Multiple Functions of the 32K and 60K Proteins in Cowpea Mosaic Virus RNA Replication

CENTRALE LANDBOUWCATALOGUS
0000 0572 3438

Promotor: Dr. A. van Kammen, hoogleraar in de moleculaire biologie

Co-promotor: Dr. ir. J. Wellink, universitair docent moleculaire biologie

NN08201, 1836

Sander A. Peters

Multiple Functions of the 32K and 60K Proteins in Cowpea Mosaic Virus RNA Replication

Proefschrift
ter verkrijging van de graad van doctor
in de landbouw- en milieuwetenschappen
op gezag van de rector magnificus,
Dr. C.M. Karssen,
in het openbaar te verdedigen
op vrijdag 7 oktober 1994
des namiddags te vier uur in de Aula
van de Landbouwuniversiteit te Wageningen.

Ontvangen
27 SEP. 1994
UB-CARDEX

ISn 365369

BIBLIOTHEEK
LANDBOUWUNIVERSITEIT
WAGENINGEN

CIP-DATA KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Peters, Sander A.

Multiple functions of the 32K and 60K proteins in cowpea mosaic virus RNA replication / Sander A. Peters. - [S.l. : s.n.]
Thesis Wageningen. - With Ref. - With summary in Dutch.
ISBN 90-5485-291-7
Subject headings: cowpea mosaic virus / RNA replication

Cover page: immunofluorescent confocal scanning laser microscope image of transfected cowpea mesophyll protoplasts containing wild-type- and mutant CPMV B-RNA encoded proteins.

The research described in this thesis was carried out at the department of Molecular Biology, Agricultural University Wageningen, The Netherlands, and was supported by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for Scientific Research (NWO). Financial support for publication of this thesis by Pharmacia Nederland BV, Diagnostics is greatfully acknowledged.

NN08201, 1836

Stellingen

1. Het door het CPMV B-RNA gecodeerde 32K eiwit is een chaperone-eiwit voor het 170K eiwit.

   Dit proefschrift.

2. De conclusie dat het ribonucleoside trifosfaat bindingsdomein in het poliovirus 2C eiwit een onmisbare functie in poliovirus RNA replicatie heeft, is voorbarig.

   Teterina *et al.*, 1992. *J. Gen. Virol.* **73**, 1977-1986.
   Mirzayan en Wimmer, 1992. *Virology* **189**, 547-555.

3. Bij de analyse van *cis*-klievingsactiviteit van het 24K proteïnase van GFLV, gaan Margis en collega's voorbij aan de mogelijkheid dat aminozuursequenties stroomopwaarts van het VPg essentiëel zijn voor de klieving tussen VPg en het 24K proteïnase.

   Margis *et al.*, 1991. *Virology* **185**, 779-787.
   Margis *et al.*, 1994. *Virology* **200**, 79-86.

4. De conclusie van Restrepo-Hartwig en Carrington dat aminozuurinserties in het 6kD domein van het TEV-polyproteïn geen invloed zouden hebben op de polyproteïn processing, zou met een *in vitro* vertaling van het volledige mutant RNA aannemelijker kunnen worden gemaakt.

   Restrepo-Hartwig en Carrington, 1994. *J. Virol.* **68**, 2388-2397.

5. Tang en collega's leveren geen enkel bewijs voor hun conclusie dat een hammerhead ribozym, gericht tegen het nucleoproteïne gen van het influenza A virus, *in vivo* het virale RNA klieft.

   Tang *et al.*, 1994. *J. Med. Virol.* **42**, 385-395.

609

6. Het door het ministerie van WVC gepropageerde beleid "zorg op maat" ten aanzien van verstandelijk gehandicapte mensen is onrechtvaardig en respectloos, als men bedenkt dat door bezuinigingen kwalitatieve- en kwantitatieve tekorten in de internaten zijn ontstaan, die hebben geleid tot het isoleren en vastketenen van deze mensen.

7. Tijdens het promotie onderzoek moet "promoter-mappen" worden vermeden.

8. De stelling Ke4, e2 en Ke6, leidt met 1. e3,... tot het veroveren van oppositie en uiteindelijk tot promotie.

Stellingen behorende bij het proefschrift
"Multiple functions of the 32K and 60K proteins
in cowpea mosaic virus RNA replication"
door Sander A. Peters,
te verdedigen op 7 oktober 1994 te Wageningen.

Voor Jenny en Tom,
aan mijn ouders.

Contents

	Page
Scope of the Investigations	9
Chapter 1: General Introduction	13
Chapter 2: A Regulatory Role for the 32K Protein in Proteolytic Processing of Cowpea Mosaic Virus	37
Chapter 3: Processing of VPg-Containing Polyproteins Encoded by the B-RNA from Cowpea Mosaic Virus	55
Chapter 4: The Cowpea Mosaic Virus B-RNA Encoded 112K Protein May Function as a VPg Precursor *in vivo*	71
Chapter 5: The Nucleoside Triphosphate Binding Motif-Containing 60K and 84K Proteins of Cowpea Mosaic Virus Bind ATP	85
Chapter 6: The NTP-Binding Motif in the Cowpea Mosaic Virus B-Polyprotein is Essential for Viral Replication	101
Chapter 7: Discussion	121
Samenvatting	135
Curriculum Vitae	141
Nawoord	142

Scope of the Investigations

Cowpea mosaic virus (CPMV) is the type member of the *comoviridae*, a group of 14 different plant viruses that have a divided genome consisting of two plus-strand RNAs. These RNAs, designated B-RNA and M-RNA, have a small protein, VPg, attached to the 5'-end and a poly(A) tail at the 3'-end and are separately packaged into icosahedral particles of 28 nm in diameter. Nucleotide sequence analysis has revealed that each RNA contains one large open reading frame. Upon infection the RNAs are translated into large polyproteins that are subsequently processed into several stable intermediate and final cleavage products.

The B-RNA and its encoded enzyme activities constitute an autonomous RNA replicon, since the B-RNA can replicate independently of M-RNA in isolated plant cells. However, B-RNA is dependent on M-RNA for cell-to-cell movement in intact plants. The development of full-length cDNA clones, of B- and M-RNA from which infectious RNA transcripts can be derived, has made it possible to study the mechanism of viral gene expression in more detail. By introducing specific mutations in B cDNA clones, several functional domains in the B-polyprotein were identified, but the understanding of the activity of each individual B-RNA encoded protein in the replicative machinery is still incomplete. At the start of the research described in this thesis the B-RNA encoded 110K, which consists of the 24K protein and the 87K core polymerase protein, has been shown to represent the viral RNA-dependent RNA-polymerase. The 60K protein has been proposed to function as a precursor for VPg, that probably has a role as a primer in the initiation of viral RNA replication. Furthermore, the 60K protein is thought to function in anchoring the viral RNA replication complex to membranes, known to be the site of viral RNA replication. Processing of the B-polyprotein is accomplished by the B-RNA encoded 24K proteinase, cleaving the viral proteins at specific Gln/Gly, Gln/Ser and Gln/Met sites. However, for efficient *trans* processing of the Gln/Met site in the M-polyprotein also the B-RNA encoded 32K protein is required.

The studies described in this thesis were concentrated on elucidating the role of the 32K and 60K proteins in the viral replication process. The work presented in chapter 2 of this thesis was directed towards the role of the 32K protein in the polyprotein processing. By employing an *in vitro* transcription/translation system, to express specifically modified cDNA clones, it was shown that the 32K protein regulates both M- and B-polyprotein processing, by interacting with the 58K domain of the 170K and 84K precursors of the 24K proteinase, thereby modulating the cleavage activity and specificity of the 24K proteinase. In chapters 3 and 4 processing of several VPg precursors has been examined. This study showed that *in vitro* processing of the 170K protein can occur via three

alternative pathways to generate 112K, 84K and 60K putative VPg precursor proteins. The 60K protein was found to be stable in this *in vitro* system, whereas the 112K and 84K proteins were processed and might function as a VPg precursor. Using a transient expression system, the 112K protein was evidently shown to function as a direct VPg precursor in cowpea protoplasts (chapter 4).
A study on the biochemical properties of the B-RNA encoded 60K and 84K proteins is described in chapters 5 and 6 of this thesis. A covalent affinity labelling assay was exploited and the 60K and 84K protein were shown to specifically bind ATP, possibly at a ribonucleoside triphosphate binding motif (NTBM) located in the 58K domain of these proteins (chapter 5). In chapter 6 the effect of mutations that were introduced in the coding region of the NTBM is described. With this study an essential role in viral RNA replication could be attributed to the NTBM.

Chapter 1

General Introduction

1.1 The nature of plus-strand RNA viruses

Viruses are among the smallest of all life forms. Their chemical composition is remarkably simple. The major constituents are protein and RNA or DNA. Evenmore simple are viroids which consists of small RNA molecules only. Viruses are obligate intracellular parasites that have evolved efficient mechanisms for replication, genome expression, packaging strategies, as well as generating genetic diversity, which allow them to survive in and to respond quickly to changing intra- and extracellular conditions. Although constraint by their coding capacities, viruses are able to apply many of the host cell functions for their own purposes of replication. They have very limited genetic information and encode primarily some specific functions required for virus multiplication.

The plus-strand RNA viruses are a diverse group that can infect prokaryotes as well as eukaryotes, both plants and animals. After infection the parental genomic RNA is translated in non-structural proteins that are necessary for viral RNA replication. One of these proteins is a RNA-dependent RNA-polymerase (RdRp), which is effective as a viral RNA replicase. The replication cycle of plus-strand RNA viruses involves the synthesis of complementary minus-strand RNA, which then in turn serves as template for synthesis of progeny plus-strand genomic RNAs. Both minus-strand RNA and plus-strand RNA are synthesized by the virus encoded RdRp. Minus-strand RNA viruses, at the other hand, have a virion associated RdRp that after infection produces first mRNA molecules from which more RdRp molecules are translated. Concomitantly, replication starts with the synthesis of full-length complementary RNAs that serve as templates for the synthesis of negative-strand genomic RNAs.

The investigations described in this thesis were focussed on the analysis of the possible functions of some of the proteins encoded by the plant virus cowpea mosaic virus (CPMV) in the replication of the RNA genome of this virus. CPMV is a representative of the plus-strand RNA viruses. As an introduction, in this chapter a comparison with other plus-strand RNA viruses will be made as background information for the interpretation and discussion of the experimental results which are described in chapters 2 to 6 hereafter.

1.2 Superfamilies of positive-strand RNA viruses

The plus-strand RNA viruses exhibit an incredible diversity of morphology, host range, genome structure and expression. During the last decade the nucleotide sequences of the genomic RNAs of many viruses have been determined. This has

allowed extensive sequence homology studies which considerably contributed to the insight into the genome organization and different strategies of expression and replication of positive-strand RNA viruses. Such studies have revealed a number of striking homologies among otherwise disparate virus groups and has resulted into the division of related virus groups into large families, known as 'superfamilies'. Available data so far indicate that most of the plus-strand RNA viruses can be divided into two superfamilies which are the 'Sindbis'-like or 'alphavirus-like' superfamily and the 'picorna-like' superfamily respectively (Goldbach, 1986; Goldbach and Wellink, 1987). The viruses of the first group are related to the animal Sindbis virus (an alphavirus) and include plant bromo-, carla-, clostero-, cucumo-, furo-, hordei-, ilar-, tobamo-, tobra-, tymo-, and potexviruses, and animal alpha- and rubiviruses (Goldbach *et al.*, 1991). They contain enveloped and non-enveloped viruses and comprise a wide variety of virion morphologies with either icosahedral, rod-shaped or bacilliform particles. All viruses belonging to this superfamily contain a cap-structure at the 5'-terminal end of the genomic RNA. The 3'-end of the genomic RNAs is less uniform and may possess a poly(A) tail, a tRNA-like structure or another structure. Criteria for the grouping of viruses in this superfamily include a similar arrangement of three conserved elements in the replicase proteins of these viruses (reviewed by Goldbach *et al.*, 1991). These are a methyltransferase-motif, possibly involved in capping of the viral genome (Dunigan and Zaitlin, 1990; Goldbach *et al.*, 1991), the 'Walker-motif' and the 'GDD-motif' which will be discussed hereafter.

The picorna-like superfamily, includes animal picorna- and caliciviruses and plant como-, poty-, nepo- and bymoviruses. (Goldbach, 1986; Domier *et al.*, 1987; Strauss *et al.*, 1990; King *et al.*, 1991). CPMV is a comovirus and therefore belongs to the picorna-like viruses. The viruses of this superfamily are much more uniform in capsid architecture, genome structure and expression strategy then found for the members of the 'Sindbis-like' superfamily. The picorna-like viruses have icosahedral virions except for potyviruses and bymoviruses, which have rod-shaped virions. Their single-stranded genomic RNAs are 3'-polyadenylated. A major difference with alpha-like viruses is that they have a small protein, VPg (viral protein genome linked), attached to the 5'-end of their RNAs. Their genomic RNAs are translated into large polypeptides, so-called polyproteins, containing different functional domains. The replication of these viruses depends entirely on the full expression of proteins contained within the polyprotein and requires the action of proteolytic enzymes to separate the functional domains. This attests to the key role that a controlled proteolytic processing plays in replication of these viruses (Kemp *et al.*, 1992). The involved proteinases are encoded by the virus

itself and cleave at specific sites to generate the functional mature proteins (Goldbach, 1987; Goldbach and Wellink, 1988; Strauss *et al.*, 1990). They all show a clear sequence homology to each other and contain a cysteine residue at their active site (Wellink and Van Kammen, 1988; Kräusslich and Wimmer, 1988; Helen *et al.*, 1989). Remarkably, these viral proteinases are related to trypsin-like serine proteinases (Gorbalenya *et al.*, 1989a), which contain a catalytic serine residue (Bazan and Fleterick, 1989). The substitution of a serine residue by a cysteine is characteristic for the proteinases in the picorna-like viruses (King *et al.*, 1991).
Homologous regions found for the non-structural proteins include furthermore a conserved 'GDD-motif' and a nucleoside triphophate binding (NTBM)-motif or 'Walker-motif'. Remarkably, the proteins with these motifs are arranged in the same order within the multidomain proteins, NH_2-NTBM-VPg-proteinase-polymerase-COOH respectively (Argos *et al.*, 1984; Gorbalenya *et al.*, 1989b).
Comparison of amino acid sequences revealed that the GDD amino acid motif flanked by hydrophobic stretches is common to all RdRps of animal-, plant- and bacterial RNA viruses. In addition, a second conserved amino acid motif S/GXXXTXXXNT/S (X represents any amino acid) is located upstream of the GDD motif (Kamer and Argos, 1984; Franssen *et al.*, 1984a; Argos, 1988; Ishihama and Nagata, 1988; Poch *et al.*, 1989). A more detailed comparison of plant viral (putative) RdRps revealed that not less than eight blocks of amino acids are conserved (Koonin, 1991a). The GDD-motif has been proposed to be involved in the polymerase function, either directly by binding ribonucleotides or by binding of Mg^{2+}-ion as a cofactor (Argos, 1988), whereas possible functions for the other motifs are not known at present.
By screening a vast number of viral genome sequences it was shown that proteins containing a nucleoside triphosphate binding motif (NTBM) are non-randomly distributed among different virus classes (Gorbalenya and Koonin, 1989). Approximately 80% of all viral plus-strand RNA genomes encode one or two NTBM-containing proteins, which have been classified into three distinct protein families on the basis of sequence analysis and secondary structure predictions (Gorbalenya, 1992). NTBM-containing proteins encoded by 'picorna-like' plant como- and nepoviruses and animal picornaviruses and the 'Sindbis-like' viruses respectively form two distinct families (Hodgeman, 1988; Gorbalenya *et al.*, 1988). Remarkably, although the NTBM-containing proteins of the picorna-like supergroup are all located in similar positions of the cistron maps, the poty- and bymoviral NTBM-containing proteins are more closely related to animal flavi- and pestiviral NTBM-containing proteins and together form a third family of viral NTBM-containing proteins (Lain *et al.*, 1989). Also, the position of the coat protein

downstream of the conserved gene set, differs in the potyviruses and bymoviruses from the other plant and animal viruses of the picorna-like superfamily. It therefore appears that poty- and bymoviruses are somewhat more distantly related to the other members of this superfamily. The genome organization and amino acid homologies described above are summarized in Fig. 1. The implications of the sequence homology found for the non-structural proteins and the similarity in genome organization will be discussed hereafter.

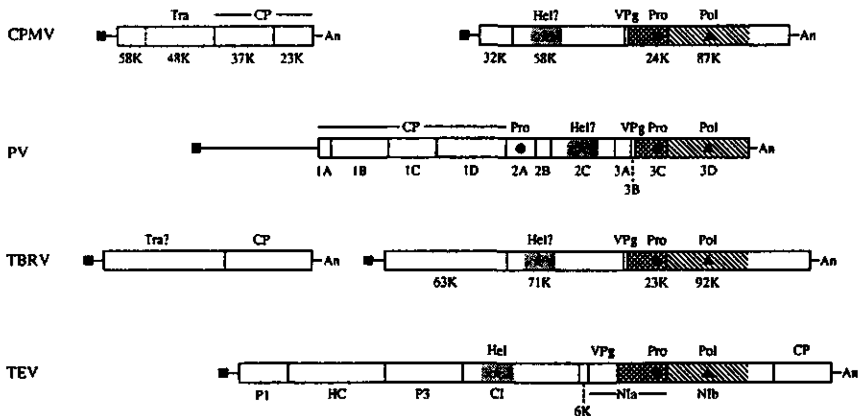


Fig. 1. Comparison of the genomic RNAs of members of the superfamily of picorna-like viruses including poliovirus (picornaviruses), cowpea mosaic virus (comoviruses), tobacco black ring virus (TBRV, nepoviruses) and tobacco etch virus (TEV, potyviruses). Sequences were determined by Kitamura *et al.*, 1981 for poliovirus; Van Wezenbeek *et al.*, 1983 and Lomonossoff and Shanks, 1984 for CPMV; Meyer *et al.*, 1986 and Grief *et al.*, 1988 for TBRV; Allison *et al.*, 1986 for TEV. Open reading frames are represented as open bars. The regions which share significant amino acid homology in the gene products are indicated by shaded areas (Domier *et al.*, 1987; Goldbach and Wellink, 1988). Abbreviations: Tra, transport function; CP, capsid protein(s); Pro, proteinase; Pol, core RNA-dependent RNA-polymerase; Hel, RNA helicase; HC-Pro, helper component-proteinase; CI, cilindrical inclusion protein; NI, nuclear inclusion protein; VPg, viral protein genome linked; An, poly(A) tail. Symbols: ★, nucleoside triphosphate binding domain; ●, proteinase domain; ▲, polymerase domain; ■, VPg.

1.3 Viral RNA replication of picorna-like viruses

The amino acid homologies for the non-structural proteins, and the similarity in genome organization, presumably reflect the employment of one and the same mechanism of RNA replication by the members of the picorna-like superfamily (Van Kammen *et al.*, 1987). Since CPMV is a member of this superfamily, I shall

discuss some specific features of the picorna-like viral RNA replication. A great deal of experimental results have been obtained from studies on polioviral RNA replication and expression and these have considerably contributed to a better understanding of the replication of viruses belonging to the picorna-like superfamily.

1.3.1 The role of membranes

For many years it has been known that an infection of cells by poliovirus stimulates phospholipid biosynthesis (Vance *et al.*, 1980), and is accompanied with the formation of membranous vesicles in the cytoplasm of infected cells, which are thought to be the site of replication (Giachetti and Semler, 1991 and references therein). The physical association of 2C and 2BC proteins with virus induced smooth surfaced membranous vesicles has been observed (Butterworth *et al.*, 1976; Bienz *et al.*, 1983) and it has therefore been suggested that the 2BC and/or 2C proteins induce the continuous formation of the vesicular membranes (Bienz *et al.*, 1987). Interestingly, the cilindrical inclusion (CI) protein of pea seedborne mosaic virus (PSbMV, potyviridae), which has sequence homology with the poliovirus 2C protein (Domier *et al.*, 1987) (see Fig. 1), is also found to be associated with smooth surfaced vesicles that upon infection appear to be induced by this protein (Calder and Ingerfeld, 1990). Cerulenin, a potent inhibitor of lipid biosynthesis, has been demonstrated to effectively block RNA synthesis in Hela cells, indicating that continuous phospholipid biosynthesis is required for successful replication of the poliovirus genome (Carrasco *et al.*, 1989; Guinea and Carrasco, 1990). A similar effect of cerulinin on encephalomyocarditis virus (EMCV), also belonging to the picornaviruses, has been reported (Perez *et al.*, 1991; Perez and Carrasco, 1991). Furthermore, the fungal metabolite brefeldin A, which is known to block the movement of transport vesicles from the ER to the Golgi apparatus, also is an inhibitor of poliovirus replication, suggesting that poliovirus replication depends on the formation and trafficking of these vesicles (Maynell *et al.*, 1992). Taking into account these experimental data, the formation and redirection of such vesicles might be a common requirement shared by the members of the 'picorna-like' virus superfamily. Bienz *et al.* (1990) showed that guanidine either directly or indirectly prevents 2C and 2C precursor proteins from associating with vesicular membranes. This result is consistent with the proposal that the 2C domain serves in anchoring the polioviral replication complex to the membranes. Indeed, treatment of membrane complexes with detergents, released polymerase activity from the membranes, capable of

elongating nascent viral RNA chains. However, initiation of replication on exogenous template could not be achieved after solubilization (Lundquist and Maizel, 1978; for a review see Kuhn and Wimmer, 1987). Apparently initiation of replication is dependent on a membrane associated replication complex (Etchinson and Ehrenfeld, 1981; Takeda *et al.*, 1986).

1.3.2 The role of VPg

The mechanism of initiation of viral RNA replication has not yet been solved. The genomic RNAs of picorna-like viruses have a small protein VPg covalently linked to their 5'-end. Analysis of genome linked proteins revealed that these proteins are bound via a phophodiester bridge to the 5'-terminal nucleotide of the RNA chain. This has suggested a function for the VPg in viral RNA replication, perhaps at the level of initiation (Wimmer, 1982; Eggen and Van Kammen, 1988). A partially purified replication complex from poliovirus infected Hela cells has been found capable of uridylating VPg *in vitro*, which could then be elongated into longer VPg-oligonucleotide products under conditions which favour RNA synthesis (Takeda *et al.*, 1986). Furthermore in poliovirus VPg is found linked to both nascent plus-strands of replicative intermediate (RI) molecules and to the 5'-terminal poly(U) of minus-strands of RI. It is therefore possible that uridylated VPg serves as a primer for the synthesis of viral RNA chains by poliovirus RdRp ($3D^{pol}$), (for reviews see Kuhn and Wimmer, 1987; Semler *et al.*, 1988). The observations that poliovirus specific RNA polymerase is strictly primer-dependent (Van Dyke and Flanegan, 1980), that small amounts of VPg-pUpU can be found in poliovirus infected cells (Crawford and Baltimore, 1983), that uridylation of VPg into VPg-pU and VPg-pUpU can be achieved with crude membrane fractions obtained from polio infected cells (Takegami *et al.*, 1986) which can be elongated into longer products (Takeda *et al.*, 1986), that anti-VPg antibodies specifically inhibit viral RNA synthesis (Baron and Baltimore, 1982), and that mutations in VPg that prevent uridylation abolish synthesis of viral genomic RNA (Reuer *et al.*, 1990; Giachetti and Semler, 1991), substantiate the protein-primed model for initiation of replication. Since VPg is found linked to the 5'-terminal uridylic acid residue of both plus- and minus-strands, it has been suggested that the mechanism of initiation of both RNA strands involves VPg. However, this proposed mechanism for initiation of viral RNA replication awaits further experimental evidence.

1.3.3 The role of the NTBM-containing proteins

Enzymes *that use* ATP *for* their *activity,* like the α- and β*-subunits* of ATP synthetases, myosin, adenylate kinase, recA, many kinases, translation elongation factors and cellular RNA helicases (e.g eIF4A) and DNA helicases (uvrB, uvrD, rep, recB, recD, recQ and human nuclear protein p68), contain a nucleoside triphosphate binding motif (NTBM), consisting of the A- and B-domain that make up the two most important elements of the NTBM (Walker *et al.*, 1982). The A-site consists of a flexible loop (P-loop) (Rossmann *et al.*, 1974; Saraste *et al.*, 1990) and comprises the consensus sequence A/GXXXXGKS/T (X represents any amino acid) involved in the binding of the pyrophosphate moiety of ATP. The B-site consists of a β-sheet strand ending with an invariant Asp residue as a part of the consensus sequence $(Z)_5$DD/E (Z represents a hydrophobic amino acid) (Rossmann *et al.*, 1974), interacting with the Mg^{2+}-ion complexed with the phosphates.

