-directed translation mixture additional cleavages were observed. Whereas after 1 h mainly 170K and 32K polypeptides were found, further incubation of the mixture for another 17 hours led to the production of 5 new proteins (cf. lanes 1h and 18h in Fig. 2). Two of these polypeptides comigrated exactly with the 110K and 87K polypeptides specified by B-RNA in vivo (Fig. 2, lane B), and two other ones were found on positions in the gel as expected for the 84K and 60K polypeptide. This result indicates that in the reticulocyte lysate the 170K polypeptide can undergo the same cleavages as found in the infected cell (Rezelman et al., 1980; Goldbach et al., 1982). Close inspection of the gel patterns in Fig. 2 reveals that some of the 87K and 84K polypeptides were already produced after 1 h of translation. This suggests that the cleavage by which they were generated proceeded more rapidly than the cleavage by which the 110K and 60K polypeptides were formed. The fifth new protein migrated slightly faster than the 32K polypeptide (Fig. 2). This polypeptide might represent the '24K' polypeptide, the existence of which has been postulated, but has never been found in vivo so far (Rezelman et al., 1980). Such a protein could arise by further cleavage of the 110K and 84K polypeptides into the 87K and 60K polypeptides, respectively. The actual size of the protein detected, as calculated from its electrophoretic mobility, is approximately 28K, which is in good agreement with the theoretically expected size.

Characterisation of the cleavage products

Although the electrophoretic mobility of the four largest polypeptides generated by the secondary cleavages provides good evidence that they were similar to the

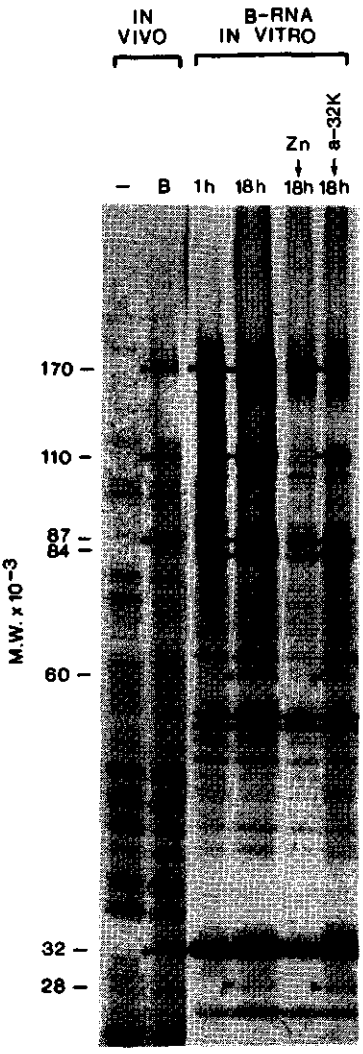


Fig. 2. Proteolytic processing of the 170K polypeptide in the reticulocyte lysate system. Lane 1h, the products obtained 1 h after addition of B-RNA; lane 18h, the products obtained 18 h after addition of B-RNA. Lanes 18h-Zn and 18h-α32K, idem, but 2 mM $ZnCl_2$ or anti-32K serum (1 μl per 3 μl translation mixture) was added 1 h after start of protein synthesis. The products were analysed in a 12.5% SDS–polyacrylamide gel. The left two lanes contain the $30000 \times g$ supernatant fractions from [^{35}S]methionine-labeled cowpea protoplasts, either uninoculated (–) or inoculated with B-components (B). The viral 84K, 60K and 58K polypeptides are not visible in lane B since these polypeptides are mainly membrane-bound, and almost exclusively found in the $30000 \times g$ pellet fraction of infected protoplasts (see for subcellular fractionation of cowpea protoplasts Rezelman et al., 1980).

110K, 87K, 84K and 60K polypeptides found in vivo, the identity of some polypeptides was further established by peptide mapping. For this purpose, the in vitro generated 170K, 110K, 84K and 60K polypeptides were sliced from an unfixed gel and partially digested with *Staphylococcus aureus* V8 protease. The resulting peptides were electrophoresed in a 20% polyacrylamide gel and compared with those

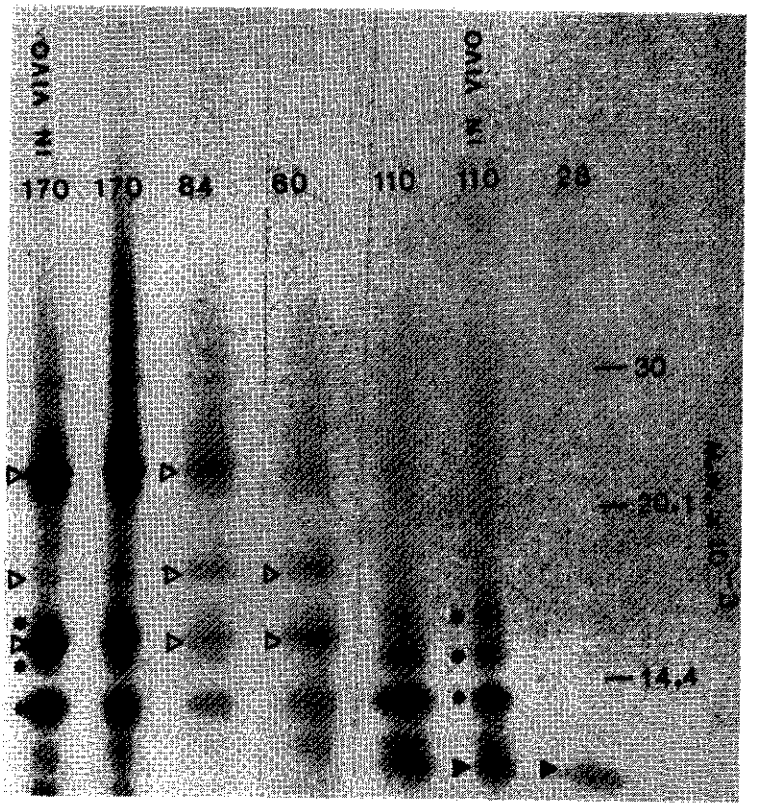


Fig. 3. Proteolytic peptide patterns of the in vitro cleavage products. The 170K, 110K, 84K, 60K and 28K in vitro cleavage products obtained after prolonged incubation of a B-RNA-directed translation mixture were sliced from an unfixed polyacrylamide gel and digested with *Staphylococcus aureus* V8 protease (for details see Materials and Methods). The resulting peptides were compared with the protease V8-generated peptides from the 170K and 110K polypeptides isolated from CPMV-infected protoplasts (lanes in vivo), using a 20% SDS–polyacrylamide gel. *, common peptides in the patterns of the 170K and 110K polypeptides; ▷, common peptides in the patterns of the 170K, 84K and 60K polypeptides; ▶, peptide specific for the 110K polypeptide. The molecular weight markers used were carbonate dehydratase (30000), soybean trypsin inhibitor (20100) and α-lactalbumin (14400).

of the 170K and 110K polypeptides isolated from infected protoplasts (Fig. 3). The peptide patterns of the in vitro produced 170K and 110K polypeptides were similar to those of their equal-sized counterparts in vivo (Fig. 3, lanes indicated in vivo). The peptide patterns of the 110K and 60K in vitro polypeptides differed, indicating that these cleavage products had unique rather than overlapping sequences. From these data it can be concluded that the 170K polypeptide was correctly cleaved into 110K and 60K polypeptides, just as happens in vivo.

Furthermore, the digests of the 60K and 84K in vitro cleavage products were very similar (Fig. 3), in full agreement with previous findings (Rezelman et al., 1980; Goldbach et al., 1982) that both polypeptides are derived from the same part of the 170K precursor.

Finally, Fig. 3 includes the proteolytic peptide 'pattern' of the in vitro generated 28K polypeptide. The digest of this polypeptide contained only one radiolabeled (methionine-containing) peptide of approximately 12K, which comigrated with a peptide unique for the digest of the 110K polypeptide (Fig. 3, peptide indicated with ▶). Since this latter peptide is absent from the proteolytic digests of both the 170K polypeptide (Fig. 3) and 87K polypeptide (not shown here, but see Fig. 3 in Rezelman et al., 1980), it must contain that part of the 110K polypeptide from which the proposed '24K' polypeptide would be derived. Therefore, the observed equal-sized V8 peptide obtained from the 28K polypeptide generated in vitro supports the idea that this cleavage product represents this postulated polypeptide.

The proteolytic activity involved in the in vitro cleavages

For preliminary characterisation of the protease involved in the cleavage of the 170K polypeptide the effect of zinc ions was examined. Addition of 2 mM $ZnCl_2$ to a translation mixture, 1 h after the start of translation, did block the 60K/110K cleavage completely and the 84K/87K cleavage partially (Fig. 2, lane indicated Zn). Although this inhibitory effect of Zn^{2+} suggests that the responsible protease is of the thiol-type, further experiments are needed to verify this. The 32K polypeptide has recently been identified as a protease, providing for the cleavage of the 95K and 105K proteins from M-RNA (Franssen et al., 1983). Since its activity is also sensitive to zinc ions (Pelham, 1979; Franssen et al., 1982) it might appear tempting to hold this protease to be responsible for the further cleavage of the 170K polypeptide. However, antiserum against the 32K polypeptide, which has been shown to neutralize the activity of this protease (Franssen et al., 1984), did not at all interfere with the processing of the 170K polypeptide (Fig. 2, lane α-32K). Thus, a different protease seems to be involved.

Discussion

In previous reports it has been shown that upon translation of CPMV B-RNA in a reticulocyte lysate the 200K primary translation product undergoes proper primary cleavage, resulting in the production of 170K and 32K polypeptides as found in vivo (Pelham, 1979; Rezelman et al., 1980; Goldbach et al., 1981). The experiments

described here demonstrate that upon prolonged incubation of the translation mixture proteolysis proceeds, the 170K polypeptide being cleaved into 110K, 87K, 84K and 60K polypeptides identical to the B-RNA-encoded proteins found in vivo. Moreover, a fifth cleavage product has been detected, whose size and peptide pattern strongly suggest that it represents the 'hypothetical' 24K polypeptide proposed by Rezelman et al. (1980). Thus, the present in vitro studies not only confirm the translational map of B-RNA as deduced from in vivo (cowpea protoplasts) analyses, but provide the first evidence that the 110K and 84K polypeptides are additionally cleaved to give the 87K and 60K polypeptides, respectively, together with a 28K polypeptide. As a consequence the 110K and 84K should be regarded as processing intermediates (although with long half-lives, both in vivo and in vitro). Recently the 32K polypeptide has been identified as the B-RNA-encoded protease involved in the cleavage of the M-RNA-encoded 95K and 105K primary translation products (Franssen et al., 1983). Our results indicate that this enzyme is not involved in the processing of the B-RNA-encoded primary product, neither in the primary cleavage (by which the 32K protease is released), nor in any of the other cleavage steps:

(1) The 32K polypeptide is not released from nascent 200K chains before these chains have reached a length of at least 110K (Fig. 1). If the 32K polypeptide would release itself autocatalytically it is hard to imagine why this does not occur as soon as its sequence has been completed.

(2) Antiserum against the 32K protease, which had previously been shown to neutralise the activity of this enzyme, does not inhibit the secondary cleavage of the 170K polypeptide (Fig. 2).

(3) The 32K protease recognizes a glutamine–methionine cleavage site within the M-RNA-encoded polypeptides (thereby releasing the 60K precursor to both capsid proteins) (van Wezenbeek et al., 1983). Inspection of the nucleotide sequence of B-RNA and of the amino acid sequence of the 200K primary translation product deduced from it (Lomonossoff and Shanks, 1983) reveals that glutamine-methionine dipeptide sequences are absent at or around the expected cleavage sites.

Therefore, a different protease seems to be involved in the processing of the B-RNA-encoded 200K polypeptide. Since all processing steps occur properly in the rabbit reticulocyte lysate (i.e. in the absence of any plant proteins) this protease must be encoded by B-RNA itself. The information available to date indicates that the 28K polypeptide might represent this protease. This would explain why the primary cleavage only takes place on nascent chains longer than approximately 110K. As shown in the translation model of Fig. 4, nascent chains shorter than 120K (32K + 60K + 28K) do not contain a complete 28K polypeptide sequence. Furthermore, within the 170K polypeptide the 28K polypeptide is located between the 60K and 87K polypeptide sequences. Therefore, if this polypeptide would be capable of cleaving itself out of the 170K polypeptide, first left and then right, or the other way around, then the simultaneous formation of the 110K, 87K, 84K and 60K polypeptides is readily explained (see Fig. 4). Experiments to verify this hypothesis are in progress.

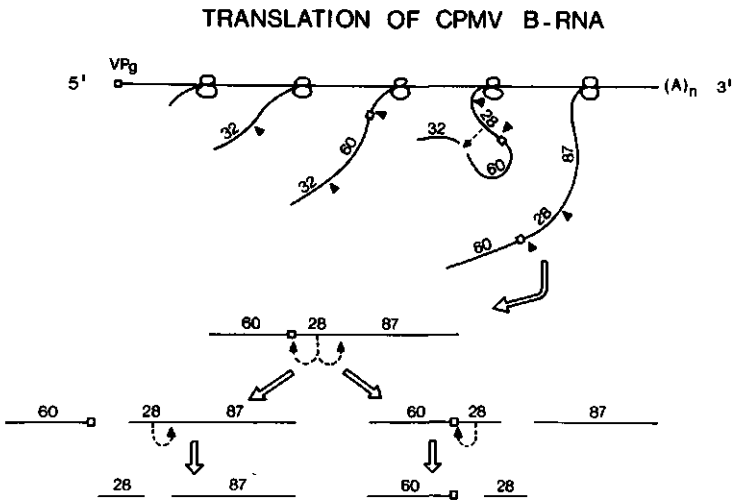


Fig. 4. Model for the proteolytic processing of the B-RNA-encoded primary translation product (cleavage sites are indicated with ▲). The primary cleavage occurs predominantly on nascent 200K chains, which should be longer than approximately 120K. The protease involved in this and following cleavages is proposed to be the 28K polypeptide. After completion of the polypeptide chain the 28K polypeptide releases itself by cleaving the 170K polypeptide chain, first left and then right, or the other way around. The cleavage step by which VPg (indicated with □) is released from its 60K precursor (Goldbach et al., 1982) is not included in the scheme. It is not known whether VPg is present on RNA chains which are used as messenger.

Acknowledgements

The authors thank Geertje Rezelman and Jan Verver for technical assistance, Richard Jackson for the gift of rabbit reticulocyte lysates, Piet Madern for art work and printing the photographs, Marie-José van Neerven for typing the manuscript, and Dr. D. Shih for sending us a manuscript dealing with the same subject prior to publication.

This work was supported by the Netherlands Foundation for Chemical Research, with financial aid from the Netherlands Foundation for the Advancement of Pure Research.

References

Ahlquist, P. and Kaesberg, P. (1979) Determination of the length distribution of poly(A) at the 3'-terminus of the virion RNAs of the EMC virus, poliovirus, rhinovirus RAV-61 and CPMV and of mouse globin mRNA. Nucleic Acids Res. 7, 1195–1204.

Daubert, S.D., Bruening, G. and Najarian, R.C. (1978) Protein bound to the genome RNAs of cowpea mosaic virus. Eur. J. Biochem. 92, 45–51.

Davies, J.W., Verver, J., Goldbach, R.W. and van Kammen, A. (1978) Efficient reverse transcription of cowpea mosaic virus RNAs. Nucleic Acids Res. 5, 4643–4661.

El Manna, M. and Bruening, G. (1973) Polyadenylate sequences in the ribonucleic acids of cowpea mosaic virus. Virology 56, 198–206.

Franssen, H., Goldbach, R., Broekhuijsen, M., Moerman, M. and van Kammen, A. (1982) Expression of middle-component RNA of cowpea mosaic virus: in vitro generation of a precursor to both capsid proteins by a bottom-component RNA-encoded protease from infected cells. J. Virol. 41, 8–17.

Franssen, H.J., Moerman, M., Rezelman, G. and Goldbach, R. (1984) Evidence that the 32,000-dalton protein encoded by the bottom-component RNA of cowpea mosaic virus is a proteolytic processing enzyme. J. Virol., in press.

Goldbach, R. and Rezelman, G. (1983) Orientation of the cleavage map of the 200-kilodalton polypeptide encoded by the bottom-component RNA of cowpea mosaic virus. J. Virol. 46, 614–619.

Goldbach, R.W., Schilthuis, J.G. and Rezelman, G. (1981) Comparison of in vivo and in vitro translation of cowpea mosaic virus RNAs. Biochem. Biophys. Res. Commun. 99, 89–94.

Goldbach, R., Rezelman, G., Zabel, P. and van Kammen, A. (1982) Expression of the bottom-component RNA of cowpea mosaic virus: evidence that the 60-kilodalton VPg precursor is cleaved into single VPg and a 58-kilodalton polypeptide. J. Virol. 42, 630–635.

Klootwijk, J., Klein, I., Zabel, P. and van Kammen, A. (1977) Cowpea mosaic virus RNAs have neither $m^7GpppN\ldots$ nor mono-, di-, or triphosphates at their 5'-ends. Cell 11, 73–82.

Koch, G., Koch, F., Bilello, J.A., Hiller, E., Schârli, C., Warnecke, G. and Weber, C. (1981) Biosynthesis, modification and processing of viral polyproteins. In: Protein Biosynthesis in Eukaryotes (Perez-Bercoff, R., ed.), pp. 275–310. Plenum Press, New York.

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) 227, 680–685.

Lomonossoff, G. and Shanks, M. (1983) The nucleotide sequence of cowpea mosaic virus B-RNA. EMBO J. 2, 2253–2258.

Palmenberg, A.C. (1982) In vitro synthesis and assembly of picornaviral capsid intermediate structures. J. Virol. 44, 900–906.

Pelham, H.R.B. (1978) Translation of encephalomyocarditis virus RNA in vitro yields an active proteolytic processing enzyme. Eur. J. Biochem. 85, 457–462.

Pelham, H.R.B. (1979) Synthesis and proteolytic processing of cowpea mosaic virus proteins in reticulocyte lysates. Virology 96, 463–477.

Pelham, H.R.B. and Jackson, R.J. (1976) An efficient mRNA-dependent translation system from reticulocyte lysates. Eur. J. Biochem. 67, 247–256.

Putnak, J.R. and Phillips, B.A. (1981) Picornaviral structure and assembly. Microbiol. Rev. 45, 287–315.

Rezelman, G., Goldbach, R. and van Kammen, A. (1980) Expression of bottom component RNA of cowpea mosaic virus in cowpea protoplasts. J. Virol. 36, 366–373.

Rueckert, R.R., Matthews, T.J., Kew, O.M., Pallansch, M., McLean, C. and Omilianowski, D. (1979) Synthesis and processing of picornaviral polyprotein. In: The Molecular Biology of Picornaviruses (Perez-Bercoff, R., ed.), pp. 113–126. Plenum Press, New York.

Shih, D.S., Shih, C.T., Zimmern, D., Rueckert, R.R. and Kaesberg, P. (1979) Translation of encephalomyocarditis virus RNA in reticulocyte lysates: kinetic analysis of the formation of virion proteins and a protein required for processing. J. Virol. 30, 472–480.

Stanley, J., Rottier, P., Davies, J.W., Zabel, P. and van Kammen, A. (1978) A protein linked to the 5' termini of both RNA components of the cowpea mosaic virus genome. Nucleic Acids Res. 5, 4505–4522.

Van Kammen, A. (1967) Purification and properties of the components of cowpea mosaic virus. Virology 31, 633–642.

Van Wezenbeek, P., Verver, J., Harmsen, J., Vos, P. and van Kammen, A. (1983) Primary structure and gene organization of the middle component RNA of cowpea mosaic virus. EMBO J. 2, 941–946.

Zabel, P., Moerman, M., van Straaten, F., Goldbach, R. and van Kammen, A. (1982) Antibodies against the genome-linked protein VPg of cowpea mosaic virus recognize a 60,000-dalton precursor polypeptide. J. Virol. 41, 1083–1088.

(Manuscript received 17 October 1983)

CHAPTER VIII

HOMOLOGOUS SEQUENCES IN NON-STRUCTURAL PROTEINS FROM COWPEA MOSAIC VIRUS AND PICORNAVIRUSES

Henk Franssen[1], Jack Leunissen[2], Rob Goldbach[1], George Lomonossoff[3], and David Zimmern[4].

1: Department of Molecular Biology, Agricultural University, De Dreijen 11, 6703 BC WAGENINGEN, The Netherlands;
2: Department of Biochemistry, University of Nijmegen, Geert Grooteplein Noord 21, 6526 EZ NIJMEGEN, The Netherlands.
3: Department of Virus Research, John Innes Institute, Colney Lane, NORWICH NR4 7UH, England;
4. Medical Research Council Laboratory of Molecular Biology, Hills Road, CAMBRIDGE, CB2 2QH, England.

EMBO J., 3, 855-861 (1984).