The presence of the NTBM unifies numerous proteins which play a crucial role in energy demanding biochemical processes such as replication, transcription, recombination and repair of DNA, membrane transport, mRNA translation, signal transduction and folding of polypeptides assisted by molecular chaperones. Secondary structure predictions suggest similarity between cellular DNA and RNA helicases and NTBM-containing proteins from picorna-like potyviruses and flavi- and pestiviruses (Gorbalenya *et al.*, 1989c). Up to seven highly conserved stretches, including both the A and B-site of the NTBM, were found in these cellular helicases and viral proteins (Gorbalenya and Koonin, 1988; Gorbalenya *et al.*, 1989c; Lain *et al.*, 1989; Koonin, 1991b). Based on this sequence homology and predicted protein structures, it was alleged that a conservation of helicase function for viral NTBM-containing proteins might underlie the observed similarities (Gorbalenya *et al.*, 1988; Gorbalenya *et al.*, 1989b). For replication of RNA viruses the need for a helicase activity does not seem obvious, as their genomes are single stranded. However, in synthesizing minus-strand RNA from the genomic viral RNA and subsequently synthesizing progeny plus-strand RNA from minus-strand templates, the replication machinery should function such that templates become time and again available for a new round of replication. Therefore the replication mechanism might involve an RNA-unwinding (helicase) activity to resolve both intramolecular base-pairing in template RNA or to prevent the formation of extensive base-pairing between template RNA and the nascent complementary strand. Indeed, the CI protein of plum pox virus (PPV, potyviruses), which has a NTBM with A and B-sites, has been shown to display an ATP-dependent RNA helicase activity (Lain *et al.*, 1990, 1991).

Remarkably, for other NTBM-containing proteins of the picorna-like superfamily sequence homology suggests a closer relation with NTBM-containing proteins of small DNA viruses (Gorbalenya and Koonin, 1989; Gorbalenya *et al.*, 1990). A number of studies on the biochemical properties of SV40 (papovaviruses) large T antigen and mutants thereof revealed that large T protein regulates the initiation of SV40 DNA replication (Auborn *et al.*, 1989; Weiner and Bradley, 1991; Wessel *et al.*, 1992; Ray *et al.*, 1992). Furthermore both DNA and RNA helicase activity could be attributed to large T-antigen of SV40 and has been shown to be essential for viral DNA replication (Stahl *et al.*, 1986; Scheffner *et al.*, 1989). Guanidine-resistant and guanidine-dependent mutants of poliovirus, foot-and-mouth disease virus (FMDV) and encephalomyocarditis virus (EMCV) have been shown to map in or near the conserved 'A' and 'B' site of the NTBM in the 2C proteins encoded by these viruses. (Pincus and Wimmer, 1986; Pincus *et al.*, 1986; Li and Baltimore, 1988; Baltera and Tershak, 1989). The antiviral activity of guanidine has been ascribed to inhibition of initiation of RNA synthesis and release of RNA chains from the replication complex (Kuhn and Wimmer, 1987; Semler *et al.*, 1988; Li and Baltimore, 1988). Although little is known about the process, it could involve a helicase activity. However, to date no direct evidence has been published indicating that the 2C or 2C precursor proteins of picorna viruses are involved in RNA unwinding activity. In addition to the conserved A and B-sites, the sequence similarity for 2C-like proteins with the other NTBM-containing proteins of the picorna-like superfamily is limited to one more conserved segment (Gorbalenya *et al.*, 1988). As yet the role of 2C-like proteins and their possible function in viral RNA replication remains uncertain.

1.4 Genome structure and organization of CPMV

CPMV has *Vigna unguiculata* (L.) (also known as cowpea, southernpea or blackeyed pea) as its natural host. The virus has been elevated to type member of the comoviridae, a group of 14 different plant viruses that have a number of characteristic molecular features in common (for a review see Eggen and Van Kammen, 1988). Other members of this group to which will be referred are red clover mottle virus (RCMV), cowpea severe mosaic virus (CPSMV) and bean pod mottle virus (BPMV). The genome of CPMV consists of two positive strand RNAs, designated B-RNA and M-RNA, that are separately packaged into icosahedral particles of approximately 28 nm in diameter. A successful infection depends on the expression of both M and B-RNA, but there is a clear distinction in functions encoded by the genomic RNAs. Whereas M-RNA encodes the structural proteins

and proteins involved in cell-to-cell movement (Wellink and Van Kammen, 1989), the viral functions needed for RNA replication are encoded by the B-RNA (Goldbach *et al.*, 1980). The capsids of both B and M-nucleoprotein components contain 60 copies of two different coat proteins, VP37 and VP23 respectively. B-RNA has a poly(A)-tract at the 3'-end of approximately 80 nucleotides and the M-RNA of about 160 nucleotides (Ahlquist and Kaesberg, 1979). The function of the poly(A) tail remains unclear, but it has been proposed that the poly(A) tail protects the RNA against degradation (Huez *et al.*, 1983; Eggen *et al.*, 1989a). For CPMV it was found that the VPg is attached to the 5'-terminal uridine of the genomic RNAs by a phosphodiester bond to the β-OH group of the serine residue at the N-terminal end of the protein (Jaegle *et al.*, 1987). The role of VPg will be further discussed in the paragraph on the replication of CPMV.

1.5 Expression of CPMV B-RNA

The B-RNA sequence with a length of 5889 nucleotides contains a 5'-nontranslated region of 206 nucleotides, a single open reading frame of 5601 nucleotides including an AUG-codon at position 207, an UAG stopcodon at position 5805, a 3' nontranslated region (NTR) of 82 nucleotides and then a 3'-poly(A) tail (Lomonossoff and Shanks, 1983). CPMV employs polyprotein processing as a strategy to regulate its gene expression. Upon infection, the B-RNA is translated into a 200K polyprotein that is rapidly processed into 32K and 170K primary cleavage products (Pelham, 1979; Rezelman *et al.*, 1980; Goldbach *et al.*, 1981). The processing of the 32K protein occurs on the nascent polypeptide chain soon after the ribosomes have traversed the coding sequence of the 24K protein, which contains proteolytic activity (Franssen *et al.*, 1984b; Peng and Shih, 1984). The 32K protein interacts with the 58K domain of the 200K polypeptide to form a complex, as is suggested by the observation that anti-serum directed against the 32K polypeptide coprecipitated the 58K protein and 60K, 84K and 170K precursors of this protein (Franssen *et al.*, 1984c). The 24K protein has been identified as the viral proteinase (Verver *et al.*, 1987; Vos *et al.*, 1988a). Mutational analysis indicates that amino acid residues His40, Glu76 and Cys166 are the active residues in the 24K proteinase (Dessens and Lomonossoff, 1991). The proteinase cleaves the viral proteins at specific Gln/Ser, Gln/Met and Gln/Gly sites (Wellink *et al.*, 1986) to yield functional mature proteins. Since only a few out of a vast number of such amino acid pairs are cleaved, it is obvious that additional determinants must play a role in the selection of the cleavage site. After the synthesis of the 170K protein has been completed, further processing gives rise to the 110K, 87K, 84K, 60K, 58K,

24K and VPg proteins. All proteins but free VPg have been observed *in vivo*, implying that the different cleavage products may perform essential functions in viral proliferation. The order of the cleavage products, which are displayed in Fig. 2, has been deduced from the coding regions, by peptide mapping and sequencing the amino terminal ends of the various processing products (Rezelman *et al.*, 1980; Goldbach and Rezelman, 1983; Zabel *et al.*, 1984; Wellink *et al.*, 1986).

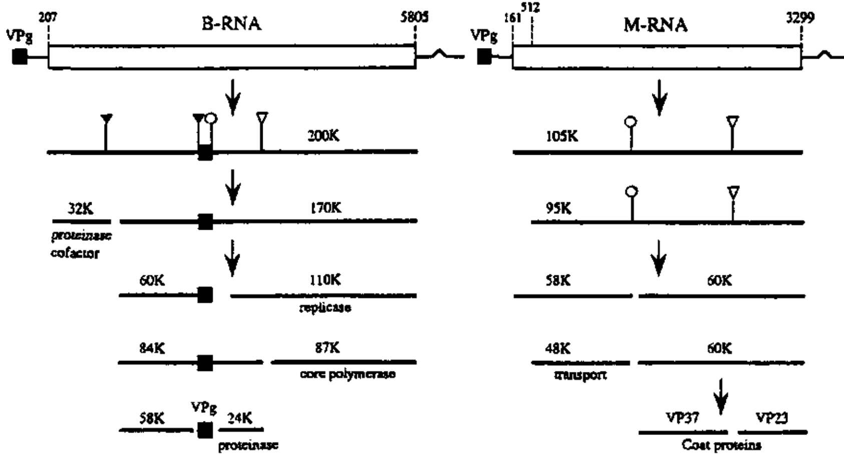


Fig. 2. Expression of CPMV M-RNA and B-RNA. The large open reading frame on both genomic RNAs are represented by an open bar. Positions of start and stop codons are indicated. The cleavage sites on the polyproteins and processing intermediates are indicated by O, Gln/Met; ▽, Gln/Gly and ▼, Gln/Ser. VPg is indicated by a black square, whereas the other proteins are represented by a solid line. Known protein functions are indicated as well.

In contrast to the primary processing of the 200K polyprotein into the 32K and 170K proteins, cleavages at processing sites in the 170K protein occur at a much slower rate (Franssen *et al.*, 1984b). Processing into 84K and 87K proteins has been shown to proceed more rapidly than the cleavage by which the 60K and 110K proteins are generated (Franssen *et al.*, 1984b). It is not clear why CPMV utilizes these alternative cleavage pathways. Possibly cowpea mosaic virus controls the release of discrete proteins from the parent polyprotein by varying rates of cleaving and in that way regulates the function of its viral gene products. However, the mechanism by which this cascade of proteolytic cleavages is ordered is still obscure.

1.6 Expression of CPMV M-RNA

The M-RNA has a length of 3481 nucleotides and consists of a 5'-nontranslated region of 160 nucleotides, an open reading frame starting from the AUG codon at position 161 to the UAG stopcodon at position 3299, a 3' NTR of 180 nucleotides and a 3'-poly(A) tail (Van Wezenbeek *et al.*, 1983). *In vitro* the M-RNA is translated into two C-terminal overlapping polyproteins of 105K and 95K that arise from initiation of translation at nucleotide position 161 and 512 respectively (Vos *et al.*, 1984; Holnes *et al.*, 1989). In studies using $ZnCl_2$, which was shown to inhibit the viral proteinase activity (Wellink *et al.*, 1987), both polyproteins have also been observed in cowpea protoplasts (Rezelman *et al.*, 1989). Evidence from *in vitro* translation studies indicates that ribosomes can bypass the AUG codon at position 161 by leaky scanning and start translation at position 512 (Verver *et al.*, 1991). Furthermore, *in vitro* the sequence between position 161 and 512 is able to support internal entry of ribosomes allowing initiation of translation at position 512 (Verver *et al.*, 1991; Thomas *et al.*, 1991). The 105K and 95K proteins are proteolytically processed into 58K/48K proteins and the 60K capsid precursor (Pelham, 1979; Franssen *et al.*, 1982; Vos *et al.*, 1988b). The latter protein is then further processed into the two capsid proteins, VP37 and VP23 (see Fig. 2). The 58K/48K and the capsid proteins are involved in cell-to-cell transport of the virus (Wellink and Van Kammen, 1989). A CPMV infection is accompanied by the appearance of typical tubular structures that penetrate through the cell wall into a neighbouring cell. These tubular structures contain viral particles suggesting the movement of the virus through these tubules. By immunogold labelling it has been shown that at least the 58K and/or 48K proteins are components of these structures (Van Lent *et al.*, 1990, 1991). Transient expression of the 48K protein in cowpea protoplasts is sufficient to induce these structures (Wellink *et al.*, 1993). The 58K and 48K proteins differ in cellular localization. Whereas the 48K protein is found both in the cytoplasm and membrane fraction (Rezelman *et al.*, 1989), the 58K protein seems to be mainly located in the nucleus (Wellink *et al.*, 1993), suggesting a functional difference for the proteins in viral proliferation. The 48K/58K proteins are generated by *trans* cleavage of the Gln/Met site in the M-polyprotein. The 24K proteinase carries out this cleavage, but also the 32K protein is required, although the latter protein bears no proteolytic activity (Vos *et al.*, 1988b). The underlying mechanism by which the 32K protein promotes this proteolytic cleavage is unclear.

1.7 Replication of CPMV

Eggen and Van Kammen (1988) have proposed a replication model for a membrane bound replication complex, in which the B-RNA encoded 60K protein associated with membranes, functions as membrane anchor for the replication complex and as a supplier for VPg. In the replication complex the 60K protein is associated with the 110K viral replicase and a putative host factor. The 110K protein contains the 24K proteinase that is needed to release the VPg from its precursor. Analogous to poliovirus, the VPg is uridylated and serves as a primer for the synthesis of RNA by the 110K replicase. The 87K region in the 110K protein represents the active core-polymerase domain and contains the conserved GDD amino acid sequence motif flanked by hydrophobic stretches and the S/GTXXXTXXXNT/S motif that is characteristic for RNA-dependent RNA-polymerases of several plant and animal viruses (Kamer and Argos, 1984; Argos, 1988; Ishihama and Nagata, 1988; Poch *et al.*, 1989; Koonin, 1991a).

This replication model has been deduced from several observations. The CPMV B-RNA with its encoded enzyme activities constitute an autonomous replicon, since it is able to replicate in cowpea protoplasts independently from M-RNA (Goldbach *et al.*, 1980). Therefore the viral proteins that are necessary for RNA replication are clearly restricted to the B-RNA encoded proteins. VPg is found covalently bound to the 5'-ends of both positive and negative strands in the replicative forms isolated from CPMV-infected fractions (Lomonossoff *et al.*, 1985). VPg has no function in translation, nor an obvious role in infectivity, since this is not abolished when the VPg is removed from the viral RNAs (Stanley *et al.*, 1978; Eggen *et al.*, 1989b). As has been suggested for poliovirus, VPg very likely has a primer function in viral RNA synthesis and is generated in an active replication complex at the initiation of RNA synthesis. Because VPg has not been observed as a free protein, VPg probably enters the RNA replication complex in a precursor form. Since VPg is a strongly polar (basic) protein, it must be delivered to the membranes by a lipophilic carrier. Probably the 58K domain of the 60K protein performs this function as in membrane fractions from CPMV infected protoplasts both the 58K and the 60K proteins were detected tightly bound to the membranes (Goldbach *et al.*, 1982). No evidence for a smaller precursor was found and therefore the 60K protein was proposed to function as a direct VPg precursor (Zabel *et al.*, 1982; Goldbach *et al.*, 1982, Dorssers *et al.*, 1984). However, this is not yet definitively established, since actual cleavage of the 60K protein into 58K and VPg has never been observed. Moreover, also a 112K protein (VPg+110K) has been found in CPMV infected cells (Dorssers *et al.*, 1983) and this protein might as well be the

precursor of VPg, but the function of this protein has not yet been further investigated.

Further support for the proposed replication model comes from the observation that CPMV has been shown to induce typical cytopathic structures in infected cells. These structures consist of amorphous electron-dense material containing B-RNA encoded non-structural proteins and numerous membranous vesicles, which, analogous to poliovirus, are thought to be the site of viral replication (Assink *et al.*, 1973; De Zoeten *et al.*, 1974; Hibi *et al.*, 1975; Wellink *et al.*, 1988). Such cytopathic structures have also been observed in CPMV B-RNA infected protoplasts and this observation suggests that induction of these vesicles is a B-RNA encoded function (Rezelman *et al.*, 1982). It appears that viral plus-strand synthesis is associated with these vesicles (Dorssers *et al.*, 1983). Purified membrane bound replication complexes containing the 110K protein are able to elongate nascent RNA chains to full-length viral RNAs (Dorssers *et al.*, 1984). The mechanism of initiation of replication, however, remains obscure. Several attempts have been undertaken to demonstrate activity of CPMV RdRp on exogenous poly(A) oligo(U)-primed template (Richards *et al.*, 1989; Van Bokhoven *et al.*, 1992), as has been observed for poliovirus (Van Dyke and Flanegan, 1980). These attempts were not successful and might reflect the strict dependence of initiation of replication on a protein-primed mechanism. However, free VPgpU or uridylation of VPg precursor proteins has never been observed. Other (viral) proteins in addition to the 110K viral replicase and VPg may be necessary to initiate viral RNA synthesis. This is suggested by the observation that in an early stage of CPMV infection actinomycine D inhibits CPMV replication, indicating that host DNA-dependent RNA synthesis is necessary for replication (for a review see Eggen and Van Kammen, 1988).

The replication model described here does not account for a specific role of the B-RNA encoded 32K protein. Yet, Vos *et al.* (1988b) demonstrated that B-RNA transcripts carrying an in frame deletion in the 32K coding region are not infectious in cowpea protoplasts, indicating that an essential function is performed by the 32K protein in the replication of CPMV RNA.

1.8 Some selected differences between CPMV and poliovirus

Unlike cowpea mosaic virus, poliovirus synthesizes its structural- and non-structural proteins as parts of one single polyprotein, which contain two distinct proteinase domains, $2A^{pro}$ and $3C^{pro}$ respectively. The autocatalytic cleavage by $2A^{pro}$ at its own N-terminus occurs already on the nascent polypeptide chain and

serves to separate the P1 capsid precursor from the rest of the polyprotein (Nicklin *et al.*, 1987). The second cleavage carried out by 2A^{pro} occurs in the 3D^{pol}, that represents the viral RNA-dependent RNA-polymerase, is not essential for a successful replication (Lee and Wimmer, 1988). The further cleavages in the P2-region and cleavages for the maturation of the 3D^{pol} are carried out by 3C^{pro} (for a review see Helen *et al.*, 1989). Analogous to CPMV, *trans* processing of the capsid precursor requires additional amino acid sequences and has been shown to be carried out by the 3CD^{pro} precursor protein (Jore *et al.*, 1988; Ypma-Wong *et al.*, 1988). Yet, cowpea mosaic virus and poliovirus use their combined polymerase-proteinase proteins for different purposes. In the 3CD^{pro} protein the 3D^{pol} domain functions as a cofactor for the 3C proteinase but does not display polymerase activity in this precursor form, whereas the corresponding 110K protein from CPMV, consisting of the 24K proteinase and the 87K core polymerase, has been shown to function as the viral replicase (Dorssers *et al.*, 1983, 1984). Moreover, CPMV uses the 32K protein as a cofactor, rather than its polymerase domain to modulate the cleavage specificity of the 24K proteinase. Although the 32K protein of CPMV and the 2A^{pro} from poliovirus are encoded in similar regions of the viral genomes and both proteins are involved in processing of the capsid precursor, no significant homology has been found between these proteins (Argos *et al.*, 1984). Moreover, their actions in the maturation of the capsid protein, as outlined above, appear to be quite different. The 32K protein appears to be a unique protein, since no counterpart among the other viruses of the picorna-like supergroup has been found yet and the 32K protein may have evolved to control the production of active proteins in a rather specific way.
Another difference between CPMV and poliovirus lies in the maturation of their VPg precursors. Three polioviral VPg precursors, 2BC3AB, 2C3AB and 3AB were found in membrane fractions of poliovirus-infected cells (Takegami *et al.*, 1983) and it has been suggested that the former two proteins are very rapidly further processed by the 3C^{pro} into 3AB (Lawson and Semler, 1992), which then in turn serves as the direct precursor to VPg (3AB) (Takegami *et al.*, 1983; Giachetti and Semler, 1991). As outlined above, a CPMV homologue to the poliovirus 3AB protein is not found, but instead CPMV probably uses the 60K protein, homologous to poliovirus 2BC3AB protein, as a direct VPg precursor, leaving the 58K protein as a stable processing product. Thus, the cascade of processing steps that lead to the production of VPg seems to differ significantly between CPMV and poliovirus.
From the dicussion described in this chapter several questions emerge: i) what are the functions of the B-RNA encoded 32K protein in polyprotein processing and viral RNA replication?, ii) which mechanism underlies the enhancement of *trans*

cleavage of the M-polyprotein by the 32K protein?, iii) which successive processing steps are necessary to make VPg available for RNA synthesis?, iv) is the NTBM in the 58K domain functionally important and what role does the NTBM have in viral RNA replication? These questions have been the incentive for the experiments of which the results are described in chapters 2 to 6 hereafter. This thesis will then be concluded with a discussion of these questions.

1.9 References

Ahlquist, P., and Kaesberg, P. (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.

Allison, R.F., Johnston, R.E., and Dougherty, W.G. (1986). The nucleotide sequence of the coding region of tobacco etch virus genomic RNA: Evidence for the synthesis of a single polyprotein. *Virology* **154**, 9-20.

Argos, P. (1988). A sequence motif in many polymerases. *Nucleic Acids Res.* **16**, 9909-9916.

Argos, P., Kamer, G., Nicklin, M.J.H., and Wimmer, E. (1984). Similarity in gene organization and homology between proteins of animal picornaviruses and a plant comovirus suggest common ancestry of these virus families. *Nucleic Acids Res.* **12**, 7251-7267.

Assink, A.M., Swaans, H., and 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.

Auborn, K., Guo, M., and Prives, C. (1989). Helicase, DNA-binding, and immunological properties of replication-defective simian virus 40 mutant T antigens. *J. Virol.* **63**, 912-918.

Baltera, R.F., and Tershak, D.R. (1989). Guanidine-resistant mutants of poliovirus have distinct mutations in peptide 2C. *J. Virol.* **63**, 4441-4444.

Baron, M.H., and Baltimore, D. (1982). Anti-VPg antibody inhibition of the poliovirus replicase reaction and production of covalent complexes of VPg-related proteins and RNA. *Cell* **30**, 745-752.

Bazan, J.F., and Fleterick, R.J. (1989). Detection of a trypsin-like serine protease domain in flavivirus and pestivirus. *Virology* **171**, 637-639.

Bienz, K., Egger, D., and Pasamontes, L. (1987). Association of polioviral proteins of the P2 genomic region with the viral replication complex and virus-induced membrane synthesis as visualized by electron microscopic immunocytochemistry and autoradiography. *Virology* **160**, 220-226.

Bienz, K., Egger, D., Rasser, Y., and Bossart, W. (1983). Intracellular distribution of poliovirus proteins and the induction of virus-specific cytoplasmic structures. *Virology* **131**, 39-48.

Bienz, K., Egger, D., Troxler, M., and Pasamontes, L. (1990). Structural organization of poliovirus RNA replication is mediated by viral proteins of the P2 genomic region. *J. Virol.* **64**, 1156-1163.

Butterworth, B.E., Shimshick, E.J., and Yin, F.H. (1976). Association of the polioviral RNA polymerase complex with phospholipid membranes. *J. Virol.* **19**, 457-466.

Calder, V.L., and Ingerfeld, M. (1990). The roles of the cylindrical inclusion protein of a potyvirus in the induction of vesicles and in cell-to-cell spread. *J. Struct. Biol.* **105**, 62-66.

Carrasco, L., Castrillo, J.L., and Otero, M.J. (1989). Modification of membrane permeability by animal viruses. *Pharmacol. Ther.* **40**, 171-212.

Crawford, N.M., and Baltimore, D. (1983). Genome-linked protein of VPg of poliovirus is present as free VPg and VPg-pUpU in poliovirus-infected cells. *Proc. Natl. Acad. Sci. USA* **80**, 7452-7455.

Dessens, J.T., and Lomonossoff, G.P. (1991). Mutational analysis of the putative catalytic triad of the cowpea mosaic virus 24K protease. *Virology* **184**, 738-746.

De Zoeten, G.A., Assink, A.M., and Van Kammen, A. (1974). Association of cowpea mosaic virus-induced double-stranded RNA with a cytopathological structure in infected cells. *Virology* **59**, 341-355.

Domier, L.L., Shaw, J.G., and Rhoads, R.E. (1987). Potyviral proteins share amino acid sequence homology with picorna-, como-, and caulimoviral proteins. *Virology* **158**, 20-27.

Dorssers, L., Van der Krol, S., Van der Meer, J., Van Kammen, A., and Zabel, P. (1984). Purification of cowpea mosaic virus RNA replication complex: Identification of a virus-encoded 110,000-dalton polypeptide responsible for RNA chain elongation. *Proc. Natl. Acad. Sci. USA* **81**, 1951-1955.

Dorssers, L., Van der Meer, J., Van Kammen, A., and Zabel, P. (1983). The cowpea mosaic virus RNA replication complex and the host-encoded RNA-dependent RNA polymerase-template complex are functionally different. *Virology* **125**, 155-174.

Dunigan, D.D., and Zaitlin, M. (1990). Capping of tobacco mosaic virus RNA. Analysis of viral-coded guanylyl transferase-like activity. *J. Biol. Chem.* **265**, 7779-7786.

Eggen, R., Verver, J., Pleij, K., Van Kammen, A., and Goldbach, R. (1989a). Analysis of sequences involved in cowpea mosaic virus RNA replication using site-specific mutants. *Virology* **173**, 456-464.

Eggen, R., and Van Kammen, A. (1988). RNA replication in comoviruses. "RNA Genetics" (P. Ahlquist, J. Holand, and E. Domingo, Eds.), Vol 1, pp 49-69. CRC Press, Boca Raton.

Eggen, R., Verver, J., Wellink, J., De Jong, A., Goldbach, R., and Van Kammen, A. (1989b). Improvements of the infectivity of in vitro transcripts from cloned cowpea mosaic virus cDNA: impact of terminal nucleotide sequences. *Virology* **173**, 447-455.

Etchinson, D., and Ehrenfeld, E. (1981). Comparison of replication complexes synthesizing poliovirus RNA. *Virology* **111**, 33-46.

Flanegan, J.B., and Baltimore, D. (1979). Poliovirus polyuridylic acid polymerase and RNA replicase have the same viral polypeptide. *J. Virol.* **29**, 352-360.

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., Goldbach, R., and Van Kammen, A. (1984). Translation of bottom component RNA of cowpea mosaic virus in reticulocyte lysate: faithful proteolytic processing of the primary translation product. *Virus Res.* **1**, 39-49.

Franssen, H., Goldbach, R., and Van Kammen, A. (1984b). Translation of bottom component RNA of cowpea mosaic virus in reticulocyte lysate: faithful proteolytic processing of the primary translation product. *Virus Res.* **1**, 39-49.

Franssen, H., Leunissen, J., Goldbach, R., Lomonossoff, G., and Zimmern, D. (1984a). Homologous sequences in non-structural proteins from cowpea mosaic virus and picornavirus. *EMBO J.* **3**, 855-861.

Franssen, H., Moerman, M., Rezelman, G., and Goldbach, R. (1984c). Evidence that the 32,000-Dalton protein encoded by bottom component RNA of cowpea mosaic virus is a proteolytic processing enzyme. *J. Virol.* **50**, 183-190.

Garcia, J.A., Schrijvers, L., Tan, A., Vos, P., Wellink, J., and Goldbach, R. (1987). Proteolytic activity of cowpea mosaic virus encoded 24K protein synthesized in *Escherichia coli*. *Virology* **159**, 67-75.

Giachetti, C., and Semler, B.L. (1991). Role of a viral membrane polypeptide in strand specific initiation of poliovirus RNA synthesis. *J. Virol.* **65**, 2647-2654.

Goldbach, R. (1987). Genome similarities between plant and animal RNA viruses. *Microbiol. Sci.* **4**, 197-202.

Goldbach, R.W. (1986). Molecular evolution of plant viruses. *Ann. Rev. Phytopath.* **24**, 289-310.

Goldbach, R., Le Gall, O., and Wellink, J. (1991). Alpha-like viruses in plants. *Semin. Virol.* **2**, 19-25.

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

Goldbach, R., Rezelman, G., and Van Kammen, A. (1980). Independent replication and expression of B-component RNA of cowpea mosaic virus. *Nature, (London)* **286**, 297-300.

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.

Goldbach, R.W., Schilthuis, J.G., and Rezelman, G. (1981). Comparison of *in vivo* and *in vitro* translation of cowpea mosaic virus RNA. *Biochem. Bioph. Res. Comm.* **99**, 89-94.

Goldbach, R., and Wellink, J. (1988). Evolution of plus-strand RNA viruses. *Intervirology* **29**, 260-267.

Gorbalenya, A.E. (1992). Host-related sequences in RNA viral genomes. *Semin. Virol.* **3**, 359-371.

Gorbalenya, A.E., Blinov, V.M., Donchenko, A.P., and Koonin, E.V. (1989b). An NTP-binding motif is the most conserved sequence in a highly diverged monophyletic group of proteins involved in positive strand RNA viral replication. *J. Mol. Evol.* **28**, 256-268.

Gorbalenya, A.E., Donchenko, A.P., Blinov, V.M., and Koonin, E.V. (1989a). Cysteine proteases of positive strand RNA viruses and chymotrypsin-like serine proteases. FEBS *Lett.* **243**, 103-114.

Gorbalenya, A.E., and Koonin, E.V. (1988). One more conserved sequence motif in helicases. *Nucleic Acids Res.* **16**, 7734.

Gorbalenya, A.E., and Koonin, E.V. (1989). Viral proteins containing the purine NTP-binding sequence pattern. *Nucleic Acids Res.* **17**, 8413-8440.

Gorbalenya, A.E., Koonin, E.V., Donchenko, A.P., and Blinov, V.M. (1988). A novel superfamily of nucleoside triphophate-binding motif containing proteins which are probably involved in duplex unwinding in DNA and RNA replication and recombination. FEBS *Lett.* **235**, 16-24.

Gorbalenya, A.E., Koonin, E.V., Donchenko, A.P., and Blinov, V.M. (1989c). Two related superfamilies of putative helicases involved in replication, recombination, repair and expression of DNA and RNA genomes. *Nucleic Acids Res.* **17**, 4713-4730.

Gorbalenya, A.E., Koonin, E.V., and Wolf, Y.I. (1990). A new super family of putative NTP-binding domains encoded by genomes of small DNA and RNA viruses. FEBS *Lett.* **262**, 145-148.

Grief, C., Hemmer, O., and Fritsch, C. (1988). Nucleotide sequence of tomato black ring virus RNA-1. *J. Gen. Virol.* **69**, 1517-1529.

Guinea, R., and Carrasco, L. (1990). Phospholipid biosynthesis and poliovirus genome replication, two coupled phenomena. *EMBO J.* **9**, 2011-2016.

Helen, C.U.T., Kräusslich, H-G., and Wimmer, E. (1989). Proteolytic processing of polyproteins in the replication of RNA viruses. *Biochem.* **28**, 9881-9890.

Hibi, T., Rezelman, G., and Van Kammen, A. (1975). Infection of cowpea mesophyll protoplasts with cowpea mosaic virus. *Virology* **64**, 308-318.

Hodgeman, T.C. (1988). A new superfamily of replicative proteins. *Nature, (London)* **333**, 22-23.

Holnes, C.L., Lomonossoff, G.P., Evans, D., and Maule, A.J. (1989). Identification of the initiation codons for translation of cowpea mosaic virus midle component RNA using site-directed mutagenesis of an infectious cDNA clone. *Virology* **172**, 311-320.

Huez, G., Cleuter, Y., Bruck, C., Van Vlooten-Doting, L., Goldbach, R., and Verduin, B. (1983). Translational stability of plant viral RNAs microinjecting into living cells. Influence of a poly(A) segment. *Eur. J. Biochem.* **130**, 205-209.

Ishihama, A., and Nagata, K. (1988). Viral polymerases. Critical reviews in Biochemistry, Vol. 23, pp 27-76. CRC Press, Boca Raton.

Jaegle, M., Wellink, J., and Goldbach, R. (1987). The genome-linked protein of cowpea mosaic virus RNA by a phosphodiester linkage to serine. *J. Gen. Virol.* **68**, 627-632.

Jore, J., De Geus, B., Jackson, R.J., Pouwels, P.H., and Enger-Valk, B.E. (1988). Poliovirus protein 3CD is the active protease for processing of the precursor for processing of the precursor protein P1 *in vitro. J. Gen. Virol.* **69**, 1627-1636.

Kamer, G. and Argos, P. (1984). Primary structural comparison of RNA dependent polymerases from plant, animal and bacterial viruses. *Nucleic Acids Res.* **12**, 7269-7282.

Kemp, G., Webster, A., and Russel, W.C. (1992). Proteolysis is a key process in virus replication. *Essays Biochem.* **27**, 1-16.

King, A.M.Q., Lomonossoff, G.P., and Ryan, M.D. (1991). Picornaviruses and their relatives in the plant kingdom. *Semin. Virol.* **2**, 11-17.

Kitamura, N., Semler, B.L., 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). Primary structure, gene organization and polypeptide expression of poliovirus RNA. *Nature, (London)* **291**, 547-553.

Koonin, E. (1991a). The phylogeny of RNA-dependent RNA polymerases of positive strand RNA viruses. *J. Gen. Virol.* **72**, 2197-2206.

Koonin, E.V. (1991b). Similarities in RNA helicases. *Nature, (London)* **352**, 290.

Kräusslich, H-G., and Wimmer, E. (1988). Viral proteinases. *Ann. Rev. Biochem.* **57**, 701-754.

Kuhn, R.J., and Wimmer, E. (1987). The replication of picornaviruses. *In* "The molecular biology of positive strand RNA viruses". D.J. Rowlands, B.W. Mahy, and M. Mayo, Eds.), pp 17-50. Academic press, New York.

Lain, S., Martin, M.T., Riechmann, J.L., and Garcia, J.A. (1991). Novel catalytic activity associated with positive-strand RNA virus infection: nucleic acid-stimulated ATPase activity of the plum pox potyvirus helicaselike protein. *J. Virol.* **65**, 1-6.

Lain, S., Riechmann, J.L., and Garcia, J.A. (1990). RNA helicase: a novel activity associated with a protein encoded by a positive strand RNA virus. *Nucleic Acids Res.* **18**, 7003-7006.

Lain, S., Riechmann, J.L., Martin, M.T., and Garcia, J.A. (1989). Homologous potyvirus and flavivirus proteins belonging to a superfamily of helicase-like proteins. *Gene* **82**, 357-362.

Lawson, M.A., and Semler, B.L. (1991). Alternate poliovirus nonstructural protein processing cascades generated by primary sites of 3C proteinase cleavage. *Virology* **191**, 309-320.

Lee, C.-K., and Wimmer, E. (1988). Proteolytic processing of poliovirus polyprotein: elimination of $2A^{pro}$-mediated, alternative cleavage of polypeptide 3CD by *in vitro* mutagenesis. *Virology* **166**, 405-414.

Li, J-P., and Baltimore, D. (1988). Isolation of poliovirus 2C mutants defective in viral RNA synthesis. *J. Virol.* **62**, 4016-4021.

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

Lomonossoff, G.P., Shanks, M., and Evans, D. (1985). The structure of cowpea mosaic virus replicative form RNA. *Virology* **144**, 351-362.

Lundquist, R.E., and Maizel, J.V. (1978). Structural studies of the RNA component of the poliovirus replication complex. *Virology* **85**, 434-444.

Maynell, L.A., Kirkegaard, K., and Klymkowsky, M.W. (1992). Inhibition of polioviral RNA synthesis by brefeldin A. *J. Virol.* **66**, 1985-1994.

Meyer, M., Hemmer, O., Mayo, M.A., and Fritsch, C. (1986). The nucleotide sequence of tomato black ring virus RNA-2. *J. Gen. Virol.* **67**, 1257-1271.

Nicklin, M.J.H., Kräusslich, H.-G., Toyoda, H., Dunn, J.J., and Wimmer, E. (1987). Poliovirus polypeptide precursors: Expression *in vitro* and processing by exogenous 3C and 2A proteinases. *Proc. Natl. Acad. Sci. USA* **84**, 4002-4006.

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

Peng, X.X., and Shih, D.S. (1984). Proteolytic processing of the proteins translated from the bottom component RNA of cowpea mosaic virus. *J. Biol. Chem.* **259**, 3197-3201.

Perez, L., and Carrasco, L. (1991). Cerulenin, an inhibitor of lipid synthesis, blocks vesicular stomatitis virus RNA replication. FEBS *Lett.* **280**, 129-133.

Perez, L., Guinea, R., and Carrasco, L. (1991). Synthesis of Semliki Forest virus RNA requires continuous lipid synthesis. *Virology* **183**, 74-82.

Pincus, S.E., Diamond, D.C., Emini, E.A., and Wimmer, E. (1986). Guanidine-selected mutants of poliovirus: mapping of point mutations to polypeptide 2C. *J. Virol.* **57**, 638-646.

Pincus, S.E., and Wimmer, E. (1986). Production of guanidine-resistant and -dependent poliovirus mutants from cloned cDNA: mutations in polypeptide 2C are directly responsible for altered guanidine sensitivity. *J. Virol.* **60**, 793-796.

Poch, O., Sauvaget, I., Delreau, M., and Tordo, N. (1989). Identification of four conserved motifs among the RNA-dependent polymerase encoding elements. *EMBO J.* **8**, 3867-3874.

Ray, S., Anderson, M.E., Loeber, G., Mc Vey, D., and Tegtmeyer, P. (1992). Functional characterization of temperature-sensitive mutants of simian virus 40 large T antigen. *J. Virol.* **66**, 6509-6516.

Reuer, Q., Kuhn, R.J., and Wimmer, E. (1990). Characterization of poliovirus clones containing lethal and nonlethal mutations in the genome-linked protein of VPg. *J. Virol.* **64**, 2967-2975.

Rezelman, G., Franssen, H.J., Goldbach, R.W., Ie, T.S., and Van Kammen, A. (1982). Limits to the independence of bottom component RNA of cowpea mosaic virus. *J. Gen. Virol.* **60**, 335-342.

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.

Rezelman, G., Van Kammen, A., and Wellink, J. (1989). Expression of cowpea mosaic virus M RNA in cowpea protoplasts. *J. Gen. Virol.* **70**, 3043-3050.

Richards, O.C., Eggen, R., Goldbach, R., and Van Kammen, A. (1989). High level synthesis of cowpea mosaic virus RNA polymerase and protease in *Escherichia coli*. *Gene* **78**, 135-146.

Rossmann, M.G., Moras, D., and Olsen, K.W. (1974). Chemical and biological evolution of a nucleotide-binding protein. *Nature, (London)* **250**, 194-199.

Saraste, M., Sibbald, P.R., and Wittinghofer, A. (1990). The P-loop - a common motif in ATP- and GTP-binding proteins. *Trends Biochem. Sci.* **15**, 430-434.

Scheffner, M., Knippers, R., and Stahl, H. (1986). RNA unwinding activity of SV40 large T antigen. *Cell* **57**, 955-963.

Semler, B.L., Kuhn, R.J., and Wimmer, E. (1988). Replication of the poliovirus genome. *In* "RNA Genetics" (P. Ahlquist, J. Holland, and E. Domingo, Eds.), Vol 1, pp 23-58. CRC Press, Boca Raton.

Stahl, H., Droge, P., and Knippers, R. (1986). DNA helicase activity of SV40 large tumor antigen. *EMBO J.* **5**, 1939-1944.

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.

Strauss, E.G., Strauss, J.H., and Levine, A.J. (1990). Virus evolution. 'Fields Virology' (B.N. Fields, D.M. Knipe, R.M. Chanock, M.S. Hirsh, J.L. Melnick, T.P. Monath and B. Roizman, Eds.), vol. 1, pp 167-190, second edition, Raven Press, New York.

Takeda, N., Kuhn, R.J., Yang, C-F., Takegami, T., and Wimmer, E. (1986). Initiation of poliovirus plus-strand RNA synthesis in a membrane complex of infected Hela cells. *J. Virol.* **60**, 43-53.

Takegami, T., Kuhn, R.J., Anderson, C.W., and Wimmer, E. (1983). Membrane-dependent uridylylation of the genome-linked protein VPg of poliovirus. *Proc. Natl. Acad. Sci. USA* **80**, 7447-7451.

Thomas, A.A.M., Ter Haar, E., Wellink, J., and Voorma, H.A. (1991). Cowpea mosaic virus middle component RNA contains a sequence that allows internal binding of ribosomes and that requires eukaryotic initiation factor 4F for optimal translation. *J. Virol.* **65**, 2953-2959.

Van Bokhoven, H., Van Lent, J.W.M., Custers, R., Vlak, J.M., Wellink, J., and Van Kammen, A. (1992). Synthesis of the complete 200K polyprotein encoded by the cowpea mosaic virus B-RNA in insect cells. *J. Gen. Virol.* **73**, 2775-2784.

Van Dyke, T.A., and Flanegan, J.B. (1980). Identification of poliovirus polypeptide p63 as a soluble RNA-dependent RNA polymerase. *J. Virol.* **35**, 732-740.

Van Kammen, A., Wellink, J., Goldbach, R., Eggen, R., and Vos, P. (1987). The use of full-length DNA copies in the study of the expression and replication of the RNA genome of cowpea mosaic virus, in Plant Molecular Biology, (D. von Wettstein and N-H., Chua, Eds.), Plenum publishing corporation, New York.

Van Lent, J., Storms, M., Van der Meer, F., Wellink, J., and Goldbach, R. (1991). Tubular structures involved in movement of cowpea mosaic virus are also formed in infected cowpea protoplasts. *J. Gen. Virol.* **72**, 2615-2623.

Van Lent, J., Wellink, J., and Goldbach, R. (1990). Evidence for the involvement of the 58K and 48K proteins in the intercellular movement of cowpea mosaic virus. *J. Gen. Virol.* **71**, 219-223.

Vance, D.E., Trip, E.M., and Paddy, H.B. (1980). Poliovirus increases phosphatidyl choline biosynthesis in Hela cells by stimulation of the rate-limiting reaction catalyzed by CTP:phosphocholine cytidyltransferase. *J. Biol. Chem.* **255**, 1064-1069.

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.

Verver, J., Goldbach, R., Garcia, J.A., and Vos, P. (1987). *In vitro* expression of a full-length DNA copy of cowpea mosaic virus B RNA: identification of the B RNA encoded 24-kd protein as a viral protease. *EMBO J.* **6**, 549-554.

Verver, J., Le Gall, O., Van Kammen, A., and Wellink, J. (1991). The sequence between nucleotides 161 and 512 of cowpea mosaic virus M RNA is able to support internal initiation of translation *in vitro*. *J. Gen. Virol.* **72**, 2339-2345.

Vos, P., Jaegle, M., Wellink, J., Van Kammen, A., and Goldbach, R. (1988b). Two viral proteins involved in the proteolytic processing of the cowpea mosaic virus polyproteins. *Nucleic Acids Res.* **16**, 1967-1985.

Vos, P., Jaegle, M., Wellink, J., Verver, J., Eggen, R., Van Kammen, A., and Goldbach, R. (1988a). Infectious transcripts derived from the full-length DNA copies of the genomic RNAs of cowpea mosaic virus. *Virology* **165**, 33-41.

Vos, P., Verver, J., Van Wezenbeek, P., Van Kammen, A., and Goldbach, R. (1984). Study of the genetic organization of a plant viral genome by *in vitro* expression of a full-length DNA copy. *EMBO J.* **3**, 3049-3053.

Walker, J.E., Saraste, M., Runswick, M.J., and Gay, N.J. (1982). Distantly related sequences in the α- and β-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* **1**, 945-951.

Weiner, B.M., and Bradley, M.K. (1991). Specific mutation of a regulatory site within the ATP-binding region of simian virus 40 large T antigen. *J. Virol.* **65**, 4973-4984.

Wellink, J., Jaegle, M., Prinz, H., Van Kammen, A., and Goldbach, R. (1987). Expression of the middle component RNA of cowpea mosaic virus *in vivo*. *J. Gen. Virol.* **68**, 2577-2585.

Wellink, J., Rezelman, G., Goldbach, R., and Beyreuther, K. (1986). Determination of the proteolytic processing sites in the polyprotein encoded by the bottom-component RNA of cowpea mosaic virus. *J. Virol.* **59**, 50-58.

Wellink, J., and Van Kammen, A. (1988). Proteases involved in the processing of viral polyproteins. *Arch. Virol.* **98**, 1-26.

Wellink, J., and Van Kammen, A. (1989). Cell-to-cell transport of cowpea mosaic virus requires both the 58K/48K proteins and the capsid proteins. *J. Gen. Virol.* **70**, 2279-2286.

Wellink, J., Van Lent, J., and Goldbach, R. (1988). Detection of viral proteins in cytopathic structures in cowpea protoplasts infected with cowpea mosaic virus. *J. Gen. Virol.* **69**, 751-755.

Wellink, J., Van Lent, J.W.M., Verver, J., Sijen, T., Goldbach, R.W., and Van Kammen, A. (1993). The cowpea mosaic virus M RNA-encoded 48-kilodalton protein is responsible for induction of tubular structures in protoplasts. *J. Virol.* **67**, 3660-3664.

Wessel, R., Schweizer, J., and Stahl, H. (1992). Simian virus 40 T-antigen DNA helicase is a hexamer which forms a binary complex during bidirectional unwinding from the viral origin of DNA replication. *J. Virol.* **66**, 804-815.

Wimmer, E. (1982). Genome linked proteins of viruses. *Cell* **28**, 199-201.

Ypma-Wong, M.-F., Dewalt, P.G., Johnson, V.H., Lamb, J.G., and Semler, B.L. (1988). Protein 3CD is the major proteinase responsible for cleavage of the P1 capsid precursor. *Virology* **166**, 265-270.

Zabel, P., Moerman, M., Lomonossoff, G., Shanks, M., and Beyreuther, K. (1984). Cowpea mosaic virus VPg: sequencing of radiochemicaly modified protein allows maping of the gene on B-RNA. *EMBO J.* **3**, 1629-1634.

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.

Chapter 2

A Regulatory Role for the 32K Protein in Proteolytic Processing of Cowpea Mosaic Virus Polyproteins

Sander A. Peters, Wilfried G.B. Voorhorst, Jan Wery,
Joan Wellink and Ab van Kammen

Department of Molecular Biology, Agricultural University, Dreyenlaan 3, 6703 HA
Wageningen, The Netherlands

Running title: Regulation of CPMV polyprotein processing

Virology 191, 81-89, 1992.

A Regulatory Role for the 32K Protein in Proteolytic Processing of Cowpea Mosaic Virus Polyproteins

Abstract

We have studied the regulation of proteolytic processing of the polyproteins encoded by cowpea mosaic virus (CPMV) M-RNA and B-RNA. For that purpose muations were introduced in full-length cDNA clones of these RNAs. RNA transcripts were translated in rabbit reticulocyte lysate and the effect of mutations on the processing was analysed. These studies revealed that the 32K protein is released from the 200K polyprotein by an intramolecular cleavage and remains associated with the 170K protein, probably by interaction with the 58K domain of the 170K protein. In this complex the conformation is such that further cleavages are very slow. This complex carries out the processing of the Gln/Met site in the M-polyprotein. The 170K produced by a B-RNA mutant that lacks the 32K coding region was efficiently processed into 110K, 87K, 60K, 58K and 24K cleavage products. Thus, the 32K protein regulates the B-polyprotein processing by slowing it down and, on the other hand enhances *trans* cleavage of the M-polyproteins at a Gln/Met site.

Introduction

The two genomic RNAs of cowpea mosaic virus (CPMV), denoted B-RNA and M-RNA are both translated into large polyproteins from which functional proteins are generated by proteolytic cleavages at specific Gln/Met, Gln/Ser and Gln/Gly sites (Fig. 1) by a virus encoded proteinase (Wellink *et al.*, 1986; Garcia *et al.*, 1987). The larger B-RNA codes for a 200K polyprotein that, upon translation, is rapidly cleaved at a Gln/Ser site into 32K and 170K proteins. This primary cleavage can occur on the nascent polypeptide chain and is achieved by the B-RNA encoded 24K proteinase (Franssen *et al.*, 1984a; Verver *et al.*, 1987). After this primary cleavage the 32K protein becomes associated with the 170K protein (Pelham, 1979; Franssen *et al.*, 1984a). Subsequently, the 170K protein can be further cleaved via 110K, 84K and 60K intermediates ultimately into 58K, VPg (viral protein genome linked), 24K and 87K proteins. *In vitro*, these cleavages are only observed after prolonged incubation (Peng and Shih, 1984; Franssen *et al.*, 1984a; Vos *et al.*, 1988a). The B-RNA is able to replicate independently from M-RNA in infected protoplasts and encodes the proteins which are necessary for RNA replication (Goldbach *et al.*, 1980). The 110K protein, which consists of the 24K proteinase and the 87K core

polymerase, has been identified as the viral RNA replicase (Dorssers *et al.*, 1984; Eggen and Van Kammen, 1988). The 32K protein is also somehow involved in the replication of the viral RNAs, since RNAs with mutations in the 32K coding region are not infectious in cowpea protoplasts (Vos *et al.*, 1988b), but the function of the 32K protein in the viral RNA replication process is unclear.

M-RNA is translated into two C-terminal overlapping polypeptides of 105K and 95K (Pelham, 1979). Both polypeptides are cleaved at a Gln/Met site into 58K and 48K proteins, involved in cell-to-cell transport of the virus (Wellink and Van Kammen, 1989), and a 60K precursor of the capsid proteins (Fig. 1) (Franssen *et al.*, 1982). Cleavage of the 60K capsid precursor at the Gln/Gly site gives rise to the capsid proteins VP37 and VP23. In reticulocyte lysate this cleavage reaction is generally not observed and has been shown to be sensitive to hemin (Bu *et al.*, 1989). Although the 24K proteinase catalyzes all cleavage reactions in the CPMV polyproteins, previous studies have demonstrated that for *trans* cleavage at the Gln/Met site in the M-polyprotein, the presence of the B-RNA encoded 32K protein is also required, although the 32K protein itself does not have proteolytic activity (Vos *et al.*, 1988a).

To gain more insight in the regulation of polyprotein processing, we have studied the expression of RNA transcripts of specifically modified cDNA clones in a rabbit reticulocyte lysate system. Previous studies indicated that all processing steps can occur correctly and that polypeptides produced in this system are identical to the B-RNA and M-RNA encoded proteins found *in vivo* (Pelham, 1979; Peng and Shih, 1984; Franssen *et al.*, 1984a; Vos *et al.*, 1988a; Bu *et al.*, 1989). In this paper we report experiments which have improved our understanding of the function of the 32K protein during B-RNA and M-RNA expression. We show that the 32K protein plays a major regulatory role in the processing of both B- and M-polyproteins.