The EMBO Journal vol.3 no.4 pp.855 – 861, 1984

Homologous sequences in non-structural proteins from cowpea mosaic virus and picornaviruses

Henk Franssen, Jack Leunissen[1], Rob Goldbach*, George Lomonossoff[2] and David Zimmern[3]*

Department of Molecular Biology, Agricultural University, De Dreijen 11, 6703 BC Wageningen, [1]Department of Biochemistry, University of Nijmegen, Geert Grooteplein Noord 21, 6526 EZ Nijmegen, The Netherlands, [2]Department of Virus Research, John Innes Institute, Colney Lane, Norwich NR4 7UH, and [3]Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

*To whom reprint requests should be sent

Communicated by A. Klug

Computer analyses have revealed sequence homology between two non-structural proteins encoded by cowpea mosaic virus (CPMV), and corresponding proteins encoded by two picornaviruses, poliovirus and foot-and-mouth disease virus. A region of 535 amino acids in the 87-K polypeptide from CPMV was found to be homologous to the RNA-dependent RNA polymerases from both picornaviruses, the best matches being found where the picornaviral proteins most resemble each other. Additionally, the 58-K polypeptide from CPMV and polypeptide P2-X from poliovirus contain a conserved region of 143 amino acids. Based on the homology observed, a genetic map of the CPMV genome has been constructed in which the 87-K polypeptide represents the core polymerase domain of the CPMV replicase. These results have implications for the evolution of RNA viruses, and mechanisms are discussed which may explain the existence of homology between picornaviruses (animal viruses with single genomic RNAs) and comoviruses (plant viruses with two genomic RNAs).

Key words: cowpea mosaic virus/picornavirus/RNA-dependent RNA polymerase/protein sequence homology/evolution

Introduction

The comoviruses are a distinctive group of plant viruses which resemble the picornaviruses of animals in several respects. Members of both groups have positive-stranded RNA genomes. Their RNAs carry a small protein, called VPg, at their 5′ termini and a poly(A) tail at their 3′ termini (for review of picornaviruses, see Rueckert *et al.*, 1979; Putnak and Phillips, 1981; Koch *et al.*, 1981; for comoviruses, see Bruening, 1977, 1981). Moreover the RNAs of both como- and picornaviruses are translated into polyproteins from which the functional proteins are derived by proteolytic cleavages. The major difference between the two groups is that picornaviruses have a single RNA while comoviruses divide their genome between two RNAs. The larger RNA of comoviruses (B-RNA) can replicate on its own in protoplasts and so must encode any functions required for replication (Goldbach *et al.*, 1980), while the smaller RNA (M-RNA) encodes the structural proteins (Franssen *et al.*, 1982). In addition, although both groups of viruses have icosahedral capsids, these differ significantly in structure.

Recently, the complete nucleotide sequence of the RNAs from cowpea mosaic virus (CPMV), typemember of the comovirus group, has been elucidated (Van Wezenbeek *et al.*, 1983; Lomonossoff and Shanks, 1983). Both B- and M-RNA (5889 and 3481 nucleotides long, excluding the poly(A) tails) contain a long, unique open reading frame, the length of which is in good agreement with the size of the primary translation products reported. Translation studies (see, for instance, Pelham, 1979; Rezelman *et al.*, 1980; Goldbach *et al.*, 1982; Goldbach and Rezelman, 1983) have revealed that B-RNA is translated into a polyprotein of ~200 K from which five final cleavage products are derived in the following order: NH_2-32 K-58 K-4 K(=VPg)-28 K-87 K-COOH. Several processing intermediates have also been detected in the infected cell (Rezelman *et al.*, 1980) which may, in view of the large amounts in which they occur, represent functional molecules. M-RNA contains two translation initiation sites (Pelham, 1979; Franssen *et al.*, 1982) and is translated into two C co-terminal polyproteins of ~105 K and 95 K, each containing the sequences of both capsid proteins VP37 and VP23 (Franssen *et al.*, 1982). They are cleaved to give products in the order NH_2-58 K/48 K-VP37-VP23-COOH.

Of the products specified by B-RNA, the 32-K polypeptide has been shown to be the protease involved in the cleavage of the overlapping polyproteins from M-RNA, thereby generating a 60-K precursor to both capsid proteins (Franssen *et al.*, 1982, 1984a). The cleavage site used by this enzyme is a glutamine-methionine dipeptide sequence (Van Wezenbeek *et al.*, 1983). The 28-K polypeptide has been proposed to represent the protease involved in the processing of the 200-K polyprotein from B-RNA itself (Franssen *et al.*, 1984b). The cleavage sites used in this polyprotein are not glutamine-methionine sequences (Franssen *et al.*, 1984b) but probably glutamine-serine sequences (Zabel *et al.*, 1984). The B-RNA-encoded 110-K polypeptide is found in extensively purified CPMV replication complexes, which are able to elongate nascent RNA chains to full-length RNA molecules *in vitro* (Dorssers *et al.*, 1984). Since this polypeptide contains the sequences of both the 28-K (the proposed protease) and the 87-K polypeptide (Rezelman *et al.*, 1980; Franssen *et al.*, 1984b) the functional organisation within the C-terminal parts of the B-RNA-coded polyprotein and the picornaviral polyprotein might be co-linear: -VPg-protease-RNA-dependent RNA polymerase. In view of the apparent similarity between the como- and picornaviruses we thought it might be instructive to search for amino acid sequence homology between both the structural and non-structural regions of the CPMV polyproteins and the polyproteins from two picornaviruses, poliovirus and foot-and-mouth disease virus (FMDV). The results of this search are striking and reveal the existence of significant sequence homology between the 87-K polypeptide from CPMV and the polymerase from both FMDV and poliovirus. In addition, the 58-K polypeptide specified by CPMV B-RNA exhibits sequence homology to polypeptide P2-X from poliovirus. The results presented here

Table 1. Presence of homologous amino acid sequences in the proteins from CPMV, poliovirus and FMDV

		CPMV polyproteins	
		M (105 K)	B (200 K)
Poliovirus	P1-region	–	–
	P2-region	–	+
	P3-region	–	+
FMDV	capsid proteins	–	–
	polymerase	–	+

Data were obtained from graphic matrix comparisons as described in Materials and methods. +, homologous sequences; –, no homologous sequences.

confirm and extend the biochemical data by demonstrating that the 87-K domain of the 110-K polypeptide, identified as a component of the CPMV replication complex, is structurally homologous to the picornavirus replicase subunit P3-4b. They thus throw a new light on the evolution of viral RNA genomes.

Results

Comparison of the polypeptides from CPMV, poliovirus and FMDV

The complete nucleotide sequence of both the Mahony and Sabin strains of poliovirus type 1 has been determined (Kitamura *et al.*, 1981; Racaniello and Baltimore, 1981; Nomoto *et al.*, 1982). Together with amino acid sequence analyses (Semler *et al.*, 1981a, 1981b; Kitamura *et al.*, 1981; Larsen *et al.*, 1982; Emini *et al.*, 1982) this has enabled the precise mapping of the viral polypeptides within the single poliovirus polyprotein (NCVPOO mol. wt. ~250 K) and the identification of the cleavage sites used to release them. Based on the primary cleavages NCVPOO has been divided into three regions: P1, from which the four capsid proteins are derived, P2, the central region, and P3, the region from which VPg, the viral protease (P3-7c) and the RNA-dependent RNA polymerase (P3-4b) are derived (Kitamura *et al.*, 1981). The nucleotide sequence of two parts of the genome of a second picornavirus FMDV has been determined, corresponding to the coding regions of the capsid proteins and the viral polymerase (Boothroyd *et al.*, 1982; Robertson *et al.*, 1983). Comparison of the amino acid sequences of the poliovirus and FMDV RNA polymerases showed that they were ~30% homologous (Robertson *et al.*, 1983). In view of the apparent conservation of amino acid sequences among picornaviral polymerases and the similarities in genomic structure and expression between comoviruses and picornaviruses, we examined whether (any of) the CPMV polypeptides were related to the polymerase or to any of the other polypeptides specified by picornaviruses. For this purpose, the amino acid sequences of the polyproteins from CPMV, poliovirus and FMDV were first compared using graphic matrix procedures (McLachlan, 1971; Maizel and Lenk, 1981; Staden, 1982; see Materials and methods).

Table 1 summarizes the comparisons made and the homology found. The overlapping 95-K and 105-K polyproteins encoded by the M-RNA of CPMV appeared not to contain any amino acid sequence homologous to the structural or non-structural proteins from poliovirus and FMDV. On the other hand, the 200-K polyprotein specified by the B-RNA

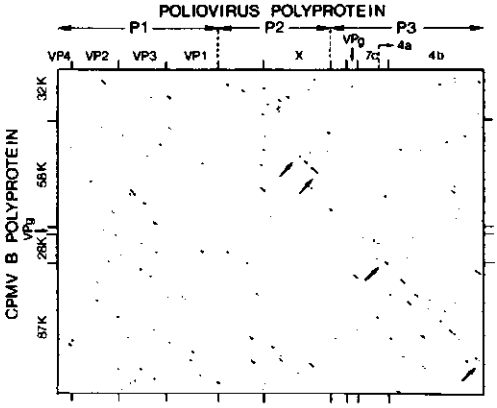


Fig. 1. Matrix displaying well matched blocks of amino acid sequence 31 residues in length from the polyproteins of poliovirus Sabin-1 strain and of CPMV B-RNA using the computer programme DIAGON. All possible pairings of 31 residue blocks from the two proteins are compared by the programme, and all those matching with a score of 340 or greater are recorded by a dot on the display at the coordinates of the midpoint of the block. A score of ≥340 would be observed in comparisons of 31 residue blocks drawn at random from a pool of amino acids of the same compositions as the real proteins with a 'double matching probability' (McLachlan, 1971) of 1.1×10^{-4}. This probability level should be sufficient to filter out most random matches (McLachlan, 1971), while the score, which corresponds to four standard deviations (S.D. = 14.12) above the mean (283), is above the three S.D.s thought to be a minimum criterion of significant relationship (Doolittle, 1981). Cleavage sites in the proteins are indicated on both axes, regions of marked homology by arrows. Sequence data are from Nomoto *et al.* (1982) and Lomonossoff *et al.* (1983).

contained regions with homology significant at the 10^{-4} level to both the P2 and P3 region of the poliovirus polyprotein and to the polymerase sequence of FMDV (Table 1). A matrix graph (Figure 1) comparing the entire polyproteins encoded by CPMV B-RNA and poliovirus RNA at this significance level revealed two extended regions of apparent homology (marked by arrows). These regions map within the P2-X and P3-4b genes of poliovirus, which encode a protein of unknown function, and the putative RNA polymerase, respectively. In CPMV the corresponding proteins, deduced from the detailed cleavage map (Goldbach *et al.*, 1982; Goldbach and Rezelman, 1983) and the precise mapping of the VPg sequence (Zabel *et al.*, 1984) are the 58-K polypeptide and the 87-K polypeptide, respectively (Figure 1). Since no sequence data are yet available from the central part of the FMDV polyprotein, neither a total comparison nor a detailed comparison with the polio P2-X and CPMV 58-K proteins could be done. The sequence of the putative FMDV replicase p56a was available, however, and pairwise comparisons with the polio P3-4b and CPMV 87-K protein sequences confirmed the known homology between polio and FMDV and showed that the CPMV 87-K protein was related to both replicases, although more distantly (Figure 2).