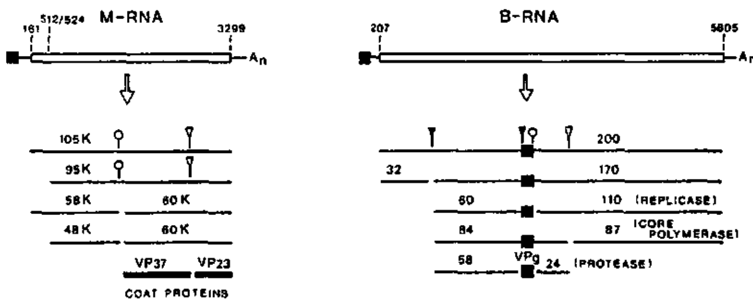


Fig. 1. Expression and genetic organization of M-RNA and B-RNA of CPMV. Both RNAs contain a single open reading frame (shown by open bars). The positions of start and stop codons are indicated. Cleavage sites are indicated by O, Gln/Met, ∇, Gln/Gly and ▼, Gln/Ser.

Materials and methods

DNA techniques

All enzymes were purchased from Gibco BRL, New England Biolabs and Pharmacia and used as described by the manufacturers. DNA fragments were isolated from agarose gels and purified with Gene Clean™ according to the manufacturer (Bio 101 Inc.). Standard recombinant DNA techniques were used for linker insertion, ligation and transformation in competent *E.coli* JM109 as described (Sambrook *et al.*, 1989). Recombinant clones were analysed by restriction enzyme mapping or nucleotide sequence analysis as described (Sanger *et al.*, 1977; Korneluk *et al.*, 1985). Oligodeoxyribonucleotides were synthesized with a cyclone DNA synthesizer (Biosearch). Oligonucleotides and single stranded recombinant M13 DNA were used for site-directed mutagenesis as described by Kunkel (1985).

Construction of mutagenized cDNA clones

As starting material for construction of the different mutant clones the full-length cDNA clones of B-RNA, pTB1G, and of M-RNA, pTM1G, were used (Eggen *et al.*, 1989). In these plasmids full-length cDNA is cloned behind a bacteriophage T7 RNA polymerase promoter. From the cDNAs RNA molecules can be transcribed which have been shown to be infectious in cowpea protoplasts (Eggen *et al.*, 1989). The positions of restriction sites and nucleotides refer to the position in the B-RNA sequence determined by Lomonossoff and Shanks (1983) and in the M-RNA sequence determined by Van Wezenbeek *et al.* (1983).
Plasmid pTBΔ32 was constructed by inserting the SalI-SstI fragment (positions -40 to 2301) from pTB1G into M13mp19. Nucleotides 210 to 1184 were deleted by site-directed mutagenesis using the phosphorylated oligonucleotide 5'-GGATAACAGGACTACTCATGTTGGGTCAAG-3' and single stranded recombinant phage DNA. The mutagenized SalI-SstI fragment was reinserted into pTB1G resulting in pTBΔ32. This plasmid lacks the entire 32K coding sequence. For a further analysis of the function of the 32K protein also pTB32S (previous pTB32*, Vos *et al.*, 1988a) was used. For construction of pTBΔ5a a 676-bp SstI-PvuII fragment (positions 1625 to 2301) was removed from pTB1G. A 7.6 kb-fragment was isolated, blunted with Klenow fragment of DNA polymerase and ligated, resulting in pTBΔ5a. This deletion in the 58K coding sequence does not affect the reading frame. For construction of pTBΔ5b a similar strategy was followed. This plasmid has a 612-bp in frame deletion in the 58K coding sequence between the SstI site at position 2301 and AvaII site at position 2913. To construct pTB58I, pTB1G was partially digested

with KpnI at position 2746. Protruding ends were removed with Klenow fragment of DNA polymerase and in the blunted KpnI site a SmaI-linker (5'-CCCCCGGGGG-3') was inserted giving pTB58KS. This plasmid was digested with SmaI. Linearized plasmids were dephosphorylated and a 414-bp BalI-BalI fragment from pTM1G (positions 1806 to 2220) was inserted resulting in pTB58I. This insertion in the 58K coding sequence does not affect the reading frame. For construction of pTBGNSII, pTBHM60 (Van Bokhoven *et al.*, 1990) was partially digested with NdeI at position 1184 and filled in with Klenow fragment of DNA polymerase. In the filled in NdeI site a SmaI-linker (5'-CCCCCGGGGG-3') was inserted resulting in pTBGNSII. Double mutant pTBGNSΔ5a was constructed by replacing the PstI-SphI fragment (positions 345 to 1336) from pTBΔ5a with the PstI-SphI fragment of pTBGNSII.

In vitro transcription and translation

For *in vitro* transcription 1 μg plasmid DNA in 50 μl was treated with 2 μl (5 mg/ml) RNAse A and incubated for 10 min at 50°. Subsequently, 1 μl (4 mg/ml) proteinase K and 2.5 μl 10% (w/v) SDS were added followed by incubation for 10 min at 37°. The DNA was purified by phenol-chloroform extraction and the recovered DNA was precipitated with polyethylene glycol and recovered again as described (Sambrook *et al.*, 1989). The recombinant plasmids were linearized with ClaI, and subsequently, linear DNA templates were used for production of run-off RNA transcripts. *In vitro* transcription and recovery of transcripts were performed as described (Eggen *et al.*, 1989). Aliquots of transcription mixture were translated in rabbit reticulocyte lysate (Green Hectares, Oregon, Wisconsin) as described (Vos *et al.*, 1988a).

Complementation assay

Translation products obtained from separately translated BΔ32-RNA and B32S-RNA were mixed in a 1:1 ratio. The mixture was incubated at 30° for 1 hr. M-polyproteins were processed with a threefold excess of either unlabelled or ^{35}S-labelled translation products from either jointly translated or separately translated BΔ32-RNA and B32S-RNA. Incubation continued at 30° for 16 hr.

Analysis of proteins

Protein samples were heated in sample buffer (10% glycerol, 5% β-mercaptoethanol, 2% SDS, 0.01% bromophenol blue, 75 mM Tris-HCl pH 6.8) for 3 min at 100° and fractionated in a SDS-polyacrylamide gel (7.5%, unless indicated

otherwise) according to Laemmli (1979). Finally gels were dried and autoradiographed with Kodak X-Omat film.

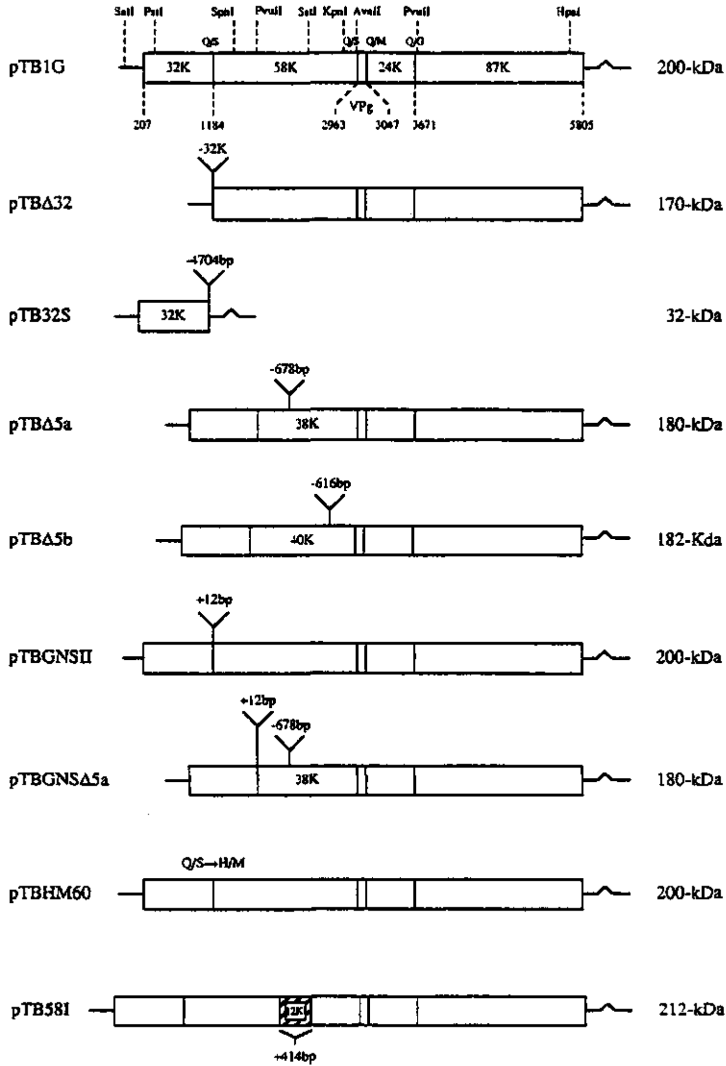


Fig. 2. Schematic representation of the various mutant cDNA clones derived from the full-length cDNA clone of CPMV B-RNA, pTB1G. The double-lined bars indicate the open reading frame and the single lines indicate the non-coding sequences. The positions of the viral proteins, the cleavage sites, and the restriction enzyme sites used in the construction of the clones are shown in the scheme of pTB1G. Mutated cleavage sites in clone pTBHM60, the size of the insertions in clones pTBGNSII, pTBGNSΔ5a, and pTB58I (hatched) and deletions in clones pTBΔ32, pTB32S, pTBΔ5a, and pTBGNSΔ5a are indicated. Predicted molecular masses of the

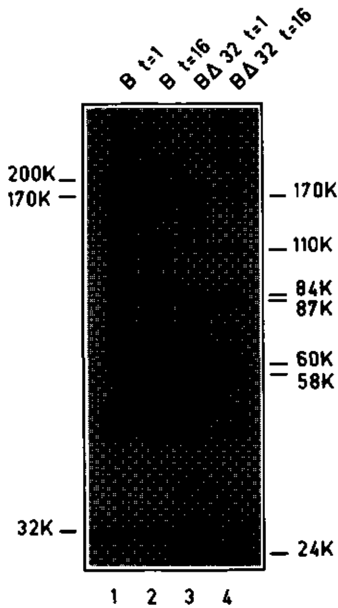
polyproteins are shown on the right. For details on the construction of the clones, see Materials and Methods.

Results

A role of the 32K protein in processing of both CPMV polyproteins

To study the possible role of the B-RNA encoded 32K protein in polyprotein processing we designed pTBΔ32, lacking the entire 32K coding sequence (see Fig. 2), and compared the processing of the primary translation product of the mutant transcripts with the processing of the 170K protein from B-RNA. B-RNA transcripts were translated into 200K polyproteins which were rapidly processed into primary cleavage products of 170K and 32K (Fig. 3A, lane 1). Processing at secondary cleavage sites in the wild-type 170K protein proceeded very slowly. Even after incubation for 16 hr the 84K and 87K cleavage products, the latter one migrating faster in the 7.5% SDS-polyacrylamide gel, and 60K and 110K cleavage products were barely visible (Fig. 3A, lane 2). In contrast the larger part of the 170K proteins from BΔ32-RNA were efficiently processed after 1 hr of incubation (Fig. 3A, lane 3) into 110K, 84K, 87K, 60K, 58K and 24K processing products which comigrated with the wild-type processing products. This suggests that the 170K protein from pTBΔ32 was processed at the same sites as wild-type 170K protein. Moreover, this result indicates that the presence of the 32K protein drastically reduces the rate of cleavage in the 170K protein.
Subsequently, the translation products from B-RNA and BΔ32-RNA were tested for their ability to process the M-polyproteins. Upon the addition of translation products from B-RNA to the 105K and 95K M-RNA translation products, the M-polyproteins were efficiently cleaved at the Gln/Met site into the 60K capsid precursor and the 58K and 48K proteins (Fig. 3B, lanes 1 and 2). Translation products from BΔ32-RNA did not process the M-polyprotein (Fig. 3B, lane 3), which is in agreement with previous data showing that for *trans* cleavage of the M-polyprotein the 32K protein is required (Vos *et al.*, 1988a). Thus, the 32K protein is involved in processing of both CPMV polyproteins.
To study the cleavage efficiency at different cleavage sites, samples were taken from a BΔ32-RNA translation mixture at various times after starting the translation. The results of the time course experiment show that full-length 170K protein was detectable after 25 min of translation (Fig. 4, lane 3). The 84K and 87K proteins could be observed also after 25 min of incubation, whereas the 60K and 110K proteins were detectable after 45 min of incubation (Fig. 4, lane 5). Furthermore, a 58K protein became visible after 60 min of translation (Fig. 4, lane

6). It appears that the cleavage generating the 84K/87K proteins is more rapid than the 60K/110K cleavage. This has also been observed for the processing of 170K from B-RNA in the presence of the 32K protein (Franssen *et al.*, 1984a).



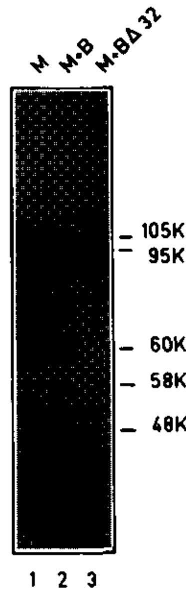


Fig. 3. (A) Involvement of the 32K protein in processing of CPMV B-polyproteins. *In vitro* translations of B-RNA (lanes 1 and 2) and BΔ32-RNA (lanes 3 and 4). Positions of B-polyprotein processing products are indicated. Translation products are shown after 1 hr (lanes 1 and 3) and 16 hr (lanes 2 and 4). (B) Involvement of the 32K protein in *trans* processing of the CPMV M-polyproteins. M-RNA was translated in presence of [^{35}S]methionine (lane 1) and after 2 hr a fourfold excess of unlabelled translation products from B-RNA (lane 2) and BΔ32-RNA (lane 3) was added. Incubation continued at 30° for 16 hr.

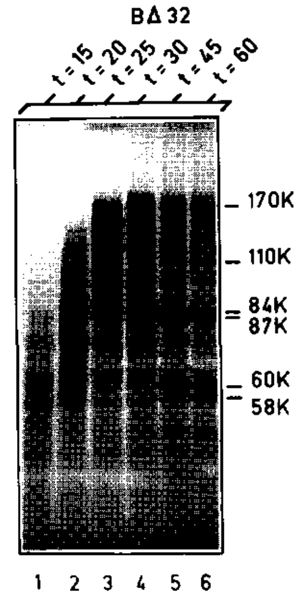


Fig. 4. Time course of translation of BΔ32-RNA. Samples were taken at 15, 20, 25, 30, 45 and 60 min after start of protein synthesis (lanes 1 to 6). To the right the positions of the 170K primary translation product and its processing products are indicated.

Involvement of the 58K domain in processing of the B-polyprotein

To study the role of the 58K domain in the proteolytic processing of B-proteins an insertion and deletions were introduced in the 58K coding sequence of pTB1G (see Fig. 2). Translation of the BΔ5a-RNA, lacking the coding sequence of the central part of the 58K protein, resulted in a 180K primary translation product. The primary processing of this protein was partly blocked (Fig. 5, lanes 8 and 9).

Translation of B58I-RNA resulted as expected in a 212K primary translation product. This mutant has an in frame insertion in the C-terminal hydrophobic domain of the 58K protein. The effect of this insertion was that the 212K polypeptide was hardly cleaved into proteins of 32K and 180K and even upon prolonged incubation an accumulation of the 32K protein could not be detected (Fig. 5, lanes 10 and 11). The 182K mutant polyprotein from BΔ5b-RNA has a deletion in the the hydrophobic C-terminal part of the 58K protein. The deletion had a drastic effect on the primary processing because release of the 32K protein from the primary translation product was not found even upon prolonged incubation (Fig. 5, lanes 12 and 13). This indicates that accurate and efficient primary processing requires a proper conformation of the polypeptide chain which clearly is disturbed in the insertion and deletion mutants. The results support the notion that the first cleavage of the 200K primary translation product into 32K and 170K is an intramolecular event. In contrast to the effect on the primary processing of polyproteins from B58I, BΔ5a, and BΔ5b-RNA processing at secondary cleavage sites was more efficient than was observed for the wild-type B-polyprotein (Fig. 5, lanes 8 to 13). This suggests that the 58K domain aids in the folding of the 170K proteins such that the rate of cleavages at secondary cleavage sites is reduced.

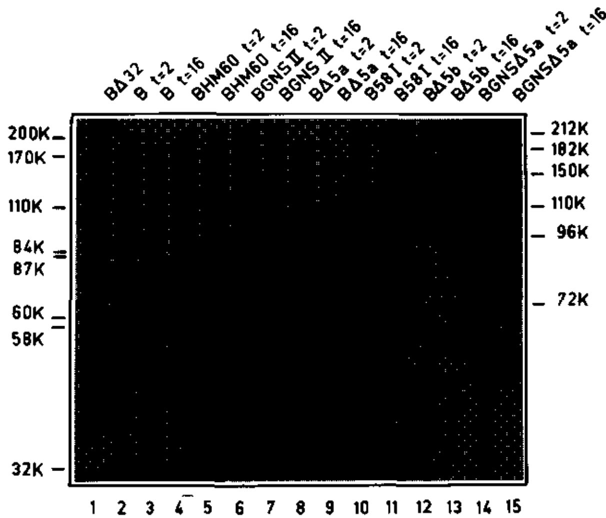


Fig. 5. Proteolytic processing of translation products containing mutations in the 58K domain of CPMV B-polyprotein. The nomenclature of the constructs, shown on top of the autoradiogram, is described under Materials and Methods. At the left side positions of the BΔ32-RNA (lane 1) and B-RNA (lanes 1 and 2) translation products and processing products are indicated. At the right side the positions of primary translation products and processing products from transcripts from the various mutagenized cDNA clones are indicated. Translation products are shown after 2 hr (lanes 2, 4, 6, 8, 10, 12, and 14) and 16 hr (lanes 3, 5, 7, 9, 11, 13, and 15).

Involvement of the 58K domain in M-polyprotein processing

In previous studies it has been demonstrated that the 24K proteinase, in presence of the 32K protein, cleaves in *trans* at the Gln/Met site in the M-polyproteins. This cleavage was successful when the 24K proteinase and the 32K protein originated from the same polyprotein. When a 32K deletion mutant was complemented with an intact 32K protein, cleavage was not observed (Vos *et al.*, 1988a). This suggested a specific interaction between the 170K and 32K in order to generate a complex active in M-polyprotein processing. To study this possibility, RNA transcripts from pTBΔ32 and pTB32S were jointly and separately translated. The separate translations were mixed to allow interaction of 32K and 170K proteins. Subsequently, translation products were tested for their ability to process the M-polyproteins. M-RNA was efficiently translated in 105K and 95K polyproteins (Fig. 6, lane 7). The addition of translation products from wild-type B-RNA resulted in complete processing into the 60K capsid precursor and 58K and 48K proteins (Fig. 6, lane 8). Translation products from BΔ32-RNA were not able to process 105K and 95K as described earlier (see Fig. 3). B32S-RNA was efficiently translated into the 32K protein (Fig. 6, lane 1). Products obtained from separately translated BΔ32-RNA and B32S-RNA could not process the M-polyproteins at the Gln/Met site (Fig. 6, lanes 5 and 6). However products obtained from jointly translated RNA transcripts processed the M-polyproteins and generated the 60K capsid precursor and 58K and 48K proteins (Fig. 6, lanes 2 and 3). Apparently, an active 32K/170K complex is generated before the translation of 170K is completed and cannot be formed from separately synthesized 32K and 170K proteins.

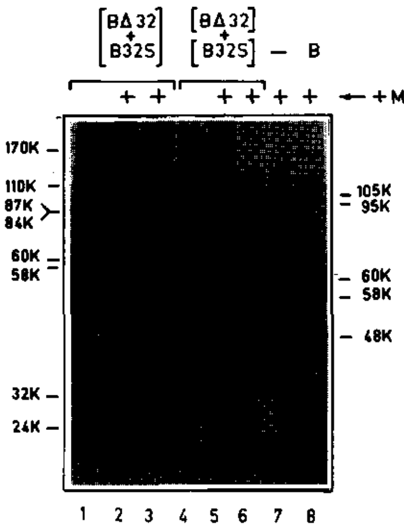


Fig. 6. Complementation for CPMV M-polyprotein processing with translation products obtained from BΔ32-RNA and B32S-RNA. M-RNA was translated in presence of [^{35}S]methionine (lane 7) and after 2 hr a fourfold excess of unlabelled (lanes 3 and 6) or ^{35}S-labelled translation products from BΔ32-RNA and B32S-RNA (lanes 2 and 5), either jointly translated (lanes 1 to 3), or separately translated (lanes 4 to 6), and with unlabelled translation products obtained from B-RNA as a control (lane 8) was added. Incubation was continued at 30° for 16 hr. The proteins were fractionated on a SDS-polyacrylamide gel (10%). Positions of translation products of BΔ32-RNA and B32S-RNA are indicated

on the left. Translation and processing products of M-RNA are indicated on the right.

To study the involvement of the 58K domain of 170K in this interaction, M-polyproteins were processed with translation products from transcripts with mutations in the 58K coding region. The addition of proteins obtained from BΔ5a-RNA with a deletion in the central part of the 58K domain resulted in M-polyprotein processing (Fig. 7, lane 4). However, transcripts from cDNA clones with a deletion (pTBΔ5b) or an insertion in the hydrophobic C-terminal part of the 58K protein (pTB58I) produced proteins which were not able to process the M-polyproteins (Fig. 7, lanes 3 and 5). This suggests that the hydrophobic region of the 58K protein is involved in the interaction with the 32K protein.

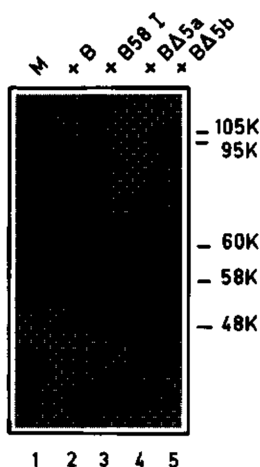


Fig. 7. Proteolytic processing of CPMV M-polyproteins by translation products of B transcripts with mutations in the 58K coding region. M-RNA was translated in presence of [^{35}S]methionine (lane 1) and after 2 hr a fourfold excess of unlabelled translation products of the mutants was added as indicated on top of the autoradiogram (lanes 2 to 5). Incubation continued at 30° for 16 hr. To the right the positions of M-RNA encoded proteins are indicated.

Release of the 32K protein from the B-polyprotein is not required for M-polyprotein processing

Cleavage of 32K from the primary translation product of BΔ5b-RNA and B58I-RNA was affected in these mutants (see Fig. 5). This might prevent the 32K protein from interacting with the 58K domain and forming a functional complex for M-polyprotein processing. To examine whether the 32K/170K cleavage has to occur first before a proper and functional proteinase complex can be formed capable of cleaving the M-polyproteins, several cleavage site mutants were constructed. In pTBHM60 the cleavage site at the 32K/170K junction has been changed from Gln/Ser into His/Met (see Fig. 2). In this mutant primary cleavage was only partly blocked and release of 32K was still rather efficient (see Fig. 5, lanes 4 and 5). Even the insertion of four amino acids between these His and Met residues (pTBGNSII) was not enough to block the primary cleavage completely although

processing at this site was less efficient than for the polyprotein from BHM60-RNA (Fig. 5, lanes 6 and 7). The primary cleavage was completely blocked in the translation product from BGNSΔ5a (Fig. 5, lanes 14 and 15), which contained in addition to an altered cleavage site a deletion in the 58K protein region, whereas the single mutations did not prevent cleavage of 32K from the primary translation product. It appeared that translation products obtained from BΔ5a-RNA, BGNSII-RNA as well as the double mutant BGNSΔ5a were able to process the M-polyproteins into the 60K capsid precursor and 58K and 48K proteins (Fig. 8, lanes 2 to 4). This proves that release of the 32K protein is not required for the formation of an active 32K/170K complex. Therefore, the failure to process M-polyproteins by the translation products from BΔ5b-RNA with a deletion in the hydrophobic C-terminal end of the 58K protein was probably due to deficient interaction of the 32K protein with the 58K protein.

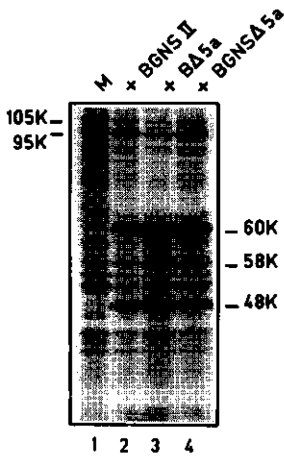


Fig. 8. Proteolytic processing of CPMV M-polyproteins by translation products from BGNSII-RNA, BΔ5a-RNA, and BGNSΔ5a-RNA. M-RNA was translated in presence of [^{35}S]methionine (lane 1) and after 2 hr a fourfold excess of unlabelled translation products of the mutants was added as indicated on top of the autoradiogram (lanes 2 to 4). Incubation continued at 30° for 16 hr. To the right the position of the M-RNA encoded proteins are indicated.

Discussion

In this paper we have shown that the B-RNA encoded 32K protein plays a major role in the regulation of processing of the polyproteins from both B-RNA and M-RNA. The initial processing event in the B-polyprotein resulting in the release of the 32K protein is an efficient process and can happen already cotranslationally. Results obtained from mutants with an insertion or a deletion in the 58K coding region indicate that the efficiency of primary cleavage is affected when the proper position of the cleavage site with respect to the active site of the 24K proteinase is changed. These observations support the notion that the primary processing of the B-polyprotein into the 32K and 170K protein is most efficient through *cis* cleavage. *Trans* processing of a 92K protein (32K+60K) at the Gln/Ser cleavage site can occur, but is far less efficient than intramolecular cleavage (Peters *et al.*, 1992). The observation that *trans* processing at the Gln/Ser site

releasing the 32K protein from a full-length B-polyprotein containing an inactive protease was inefficient also supports this notion (Vos *et al.*, 1988a).
Primary cleavage at the Gln/Ser site is not prevented when this site is changed into a His/Met site or by insertion of four amino acids between the His and Met residues. Release of the 32K protein in the His/Met mutant was still efficient and was achieved through cleavage between the His and Met residues. This was verified by amino acid sequence analysis of the 170K cleavage product (data not shown). The 170K cleavage product from BGNSII-RNA has not been sequenced, but considering the mobility of the 32K protein on SDS-polyacrylamide gel, it is likely that cleavage has occurred between His and Thr residues. These results indicate that primary cleavage site recognition requirements in the 200K polyprotein are not stringently determined by the specific Gln/Ser dipeptide sequence. The conformation around the cleavage site and the accessibility of the site are probably also important for an efficient *cis* cleavage of the 32K protein. Apparently during synthesis of the B-polyprotein folding of the polypeptide chain drives the active site of the 24K proteinase and the primary cleavage site together and favours a rapid intramolecular cleavage of the 32K protein in the nascent polypeptide chain before other cleavages in the polyprotein occur.
We have clearly demonstrated the inhibitory function of the B-RNA encoded 32K protein in the B-polyprotein processing. The 170K protein of BΔ32-RNA (in the absence of 32K) was processed within 1 hr, although a small part of the 170K protein remained intact, even upon prolonged incubation. Presumably these 170K proteins were misfolded during translation in the reticulocyte lysate and were processed either very slowly or not at all. The 170K protein from B-RNA (in presence of the 32K protein) was hardly processed at secondary cleavage sites even upon prolonged incubation. However, it is important to note that the fraction of 170K from B-RNA that was processed varied from experiment to experiment and differed between different reticulocyte lysate preparations (Franssen *et al.*, 1984a; Peng and Shih, 1984; Vos *et al.*, 1988a). Presumably the 32K functions by interacting with the 58K domain of the 170K protein. This is based on the observation that polyproteins with an insertion or deletion in the 58K portion were more efficiently processed at secondary cleavage sites than the wild-type B-polyprotein, and on previous observations that in infected protoplasts the 32K protein is found associated with 58K, 60K, 84K and 170K proteins which all have the 58K domain in common (Franssen *et al.*, 1984b). Moreover, results obtained with BΔ4-RNA, which lacks almost the complete 87K coding region, show that the 87K domain of the 170K protein has no obvious role in the processing (Peters *et al.*, 1992; Vos *et al.*, 1988a).

At present it is not known by what mechanism 32K regulates the processing of 170K at secondary cleavage sites. It is possible that the association between 32K and the 58K domain results in a conformation of the 170K protein such that secondary cleavage sites are not accessible for the 24K proteinase. *In vivo* the 170K protein also appears to be very slowly processed, since it is the most abundant protein in fractions prepared from CPMV-infected cowpea leaves (G. Rezelman unpublished observation). This delayed processing might represent a step in the regulation of production of active proteins. Possibly BΔ32-RNA is not infectious in cowpea protoplasts (data not shown) because assembly of a functional viral RNA replication complex requires the interaction between the 32K and 170K proteins. The role we propose for the 32K protein is to block a certain folding pathway for the 170K protein that could lead to the formation of nonfunctional structures as a result of premature cleavages in the 170K protein. In this sense we feel that the 32K protein is related to a class of proteins which have been described and defined by Ellis and Hemmingsen (1989) and Ellis (1991) as molecular chaperones.

Whereas the complexed 32K protein arrests the *cis* proteolytic cleavage of B-polyproteins, it is required for *trans* cleavage of the M-polyprotein at the Gln/Met site by the 24K proteinase. M-polyproteins were efficiently processed by B-RNA translation products, but in absence of 32K M-polyprotein processing could not be achieved. The data we have presented indicate that a 32K/170K complex cleaves the M-polyproteins at the Gln/Met site. Also proteins from BΔ4-RNA were able to process M-polyproteins (Vos *et al.*, 1988a). This shows that the 87K domain does not play an essential role in the formation of a functional proteinase complex. Translation products from BΔ5a-RNA with a deletion in front of the hydrophobic region in the C-terminal part of the 58K protein were able to process M-polyproteins. However, translation products from BΔ5b-RNA lacking the hydrophobic region, and from B58I-RNA with an insertion in the hydrophobic region, did not process M-polyproteins. These observations indicate that the 32K protein interacts with this hydrophobic region in the 58K protein. It appears that the conformation in the 170K protein that is needed to generate the proteinase for *trans* processing of M-polyproteins is directed by the 32K protein. Remarkebly, the results obtained with BGNSΔ5a showed that a complex active in M-polyprotein processing can be formed without the release of the 32K protein from the B-polyprotein.

The complementation assay for the 32K and 170K protein showed that for M-polyprotein processing both proteins do not necessarily need to originate from the same precursor protein. However, only in experiments in which translation products from jointly translated B32S-RNA and BΔ32-RNA were used, processing of

M-polyproteins at the Gln/Met site has been observed, whereas it has not been possible to reconstitute activity from separately translated molecules. This observation indicates that interaction, required for cleavage, has to take place between the 32K protein and a growing 170K polypeptide chain. Perhaps this reflects a critical time point for interaction and again points towards a dependence of proteolytic activity on a proper protein conformation.
Recently, it was reported that for processing of the Dengue Virus (DEN) polyprotein (flaviviridae) two non-structural proteins NS2B and NS3 are required for *trans* cleavages and it was proposed that NS2B can act as a cofactor to promote the proteolytic activity of NS3 (Falgout *et al.*, 1991). In this respect it resembles the "two component proteinase" activity required for CPMV M-polyprotein processing.

Acknowledgements

We thank Konrad Beyreuther for the amino acid sequence analysis, Hans van Bokhoven, Richard Milek, Gerhard van Prooijen, Gert Scheper and Ruud Verkerk for providing us with cDNA clones, and Piet Madern and Peter van Druten for artwork and photography. This work was supported by the Netherlands Foundation of Chemical Research (SON) with financial aid from the Netherlands Organization for Scientific Research (NWO).

References

Bu, M., and Shih, D.S. (1989). Inhibition of proteolytic processing of the polyproteins of cowpea mosaic virus by hemin. *Virology* **173**, 348-351.

Dorssers, L., Van der Krol, S., Van der Meer, J., Van Kammen, A., and Zabel, P. (1984). Purification of cowpea mosaic virus RNA replication complex: Identification of a virus-encoded 110,000-dalton polypeptide responsible for RNA chain elongation. *Proc. Natl. Acad. Sci. USA.* **81**, 1951-1955.

Eggen, R., and Van Kammen, A. (1988). RNA replication in comoviruses. *In* "RNA Genetics" (P. Ahlquist, J. Holland, and E. Domingo, Eds.), Vol. 1, pp. 49-69. CRC Press, Boca Raton.

Eggen, R., Verver, J., De Jong, A., Goldbach, R., and Van Kammen, A. (1989). Improvements of the infectivity of *in vitro* transcripts from cloned cowpea mosaic virus cDNA: impact of terminal nucleotide sequences. *Virology* **173**, 447-455.

Ellis, J.R. (1991). Chaperone function: cracking the second half of the genetic code. *Plant J.* **1**, 9-13.

Ellis, J.R., and Hemmingsen, M. (1989). Molecular chaperones: proteins essential for the biogenesis of some macromolecular structures. *Trends Biochem. Sci.* **14**, 339-342.

Falgout, B., Pethel, M., Zhang, Y., and Lai, C. (1991). Both nonstructural proteins NS2B and NS3 are required for the proteolytic processing of Dengue virus nonstructural proteins. *J. Virol.* **65**, 2467-2475.

Franssen, H., Goldbach, R., and Van Kammen, A. (1984a). Translation of bottom component RNA of cowpea mosaic virus in reticulocyte lysate: faithful proteolytic processing of the primary translation product. *Virus Res.* **1**, 39-49.

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., Moerman, M., Rezelman, G., and Goldbach, R. (1984b). Evidence that the 32.000-Dalton protein encoded by bottom component RNA of cowpea mosaic virus is a proteolytic processing enzyme. *J. Virol.* **50**, 183-190.

Garcia, J.A., Schrijvers, L., Tan, A., Vos, P., Wellink, J., and Goldbach, R. (1987). Proteolytic activity of cowpea mosaic virus encoded 24K protein synthesized in *Escherichia coli*. *Virology* **159**, 67-75.

Goldbach, R., Rezelman, G., and Van Kammen, A. (1980). Independent replication and expression of B-component RNA of cowpea mosaic virus. *Nature (London)* **286**, 297-300.

Korneluk, R.G., Quan, G., and Gravel, R.A. (1985). Rapid and reliable dideoxy sequencing of double-stranded DNA. *Gene* **40**, 317-323.

Kunkel, T.A. (1985). Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl. Acad. Sci. USA* **82**, 488-492.

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

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

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

Peng, X.X., and Shih, D.S. (1984). Proteolytic processing of the proteins translated from the bottom component RNA of cowpea mosaic virus. *J. Biol. Chem.* **259**, 3197-3201.

Peters, S.A., Voorhorst, W.G.B., Wellink, J., and Van Kammen, A. (1992). Processing of VPg-containing polyproteins encoded by the B-RNA from cowpea mosaic virus. *Virology* **191**, 90-97.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). "Molecular Cloning: A laboratory manual", 2nd ed. Cold Spring Harbour Laboratory, Cold Spring Harbor, New York.

Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467.

Van Bokhoven, H., Wellink, J., Usmany, M., Vlak, J.M., Goldbach R., and Van Kammen, A. (1990). Expression of plant virus genes in animal cells: high-level synthesis of cowpea mosaic virus B-RNA-encoded proteins with baculovirus expression vectors. *J. Gen. Virol.* **71**, 2509-2517.

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.

Verver, J., Goldbach, R., Garcia, J.A., and Vos, P. (1987). *In vitro* expression of a full-length DNA copy of cowpea mosaic virus B RNA: identification of the B RNA encoded 24-kd protein as a viral protease. *EMBO J.* **6**, 549-554.

Vos, P., Jaegle, M., Wellink, J., Van Kammen, A., and Goldbach, R. (1988a). Two viral proteins involved in the proteolytic processing of the cowpea mosaic virus polyproteins. *Nucleic Acids Res.* **16,** 1967-1985.

Vos, P., Jaegle, M., Wellink, J., Verver, J., Eggen, R., Van Kammen, A., and Goldbach, R. (1988b). Infectious RNA transcripts derived from full-length DNA copies of the genomic RNAs of cowpea mosaic virus. *Virology* **165,** 33-41.

Wellink, J., Rezelman, G., Goldbach, R., and Beyreuther, K. (1986). Determination of the proteolytic processing sites in the polyprotein encoded by the bottom-component RNA of cowpea mosaic virus. *J. Virol.* **59,** 50-58.

Wellink, J., and Van Kammen, A. (1989). Cell-to-cell transport of cowpea mosaic virus requires both the 58K/48K proteins and the capsid proteins. *J. Gen. Virol.* **70,** 2279-2286.

Chapter 3

Processing of VPg-Containing Polyproteins Encoded by the B-RNA from Cowpea Mosaic Virus

Sander A. Peters, Wilfried G.B. Voorhorst,
Joan Wellink, and Ab van Kammen

Department of Molecular Biology, Agricultural University, Dreijenlaan 3, 6703 HA Wageningen, The Netherlands

Running title: Expression of B-RNA from CPMV

Virology 191, 90-97, 1992.

Processing of VPg-containing polyproteins encoded by the B-RNA from Cowpea Mosaic Virus

Abstract

To study the processing of putative VPg precursors the expression of specific mutant transcripts derived from a full-length cDNA clone of cowpea mosaic virus (CPMV) B-RNA was examined in a rabbit reticulocyte lysate system. This study revealed that the 170K protein produced by a B-RNA mutant that lacks the 32K coding region was efficiently processed by mainly intramolecular cleavages at three different sites into three sets of proteins of 60K+110K, 84K+87K and 58K+112K. Further cleavage of the 60K protein into 58K and VPg has not been observed in this *in vitro* system. The 84K protein can be further processed by an intramolecular cleavage reaction via two alternative pathways either into 26K (VPg+24K) and 58K proteins or into 24K and 60K proteins. VPg can be released from the 112K (VPg+110K) precursor either directly or via the 26K intermediate. Immunoblot analysis showed that the 112K protein is present in CPMV-infected plant cells indicating that the *in vitro* observations may hold true *in vivo*.

Introduction

Cowpea mosaic virus (CPMV) expresses its genetic information via synthesis of large polyproteins from which functional proteins are generated by proteolytic cleavages at specific Gln/Met, Gln/Ser, and Gln/Gly sites by the viral 24K proteinase (Wellink *et al.*, 1986; Garcia *et al.*, 1987). A detailed processing pathway for the B-RNA encoded 200K polyprotein has been established (see Fig. 1, Peters *et al.*, 1992) based on *in vitro* translation studies (Pelham, 1979; Goldbach *et al.*, 1981; Franssen *et al.*, 1984; Peng and Shih, 1984; Verver *et al.*, 1987; Vos *et al.*, 1988) and *in vivo* observations (Rezelman *et al.*, 1980; Goldbach *et al.*, 1982; Zabel *et al.*, 1982; Goldbach and Rezelman, 1983; Wellink *et al.*, 1986, 1987). A primary cleavage generates the 32K and 170K proteins. Subsequently, the 170K protein can be further cleaved via 110K, 84K and 60K intermediates ultimately into 58K, VPg (viral protein genome linked), 24K, and 87K proteins. *In vitro*, these secondary cleavages are only observed after prolonged incubation (Peng and Shih, 1984; Franssen *et al.*, 1984; Vos *et al.*, 1988) and are regulated by the 32K protein which is associated with the 170K protein in a complex (Peters *et al.*, 1992).

The B-RNA can replicate in infected protoplasts independently from M-RNA and encodes the proteins which are necessary for RNA replication (Goldbach *et al.*, 1980). The 110K protein which consists of the 24K proteinase and the 87K core polymerase has been identified as the viral RNA replicase (Dorssers *et al.*, 1984; Eggen and Van Kammen, 1988). The function of the 60K protein in the replication process is unclear. The 58K domain of the 60K protein contains amino acid sequence elements that show homology to cellular and viral helicases and maybe functions as a helicase in CPMV-RNA replication (Dever *et al.*, 1987; Gorbalenya and Koonin, 1989; Gorbalenya *et al.*, 1990). Furthermore the 60K protein contains VPg, the small protein covalently linked to the 5'-end of the viral RNAs, and has been proposed to be the direct precursor of VPg (Zabel *et al.*, 1982; Goldbach *et al.*, 1982). However, actual cleavage of the 60K protein into VPg and the 58K protein has not been observed in any system including standard reticulocyte lysate conditions (Franssen *et al.*, 1984; Vos *et al.*, 1988). VPg has been proposed to play a role in the initiation of viral RNA synthesis (Eggen and Van Kammen, 1988) but evidence for such a role is lacking.

In this paper we report experiments which suggest that the 60K protein is not the direct precursor of VPg. Alternatively, a 112K protein, which can be generated from the 170K protein via cleavage of the Gln/Ser site at the 58K/VPg junction, might function as a direct precursor of VPg. An extended model for the expression of CPMV B-RNA is presented.

Materials and methods

DNA techniques

All enzymes were purchased from Gibco BRL, New England Biolabs, and Pharmacia and used as described by the manufacturers. DNA fragments were isolated from agarose gels and purified with Gene Clean™ according to the manufacturer (Bio 101 Inc.). Standard recombinant DNA techniques were used for linker insertion, ligation and transformation in competent *E.coli* JM109 as described (Sambrook *et al.*, 1989). Recombinant clones were analysed by restriction enzyme mapping or nucleotide sequence analysis as described (Sanger *et al.*, 1977; Korneluk *et al.*, 1985). Oligodeoxyribonucleotides were synthesized with a cyclone DNA synthesizer (Biosearch).

Construction of mutagenized cDNA clones

As starting material for construction of the different mutant clones the full-length cDNA clone of B-RNA, pTB1G, was used (Eggen *et al.*, 1989). In this plasmid full-length cDNA is cloned behind a bacteriophage T7 RNA polymerase promoter. From the cDNA RNA molecules can be transcribed which have been shown to be infectious in cowpea protoplasts (Eggen *et al.*, 1989). The positions of restriction sites and nucleotides refer to the position in the B-RNA sequence determined by Lomonossoff and Shanks (1983).
The construction of cDNA clone pTBΔ32 has been described recently by Peters *et al.* (1992) and pTMB110 by Vos *et al.*, (1988). For construction of pTB60 and pTB92, the SstI-BamHI fragment (positions 2301 to 3857) from pTBΔ32 and pTB1G respectively was replaced by the SstI-BamHI fragment of pTBHM60stop (Van Bokhoven *et al.*, 1990). Both pTB60 and pTB92 have an in-frame stop codon just behind the 60K coding sequence. To construct pTBΔ32R87, pTBΔ32 was digested with Bsu36I and XbaI at positions 3545 and 4119. The small fragment (nucleotides 3545 to 4119) was replaced with the Bsu36I-XbaI fragment of pEXC-3C-110, QG→R (Richards *et al.*, 1989) resulting in pTBΔ32R87. Plasmid pTBΔ32HM24 was generated by exchanging the SalI-SstI fragment of pTBΔ32 (positions -40 to 2301) with the SalI-SstI fragment of pTBHM110 (Van Bokhoven *et al.*, 1990). For a further analysis on the processing of the 84K protein also pTBΔ4, in which a 1961-bp PvuII-HpaI fragment (positions 3700 to 5661) was removed (previous pTB114Δ4, Vos *et al.*, 1988), was also used in this study. For construction of pTB84, the SalI-SstI fragment (positions -40 to 2301) from pTBΔ4 (previously pTB114Δ4, Vos *et al.*, 1988) was exchanged with the SalI-SstI fragment of pTBΔ32. To inactivate the 24K proteinase, pTBΔ32 was partially digested with SphI at position 3156 and protruding ends were removed with Klenow fragment of DNA polymerase. A SmaI-linker (5'-CCCCCGGGGG-3') was inserted in the blunted SphI site resulting in pTB24SSΔ32. Deletion of the PvuI-HpaI fragment (positions 3700 to 5661) of pTB24SSΔ32 generated pTB84N.

In vitro transcription and translation

The procedures for *in vitro* transcription and translation were the same as described before (Peters *et al.*, 1992).

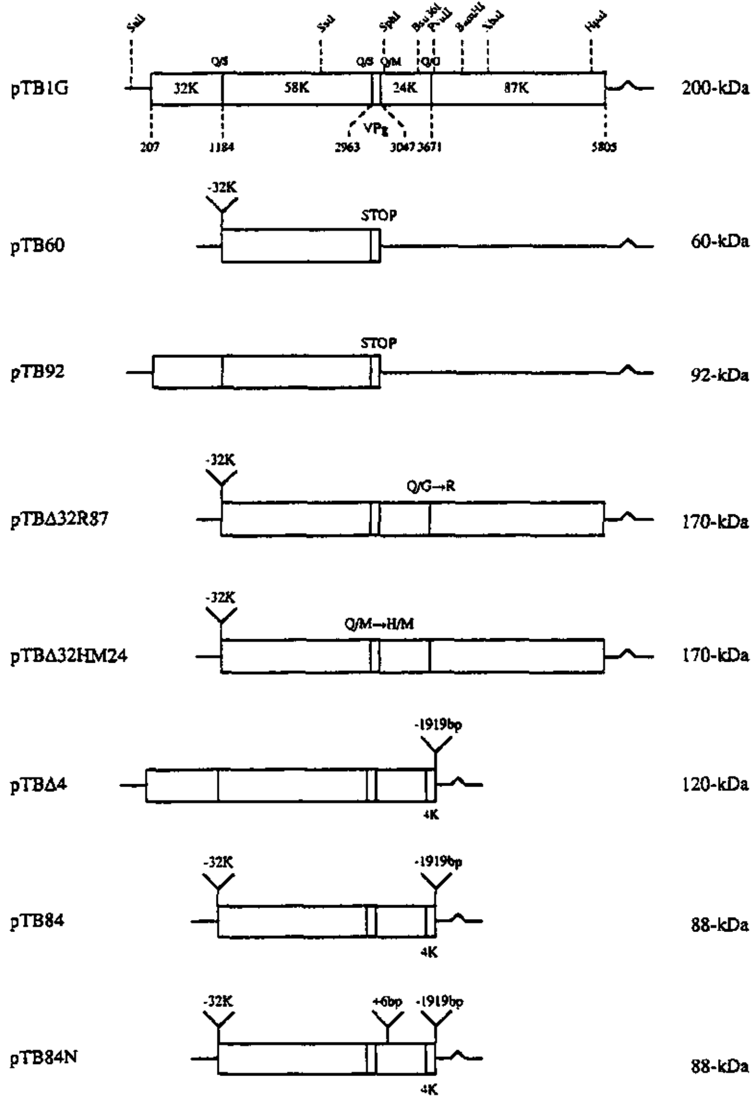


Fig. 1. Schematic representation of the mutant cDNA clones used in this study, derived from the full-length cDNA clone of CPMV B-RNA, pTB1G. The double-lined bars indicate the open reading frame and the single lines indicate the non-coding sequences. The positions of the viral proteins, the cleavage sites, and the restriction enzyme sites used to construct the clones are shown in the scheme of pTB1G. Mutated cleavage sites in pTBΔ32 and pTBΔ32HM24, introduced stop codons in pTB60 and pTB92, the size of the insertion in pTB84N, and deletions in pTB60, pTBΔ4, pTB84 and pTB84N are indicated. Predicted molecular masses of the polyproteins are indicated. For details on the construction of the clones see Materials and Methods.

Immunological methods and gel electrophoresis

In vitro translated proteins were immunoprecipitated using rabbit anti-24K (Wellink *et al.*, 1987) and anti-VPg serum (Eggen *et al.*, 1988) as described (Franssen *et al.*, 1982). Extracts from CPMV-infected and mock-inoculated cowpea leaves were prepared as described (Franssen *et al.*, 1982). Protein samples were heated in sample buffer (10% glycerol, 5% β-mercaptoethanol, 2% SDS, 0.01% bromophenol blue, 75mM Tris-HCl pH 6.8) for 3 min at 100° and fractionated in a SDS-polyacrylamide gel (7.5%, unless indicated otherwise) according to Laemmli (1979). Gels were either dried and autoradiographed with Kodak X-Omat film or analysed by immunoblotting using anti-VPg or anti-24K serum with anti-rabbit IgG-alkaline phosphatase (Promega Biotec) as a second antibody as described (Blake *et al.*, 1984).

Results

The 60K protein is not the precursor for VPg *in vitro*

Upon processing of the 170K translation product of BΔ32-RNA transcribed from pTBΔ32, the 58K protein was produced in a substantial amount (Fig. 2, lane 1, Peters *et al.*, 1992). As shown in the processing scheme of Fig. 1 (Peters *et al.*, 1992), the 58K protein can be produced from different processing intermediates by cleavage at a Gln/Ser site. If 58K arises from the 60K precursor, VPg will be released at the same time. To examine whether the 60K protein is indeed the precursor of VPg and the 58K protein, we designed pTB60 and pTB92 (see Fig. 1). The primary translation products from B60-RNA and B92-RNA were processed with the translation products from MB110-RNA (Vos *et al.*, 1988) in which the 24K proteinase mainly occurs as a free protein and, on the other hand, with the translation products from B-RNA or BΔ32-RNA in which the 24K proteinase activity is contained in 170K, 110K, or 84K precursor polypeptides. *In vitro* translation of RNA transcripts from pTB60 resulted in accumulation of a 60K protein comigrating with the 60K protein from pTBΔ32 (Fig. 2, lane 2). A protein migrating slightly slower than the 58K protein from pTBΔ32 was also produced. This protein was degraded upon prolonged incubation and presumably resulted from internal initiation of translation on the B60-RNA. Incubation with various proteinases did not result in cleavage of the 60K into 58K and VPg, although a slight decrease in the amount of 60K was observed (Fig. 2, lanes 2 to 5). This is probably the result of instablity of the 60K protein in the reticulocyte lysate. RNA transcripts from pTB92 were efficiently translated into 92K (32K+60K) proteins

(Fig. 2, lane 6). This protein was partially cleaved at the Gln/Ser site into 60K and 32K protein after incubation with the translation products from B-RNA, BΔ32-RNA or MB110-RNA (Fig. 2, lanes 7 to 9). The latter incubation resulted in the more efficient release of 32K and 60K proteins (Fig. 2, lane 9). However, the majority of the 92K protein was not cleaved indicating that *trans* processing of this protein at the Gln/Ser site was not very efficient. No further processing of the 60K protein, generated through cleavage of the 92K protein into 58K and VPg, was observed.