Alignment of the 87-K product from CPMV and the polymerases from poliovirus and FMDV

The homology between the B-RNA-encoded 87-K polypeptide from CPMV and the RNA-dependent RNA polymerases from FMDV and poliovirus has been further analysed by

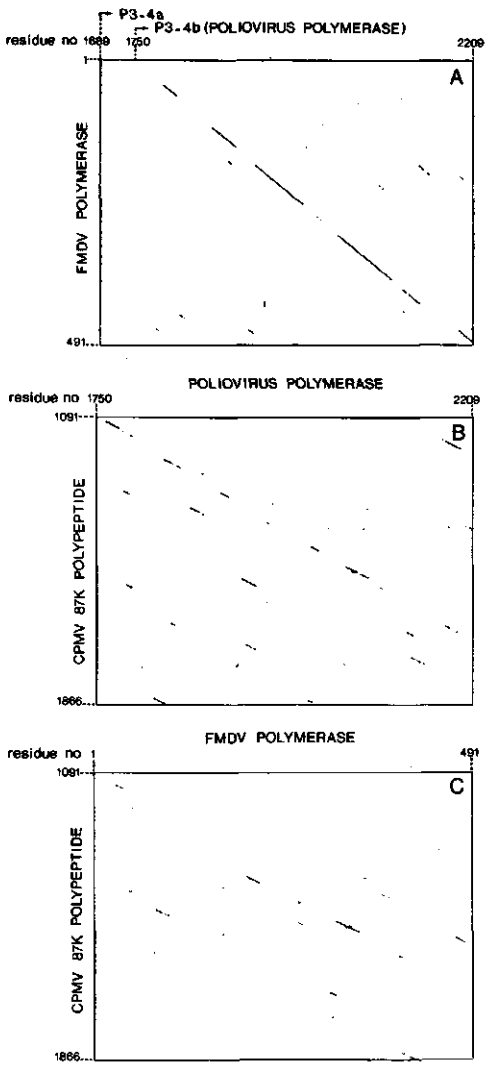


Fig. 2. DIAGON-generated matrices comparing the CPMV 87-K polypeptide and picornaviral polymerases. Block length compared was 31 (as in Figure 1), but the threshold scores used were 337 (corresponding to a probability of 3×10^{-9}). (**A**) Comparison of the two polymerases from poliovirus and FMDV. (**B**) Comparison of the CPMV 87-K polypeptide and polioviral polymerase. (**C**) Comparison of the CPMV 87-K polypeptide and FMDV polymerase. The sequence data from FMDV were from Robertson *et al.* (1983).

alignment of their amino acid sequences. Figure 3 shows the aligned sequences and illustrates the existence of homology over a region comprising the whole picornaviral polymerase sequence. The 87-K product from CPMV and the polioviral polymerase P3-4b are 20.9% homologous (113 identical residues out of a total of 541 positions, including the gaps). Likewise, the homology between the aligned sequences of the FMDV enzyme and the 87-K polypeptide was estimated as 20.9% (over 545 positions). The homology between the CPMV polyprotein and the P3-region of the polioviral polyprotein appears not to be limited to the polio polymerase sequence (P3-4b), but extends at the amino-terminal side beyond the glutamine-glycine cleavage site into the sequence of the protease (P3-7c) (Figure 3). Strikingly, the marked homology stops abruptly near position 1688 within the P3-7c region, i.e., at the glutamine-glycine cleavage site by which polypeptide P3-4a, a presumed precursor to the polymerase, is generated (Etchison and Ehrenfeld, 1980; Semler *et al.*, 1983).

It should be noted that the glutamine-glycine dipeptide cleavage sites used to release polypeptides P3-4a and P3-4b from the polioviral polyprotein exactly match with glutamine-serine and glutamine-glycine dipeptide sequences within the CPMV protein sequence, respectively (Figure 3). Whether one of these dipeptide sequences is used to release the 87-K polypeptide remains to be verified. Moreover, the carboxy-terminal residue of P3-4b (phenylalanine at position 2211) again matches with a glutamine-glycine sequence in the CPMV protein. This dipeptide sequence is unlikely to be used as a cleavage site. Cleavage at this position would generate polypeptides of −67 K and 20 K from the 87-K product, which have never been observed.

Alignment of the 58-K polypeptide from CPMV and product P2-X from poliovirus

As can be deduced from the matrix graph of Figure 1, only the central regions of the CPMV 58-K polypeptide and P2-X from poliovirus exhibit clear homology, but this homology is the most marked in the entire polyproteins (significant at the 10^{-5} level). Figure 4 shows the alignment of the amino acid sequences in these regions. Over a length of 143 residues the sequences can be matched to an homology of 30% with only five short gaps. The alignment contains two shorter regions (residue positions 521−545 and 572−598 in the CPMV sequence) exhibiting even greater homology (44% and 57%). On the other hand, since the similarity is limited to a region of only 143 residues, the overall homology between the 58-K CPMV product (553 amino acids) and P2-X (329 amino acids) is substantially lower than that observed between the 87-K polypeptide and P3-4b.

Genetic map of the CPMV RNA genome

Its marked similarity to the RNA-dependent RNA polymerases from two different picornaviruses strongly suggests that the 87-K polypeptide represents a 'core polymerase' domain of the CPMV replicase, functionally homologous to the picornavirus core polymerase (Flanegan and Baltimore, 1979; Baron and Baltimore, 1982; Semler *et al.*, 1983). The 28-K polypeptide, adjoining this polypeptide in the CPMV B polyprotein, has been proposed to represent a protease (Franssen *et al.*, 1984b), while we have found sequence homology between the CPMV 58-K polypeptide and P2-X from poliovirus. All these results are consistent with the idea that the genetic organisation of the 200-K polyprotein from CPMV is colinear with that of the carboxy-terminal half of the picornaviral polyprotein (Figure 5). B-RNA seems to encode one additional function, apparently the consequence of the divid-

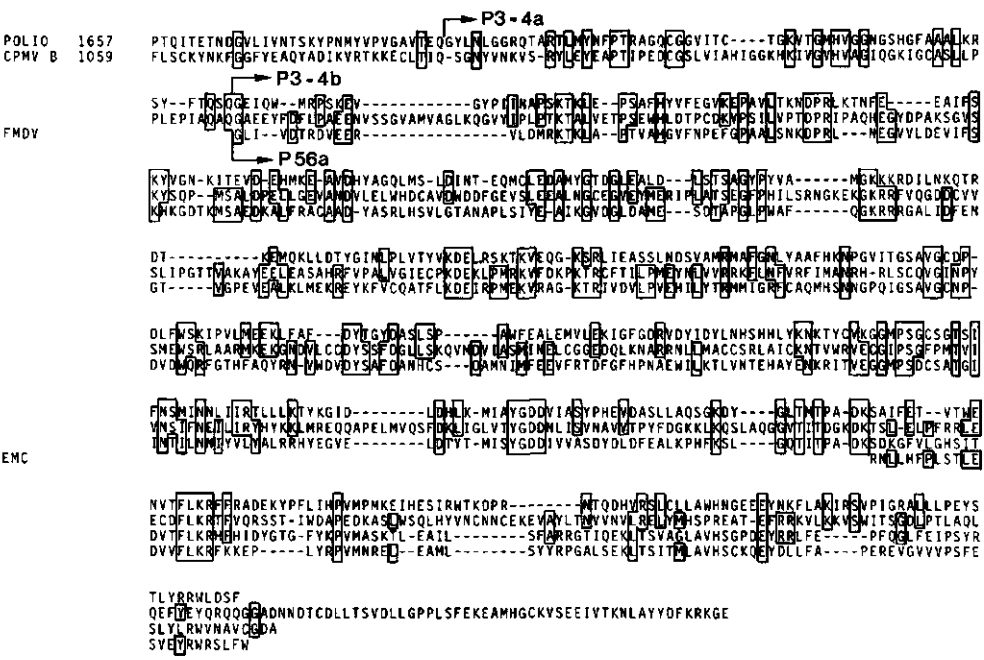


Fig. 3. Alignments of the 87-K polypeptide from CPMV and the polymerases from FMDV and poliovirus. The sequences compared start for the poliovirus and CPMV B polyproteins at positions 1657 and 1059, respectively. For FMDV only the sequence of the polymerase (P56a) has been shown (Robertson *et al.*, 1983), and for EMC the published (Drake *et al.*, 1982) sequence of the carboxy-terminal part of this enzyme. Alignment was based on the published FMDV-poliovirus alignment of Robertson *et al.* (1983) plus additional alignment using DIAGON and by inspection. Residues in the CPMV 87-K polypeptide which are identical to those in either picornaviral protein are boxed, and gaps are indicated with lines. The best matches are found where the sequences of both picornaviral proteins are most conserved. The picornaviral proteins display additional homologies to each other (not shown here, but see Robertson *et al.*, 1983 and Figure 2A).

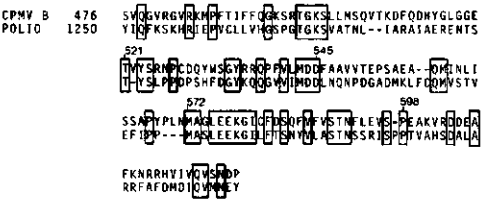


Fig. 4. Alignment of the 58-K polypeptide encoded by CPMV B-RNA and P2-X from poliovirus. The regions compared start at position 1250 in the poliovirus sequence and at position 476 in the CPMV sequence.

ed genome of CPMV: the 32-K protease involved in the processing of the polyproteins from M-RNA (Franssen *et al.*, 1984a).