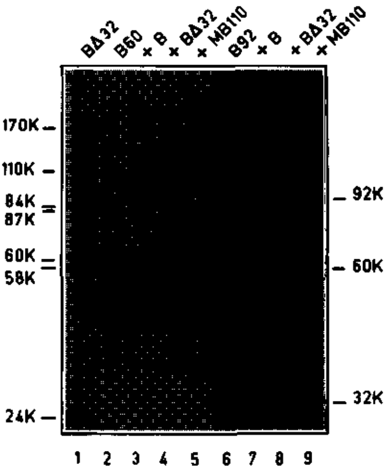


Fig. 2. *Trans* cleavage of 60K and 92K proteins from CPMV. B60-RNA (lanes 2 to 5) and B92-RNA (lanes 6 to 9) were translated in presence of [^{35}S]methionine. After 2 hr a fourfold excess of unlabelled translation products was added obtained from translations during 1 hr of B-RNA (lane 3 and 7), BΔ32-RNA (lane 4 and 8), and MB110-RNA (lane 5 and 9). Incubation was continued for 16 hr at 30°. As a marker positions of proteins of BΔ32-RNA (lane 1) are indicated on the left. Positions of processing products are indicated on the right.

Since processing of the 170K protein from BΔ32-RNA, in which the 32K coding region has been deleted, readily produced the 58K protein we investigated whether the 58K protein produced in the processing of the 170K precursor arose by cleavage of a 60K protein, or, alternatively, was generated by cleavage from the 84K protein or any other processing intermediate. For that purpose we constructed mutant pTBΔ32R87 in which the Gln/Gly site at the right border of the 24K proteinase has been changed into a single Arg residue to prevent cleavage of the 170K protein into 84K and 87K proteins. Indeed accumulation of 87K, 84K, and 24K proteins was not observed. Upon translation of BΔ32R87-RNA, the 170K protein was detectable after 25 min of translation (Fig. 3, lane 2) just as with translation of BΔ32-RNA (Peters *et al.*, 1992). Ten minutes later, 60K and 110K proteins were detectable (Fig. 3, lane 3), whereas after 90 min of translation 58K and 112K proteins appeared (Fig. 3, lane 6). The 58K and 60K proteins comigrated with the 58K and 60K proteins from BΔ32-RNA, whereas the 110K protein, probably because of the mutation in this protein, migrated slightly faster than the 110K protein from pTBΔ32. Incubation of the translation mixture with the 24K proteinase from MB110-RNA (Fig. 3, lanes 9 and 10) did not result in a significant decrease of 60K or an increase in 58K proteins. This indicates that the 60K proteins

generated through cleavage of the Gln/Met site in the 170K proteins obtained from BΔ32R87-RNA were not further processed *in trans* at the Gln/Ser site into 58K and VPg. Furthermore, these results suggest that the 170K protein can be processed also via a pathway in which the first cleavage takes place at the Gln/Ser site at the 58K/VPg junction. The 58K protein was produced in a larger amount than the 112K protein (Fig. 3, lanes 2 to 8). It therefore appears that the 112K protein was further cleaved into VPg and 110K proteins.

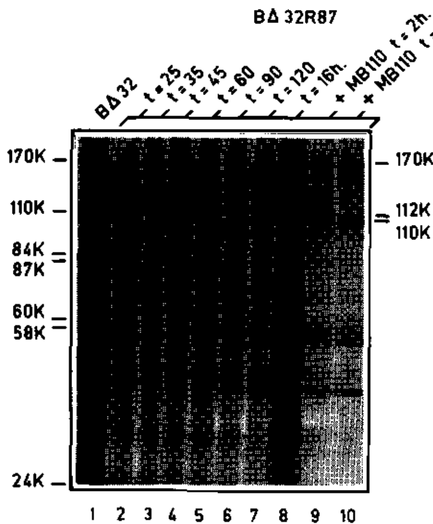


Fig. 3. Time course of translation and processing of 170K proteins obtained from BΔ32R87-RNA. Samples were taken at 25, 35, 45, 60, 90, 120 min and 16 hr after start of protein synthesis (lanes 2 to 8). After 2 hr of incubation a fourfold excess of unlabelled translation products from MB110-RNA was added and incubation was continued at 30° for 2 hr (lane 9), and 16hr (lane 10). Marker proteins from BΔ32-RNA (lane 1) are indicated on the left.

The 84K protein is preferentially cleaved *in cis* via two alternative pathways

In the previous section, we mentioned the possibility that the 84K protein serves as a precursor for VPg. In earlier processing studies by Vos *et al.* (1988) it was found that the 84K protein was not further processed. That result was obtained for a 120K protein (32K+84K+4K) encoded by BΔ4-RNA (previously pTB114Δ4), which was cleaved at the Gln/Ser site into 32K and 88K (84K+4K) proteins. It is possible that the inhibitory effect of the 32K protein on processing of the 84K protein is responsible for this result (Peters *et al.*, 1992). We have therefore constructed a number of new mutants in order to further examine the possible role of the 84K protein as a precursor of VPg. In cleavage site mutant pTBΔ32HM24 (see Fig. 1), the Gln/Met site at the left border of the 24K proteinase has been changed into a His/Met site in order to prevent cleavage of the 170K protein into 60K and 110K proteins. Mutants pTB84 and pTB84N (see Fig. 1) were constructed to study *cis* and *trans* processing of the 84K protein. Upon translation of RNA transcripts from pTBΔ32HM24 into 170K proteins, the accumulation of 84K and 87K proteins was

detected within 1 hr, comparable to that observed for BΔ32-RNA (Fig. 4A, lanes 1 and 3). This cleavage site mutation did not fully prevent processing of 170K into 60K and 110K proteins, but cleavage ocurred at a much lower rate than was observed for BΔ32. Considerable amounts of 58K and 112K proteins accumulated upon prolonged incubation and exceeded the amount of 110K and 60K proteins (Fig. 4A, lane 2). The identity of the 112K protein (VPg+110K) was confirmed by immunoprecipitation using antisera directed against the 24K polypeptide and VPg (data not shown). These 112K and 58K proteins must be generated from the 170K translation product through cleavage of the Gln/Ser site at the 58K/VPg junction. Furthermore, a decrease in the amount of 84K and an increase of 58K and 26K proteins was observed, whereas the amount of 87K protein did not increase (Fig. 4A, lane 2). The 26K protein reacted with both anti-24K and anti-VPg serum and therefore likely consists of VPg and the 24K proteinase (data not shown). It therefore appears that the 84K protein was further processed into 26K and 58K proteins through cleavage of the Gln/Ser bond at the 58K/VPg junction. Furthermore, for the BΔ32-RNA translation a decrease in the amount of 84K protein and an increase of 24K protein was observed whereas the 26K protein did not accumulate (Fig. 4A, lanes 3 and 4). This suggests that the 26K protein is rapidly further processed into VPg and 24K proteins.

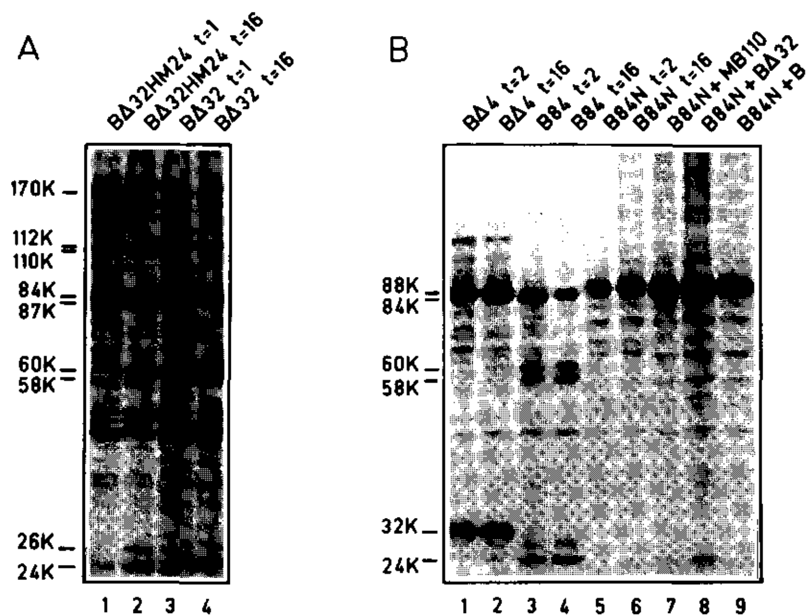


Fig. 4. Identification of CPMV B-RNA encoded 84K proteins as a precursor of VPg. (A) *In vitro* translations of RNA transcripts from pTBΔ32HM24 (lanes 1 and 2) and pTBΔ32 (lanes 3 and 4). Translations are shown after 1 hr and 16 hr incubation as indicated on top of the

autoradiogram. Positions of the translation and processing products are indicated on the left. Marker proteins from a BΔ32-RNA translation are indicated on the right. (B) *In vitro* translations of BΔ4-RNA (lanes 1 and 2), B84-RNA (lanes 3 and 4), and B84N-RNA (lanes 5 to 9). Positions of the translation and processing products are shown on the left. After 2 hr of incubation a fourfold excess of unlabelled translation products from MB110-RNA (lane 7), BΔ32-RNA (lane 8) or B-RNA (lane 9) was added to labelled translation products from B84N-RNA. Incubation continued at 30° for 16 hr. Translations are shown after 2 hr and 16 hr incubation as indicated on top of the autoradiogram.

Additional evidence for the processing of the 84K protein was obtained by comparing the proteolytic processing of polyproteins from BΔ4, B84, and B84N-RNA. Translation of BΔ4-RNA, lacking almost the complete coding sequence of the 87K protein, resulted in a 120K product (32K+84K+4K) which was rapidly processed into 32K and 88K (84K+4K) proteins (Fig. 4B, lane 1). Further incubation resulted in cleavage of the 88K polypeptide at the Gln/Gly site releasing the 4K peptide from the C-terminal end of the 84K protein, but further cleavage was not detected (Fig. 4B, lane 2 and Vos *et al.*, 1988). The 84K protein was also generated by cleavage at the Gln/Gly site of the 88K (84K+4K) primary translation product from B84-RNA, although more rapidly compared to that observed in BΔ4 (Fig. 4B, lanes 1 and 3). In addition, cleavage of the Gln/Ser and the Gln/Met site in the 88K and 84K protein generated the 60K, 58K, and the 24K proteins, whereas the 26K protein was not observed (Fig. 4B, lanes 3 and 4). It therefore appears that processing of 84K occurs either into 26K and 58K or into 24K and 60K proteins. Furthermore, these results again suggests that the 26K protein is further processed into VPg and 24K proteins. A small amount of what appears to be a 28K protein was also produced. This protein reacted with anti-24K serum only (data not shown) and probably consists of the 24K proteinase and the 4K peptide at the C-terminal end of the proteinase. To investigate *trans* processing of the 84K protein, translation products of B84N-RNA, which do not contain an active 24K proteinase, were incubated with translation products obtained from MB110, BΔ32, and B-RNA which show strong processing activity. The results showed that *trans* processing of the 88K primary translation product of B84N-RNA generated the 84K protein and a 4K peptide, but the 84K was not further processed into 60K, 58K and 24K proteins (Fig. 4B, lanes 5 to 9). Therefore the processing of 84K occurs only *in cis*. Moreover, the observation that 84K is only further processed in the absence of 32K again shows that 32K can somehow arrest the secondary processing of the B-proteins (Peters *et al.*, 1992).

Identification of 112K polyproteins in crude membrane fractions

To determine whether the 112K and 26K processing intermediates can also be found *in vivo*, an immunoblot analysis of proteins prepared from CPMV-infected and mock inoculated cowpea leaves was performed. In this analysis antisera directed against the 24K proteinase and the VPg were used. Figure 5 shows that 170K, 112K, and 110K, present in fractions prepared from CPMV-infected cowpea leaves, were immunoreactive with anti-24K serum (Fig. 5, lane 3). Using anti-VPg serum, viral proteins of 170K, 112K, and 60K were detected (Fig. 5, lane 2). These results indicate that the Gln/Ser cleavage generating 112K and 58K proteins also occurs *in vivo*. So far we have not been able to detect a 26K protein in infected cells. With anti-VPg serum a protein with a molecular mass of approximately 90-kDa was detected but the origin of this protein is unclear.

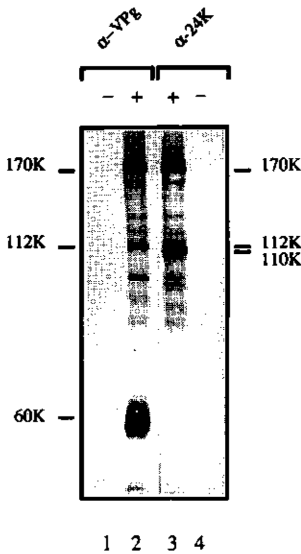


Fig. 5. Immunoblot analysis of proteins prepared from CPMV-infected (+) (lanes 2 and 3) and mock inoculated cowpea leaves (-) (lanes 1 and 4). Analysis was performed using anti-24K serum (lanes 1 and 4) or anti-VPg serum (lanes 1 and 2) as primary antibody and anti-rabbit alkaline phosphatase conjugate as the second antibody. Immunoreactive proteins are indicated on the right and left side.

Discussion

In this paper we have studied the processing of several putative VPg precursors. Under the conditions studied no evidence for the *trans* cleavage of 60K at the Gln/Ser site into 58K and VPg was found. This is in line with the finding that all cleavages in the B-polyprotein most efficiently occur *in cis*. These observations are not in agreement with the proposed role of the 60K protein as direct VPg precursor. In previous studies immunoreactive 60K proteins were detected in preparations of membrane bound replication complexes (Dorssers *et al.*, 1984) and

at present it cannot be excluded that the 60K protein is able to supply the VPg only in a membrane bound replication complex.
Whereas 58K is not generated through *trans* cleavage of 60K, the results indicate that the 84K protein preferentially is cleaved *in cis* via two alternative pathways, either into 58K and 26K (24K+VPg) or into 60K and 24K proteins. The 26K protein could only be observed as a processing intermediate during translation of BΔ32HM24-RNA when cleavage at the VPg/24K junction was prevented. It is therefore likely that 26K proteins are usually very rapidly further processed at the Gln/Met site into VPg and 24K proteins. Furthermore, processing of 170K proteins from BΔ32HM24-RNA and BΔ32R87-RNA, for which the 60K/110K cleavage and the 24K/87K cleavage was prevented, respectively, revealed a third processing pathway in which 58K and 112K (VPg+110K) proteins were generated. The 112K protein accumulated in relatively low amounts during BΔ32R87-RNA translation, however, during translation of BΔ32HM24-RNA the amount of 112K protein exceeded even the amount of 110K protein. This suggests that cleavage of 112K occurs preferentially into VPg and 110K rather than into 26K and 87K proteins. However, we cannot rule out a possible effect of the mutations on processing at the non mutant cleavage sites. Figure 6 summarizes our current knowledge of the processing pathways of the 200K B-polyprotein.
We have not, however, been able to detect free VPg in the lysate. Degradation of VPg might be a plausible explanation. Indeed earlier studies on stability of free VPg have shown that this protein is rapidly degraded in rabbit reticulocyte lysate (De Varennes *et al.*, 1986). Futhermore, VPg will not be very efficiently labelled since it contains only one methionine residue and hence detection of this protein will be difficult.
A 112K protein immunoreactive with anti-VPg and anti-24K sera was detected in fractions of CPMV-infected cowpea leaves, whereas we have not been able to detect a 26K protein. This confirms earlier results obtained with purified CPMV replication complexes in which a 112K protein was also detected (Dorssers *et al.*, 1984). It is very likely that VPg is supplied in the active replication complex at the initiation of RNA synthesis. Considering this possibility, the 112K protein, which has so far been neglected in the processing of the B-polyprotein, may play an important role as VPg precursor in this process.

Acknowledgements

We thank Hans van Bokhoven, Diederik de Bruijn, Gert Brunekreef, Richard Milek, and Jan Wery for providing us with cDNA clones, Piet Madern and Peter van Druten for artwork and photography. This work was supported by the Netherlands Foundation of Chemical Research (SON) with financial aid from the Netherlands Organization for Scientific Research (NWO).

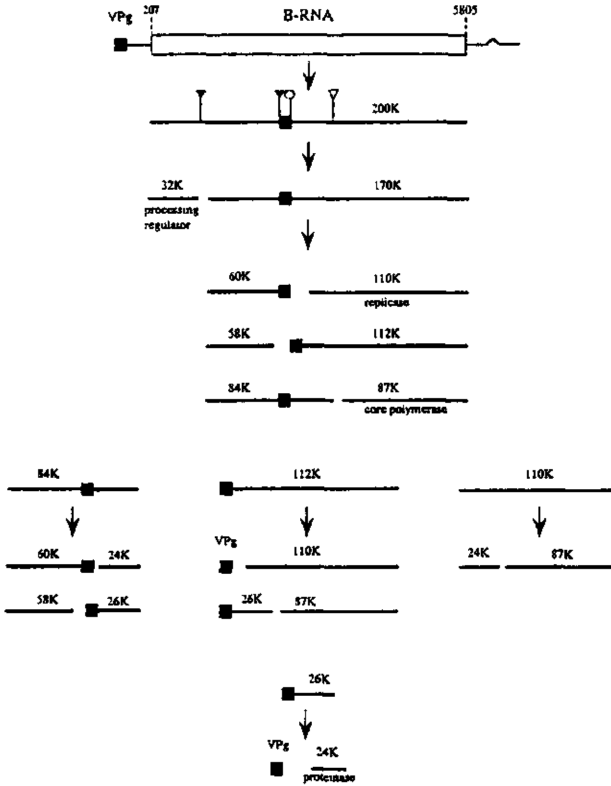


Fig. 6. Extended model for the expression of CPMV B-RNA. The various cleavage sites on the 200K protein are indicated by ○, Gln/Met, ▽, Gln/Gly and ▼, Gln/Ser. All cleavages in the B-polyprotein occur most efficiently *in cis*. B-RNA encodes a 200K polyprotein which is processed into 32K and 170K proteins. The 170K protein can be processed at three different sites into 58K+112K, 60K+110K, and 84K+87K proteins. The 110K intermediate can be processed into 24K and 87K proteins. Cleavage of the 84K protein generates either 26K+58K or 24K+60K proteins. VPg can be generated directly via cleavage of the 112K protein into VPg and 110K proteins or via the 26K intermediate into VPg and 24K proteins. So far evidence for cleavage of 112K either into VPg and 110K or 26K and 87K, and 26K into VPg and 24K is only indirect.

References

Blake, M.S., Johnston, K.H., Russell-Jones, G.J., and Gotschlich, E.C. (1984). A rapid, sensitive method for detection of alkaline phosphate-conjugated antibody on Western blots. *Anal. Biochem.* **136**, 175-179.

De Varennes, A., Lomonossoff, G.P., Shanks, M., and Maule, A.J. (1986). The stability of cowpea mosaic virus VPg in reticulocyte lysates. *J. Gen. Virol.* **67**, 2347-2354.

Dever, T.E., Glynias, M.J., and Merrick, W.C. (1987). GTP-binding domain: Three consensus sequence elements. *Proc. Natl. Acad. Sci. USA.* **84**, 1814-1818.

Dorssers, L., Van der Krol, S., Van der Meer, J., Van Kammen, A., and Zabel, P. (1984). Purification of cowpea mosaic virus RNA replication complex: Identification of a virus-encoded 110,000-dalton polypeptide responsible for RNA chain elongation. *Proc. Natl. Acad. Sci. USA.* **81**, 1951-1955.

Eggen, R., and Van Kammen, A. (1988). RNA replication in comoviruses. *In* "RNA Genetics" (P. Ahlquist, J. Holland, and E. Domingo, Eds.), Vol. 1, pp. 49-69. CRC Press, Boca Raton.

Eggen, R., Kaan, A., Goldbach, R., and Van Kammen, A. (1988). Cowpea mosaic virus RNA replication in crude membrane fractions from infected cowpea and *Chenopodium amaranticolor*. *J. Gen. Virol.* **69**, 2711-2720.

Eggen, R., Verver, J., de Jong, A., Goldbach, R., and Van Kammen, A. (1989). Improvements of the infectivity of *in vitro* transcripts from cloned cowpea mosaic virus cDNA: impact of terminal nucleotide sequences. *Virology* **173**, 447-455.

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., Goldbach, R., and Van Kammen, A. (1984). Translation of bottom component RNA of cowpea mosaic virus in reticulocyte lysate: faithful proteolytic processing of the primary translation product. *Virus Res.* **1**, 39-49.

Garcia, J.A., Schrijvers, L., Tan, A., Vos, P., Wellink, J., and Goldbach, R. (1987). Proteolytic activity of cowpea mosaic virus encoded 24K protein synthesized in *Escherichia coli*. *Virology* **159**, 67-75.

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., Rezelman, G., and Van Kammen, A. (1980). Independent replication and expression of B-component RNA of cowpea mosaic virus. *Nature (London)* **286**, 297-300.

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.

Goldbach, R., 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.

Gorbalenya, A.E., and Koonin, E.V. (1989). Viral proteins containing the purine NTP-binding sequence pattern. *Nucleic Acids Res.* **17**, 8413-8440.

Gorbalenya, A.E., Koonin, E.V., and Wolf, Y.I. (1990). A new superfamily of putative NTP-binding domains encoded by genomes of small DNA and RNA viruses. FEBS *Lett.* **262**, 145-148.

Korneluk, R.G., Quan, G., and Gravel, R.A. (1985). Rapid and reliable dideoxy sequencing of double-stranded DNA. *Gene* **40**, 317-323.

Laemmli, U.K. (1970). Cleavage of structural proteins during assembly of the head of the bacteriophage T_4. *Nature (London)* **227**, 680-685.

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

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

Peng, X.X., and Shih, D.S. (1984). Proteolytic processing of the proteins translated from the bottom component RNA of cowpea mosaic virus. *J. Biol. Chem.* **259**, 3197-3201.

Peters, S.A., Voorhorst, W.G.B., Wery, J., Wellink, J., and Van Kammen, A. (1992). A regulatory role for the 32K protein in proteolytic processing of cowpea mosaic virus polyproteins. Accepted for publication in *Virology*.

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

Richards, O.C., Eggen, R., and Van Kammen, A. (1989). High-level synthesis of cowpea mosaic virus RNA polymerase and protease in *Escherichia coli*. *Gene* **78**, 135-146.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). "Molecular Cloning: A laboratory manual", 2nd ed. Cold Spring Harbour Laboratory, Cold Spring Harbor, New York.

Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467.

Van Bokhoven, H., Wellink, J., Usmany, M., Vlak, J.M., Goldbach R., and Van Kammen, A. (1990). Expression of plant virus genes in animal cells: high-level synthesis of cowpea mosaic virus B-RNA-encoded proteins with baculovirus expression vectors. *J. Gen. Virol.* **71**, 2509-2517.

Verver, J., Goldbach, R., Garcia, J.A., and Vos, P. (1987). *In vitro* expression of a full-length DNA copy of cowpea mosaic virus B RNA: identification of the B RNA encoded 24-kd protein as a viral protease. *EMBO J.* **6**, 549-554.

Vos, P., Jaegle, M., Wellink, J., Van Kammen, A., and Goldbach, R. (1988). Two viral proteins involved in the proteolytic processing of the cowpea mosaic virus polyproteins. *Nucleic Acids Res.* **16**, 1967-1985.

Wellink, J., Jaegle, M., and Goldbach, R. (1987). Detection of a novel protein encoded by the bottom component RNA of cowpea mosaic virus, using antibodies raised against a synthetic peptide. *J. Virol.* **61**, 236-238.

Wellink, J., Rezelman, G., Goldbach, R., and Beyreuther, K. (1986). Determination of the proteolytic processing sites in the polyprotein encoded by the bottom-component RNA of cowpea mosaic virus. *J. Virol.* **59**, 50-58.

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.

Chapter 4

The Cowpea Mosaic Virus B-RNA Encoded 112K Protein May Function as a VPg Precursor *in vivo*

Sander A. Peters[1#], Ingeborg M. Kooter[1], Jan Verver[1], Johannes T. Dessens[2†], Jean-Michel Mesnard[1], Joan Wellink[1], George P. Lomonossoff[2], and Ab van Kammen[1]

1. Department of Molecular Biology, Agricultural University, Dreyenlaan 3, 6703 HA Wageningen, The Netherlands, 2. Department of Virus Research, John Innes Centre, Colney Lane, Norwich, NR4 7UH, United Kingdom

Running title: Processing of the VPg-containing 112K protein

Present address: Department of Virology, Institute of Medical Microbiology, State University, P.O. Box 320, 2300 AH Leiden, The Netherlands

† Present address: Station de Pathologie Végetale, INRA, Route de St.-Cyr, 78026 Versailles Cedex, France

Submitted

The Cowpea mosaic Virus B-RNA Encoded 112K Protein May Function as a VPg Precursor *in vivo*

Abstract

The cowpea mosaic virus B-RNA encoded 112K (VPg + 110K) protein has been proposed to function as a direct precursor of VPg. We have now examined the processing of the 112K protein in a rabbit reticulocyte lysate system and in cowpea mesophyll protoplasts using a transient expression system. The results show that the 112K protein was not processed at the Gln/Met site between the VPg and the 110K protein *in vitro*. Cleavage of the 112K protein in protoplasts occurs via two alternative pathways either into VPg and 110K (24K + 87K) or into 26K (VPg + 24K) and 87K proteins. The 26K protein is probably further cleaved into VPg and 24K proteins. Processing of the *in vitro* translated 112K protein was not enhanced with protoplast extracts suggesting that proteolytic activity for the 112K protein is dependent on a conformation determined by factors present at a cotranslational level. The results support a model in which the 112K protein functions as the precursor of VPg during initiation of replication.

Cowpea mosaic virus (CPMV) is a positive strand RNA plant virus with a bipartite genome. Both M-RNA and B-RNA have a small protein, VPg (viral protein genome linked), at their 5'-end, are polyadenylated at the 3'-end and are translated into large polyproteins (for a review see Eggen and Van Kammen, 1988). The B-RNA is translated into a 200K polyprotein and is processed by the viral 24K proteinase at specific Gln/Met, Gln/Gly and Gln/Ser sites to produce functional proteins (Wellink *et al.*, 1986; Garcia *et al.*, 1987) (see Fig. 1). The VPg coding sequence has been mapped on the open reading frame of B-RNA and comprises 28 amino acids (Stanley *et al.*, 1980; Zabel *et al.*, 1984; Wellink *et al.*, 1986). The β-OH group of the N-terminal serine residue of VPg is attached to the 5'-terminal uridyl residue of the genomic RNAs via a phophodiester bond (Stanley *et al.*, 1978; Jaegle *et al.*, 1987). A role in translation or infectivity for VPg is not very likely since *in vitro* transcripts lacking the protein, and viral RNA from which the VPg has been proteolytically removed, are infectious and retain their messenger activity (Stanley *et al.*, 1978; Vos *et al.*, 1988a; Eggen *et al.*, 1989). VPg is found covalently linked to the 5'-end of both positive and negative strands in the replicative forms isolated from CPMV-infected plants (Lomonossoff *et al.*, 1985) and has been proposed to have a function as a primer in the initiation of viral RNA synthesis (Wimmer, 1982; Eggen and Van Kammen, 1988).

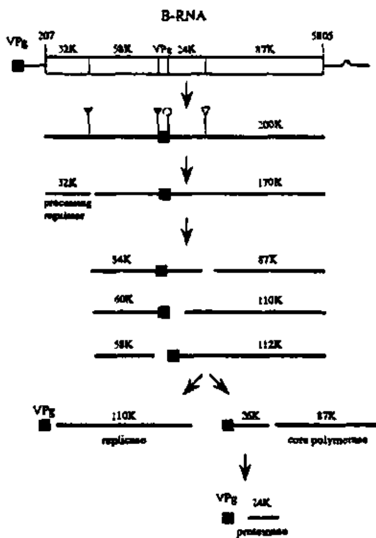


Fig. 1. Genetic organisation and expression of CPMV B-RNA. The RNA contains a long open reading frame shown by the open bar. The positions of the start and stop codons are indicated on the RNA. The cleavage sites are indicated on the 200K polyprotein by (▼) Gln/Ser, (O) Gln/Met and (∇) (Gln/Gly). VPg is represented by a black box. Processing of the 112K protein occurs via two pathways either into VPg + 110K or 26K + 87K proteins. The 110K protein is a stable cleavage product. The 26K protein is probably further cleaved into VPg and the 24K protein, although evidence for this cleavage is indirect.

In order to further understand the role of putative VPg precursors in viral RNA replication, elucidation of the successive processing steps that lead to the release of VPg from the 200K B-polyprotein is essential. The B-RNA encoded 60K protein has been proposed to be the direct precursor for VPg (Zabel *et al.*, 1982; Goldbach *et al.*, 1982). However, studies on processing of putative VPg precursors *in vitro* revealed that *trans* cleavage of the Gln/Ser site between the 58K and the VPg protein is extremely inefficient (Peters *et al.*, 1992a). Furthermore, kinetic studies on the processing of the 170K protein have revealed that the 60K and the 58K protein appear to be produced simultaneously, a pattern of accumulation inconsistent with their having precursor-product relationship (Dessens and Lomonossoff, 1992; Peters *et al.*, 1992b). Moreover, large amounts of the 60K protein are found in fractions prepared from CPMV infected cowpea leaves (Peters *et al.*, 1992a) which do not point towards a VPg precursor function for the 60K protein. In line with the finding that *in vitro* cleavages in the B-polyprotein occur most efficiently *in cis*, the 112K processing intermediate (VPg+110K) was proposed to function as a VPg precursor, either directly or via a 26K (VPg+24K) processing intermediate, although evidence for such cleavages was only indirect (Peters *et al.*, 1992a). Therefore, we have further investigated the processing of the 112K (VPg+110K) protein at the Gln/Met site between the VPg and the 110K protein. For this purpose RNA encoding the 112K protein was transcribed from clone pJII-B7 (Dessens and Lomonossoff, 1992) and translated in rabbit reticulocyte lysate in the presence of [^{35}S]methionine in a time course experiment as previously described (Peters *et al.*, 1992b). As a control, transcripts

from pTB114Δ3 were used (Vos *et al.*, 1988). These transcripts encode a protein of about 125K consisting of 46 amino acids from the N-terminus of the 32K protein, 50 amino acids of the C-terminus of the 58K protein and the 112K protein. Products were fractionated in a 7.5% SDS-polyacrylamide gel according to Laemmli (1970).

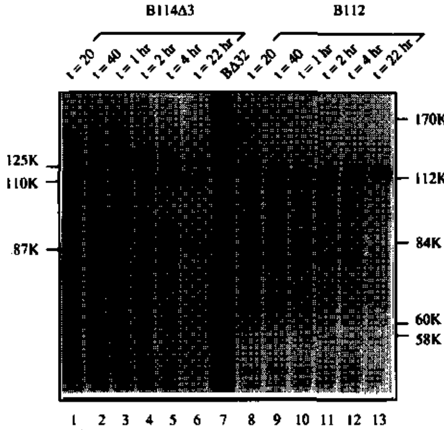


Fig. 2. Time course of an *in vitro* translation of B114Δ3 RNA (lanes 1 to 6) and B112 RNA (lanes 8 to 13). The 110K and 112K proteins were resolved by prolonged electrophoresis of the products on a 7.5% SDS-polyacrylamide gel. Products are shown after different times of translation as indicated on the top of the figure. Lane 7 contains the products of an *in vitro* translation of BΔ32-RNA that served as a marker.

Translation of B114Δ3 RNA resulted into a 125K product that was rapidly cleaved into 110K, 87K, 38K, and 24K proteins as descibed previously (Fig. 2, lanes 1 to 6; Vos *et al.*, 1988). (The 24K and 38K proteins are not visible in figure 2 because they have migrated from the gel). Translation of B112-RNA from pJII-B4 gave rise to the 112K protein, which was not processed further even upon prolonged incubation (Fig. 2, lanes 8 to 13). In previous experiments cleavage of the 112K protein at the Gln/Gly site into 26K and 87K proteins has been observed (Dessens and Lomonossoff, 1992). The only diffference in the two experiments is the source of the reticulocyte lysates. Obviously, components in the lysate influence cleavage efficiency.

Overall these results do not support a role for the 112K protein as a VPg precursor *in vitro*. Furthermore, since the 125K protein from B114Δ3 RNA is not cleaved into the 112K protein, the results suggest that cleavage at the Gln/Ser site between the 58K and the VPg requires a conformation determined by even further upstream sequences than those previously determined by Dessens and Lomonossoff (1992) to be necessary for cleavage at the Gln/Gly site between the 24K and 87K proteins. Thus far, *in vitro* cleavage at the Gln/Ser site between the 58K and 112K proteins has only been observed with the 170K protein (Peters *et al.* 1992a; Dessens and Lomonossoff, 1992).

To study whether the processing of the 112K protein occurs *in vivo*, the 112K protein was transiently expressed in cowpea protoplasts. For that purpose a 1144-bp *Hind*III-*Xho*I fragment from pTM58SΔ3 (Van Bokhoven *et al.*, 1993a) was

replaced with a 1106-bp *Hind*III-*Xho*I fragment of pJII-B7 which contains the first 1014 nucleotides of the 112K coding region, in which the serine codon at the beginning of the VPg (position 2964) has been changed into an ATG codon. From the resulting plasmid pTM58Ω' a *Bgl*II-*Xho*I fragment was isolated, which was inserted together with a 1932-bp *Xho*I-*Eco*RI fragment from pJII-B7 into the *Bgl*II-*Eco*RI digested vector pMON999 (Van Bokhoven *et al.*, 1993b) to produce transient expression vector pMB112Ω'. This plasmid was digested with *Bgl*II and *Sal*I to remove M-RNA and TMVΩ' sequences that interfered with transient expression (data not shown). The larger fragment was then religated to generate transient expression vector pMB112. The construction of transient expression vectors pMB87 and pMB110, which direct the transient expression of the 110K and 87K coding sequence respectively, has been described (Van Bokhoven *et al.*, 1993b). Protoplasts were either mock-inoculated, CPMV-RNA infected or transfected with pMB87, pMB110 and pMB112. The transfection and incubation of cowpea protoplasts were performed as described (Van Bokhoven *et al.*, 1993b). Protoplast extracts were prepared according to Franssen *et al.* (1982). Synthesis and processing of viral proteins were monitored by immunoblotting using anti-VPg (Eggen *et al.*, 1988), anti-24K (Wellink *et al.*, 1987) or anti-110K serum (Van Bokhoven *et al.*, 1992) as a primary antibody and anti-rabbit IgG-alkaline phophatase as a second antibody according to Blake *et al.* (1984).

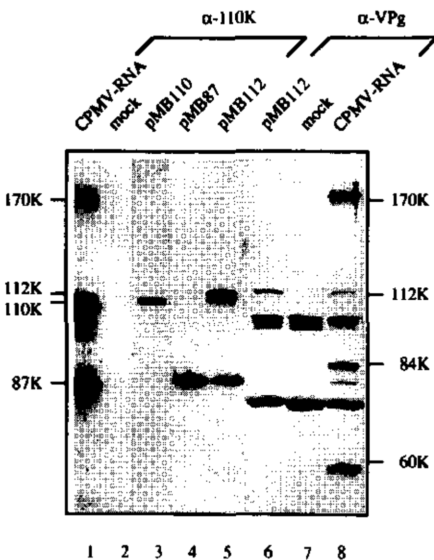


Fig. 3. Immunoblot analysis of CPMV proteins transiently produced in cowpea protoplasts. Protoplasts were harvested 16 hr post transfection to prepare a 30,000 g supernatant fraction which was fractionated in a 7.5% SDS-polyacrylamide gel. The immunological detection was performed with anti-110K (lanes 1 to 5), or anti-VPg serum (lanes 6 to 8). The nomenclature of the transient expression vectors, shown on top of the immunoblot is described in the text. Proteins from extracts of CPMV-RNA infected (lanes 1 and 8) or mock-inoculated protoplasts (lanes 2 and 7) were used as a control and immunoreactive proteins are indicated at the right and the left side.

In extracts from CPMV-RNA infected protoplasts the 170K, 110K and 87K proteins were identified with anti-110K serum (Fig. 3, lane 1) The 60K, 84K, a low amount of

112K proteins, and a protein with a molecular weight of approximately 80-kDa were detected with anti-VPg serum (Fig. 3, lane 8). The origin of the latter protein is unclear. Constructs pMB87 and pMB110 efficiently directed the synthesis of 87K and 110K proteins in cowpea protoplasts. These proteins were immunoreactive with anti-110K serum and comigrated with the marker proteins obtained from CPMV-RNA infected protoplasts (Fig. 3, lanes 1 to 4). Apparently the 110K protein is a stable protein *in vivo*, since 87K and 24K cleavage products were not detected (Fig. 3, lane 3) a result consistent with that obtained *in vitro* (Dessens and Lomonossoff, 1992). In protoplasts transfected with pMB110 a second protein migrating just below the 110K protein was produced. This protein probably arose as a result of downstream initiation of translation at AUG codon 3079, 30 nucleotides downstream of the first AUG codon of the 110K coding sequence (position 3049 in B-RNA). Construct pMB112 efficiently directed the synthesis of the 112K protein, which was identified with anti-110K (Fig. 3, lane 5) and anti-VPg serum (Fig. 3, lane 6). This protein migrated just above the 110K protein and comigrated with the 112K protein in CPMV-infected protoplasts (Fig. 3, lanes 4, 5 and 8). Surprisingly, in addition to the 112K protein, also 110K and 87K products in approximately equal amounts were observed (Fig. 3, lane 5).

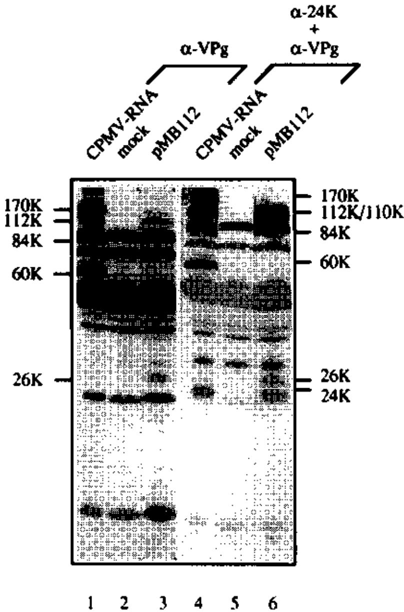


Fig. 4. Immunoblot analysis of CPMV proteins transiently produced in cowpea protoplasts. Protoplast extracts were prepared as described in the legend of Fig. 3 and were separated in a 15% SDS-polyacrylamide gel. The immunological detection was done with either anti-VPg (lanes 1 to 3) or with both anti-VPg and anti-24K serum (lanes 4 to 6). Vector pMB112, shown on top of the immunoblot (lanes 3 and 6), directs the transient expression of the 112K coding sequence. Proteins from CPMV-RNA (lane 1 and 4) infected and mock-inoculated protoplasts (lane 2 and 5) were used as a control. Immunoreactive viral proteins are indicated at the right and left side.

Further analysis revealed that 26K and 24K proteins, which were detected with anti-24K and anti-VPg serum, were also present in this fraction (Fig. 4, lanes 3 and 6). This suggests that the 112K protein is processed *in vivo* via two alternative

pathways, either at the Gln/Met site into VPg and 110K proteins or at the Gln/Gly site into 26K and 87K proteins (see Fig. 1). The 26K protein is visible as a doublet in these gels. Possibly the smaller protein is a result of initiation of translation at the AUG codon at position 2987, which after processing would result in a 26K protein that lacks the first 8 amino acids.

The results suggest that the presence of the VPg sequence can activate cleavage at the 24K/87K junction *in vivo*, a result consistent with that observed *in vitro* (Dessens and Lomonossoff, 1992). The 112K protein, when processed from the 170K protein, theoretically contains an N-terminal serine residue, whereas the *in vitro* and transiently produced 112K protein each possess an N-terminal methionine residue. Apparently, the processing of VPg does not depend on the presence of an N-terminal serine residue. The 24K protein probably accumulated through cleavage of the 26K precursor rather than through cleavage of the 110K protein, based on the observation that the 110K protein was shown to be a stable product (Fig. 3, lane 3). Probably the 26K protein is rapidly further processed into VPg and 24K proteins during an normal infection, since the 26K protein was not detected in CPMV-RNA infected protoplasts (Fig. 4, lanes 1 and 4). On the other hand, it is possible that in CPMV-infected protoplasts processing of the 112K into 110K and VPg proteins under RNA replicative conditions is favoured over processing into 26K and 87K proteins. Its even possible that RNA is directly needed to trigger release of VPg.

Furthermore, we have not detected VPg in extracts from either pMB112 transfected or CPMV-RNA infected protoplasts probably because this protein was rapidly degraded by cellular enzymes. To test this synthetic VPg (Eggen *et al.*, 1988) was incubated for 1 hr with a mock-inoculated protoplast extract and, subsequently, analysed by SDS-polyacrylamide gelelectroforesis and western blotting. Under these conditions VPg was rapidly degraded whereas other viral proteins were stable (data not shown). VPg incubated in homogenisation buffer remained stable. Previous experiments have also indicated the instability of released VPg in rabbit reticulocyte lysate (De Varennes *et al.*, 1986).

To exclude the possibility that the 110K protein present in protoplasts transfected with pMB112 is the result of internal initiation of translation at the AUG codon at position 3049 at the beginning of the 110K open reading frame, two new constructs were made. In pMB112SS and pMB112QG→R the KpnI-XhoI fragment of pMB112 (positions 3134-3978 of B-RNA) has been exchanged for the same fragment of pTB24SSΔ32 and pTBΔ32R respectively (Peters *et al.*, 1992a). Plasmid pMB112SS contains a 6 nucleotide insertion at position 3156 of B-RNA that results in an inactive proteinase, whereas pMB112QG→R contains a mutation that alters

the cleavage site between the 24K and 87K proteins from Gln/Gly into Arg (Peters *et al.*, 1992a).

Protoplasts were transfected with these constructs and their protein content was analysed after 40 hr of incubation by immunoblotting (Fig. 5). The blot was first treated with anti-VPg serum (not shown) and, subsequently, with anti-110K serum (Fig. 5). Construct pMB112SS directed the synthesis of only one virus-specific protein, a 112K protein, that reacted with anti-VPg serum (not shown) and anti-110K serum (Fig. 5, lane 2). In pMB112QG→R transfected cells two virus-specific proteins can be detected with anti-110K serum (Fig. 5, lane 1). Only the upper band reacted with anti-VPg serum (not shown) identifying this band as the 112K protein. Because of the cleavage site mutation processing of the 112K protein into 26K and 87K proteins is clearly prevented whereas cleavage into VPg and 110K protein can still take place. In this experiment the protoplasts were havested 40 hr after the transfection. Under these conditions almost all of the non-mutant 112K protein is cleaved into 110K and 87K proteins (Fig. 5, lane 3). Note that because of the small insertions and deletions in the mutant 112K and 110K proteins it is not useful to compare the migration of these proteins in the gel.

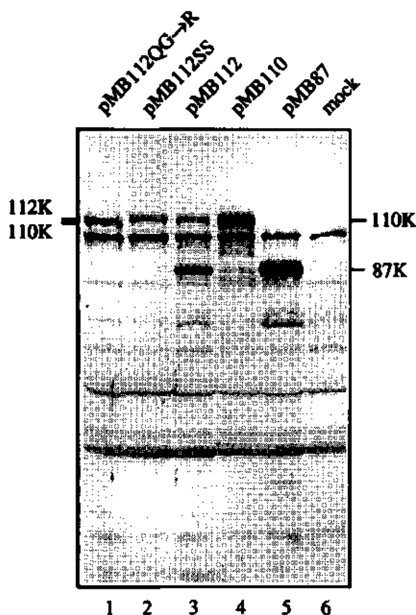


Fig. 5. Immunoblot analysis of proteins produced by transient expression of pMB112 mutants. Protoplasts were transfected with constructs as indicated on the top of the figure and harvested after 40 hr of incubation. A total protoplast fraction was separated on an 8% SDS-polyacrylamide gel. The immunological detection was performed with anti-VPg serum and subsequently with anti-110K serum.

These experiments exclude the possibility that the 110K protein in protoplasts transfected with pMB112 arose as a result of downstream initiation of translation at the first AUG codon of the 110K coding sequence (position 3049 in the B-RNA). Furthermore, the AUG 3049 codon is not in a favourable context to function as an

initiator codon according to the rules determined by Kozak (1986) and Lütcke *et al.* (1987). This notion is supported by the observation that expression of the 110K coding sequence from pMB110 generates two proteins of roughly equal amounts through initiation of translation at position 3049 and 3079 (Fig. 2, lane 3). However, such a smaller product was not observed when the 112K coding sequence was expressed, suggesting that ribosomes have already initiated translation upstream position 3049 in the 112K coding sequence. In addition, the AUG codon 3049 is preceded by one in frame AUG codon (position 2987) and one out of frame AUG codon (position 3010), the latter one in a more favourable context, which the ribosomes have to bypass in order to start translation at position 3049. The *in vitro* experiments also showed that downstream initiation of translation on AUG 3049 of the B112-RNA is not favoured (see Fig. 1, lanes 3 and 4).
Processing of the *in vitro* translated 112K protein into VPg and 110K could not be activated, either with a crude extract or a supernatant fraction prepared from either uninfected or CPMV-infected protoplasts, nor was activation observed in presence of a microsomal fraction, whereas *in vitro* translated 105K and 95K proteins from M-RNA were efficiently processed *in trans* after incubation with CPMV-infected fractions (data not shown). Lack of cleavage at the Gln/Met site in the 112K protein *in vitro* probably is therefore not due to inhibiting factors, but rather the result of a conformation that does not allow cleavage in the 112K protein. Apparently the transiently produced 112K protein has obtained a more favourable conformation for processing. We have no indication that this would be the result of a membranous environment. The observed activation of processing of the 112K protein in protoplasts might rather be due to the influence of salt concentrations, pH or host-ribosomal factors that differ between the reticulocyte lysate and protoplast system. Since protoplast extracts failed to compensate the processing defect of the *in vitro* translated 112K protein at a post-translational level, it seems likely that such factors already act at a cotranslational level to force the 112K protein into a conformation such that Gln/Ser and Gln/Gly sites become more accessible to proteolytic cleavage. In conclusion, the results reported here support our earlier proposal that the 112K protein functions as a VPg precursor (Dorssers *et al.*, 1984; Peters *et al.*, 1992a). If VPg is involved in the protein-primed mechanism to initiate RNA replication, then the 112K viral replicase precursor is a likely candidate to start this event via *cis* cleavage into VPg and the 110K protein. Subsequently, or maybe concomitantly, VPg becomes uridylated and might serve as a primer that is elongated by the 110K viral replicase. This strategy has also been proposed for the initiation of replication of polioviral RNA (Kuhn and Wimmer, 1987) and would resemble the events that have been shown to lead to

the initiation of DNA synthesis for viruses such as adenovirus and bacteriophage ϕ29 (for a review see Salas, 1983).

Acknowledgements

The authors like to thank fotolokatie "de Dreyen" for photography. This work was supported by the Netherlands Foundation of Chemical Research (SON) with financial aid from the Netherlands Organization for Scientific Research (NWO) and a John Innes Foundation studentship.

References

Blake, M.S., Johnston, K.H., Russell-Jones, G.J., and Gotschlich, E.C. (1984). A rapid, sensitive method for detection of alkaline phosphate-conjugated antibody on Western blots. *Anal. Biochem.* **136**, 175-179.

De Varennes, A., Lomonossoff, G.P., Shanks, M., and Maule, A.J. (1986). The stability of cowpea mosaic virus VPg in reticulocyte lysates. *J. Gen. Virol.* **67**, 2347-2354.

Dessens, J.T., and Lomonossoff, G.P. (1991). Mutational analysis of the putative catalytic triad of the cowpea mosaic virus 24K protease. *Virology* **184**, 738-746.

Dessens, J.T., and Lomonossoff, G.P. (1992). Sequence upstream of the 24K protease enhances cleavage of the cowpea mosaic virus B RNA-encoded polyprotein at the junction between the 24K and 87K proteins. *Virology* **189**, 225-232.

Dorssers, L., Van der Krol, S., Van der Meer, J., Van Kammen, A., and Zabel, P. (1984). Purification of cowpea mosaic virus RNA replication complex: Identification of a virus-encoded 110,000-dalton polypeptide responsible for RNA chain elongation. *Proc. Natl Acad. Sci. USA* **81**, 1951-1955.

Eggen, R., and Van Kammen, A. (1988). RNA replication in comoviruses. *In* "RNA Genetics" (P. Ahlquist, J. Holland, and E. Domingo Eds.), Vol. 1, pp. 49-69, CRC Press, Boca Raton.

Eggen, R., Kaan, A., Goldbach, R., and Van Kammen, A. (1988). Cowpea mosaic virus RNA replication in crude membrane fractions from infected cowpea and *Chenopodium amaranticolor*. *J. Gen. Virol.* **69**, 2711-2720.