Discussion

Functional implications

This study revealed sequence homology between two of the non-structural polypeptides from the plant virus CPMV and corresponding proteins from animal picornaviruses. The most extensive homology was found between the 87-K polypeptide specified by CPMV B-RNA and the picornaviral RNA-dependent RNA polymerase. Therefore, it is reasonable to assume that the 87-K polypeptide also has a function related to that of the RNA-dependent RNA polymerase. Indeed, Dorssers *et al.* (1984) have recently reported on the presence of the B-RNA-encoded 110-K polypeptide in purified CPMV replication complexes. This product contains the sequences of both the 87-K polypeptide and the 28-K polypeptide (Rezelman *et al.*, 1980; Franssen *et al.*, 1984b). However, no mature 87-K polypeptide was detected in such preparations. This may suggest that the (87-K) polymerase is only active in its 110-K 'precursor' form. Another possibility is that the presence or absence of the 28-K polypeptide sequence is irrelevant to the activity of the 87-K domain after initiation, but that the cleavage of the 110-K polypeptide into the 87-K and 28-K polypeptides proceeds slowly. Since the 28-K polypeptide is probably the protease involved in the processing of the 200-K polyprotein (Franssen *et al.*, 1984b) one or more proteolytic steps essential for RNA replication might be considered. This idea is implicit in the hypothesis that VPg acts as a primer in the RNA replication

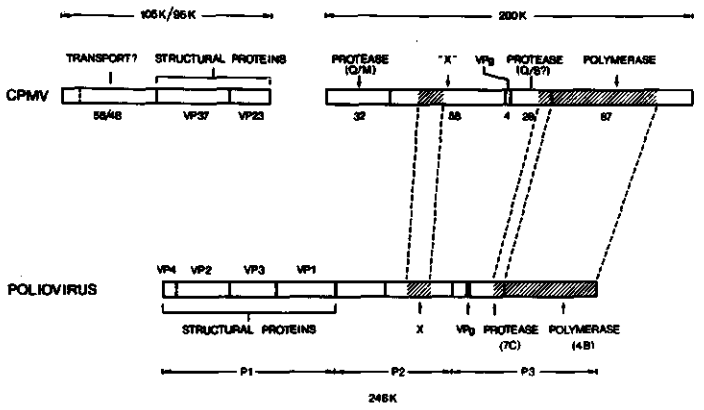


Fig. 5. Comparison of the functional organisation of the polyproteins from CPMV and poliovirus. M-RNA of CPMV specifies two overlapping polyproteins of sizes 95 K and 105 K, from which the capsid proteins VP37 and VP23 are derived, together with two polypeptides of 58 K and 48 K (Franssen *et al.*, 1982). The latter products have been proposed to play a role in viral RNA transport (Rezelman *et al.*, 1982). B-RNA of CPMV specifies a 200-K polyprotein, which has now been shown to contain two regions homologous to sequences in the poliovirus protein (shaded regions). Five different mature proteins are proposed to be derived from the B polyprotein: a 32-K protease, involved in the cleavage of the M-RNA-encoded polyproteins (Franssen *et al.*, 1984a); a 58-K protein, which – in view of the homology found – might be the counterpart of protein X from poliovirus; VPg, which may play a role in RNA replication; a 28-K protein, which has been proposed to represent a protease involved in the cleavage of the B protein (Franssen *et al.*, 1984b); and a 87-K protein which has now been proposed to be the core polymerase domain of the 110-K polypeptide, previously detected in CPMV replication complexes (Dorssers *et al.*, 1984). The precise position of the cleavage site between the 28-K and 87-K proteins has remained unknown so far. The VPg sequence has recently been mapped in the 200-K sequence (Zabel *et al.*, 1984).

cycle which is proteolytically released from its precursor at the very moment of initiation of an RNA chain (for review, see Wimmer, 1982), but it is not clear why a protease domain should be retained in an enzyme preparation active only in elongating pre-existing chains (Dorssers *et al.*, 1984).

It should be noted that the tentative assignment of protease activity to the 28-K polypeptide (Franssen *et al.*, 1984b) is now further supported by the observation that the region in the B polyprotein homologous to the picornaviral polymerase is preceded by a region homologous to the carboxy-terminal part of P3-7c, the protease from poliovirus (Figure 3). However, the precise cleavage site by which the 28-K and 87-K polypeptides are separated is not known and detailed amino acid sequence analyses are necessary to verify whether this homology resides in the 28-K product. If this is the case then a glutamine-glycine sequence (at position 1155) is used to generate the 87-K polypeptide and not a glutamine-serine sequence (at position 1091), as was proposed to be the preferred cleavage site in the B polyprotein (Zabel *et al.*, 1984). Alternatively, both sites may be used as cleavage points which would further enhance the analogy with poliovirus.

A second region of homology has been mapped within the 58-K polypeptide from CPMV and P2-X from poliovirus, respectively (Figure 4). The function of P2-X has not been biochemically defined, but it is interesting to note that in Cooper's genetic map of poliovirus (Cooper, 1968) mutations affecting double-stranded RNA synthesis occur near the middle of the map. While the co-linearity of the genetic and physical maps has not been proven, it is possible that the 58-K polypeptide from CPMV and P2-X are also proteins involved in RNA replication, and that the X-VPg-protease-core polymerase genes represent a module of RNA replicating functions that have retained a similar organisation in both CPMV and picornaviruses (Figure 5).

Evolutionary implications

Our findings, together with the recent observation that a product encoded by the plant DNA virus, cauliflower mosaic virus, exhibits homology to retroviral reverse transcriptase (Toh *et al.*, 1983), and the discovery that three different plant RNA viruses (AlMV, BMV and TMV) encode proteins homologous to polypeptide ns72 from Sindbis virus (Haseloff *et al.*, unpublished data) throw a new light on the evolution of viruses. How can viruses infecting organisms from two different eukaryotic kingdoms encode polypeptides with similar sequences? And, in the case of the comovirus and picornavirus groups, how can the structural and genetic organisation of their genomes be similar? At least three possibilities may be considered to explain the similarities between comoviruses and picornaviruses. (i) Convergent evolution. Different viruses that originated independently may encode proteins interacting with highly conserved host machinery and solely for this reason may evolve related tertiary and hence primary structures. (ii) Common viral ancestry. Comoviruses and picornaviruses may both derive from a common viral ancestor. If such a common ancestor arose before or at the time of divergence of the plant and animal kingdoms the residual conservation of structure must reflect extreme selective pressures, considering the rapid rate of mutation in RNA genomes (Domingo *et al.*, 1978; Holland *et al.*, 1982). A common ancestor need not have pre-dated the evolutionary separation of plant and animal cells however, since viruses such as wound tumour virus (a plant reovirus) and potato yellow dwarf virus (a plant rhabdovirus) can replicate both in plants and in the cells of their insect vectors (for review, see Matthews, 1981) (it should be noted that comoviruses have beetle vectors). A virus ancestral to both como- and picornaviruses, which could replicate in both kinds of host, may therefore have existed relatively recently, subsequently giving

rise to more specialised descendants. (iii) Transduction of conserved host genes. Both comoviruses and picornaviruses may have (independently) transduced equivalent, conserved host genes to apply them for their own multiplication. In this context it is worthwhile mentioning that RNA-dependent RNA polymerases – the functions of which are so far unknown – have been reported from a number of plant species, among them cowpea (*Vigna unguiculata*) (Dorssers *et al.*, 1982; Van der Meer *et al.*, 1983).

Both the second and third possibilities seem difficult to imagine without invoking some form of recombination, possibly at the RNA level. Given that the comovirus B-RNA lacks genes corresponding to the picornavirus capsid proteins, the minimum rearrangement necessary would be a deletion of the capsid protein genes from a picornaviral-like RNA, perhaps in the manner known to be a preferred pathway of poliovirus DI RNA formation (Cole and Baltimore, 1973) followed by the acquisition of the entire M-RNA from another source. This scheme still leaves unclear the source of the 32-K protease, a B-RNA gene required for M-RNA polyprotein maturation. Most other schemes one might imagine would necessitate a more general form of recombination than straightforward internal deletion. It is particularly interesting in this regard that there is longstanding genetic evidence for recombination in picornaviruses (Cooper, 1977) which has more recently received additional biochemical support (King *et al.*, 1982a, 1982b; Romanova *et al.*, 1980; Tolskaya *et al.*, 1983).

It is not yet possible to judge which one, or which combination of these mechanisms is most likely but we anticipate that further experimental work may indicate where the balance of probability lies. Even though evolutionary debates are necessarily conjectural to a degree, we suspect that more insight into this problem may shed light on questions beyond the immediate confines of RNA viral evolution.

Materials and methods

Nucleotide sequences of viral RNAs

The amino acid sequences of the polypeptides compared were derived from the following RNA sequences: CPMV M-RNA: Van Wezenbeek *et al.*, 1983; CPMV B-RNA: Lomonossoff and Shanks, 1983; poliovirus RNA (Sabin 1 strain): Nomoto *et al.*, 1982, retrieved from the EMBL-database; FMDV RNA: Boothroyd *et al.*, 1982; Robertson *et al.*, 1983.

Matrix comparisons and sequence alignments

Initial searches for amino acid sequence homology were made using graphic matrix methods. Two independent searches were made. In one search, using a Fortran 77 programme (unpublished) basically as described by Gibbs and McIntyre (1970) run on a Hitachi AS 9000 computer, all matches of three or more out of five amino acids were scored in all possible pairwise alignments of five amino acid blocks from the proteins being compared. In another search, using the programme DIAGON (Staden, 1982) based on the procedure of McLachlan (1971) run on a VAX 11/780 computer, blocks of 31 amino acids were compared for their match of both identical and related amino acids. (DIAGON is routinely used with a block setting of 21, we increased this to 31 for these comparisons because we have found that a longer block length reduces the background in some comparisons of distantly related sequences, although we have not explored this effect systematically.) The scoring system for this comparison is based on the observed frequency of substitution of one amino acid by another in a number of protein families (Schwartz and Dayhoff, 1978). Only scores above a preset threshold appear on the display, the threshold scores used and their probabilities of occurrence in pairs of 31 residue random sequences are given in the text and figure legends. In both procedures matching blocks are marked by a dot on the display at coordinates corresponding to the midpoint of the blocks in the sequences of the respective proteins.

Detailed alignments were made using the interactive facility of DIAGON and by inspection. Gaps were restricted so as not to exceed the common upper bound of 4/100 residues found in families of distantly related proteins (Doolittle, 1981).

Acknowledgements

We thank Ab van Kammen for critical reading of the text, Piet Madern for printing the photographs and Marie-José van Neerven/Bea Verhaar for typing the manuscript.

References

Baron,M.H. and Baltimore,D. (1982) *J. Virol.*, **43**, 969-978.
Boothroyd,J.C., Harris,T.J.R., Rowlands,D.J. and Lowe,P.A. (1982) *Gene*, **17**, 153-161.
Bruening,G. (1977) in Fraenkel-Conrat,H. and Wagner,R.R. (eds.), *Comprehensive Virology*, Vol. **11**, Plenum Press, NY, pp. 55-141.
Bruening,G. (1981) in Marcus,E. (ed.), *The Biochemistry of Plants*, Vol. **6**, Academic Press, NY, pp. 571-563.
Cole,C.N. and Baltimore,D. (1973) *J. Microbiol.*, **26**, 325-343.
Cooper,P.D. (1968) *Virology*, **35**, 584-596.
Cooper,P.D. (1977) in Fraenkel-Conrat,H. and Wagner,R.R. (eds.), *Comprehensive Virology*, Vol. **9**, Plenum Press, NY, pp. 133-207.
Domingo,E., Sabo,D., Taniguchi,T. and Weissman,C. (1978) *Cell*, **13**, 735-744.
Doolittle,R.F. (1981) *Science (Wash.)*, **214**, 149-159.
Dorssers,L., Zabel,P., Van der Meer,J. and Van Kammen,A. (1982) *Virology*, **116**, 236-249.
Dorssers,L., Van der Krol,S., Van der Meer,J., Van Kammen,A. and Zabel, P. (1984) *Proc. Natl. Acad. Sci. USA*, in press.
Drake,N.L., Palmenberg,A.C., Ghosh,A., Omilianowski,D.R. and Kaesberg,P. (1982) *J. Virol.*, **41**, 726-729.
Emini,E.A., Elzinga,M. and Wimmer,E. (1982) *J. Virol.*, **42**, 194-199.
Etchison,D. and Ehrenfeld,E. (1980) *Virology*, **107**, 135-143.
Flanegan,J.B. and Baltimore,D. (1979) *J. Virol.*, **29**, 352-360.
Franssen,H., Goldbach,R., Broekhuijsen,M., Moerman,M. and Van Kammen,A. (1982) *J. Virol.*, **41**, 8-17.
Franssen,H., Moerman,M., Rezelman,G. and Goldbach,R. (1984a) *J. Virol.*, in press.
Franssen,H., Goldbach,R. and Van Kammen,A. (1984b) *Virus Research*, Vol. **1**, in press.
Gibbs,A.J. and McIntyre,G.A. (1970) *Eur. J. Biochem.*, **16**, 1-11.
Goldbach,R., Rezelman,G. and Van Kammen,A. (1980) *Nature*, **286**, 297-300.
Goldbach,R., Rezelman,G., Zabel,P. and Van Kammen,A. (1982) *J. Virol.*, **42**, 630-635.
Goldbach,R. and Rezelman,G. (1983) *J. Virol.*, **46**, 614-419.
Holland,J., Spindler,K., Horodyski,F., Grabau,E., Nichol,S. and VandePol, S. (1982) *Science (Wash.)*, **215**, 1577-1585.
King,A.M.Q., McCahon,D., Slade,W.R. and Newman,J.W.I. (1982a) *J. Virol.*, **41**, 66-77.
King,A.M.Q., McCahon,D., Slade,W.R. and Newman,J.W.I. (1982b) *Cell*, **29**, 921-928.
Kitamura,N., Semler,B., Rothberg,P.G., Larsen,G.R., Adler,C.J., Dorner, A.J., Emini,E.A., Hanecak,R., Lee,J.J., Van der Werf,S., Anderson, C.W. and Wimmer,E. (1981) *Nature*, **291**, 547-553.
Koch,G., Koch,F., Bilello,J.A., Hiller,E., Schärli,C., Warnecke,G. and Weber,C. (1981) in Perez-Bercoff,R. (ed.), *Protein Biosynthesis in Eukaryotes*, Plenum Press, NY, pp. 275-310.
Larsen,G.R., Anderson,C.W., Dorner,A.J., Semler,B.L. and Wimmer,E. (1982) *J. Virol.*, **41**, 340-344.
Lomonossoff,G. and Shanks,M. (1983) *EMBO J.*, **2**, 2253-2258.
Maizel,J.V., Jr. and Lenk,R.P. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 7665-7669.
Matthews,R.E.F. (1981) *Plant Virology*, second edition, published by Academic Press, NY.
McLachlan,A.D. (1971) *J. Mol. Biol.*, **61**, 409-424.
Nomoto,A., Omata,T., Toyoda,H., Kuge,S., Horie,H., Kataoka,Y., Genba, Y., Nakano,Y. and Imura,N. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 5793-5797.
Pelham,H.R.B. (1979) *Virology*, **96**, 463-477.
Putnak,J.R. and Phillips,B.A. (1981) *Microbiol. Rev.*, **45**, 287-315.
Racaniello,V.R. and Baltimore,D. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 4887-4891.
Rezelman,G., Goldbach,R. and Van Kammen,A. (1980) *J. Virol.*, **36**, 366-373.
Rezelman,G., Franssen,H.J., Goldbach,R.W., Ie,T.S. and Van Kammen,A. (1982) *J. Gen. Virol.*, **60**, 335-342.

Robertson,B.H., Morgan,D.O., Moore,D.M., Grubman,M.J., Card,J., Fischer,T., Weddell,G., Dowbenko,D. and Yansura,D. (1983) *Virology*, **126**, 614-623.
Romanova,L.I., Tolskaya,E.A., Kolesnikova,M.S. and Agol,V.I. (1980) *FEBS Lett.*, **118**, 109-112.
Rueckert,R.R., Matthews,T.J., Kew,O.M., Pallansch,M., McLean,C. and Omilianowski,D. (1979) in Perez-Bercoff,R. (ed.), *The Molecular Biology of Picornaviruses*, Plenum Press, NY, pp. 113-126.
Schwartz,R.M. and Dayhoff,M.O. (1978) in Dayhoff,M.O. (ed.), *Atlas of Protein Sequence and Structure*, Vol. 5, Suppl.3, National Biomedical Research Foundation, Washington, pp. 353-358.
Semler,B.L., Anderson,C.W., Kitamura,N., Rothberg,P.G., Wishart,W.L. and Wimmer,E. (1981a) *Proc. Natl. Acad. Sci. USA*, **78**, 3464-3468.
Semler,B.L., Hanecak,R., Anderson,C.W. and Wimmer,E. (1981b) *Virology*, **114**, 589-594.
Semler,B.L., Hanecak,R., Dorner,L.F., Anderson,C.W. and Wimmer,E. (1983) *Virology*, **126**, 624-635.
Staden,R. (1982) *Nucleic Acids Res.*, **10**, 2951-2961.
Tolskaya,E.A., Romanova,L.I., Kolesnikova,M.S. and Agol,V.I. (1983) *Virology*, **124**, 121-132.
Toh,H., Hayashida,H. and Miyata,T. (1983) *Nature*, **305**, 827-829.
Van der Meer,J., Dorssers,L. and Zabel,P. (1983) *EMBO J.*, **2**, 233-237.
Van Wezenbeek,P., Verver,J., Harmsen,J., Vos,P. and Van Kammen,A. (1983) *EMBO J.*, **2**, 941-946.
Wimmer,E. (1982) *Cell*, **28**, 199-201.
Zabel,P., Moerman,M., Lomonossoff,G., Shanks,M. and Beyreuther,K. (1984) *EMBO J.*, **3**, in press.

Received on 9 January 1984

CHAPTER IX

CONCLUDING REMARKS

At the beginning of the research reported in this thesis it was known that both B-RNA and M-RNA of CPMV are translated into large polyproteins. These primary translation products were clearly subjected to proteolytic cleavage yielding proteins with sizes in the range of 170K to 4K, which could be identified in CPMV-infected cowpea protoplasts. We made it our object to identify the proteolytic activities involved in the processing of the primary translation products and to define their specificity. For this purpose we used cell-free translation systems in combination with a mesophyll protoplast system. It had been proven that the 170K and 32K proteins produced from B-RNA *in vitro* are identical to equal-sized virus-specific products found *in vivo* (Rezelman *et al.*, 1980), indicating that the *in vitro* translation of B-RNA reflects the *in vivo* process. For the expression of M-RNA however, such indication was not available.

Since in B-component-infected protoplasts a large number of B-RNA-specific proteins are found, which all arise by processing of a 200K primary translation product, the presence of a specific proteolytic activity in B component-inoculated protoplasts was supposed. Indeed, if cellular extracts from B component-infected protoplasts were added to the 105K and 95K *in vitro* translation products from M RNA, these were specifically cleaved into 60K, 58K and 48K polypeptides, a process which could not be achieved by addition of cellular extracts from non-infected protoplasts (Chapter III). This finding demonstrated that B-RNA indeed specifies a protease which is capable to process at least the M-RNA encoded proteins. We have demonstrated that the 60K cleavage product represents the precursor to the two coat proteins VP37 and VP23, respectively (Chapter III). This observation suggested that, as with B-RNA, the *in vitro* translation of M-RNA also reflects the course of its *in vivo* expression, even if none of the products, except for the coat proteins, have been detected *in vivo* yet. We were able to map the 48K, 58K and 60K polypeptides on M-RNA and thus provided evidence that the primary translation products share identical carboxy-terminal ends but

differ in their amino-terminal sequences. The location of the genes of these proteins on M-RNA was confirmed when the nucleotide sequence of M-RNA, elucidated by Van Wezenbeek *et al.* (1983), became available at a later stage in our research. As the *in vitro* processing of the M-RNA encoded polyproteins was also achieved with *in vitro* translation products of B-RNA, which mainly consisted of 32K and 170K proteins, we have concentrated our efforts on distinguishing which of these polypeptides harbors the proteolytic activity. Since we had the disposal of antisera directed against purified 32K and 170K polypeptides we investigated which of these sera was able to inhibit the proteolytic processing. This approach led to the identification of the 32K polypeptide as the protease involved in the primary cleavage of the 105K and 95K polyproteins produced by M-RNA (Chapter V).

Amino- and carboxy-terminal amino acid sequences of the coat proteins were determined in order to map the position of the 60K precursor in the nucleotide sequence of M RNA more precisely and to elucidate the dipeptide sequence utilised by the 32K protease. Comparison of the experimentally obtained amino acid data with the amino acid sequence derived from the M-RNA nucleotide sequence allowed us to conclude that this protease recognises a glutamine-methionine dipeptide sequence while a glutamine-glycine dipeptide sequence is utilised to cleave the 60K polypeptide further into coat proteins VP37 and VP23 (Chapter VI). We have not been able to identify the protease repsonsible for the latter cleavage. In CPMV-infected leaves, however, an activity was detected which cleaves VP23 specifically from the primary translation products, but not from the 60K coat protein precursor. Since this activity was rather weak, it was not further analysed (Chapter III). Thus it remains a goal for the future to determine whether this activity is virus- or host-encoded.

Having established the CPMV encoded protease involved in the processing of the primary translation products of M-RNA, it was obvious to examine whether the 32K protease played any role in the processing of the B-RNA encoded polyprotein. This study became possible when we observed that upon longer incubation, the 170K polypeptide was further processed *in vitro* into 170K, 87K, 84K and 60K polypeptides identical to the

in vivo detected B-RNA specific proteins. In addition, among the *in vitro* cleavage products we detected a 28K polypeptide not found *in vivo* so far. Since antiserum against the 32K protease did not interfere with the processing of the 170K polypeptide, the existence of another B-RNA-encoded proteolytic activity had to be assumed. In an attempt to elucidate the proteolytic activity involved in the processing of the B-RNA encoded polyprotein the cleavage of the 200K protein into 32K and 170K polypeptides was studied by time course experiments. Using this approach it was shown that the 32K polypeptide is only released from growing polypeptide chains when they exceed a size of 120K. This suggests that a B-RNA-encoded proteolytic activity comes available when the growing polypeptide chain has reached a size enclosing the 28K polypeptide, which led us to propose that these sequences contain a second proteolytic activity (Chapter VII).

In the meantime Zabel *et al.* (1984) elucidated the amino acid sequence of VPg which allowed the mapping of this polypeptide on the B-RNA nucleotide sequence, recently determined by Lomonossoff and Shanks (1983). As a result it is suggested that VPg is released from its precursor by cleavages at glutamine-serine dipeptide sequences. This same dipeptide sequence is probably also utilised to cleave the primary translation product into 32K and 170K polypeptides, since inspection of the amino acid sequence of the 200K primary translation product as derived from the B-RNA nucleotide sequence reveals the presence of a glutamine-serine dipeptide sequence but not of a glutamine-methionine dipeptide sequence in the region where the cleavage is expected to occur. This observation provided further evidence for the involvement of a proteolytic activity with a cleavage specificity different to that of the 32K protease in the processing of the B-RNA encoded polyprotein. The 28K polypeptide was proposed to be the best candidate for bearing this activity. Supporting evidence for this assumption was obtained as the amino acid sequence homology studies on the CPMV B-RNA encoded 200K polyprotein and the polyprotein encoded by poliovirus scored for the presence of amino acid sequence homology between this 28K polypeptide and the polioviral protease P3-7c (Chapter VIII).

As the organisation of the genes on poliovirus RNA, i.e. VPg-protease

(P3-7c)-polymerase, and B-RNA, i.e. VPg-28K-core polymerase (Chapter VIII), thus apparently are comparable, this provides evidence that the 28K polypeptide encloses proteolytic activity.

The preliminary experimental data indicate that the cleavage specificity of the 28K protease is a glutamine-serine dipeptide sequence. It has not yet been determined whether all cleavages occur on these sites. For instance, the sequence of the 200K polyprotein encloses behind the VPg-28K cleavage site a glutamine-serine dipeptide sequence that upon utilisation would lead to the release of a protein with a calculated molecular weight of 16K, instead of the apparent molecular weight of 28K as determined by SDS-polyacrylamide gel electrophoresis. On the other hand by thorough inspection of the described amino acid sequence homology between poliovirus and CPMV encoded proteins we found support to the idea that also a glutamine-glycine dipeptide sequence might be used in the cleavage of the 170K polypeptide. Assuming that this site is used, which is supported by an optimal amino acid sequence homology between the 28K protease and the polioviral protease P3-7c, now achieved, a 24K protein would be generated, which is a better approach to the apparent molecular weight of 28K. In view of this we propose that one of the cleavage sites on the 200K B-RNA encoded polypeptide is a glutamine-glycine dipeptide sequence. Considering that this site is recognised in the processing of the 200K B-RNA-encoded polyprotein, it is remarkable that a similar site is utilised to generate mature capsid proteins from their common precursor encoded by M-RNA. It is not clear whether a third protease is involved or that the conformation of the polypeptide is such that at these positions either the 28K or the 32K protease has an altered cleavage specificity. The information available indicate that B-RNA encodes two different proteolytic activities: one involved in the processing of the B-RNA encoded polyprotein itself, and a second involved in the processing of the M-RNA encoded polyproteins. The 32K polypeptide which is charged with the latter process is cleaved off from the 200K polyprotein early in the synthesis of the polyprotein. After this cleavage the 32K polypeptide seems to remain complexed to other B-RNA encoded proteins. Whether this reflects a functional significance or represents an artefact of the isolation pro-

cedure is not clear.

For the encapsidation of each RNA 60 copies of the coat proteins containing precursor have to be produced. Since we detected neither this 60K precursor nor the 105K and 95K primary translation products from M-RNA *in vivo*, the proteolytic cleavages leading to the mature capsid proteins are very rapidly catalysed which prevents the accumulation of the precursors *in vivo*. In contrast, intermediate products in the processing of the B-RNA encoded 200K polyprotein are abundantly found in the infected cell. The processing of this polyprotein does not appear to be a straightforward catalytic process, but possibly occurs in close relation with the viral RNA replication cycle.

Processing might be restricted to the formation of replication complexes, in a way that each cleavage leads to the next step in the replication process irreversibly. In this respect it remains to be tested if the 28K polypeptide, completely released from the precursor protein, is still an active protease. The combination of data on the complete nucleotide sequence, the mapping of proteins on this sequence and the identification of various functions encoded on the RNA has enabled us to compare the organisation of genes and functions involved in viral RNA replication between CPMV and picornavirus RNAs. A fascinating result is the homology found between the picornaviral and comoviral polymerases, suggesting a common ancest in the evolution. It will be of interest if these sequence homology studies can be extended to the proteins encoded by other plant viruses. In this respect nepoviruses are very promising candidates as they resemble comoviruses in various aspects.

References

- Goldbach, R.W., Schilthuis, J.G., and Rezelman, G. (1981). Biochem. Biophys. Res. Commun., 99, 89-94.
- Hanecak, R., Semler, B.L., Anderson, C.W., and Wimmer, E. (1982). Proc. Natl. Acad. Sci. USA 79, 3973-3977.
- Lomonossoff, G., and Shanks, M. (1983). EMBO J., 2, 2253-2258.
- Rezelman, G., Goldbach, R., and Van Kammen, A. (1980). J. Virol. 36, 366-373.
- Van Wezenbeek, P., Verver, J., Harmsen, J., Vos, P., and Van Kammen, A. (1983). EMBO J., 2, 941-946.
- Zabel, P., Moerman, M., Lomonossoff, G., Shanks, M., and Bayreuther, K. (1984). EMBO J., in press.

CHAPTER X

SAMENVATTING

Aan het begin van het onderzoek, waarover in dit proefschrift gerapporteerd wordt, was bekend dat de twee RNA moleculen die samen het genoom van CPMV vormen vertaald worden in grote primaire translatieprodukten, z.g. polyproteinen, waaruit de functionele eiwitten door proteolytische klievingen vrijkomen. Er was vastgesteld dat het grootste RNA, B-RNA (molecuulgewicht 2,02 x 10^6), codeert voor een polyprotein met een molecuulgewicht van ongeveer 200.000 (200K) en het andere RNA, M-RNA (molecuulgewicht 1,22 x 10^6), voor twee polyproteinen met molecuulgewichten van 105.000 (105K) en 95.000 (95K).

De doelstelling van mijn onderzoek was de proteolytische activiteiten die betrokken zijn bij de splitsingen van de door CPMV gecodeerde eiwitten te karakteriseren en hun specificiteit vast te stellen. Daarbij hebben we gebruik gemaakt van cel-vrije eiwitsynthetiserende systemen en een cowpea mesophyll protoplasten systeem. Er was al aangetoond dat bij vertaling *in vitro* het door B-RNA gecodeerde 200K gesplitst kon worden in 170K en 32K eiwitten. Ook in met CPMV geinfecteerde protoplasten worden deze 170K en 32K eiwitten aangetroffen. Dit wijst erop dat de vertaling van B-RNA *in vitro* overeenkomt met de expressie van dit RNA *in vivo*. Echter, voor de expressie van M-RNA ontbraken zulke aanwijzingen toen nog.

Aangezien in met B-componenten geïnoculeerde protoplasten een groot aantal B-RNA specifieke eiwitten aangetroffen worden, die allen ontstaan door klievingen binnen het primaire translatie produkt (Rezelman *et al.*, 1980), mocht verondersteld worden dat er in geinfekteerde protoplasten een enzym met een specifieke proteolytische activiteit voorkomt. Inderdaad bleek (hoofdstuk III) dat extracten van met B-component geïnfekteerde protoplasten in staat zijn om althans de 105K en 95K *in vitro* produkten van M-RNA specifiek te klieven, en wel in eiwitten met molecuulgewichten van 60.000 (60K), 58.000 (58K) en 48.000 (48K). We konden aantonen dat het 60K eiwit de beide manteleiwitten VP37 en VP23 bevat en daarom als een precursor van de manteleiwitten beschouwd kan worden. Hoewel *in vivo* alleen de beide manteleiwitten aangetroffen worden en de 60K, 58K en 48K eiwitten tot nu toe *in vivo* niet gedetecteerd zijn,

wijst het feit dat door specifieke splitsing van de *in vitro* translatie produkten van M-RNA een 60K precursor van beide manteleiwitten kan worden verkregen erop dat de expressie van M-RNA *in vivo* eveneens langs deze weg verloopt. We hebben de volgorde van de coderende regios voor de 48K, 58K en 60K eiwitten binnen het M-RNA kunnen vaststellen. Op grond hiervan kon vastgesteld worden dat de primaire translatieprodukten gemeenschappelijk carboxy uiteinden hebben en verschillende amino uiteinden. Aangezien het *in vitro* translatieprodukt van B-RNA, dat voornamelijk uit 32K en 170K eiwitten bestaat, eveneens in staat was om de *in vitro* translatieprodukten van M-RNA te klieven, leek het aannemelijk dat een van deze eiwitten de bewuste proteolytische activiteit zou bezitten. Omdat onze groep beschikte over antisera gericht tegen de gezuiverde virus-specifieke 32K en 170K eiwitten hebben we onderzocht of één van deze antisera de proteolytische splitsing van de door M-RNA gecodeerde polyproteinen kon verhinderen (hoofdstuk V). Op deze manier hebben we kunnen aantonen dat het 32K eiwit het protease is dat verantwoordelijk is voor de primaire splitsing van de 105K en 95K eiwitten.

Door bepaling van de amino- en carboxy eindstandige aminozuurvolgorde van beide manteleiwitten was het mogelijk om de positie van de genen van deze eiwitten binnen de inmiddels bekend geworden nucleotidenvolgorde van M-RNA (Van Wezenbeek *et al.*, 1983) precies te bepalen. Vergelijking van de experimentele gegevens met de aminozuurvolgorde van het primaire translatieprodukt, zoals die uit de nucleotidenvolgorde van M-RNA afgelezen kan worden, maakte het mogelijk de peptide binding te identificeren die specifiek door het 32K protease gesplitst wordt (hoofdstuk VI). Dit bleek een glutamine-methionine binding te zijn. Anderzijds wordt een glutamine-glycine binding gesplitst bij de verdere klieving van het 60K precursor eiwit in de beide manteleiwitten. We zijn er niet in geslaagd de identiteit van het protease dat deze laatste klieving bewerkstelligt vast te stellen. Hoewel in CPMV geinfekteerde bladeren een proteolytische activiteit aangetoond werd, die het manteleiwit VP23 specifiek van de 105K en 95K eiwitten afsplitst, was deze activiteit te zwak om verder onderzoek mogelijk te maken (Hoofdstuk III). Het is daarom nog steeds niet bekend of deze activiteit door het virus dan wel door de waardplant gecodeerd wordt.

Nadat we op de hierboven beschreven wijzen het protease geidentificeerd hadden dat voor de eerste specifieke splitsing van de primaire translatieprodukten van M-RNA zorgt, lag het voor de hand om na te gaan of het 32K protease ook een rol speelt bij de splitsing van het primaire translatieprodukt van B RNA. Dit kon bestudeerd worden nadat eenmaal geconstateerd was dat door langdurige incubatie van de *in vitro* translatieprodukten van B-RNA het 170K eiwit verder gesplitst wordt in 110K, 87K, 84K en 60K eiwitten die ook *in vivo* door het virus geproduceerd worden. Onder de *in vitro* gevormde splitsingsprodukten bleek ook een eiwit van molecuulgewicht van 28.000 (28K) voor te komen, dat tot nu toe niet *in vivo* waargenomen is (hoofdstuk VII). Antiserum gericht tegen het 32K protease bleek de *in vitro* klieving van het 170K eiwit niet te remmen. Daaruit hebben we voorzichtig geconcludeerd dat het 32K protease niet bij dit proces betrokken is. Dit zou betekenen dat er zich onder de door B-RNA gecodeerde eiwitten een tweede protease moet bevinden. Om meer inzicht te krijgen in de manier waarop de B-RNA-gecodeerde eiwitten uit hun 200K precursor vrijkomen is in eerste instantie onderzocht hoe de klieving van het 200K eiwit in 32K en 170K in de tijd verloopt (hoofdstuk VII). We hebben aldus vastgesteld dat bij vertaling van B-RNA *in vitro* het 32K eiwit pas uit groeiende polyprotein ketens vrij komt als deze ketens een lengte bereikt hebben die overeenkomt met een molecuulgewicht van 120.000. Dit wijst erop dat binnen het door B-RNA gecodeerde polyprotein een proteolytische activiteit werkzaam wordt zodra de groeiende keten de volledige sequentie van het 28K eiwit bevat. Op grond hiervan hebben wij gepostuleerd dat het 28K eiwit de tweede door B-RNA gecodeerde protease is (hoofdstuk VII). Intussen hadden Zabel *et al.* (1984) de volgorde van een aantal aminozuren in VPg (molecuulgewicht 4.000) bepaald waardoor zij in staat waren om de VPg-coderende regio op de nucleotidenvolgorde van B-RNA te lokaliseren, die inmiddels door Lomonossoff en Shanks (1983) opgehelderd was. Uit deze kartering bleek dat VPg uit zijn precursor vrijkomt door splitsing van glutamine-serine peptide bindingen aan zowel het N-terminale als het C-terminale uiteinde. Deze waarnemingen ondersteunden de gedachte dat niet het 32K protease, dat glutamine-methionine bindingen herkent, maar een ander protease betrokken is bij de splitsingen

in het B-RNA gecodeerde 200K eiwit. Als meest kansrijke kandidaat voor deze funktie werd door ons het 28K eiwit aangewezen (hoofdstuk VII).

Verdere steun voor deze veronderstelling werd verkregen uit onderzoek naar het voorkomen van overeenkomsten in aminozuurvolgorde tussen het door B-RNA gecodeerde 200K eiwit en het polyprotein gecodeerd door poliovirus. Er bleek namelijk een aanzienlijke homologie in de aminozuur volgorde te bestaan tussen het 28K eiwit en het poliovirus specifieke protease P3-7c. Aangezien verder de organisatie van genen op het poliovirus RNA (nl. VPg-protease (P3-7c) polymerase) en op B-RNA (nl. VPg-28K-core polymerase) vergelijkbaar blijkt te zijn (hoofdstuk VIII) lijkt het zeer aannemelijk dat het 28K eiwit inderdaad proteolytische activiteit bezit. De gegevens die tot nu toe verkregen zijn, hoewel nog onvoldoende voor een onomstotelijk bewijs, geven aan dat het 28K protease binnen het 200K polyprotein specifiek glutamine-serine peptide bindingen kan splitsen. Het is nog niet duidelijk of alle splitsingen bij deze peptide bindingen optreden. Zo blijkt er in de aminozuurvolgorde van het 200K eiwit, zoals die afgeleid kan worden uit de nucleotidenvolgorde van B-RNA, een glutamine-serine dipeptide voor te komen achter de VPg-28K klievingsplaats, die mogelijk gebruikt kan worden om 170K te splitsen in 84K en 87K eiwitten. Echter als splitsing van deze binding gebruikt zou worden in de processing van het 170K eiwit dan zou een eiwit met een molecuuulgewicht van 16.000 (16K) ontstaan, terwijl op grond van SDS-polyacrylamide gel electrophorese het molecuulgewicht voor dit eiwit op 28.000 geschat is. Daar staat tegenover dat uit nauwkeurige bestudering van de beschreven (hoofdstuk VIII) overeenkomst in aminozuurvolgorde tussen het door poliovirus- en door CPMV gecodeerde polyproteinen blijkt dat mogelijk een glutamine-glycine peptide binding gebruikt wordt om het 170K in de 84K en 87K eiwitten te splitsen. Dit zou een eiwit van 24K opleveren, wat een betere benadering van de experimenteel vastgestelde grootte (28K) is, met optimale homologie ten opzichte van het door poliovirus gecodeerde protease P3-7c. Het is daarom aantrekkelijk om te veronderstellen dat er in het door B-RNA gecodeerde polyprotein naast glutamine-serine klievingsplaatsen tenminste één glutamine-glycine klievingsplaats voorkomt. Het is duidelijk dat dit nog verder uitgezocht dient te worden.

Doorgaande op deze gedachte is het echter wel opvallend dat zowel in het door B-RNA gecodeerde polyprotein als in de door M-RNA gecodeerde polyproteinen een glutamine-glycine volgorde zou voorkomen, waar splitsing optreedt. Het is niet duidelijk of een derde protease betrokken is bij splitsing van deze bindingen of dat een van de beide virusspecifieke proteasen, 32K of 28K, onder bepaalde omstandigheden een andere splitsings specificiteit heeft. Om dit na te gaan zal de specificiteit van de virale proteasen verder onderzocht moeten worden.

Resumerend kan geconcludeerd worden dat B-RNA codeert voor twee verschillende proteolytische enzymen. Eén is verantwoordelijk voor de proteolytische splitsing van het door B-RNA gecodeerde polyprotein en het andere voor de klieving binnen de M-RNA gecodeerde polyproteinen. Het 32K eiwit dat bij dit laatst genoemde proces betrokken is wordt reeds in een vroeg stadium van de synthese van het 200K eiwit van dit polyprotein afgesplitst. Het blijkt dat dit 32K eiwit gecomplexeerd blijft aan andere door B-RNA gecodeerde eiwitten (hoofdstuk V). Of het voorkomen van het 32K eiwit in complexen enige functionele betekenis heeft of dat dit het gevolg is van de isolatieprocedure is vooralsnog niet duidelijk.

Om ieder geproduceerd RNA molecuul van een eiwitmantel te voorzien zijn 60 kopieën van de precursor voor beide manteleiwitten nodig. Aangezien we tot nu toe noch dit 60K precursor-eiwit noch de primaire translatieprodukten van M-RNA *in vivo* aan hebben kunnen tonen moeten we aannemen dat het 32K protease in een snelle reactie de vorming van 60K precursors katalyseert, waarna deze vervolgens eveneens zeer snel gesplitst worden in rijpe manteleiwitten. Daarentegen worden de tussenprodukten in het splitsingsschema van het door B-RNA gecodeerde 200K polyprotein in aanzienlijke hoeveelheden aangetroffen in met CPMV geïnfecteerde cellen. Wellicht ontstaan de diverse eindprodukten uit het 200K polyprotein niet in een rechttoe-rechtaan proces, maar zijn de verschillende klievingen van de door B-RNA gecodeerde eiwitten nauw gerelateerd aan processen in de virale RNA replicatie cyclus. Mogelijk wordt de te volgen route in het splitsingsschema van het 200K polyprotein bepaald door het te vormen replicatie complex, en wel op een zodanige wijze dat iedere klievingsstap onvermijdelijk tot de volgende stap in het replicatie proces leidt. In dit verband is het zeker van

belang om na te gaan of het 28K polypeptide als vrij voorkomend eiwit nog steeds als protease actief is.

De combinatie van gegevens over de nucleotidenvolgorde, de localisatie van eiwitten op deze volgorde en de opheldering van de functie van verschillende virale eiwitten, heeft ons in staat gesteld om een vergelijking te maken tussen CPMV en picornavirussen voor wat betreft hun genetische organisatie (hoofdstuk VIII). Een opmerkelijk resultaat dat hieruit naar voren is gekomen, is het feit dat er een grote overeenkomst in aminozuurvolgorde bestaat tussen de RNA-afhankelijke RNA polymerasen van picornavirussen en CPMV, hetgeen suggereert dat deze virussen mogelijk uit dezelfde voorouders geëvolueerd zijn.

Het zou zeer interessant zijn als deze studies naar overeenkomsten in aminozuurvolgorde uitgebreid konden worden naar de eiwitten van andere plantevirussen. Wat dat betreft lijken de nepovirussen interessante kandidaten te zijn omdat zij in vele opzichten op comovirussen lijken.

CURRICULUM VITAE

Henk Franssen werd op 16 november 1955 te Melick (L) geboren. In 1974 behaalde hij het diploma gymnasium B aan het Bisschoppelijk College afdeling Schöndeln te Roermond. Daarna studeerde hij scheikunde aan de Katholieke Universiteit te Nijmegen. In september 1977 werd het kandidaatsexamen S2 afgelegd en in oktober 1980 slaagde hij voor het doctoraalexamen. Hoofdrichting was Biochemie (Prof.dr. H.P.J. Bloemers) en tot de tijdsduur van een hoofdrichting uitgebreid bijvak Biofysische Chemie (Prof.dr. C.W. Hilbers).

Van 1 november 1980 tot 31 oktober 1983 was hij in dienst van de Landbouwhogeschool te Wageningen, werkzaam op de vakgroep Moleculaire Biologie (Prof.dr. A. van Kammen). Hier verrichtte hij onderzoek aan de expressie van cowpea mosaic virus, waarover in dit proefschrift gerapporteerd wordt.

Van 1 januari tot 1 september 1984 was hij in tijdelijke dienst werkzaam op het Biochemisch Laboratorium van de Universiteit van Nijmegen in de werkgroep van Prof.dr. H.P.J. Bloemers.

Vanaf 1 september 1984 is hij in dienst getreden van de Landbouwhogeschool in Wageningen, gefinancierd middels een contract met Agrigenetics.