Eggen, R., Verver, J., Wellink, J., De Jong, A., Goldbach, R., and Van Kammen. (1989). Improvements of the infectivity of *in vitro* transcripts from cloned cowpea mosaic virus cDNA: impact of terminal nucleotide sequences. *Virology* **173**, 447-455.

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.

Garcia, J.A., Schrijvers, L., Tan, A., Vos, P., Wellink, J., and Goldbach, R. (1987). Proteolytic activity of cowpea mosaic virus encoded 24K protein synthesized in *Escherichia coli*. *Virology* **159**, 67-75.

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.

Jaegle, M., Wellink, J., and Goldbach, R. (1987). The genome-linked protein of cowpea mosaic virus is bound to the 5' terminus of virus RNA by a phosphodiester linkage to serine. *J. Gen.Virol.* **68**, 627-632.

Kozak, M. (1986). Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **44**, 283-292.

Kuhn, R.J., and Wimmer, E. (1987). The replication of picornaviruses. *In* "The molecular biology of the positive strand RNA viruses", (D.J. Rowlands, M.A. Mayo and B.W.J. Mahy Eds.), pp 17-51, Academic Press, New York and London.

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

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

Lomonossoff, G., Shanks, M., and Evans, D. (1985). The structure of cowpea mosaic virus replicative form RNA. *Virology* **144**, 351-362.

Lütcke, H.A., Chow, K.C., Mickel, F.S., Moss, K.A., Kern, H.F., and Scheele, G.A. (1987). Selection of AUG initiation codons differs in plants and animals. *EMBO J.* **6**, 43-48.

Peters, S.A., Voorhorst, W.G.B., Wellink, J., and Van Kammen, A. (1992a). Processing of VPg-containing polyproteins encoded by the B-RNA from cowpea mosaic virus. *Virology* **191**, 90-97.

Peters, S.A., Voorhorst, W.G.B., Wery, J., Wellink, J., and Van Kammen, A. (1992b). A regulatory role for the 32K protein in proteolytic processing of cowpea mosaic virus polyproteins. *Virology* **191**, 81-89.

Salas, M. (1983). A new mechanism for the initiation of replication of Φ29 and adenovirus DNA: priming by the terminal protein. *Curr. Top. Microbiol. Immunol.* **109**, 89-106.

Stanley, J., Goldbach, R., and Van Kammen, A. (1980). The genome-linked protein of cowpea mosaic virus is coded by RNA from the bottom component. *Virology* **106**, 180-182.

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 Bokhoven, H., Van Lent, J.W.M., Custers, R., Vlak, J.M., Wellink, J., and Van Kammen, A. (1992). Synthesis of the complete 200K polyprotein encoded by cowpea mosaic virus B-RNA in insect cells. *J. Gen. Virol.* **73**, 2775-2784.

Van Bokhoven, H., Le Gall, O., Kasteel, D., Verver, J., Wellink, J., and Van Kammen, A. (1993a). *Cis* and *trans* acting elements in cowpea mosaic virus RNA replication. *Virology* in press.

Van Bokhoven, H., Verver, J., Wellink, J., and Van Kammen. (1993b). Protoplasts transiently expressing the 200K coding sequence of cowpea mosaic virus B-RNA support replication of M-RNA. *J. Gen. Virol.* in press.

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.

Vos, P., Jaegle, M., Wellink, J., Verver, J., Eggen, R., Van Kammen, A., and Goldbach, R. (1988a). Infectious RNA transcripts derived from full-length DNA copies of the genomic RNAs of cowpea mosaic virus. *Virology* **165**, 33-41.

Vos, P., Verver, J., Jaegle, M., Wellink, J., Van Kammen, A., and Goldbach, R. (1988b). Two viral proteins involved in the proteolytic processing of the cowpea mosaic virus polyproteins. *Nucleic Acids Res.* **16**, 1967-1985.

Wellink, J., Jaegle, M., and Goldbach, R. (1987). Detection of a novel protein encoded by the bottom component RNA of cowpea mosaic virus, using antibodies raised against a synthetic peptide. *J.Virol.* **59**, 50-58.

Wellink, J., Rezelman, G., Goldbach, R., and Beyreuther, K. (1986). Determination of the proteolytic processing sites in the polyprotein encoded by the bottom-component RNA of cowpea mosaic virus. *J.Virol.* **59**, 50-58.

Wimmer, E. (1982). Genome-linked proteins of viruses. *Cell* **28**, 199-201.

Zabel, P., Moerman, M., Lomonossoff, G., Shanks, M., and Beyreuther, K. (1984). Cowpea mosaic virus VPg: sequencing of radiochemically modified protein allows mapping of the gene on B RNA. *EMBO J.* **3**, 1629-1634.

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.

Chapter 5

The Nucleoside Triphosphate Binding Motif-Containing 60K and 84K Proteins of Cowpea Mosaic Virus bind ATP

Sander A. Peters, Joan Wellink, and Ab van Kammen

The nucleoside triphosphate binding motif-containing 60K and 84K proteins of cowpea mosaic virus bind ATP.

Abstract

The ribonucleotide binding properties of nucleoside triphosphate binding motif (NTBM) containing proteins encoded by cowpea mosaic virus (CPMV) B-RNA were studied by assaying for binding of periodate oxidized [α-^{32}P]ATP and [α-^{32}P]GTP using partially purified viral proteins from CPMV-infected plants. Two groups of labelled viral proteins were detected. The 170K, 112K, 110K and 87K proteins had bound ATP and GTP to a site in the 87K core polymerase domain. The NTBM-containing 60K and 84K proteins were unambiguously identified as ATP binding proteins, whereas binding of GTP was not observed with these proteins. These results are a first indication that the NTBM located in the N-terminal part of the 58K domain of these proteins is a functional binding site for ATP.

Introduction

The genetic information of cowpea mosaic virus (CPMV) is contained in two plus-sense single stranded RNAs, designated M-RNA and B-RNA, that have a small protein, VPg, attached to the 5'-end and are polyadenylated at the 3'-end. Upon infection the genomic RNAs are translated into large polyproteins from which functional proteins are generated by cleavages at specific sites by the viral encoded 24K proteinase (Wellink *et al.*, 1986; Garcia *et al.*, 1987) (see Fig. 1).

The B-RNA is able to replicate independently of M-RNA and encodes the information for viral RNA replication (Goldbach *et al.*, 1980), but the precise action of each individual B-RNA coded protein is not understood. Further insight into the role of the B-proteins in CPMV-RNA replication will be obtained when the biochemical activities of each of the proteins is known. The occurrence of conserved amino acid sequence motifs can help in the identification of the activity of functionally uncharacterized proteins. The 87K protein encoded by CPMV B-RNA (Fig. 1) contains a conserved GDD amino acid sequence flanked by hydrophobic residues that is characteristic for many RNA-dependent RNA-polymerases (Kamer and Argos, 1984; Argos, 1988; Ishihama and Nagata, 1988), and a second motif S/GTXXXTXXXNT/S that is well conserved in RNA-dependent RNA-polymerases of several plant and animal viruses (Kamer and Argos, 1984). Indeed, the 110K protein (24K+87K) was shown to occur in purified replication complexes

and has been shown to possess RNA-dependent RNA-polymerase activity (Dorssers *et al.*, 1984).

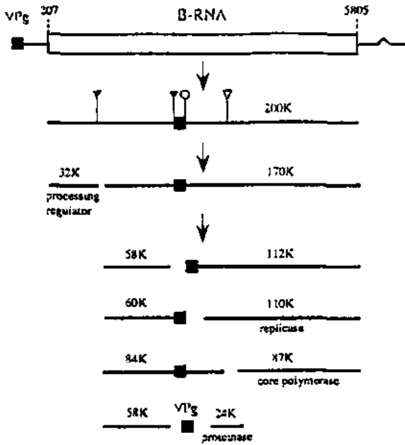
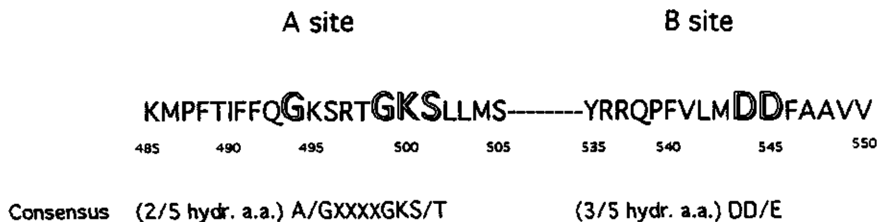


Fig. 1. Expression and genetic organization of B-RNA from CPMV. The RNA contains a long open reading frame (shown by open bar). The postions of the start and stop codon are indicated. Cleavage sites are indicated by (○) Gln/Met, (▽) Gln/Gly, and (▼) Gln/Ser. B-RNA encodes a 200K polyprotein which is processed into 32K and 170K proteins. The 170K is processed at three different sites into three sets of proteins of 58K+112K, 60K+110K, and 84K+87K (Peters *et al.*, 1992a). Final cleavage products are 58K, VPg, 24K and 87K proteins.

The 24K protein is the viral proteinase and the 87K protein represents the core polymerase of the CPMV replicase (Eggen and Van Kammen, 1988). Besides the RNA polymerase protein, positive-strand RNA viruses encode proteins with a typical nucleoside triphosphate binding motif (NTBM). This motif was first recognized and described by Walker *et al.*, (1982), and is a highly conserved element for different viruses (Gorbalenya *et al.*, 1989a). Such an NTBM has been identified in the N-terminal part of the 58K domain of the CPMV B-RNA encoded 200K polyprotein (Dever *et al.*, 1987; Gorbalenya and Koonin, 1989) (see Fig. 2). The amino acid sequence motif of the NTBM is characteristic for many ATP and GTP-utilizing proteins and consists of two separate parts, referred to as the 'A'-site and 'B'-site respectively. The consensus sequence A/GXXXXGKS/T, preceeded by 3 to 5 hydrophobic amino acids, is found in the A-site (where X stands for any amino acid), whereas conserved DD or DE residues, preceeded by 2 to 5 hydrophobic amino acids are located in the B-site (Walker *et al.*, 1982; Halliday, 1984; Möller and Amons, 1985). X-ray crystallography data indicate that the A-site binds the phosphoryl-moiety of a ribonucleotide (Möller and Amons, 1985; Bradley *et al.*, 1987; De Vos *et al.*, 1988), whereas the negatively charged aspartic acid residues chelate a Mg-cation complexed with the same phosphate groups (La Cour *et al.*, 1985; Fry *et al.*, 1986; Jurnak, 1988). In addition, a third consensus element NKXD, which is located 40 to 80 amino acids downstream from the B-site, has been proposed to be specific for GTP-utilizing proteins (Dever *et al.*, 1987). This third consensus element specific for GTP-utilizing proteins appears to be absent in the 58K domain (Dever *et al.*, 1987). For a first attempt to determine the precise function of the NTBM in the CPMV B-polyprotein, we have investigated the

ribonucleotide binding properties of B-RNA encoded proteins. Our approach has been to demonstrate directly the presence of a functional NTBM by covalent affinity-labelling with periodate-oxidized ribonucleotides. In this paper we demonstrate that NTBM-containing 60K and 84K proteins specifically bind ATP, whereas affinity for GTP is not observed.



Consensus (2/5 hydr. a.a.) A/GXXXXGKS/T (3/5 hydr. a.a.) DD/E

Fig. 2. Schematic representation of the nucleoside triphosphate binding motif (NTBM) in the 58K protein of CPMV B-RNA. The amino acid numbers correspond to the position from the amino terminus of the 200K B-polyprotein. Amino acids are given in the single-letter code. The NTBM consists of the N-terminal A-site and the C-terminal B-site. Conserved amino acids are shown outlined. The consensus sequence found in the A and B-site is displayed below, where X stands for any amino acid.

Materials and Methods

Preparation of subcellular fractions from cowpea leaves and cowpea mesophyll protoplasts

Extracts were prepared from mock-inoculated and CPMV-infected cowpea leaves as described previously (Zabel *et al.*, 1976) with minor modifications. Briefly, portions of 10 g primary cowpea leaves supplemented with 35 ml homogenization buffer (10 mM Tris-HAc pH 7.5, 10 mM KAc, 1 mM EDTA, 10 mM DTT and 1 mM PMSF) were ground in a morter on ice. The homogenate was filtered and centrifuged at 1,000 × g for 15 min at 4°. Subsequently, the supernatant was adjusted to 20% glycerol and centrifuged at 30,000 × g for 30 min at 4°. The pellet fraction (F4) was homogenized with 10 ml TGKEDP buffer (50 mM Tis-HAc pH 8.2, 50 mM KAc, 1 mM EDTA, 10 mM DTT, 1 mM PMSF, 25% glycerol) in a dounce homogenizer at 4°. The homogenate was centrifuged again at 30,000 × g for 30 min at 4°. The resulting supernatant fraction (F6) was used either for SDS-polyacrylamide gel electrophoresis or for a covalent affinity labelling assay. Approximately 1.10^6 cowpea mesophyll protoplasts were transfected with 1 µg CPMV-RNA or mock-inoculated as described (Van Bokhoven *et al.*, 1993) and incubated with 0.5 mCi $[^{32}P]KH_2PO_4$ (Amersham International, U.K.). Protoplasts were harvested 20 hr

post transfection and subjected to fractionation by centrifugation at 30,000 × g for 30 min. Extracts from infected and uninfected protoplasts were prepared as described (Franssen *et al.*, 1982).

Labelling with periodate-oxidized ribonucleotides (ox-NTP)

Ox-[α-^{32}P]ATP and ox-[α-^{32}P]GTP were prepared as described (Clertant and Cuzin, 1982). Briefly, 2 µl [α-^{32}P]ATP or [α-^{32}P]GTP (Amersham International, U.K., 3000 Ci/mmol) were mixed with 2 µl 2 mM $NaIO_4$ and 2 µl 2 mM HCl in a total volume of 8 ml and incubated for 20 min at 4° in the dark. Preparations were then adjusted to 7.5 mM $MgCl_2$, 1 mM DTT, 1 mM EGTA, 50 mM Tris-HCl pH 7.5 in a total volume of 20 µl and directly used for labelling reaction. Portions of 3 µl protein sample of F6 fractions prepared from CPMV-infected and mock-inoculated cowpea leaves were mixed with 5 µl of ox-NTP in a final volume of 20 µl. Incubation continued for 16 hr at 0° in the presence of 8 mM $NaCNBH_3$ (Fluka) in the dark (Rayford *et al.*, 1985). As a control 3 µl protein sample of fraction F4 were mixed with 1 µl [γ-^{32}P]ATP (Amersham International, U.K., > 5000 Ci/mmol) and incubated during 10 min at 30° in 50 mM Hepes pH 8.0, 5 mM $MgCl_2$, 2 mM DTT in a final volume of 15 µl. Labelled fractions were either separated on an SDS-polyacrylamide gel or subjected to immunoprecipitation.

Immunoprecipitation and westernblotting

Unlabelled and labelled viral proteins were heated in sample buffer (10% glycerol, 5% β-mercaptoethanol, 2% SDS, 0.01% bromophenolblue, 75 mM Tris-HCl pH 6.8) for 3 min at 100°. Unlabelled viral proteins were fractionated in an 7.5% SDS-polyacrylamide gel according to Laemmli (1970). The gels were analysed by silver staining or by immunoblotting using anti-VPg (Eggen *et al.*, 1988), anti-24K (Wellink *et al.*, 1987), or anti-110 serum (Van Bokhoven *et al.*, 1992) as a primary antibody and anti-rabbit IgG alkaline phosphatase (Promega Biotec) as a second antibody as described (Blake *et al.*, 1984). For immunoprecipitation of labelled viral proteins 10 to 15 µl sample was adjusted to PBS-TDS buffer (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 200 µl and was incubated for 3 hr at 4° in presence of 5 µl preimmune serum, anti-VPg, anti-24K, or anti-110K serum. Additionally, 20 µl of a 10% suspension of *Staphylococcus aureus* cells (Calbiochem corporation) in PBS-TDS buffer supplemented with 10 mg/ml BSA were added and incubation continued for 1 hr at 4°. Samples were centrifuged and pellet material was washed three times with 1 ml ice cold PBS-TDS buffer. Pellet material was resuspended in 40 µl

sample buffer and heated for 3 min at 100°. The supernatant fraction was separated on a 7.5% SDS-polyacrylamide gel. Gels were dried and autoradiographed with Kodak X-Omat film.

Results and Discussion

Identification of CPMV B-proteins in subcellular fractions

Viral proteins were solubilized from a 30,000 × g pellet fraction (F4). An adventitious effect of the solubilization was that several proteins of host origin and membranes that interfered with the covalent affinity-labelling assay were removed. Figure 3 displays the immunological detection of the viral proteins that are found in the soluble fraction (F6).

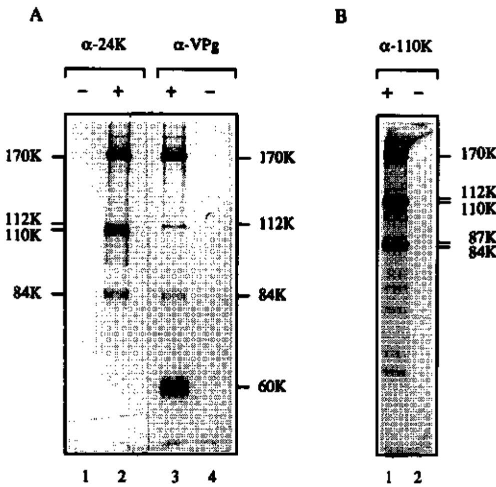


Fig. 3. Immunoblot analysis of proteins prepared from CPMV-infected (+ indicated) (panel A, lanes 2 and 3, panel B, lane 1) and mock-inoculated (- indicated) cowpea leaves (panel A, lanes 1 and 4, panel B lane 2). Analysis was performed with anti-24K serum (panel A, lanes 1 and 2), anti-VPg serum (panel A, lanes 3 and 4) and anti-110K serum (panel B, lanes 1 and 2). Immunoreactive proteins are indicated at the left and right side.

The 170K, 110K, and small amounts of 84K and 112K proteins, the latter one migrating just above the 110K protein, were detected with anti-24K serum (Fig. 3A, lane 2). The 170K, 112K, 84K, and 60K proteins were identified with anti-VPg serum (Fig. 3A, lane 3). Using anti-110K serum the 170K, 112K, 110K, 87K, and 84K proteins were detected (Fig. 3B, lane 1). Besides viral proteins, many host proteins are also present in these preparations as observed by silver staining of gels (data not shown). Immunoreactive proteins were not observed in F6 fractions prepared from mock-inoculated, healthy cowpea leaves (Fig. 3A, lanes 1 and 4, Fig. 3B, lane 2), indicating that the F6 fractions are free from host proteins that could interfere with the immunological detection of the viral proteins.

Ribonucleotide binding by CPMV B-proteins

The ribonucleotide binding properties of CPMV B-proteins were examined using a covalent affinity-labelling assay with periodate-oxidized [α-^{32}P]-labelled ribonucleotides. The periodate cleaves the ribose moiety between the C-2' and C-3' position and hydroxyl groups are oxidized to aldehyde groups (Morrison and Boyd, 1974). Binding of an ox-NTP to a protein can result in the covalent attachement of the ribonucleotide to its target via an aminoacyl-carbonyl condensation reaction with an amino group at the ε-position of the Lys500 residue in the A-site of the NTBM. The [α-^{32}P]-labelled reaction product is an imine or unstable Schiff's base, which is stabilized by reduction with cyanoborohydride (Clertant and Cuzin, 1982; Rayford *et al.*, 1985). Viral proteins F6 fractions from CPMV infected and healthy cowpea leaves were mixed with ox-[α-^{32}P]ATP or ox-[α-^{32}P]GTP and after labelling the proteins were separated on a SDS-polyacrylamide gel.

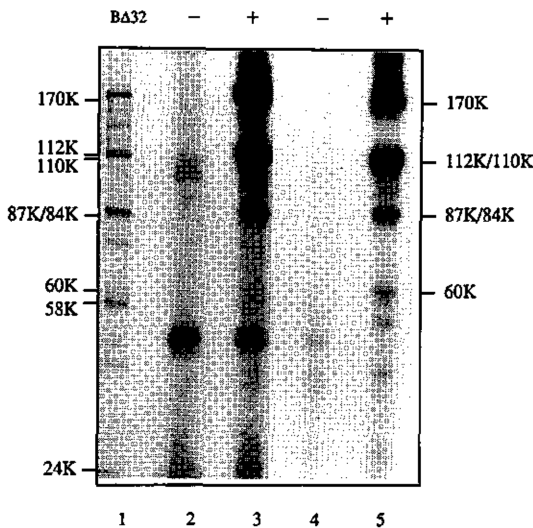


Fig. 4. Covalent affinity labelling with periodate-oxidized ribonucleotides of cowpea mosaic virus proteins. Proteins from CPMV-infected (+) and mock-inoculated (-) cowpea leaves were partially purified and incubated with ox-[α-^{32}P]-GTP (lanes 2 and 3) or ox[α-^{32}P]-ATP (lanes 4 and 5) and separated in a SDS-polyacrylamide gel. Positions of labelled products are indicated on the right. Marker proteins obtained from BΔ32-RNA translations (lane 1) (Peters *et al.*, 1992b) are indicated on the left.

In the F6 fraction from CPMV-infected leaves 170K, 110K, and 87K proteins were labelled with both ox-GTP (Fig. 4, lane 3) and ox-ATP (Fig. 4, lane 5). In addition a 60K protein was labelled with ox-ATP (Fig. 4, lane 5) but not with ox-GTP (Fig. 4, lane 3). Among the host proteins present in these preparations, only a protein of approximately 55-kDa was labelled with ox-GTP (Fig. 4, lanes 2 and 3), whereas labelling of this protein with ox-ATP did not occur (Fig. 4, lanes 4 and 5). The results indicate that the reaction conditions were such that labelling of proteins resulted from specific binding of oxidized ribonucleotides and not by a non-specific cross-linking reaction to an exposed amino group. As controls CPMV-

infected and mock-inoculated protoplasts were incubated with $[^{32}P]KH_2PO_4$, and an F4 fraction from infected cowpea plants was incubated with $[\gamma\text{-}^{32}P]ATP$. Labelling of viral proteins after these incubations was not observed (data not shown), indicating that non-oxidized ribonucleotides were not covalently bound by viral proteins, nor that labelling resulted from phosphorylation through kinase activities.

The ox-ATP and ox-GTP binding proteins were further identified by immunoprecipitation with anti-VPg, anti-110K, anti-24K, or preimmune serum prior to SDS-polyacrylamide gel electrophoresis. SDS-PAGE of the immunoprecipitates revealed that two groups of viral proteins appeared to be labelled with ox-$[\alpha\text{-}^{32}P]$ATP. The NTBM-containing 60K and 84K (60K+24K) proteins were identified by immunoprecipitation with anti-VPg (Fig. 5, lane 2) and anti-110K (Fig. 5, lane 5). In addition the 87K, 110K, 112K and 170K proteins were identified by immunoprecipitation with anti-VPg (Fig. 5, lane 2) and anti-110K (Fig. 5, lane 5). This second group of labelled viral proteins had bound ox-ATP probably to a site in the 87K polymerase domain contained in these proteins. ATP-labelled products were neither immunoprecipitated from extracts of uninfected cowpea leaf tissue (Fig. 5, lanes 3, 6 and 8), nor from extracts of CPMV infected leaves with preimmune serum (Fig. 5, lane 7).

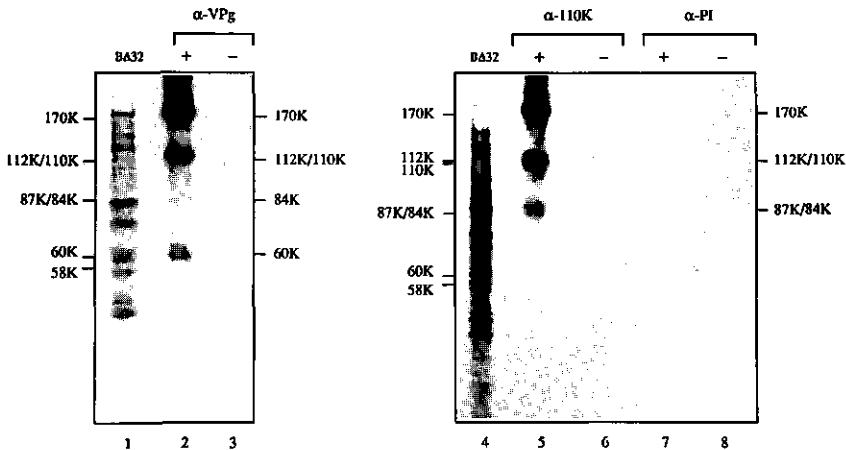


Fig. 5. SDS-polyacrylamide gel electrophoresis of labelled proteins from immunoprecipitates of extracts incubated with ox-$[\alpha\text{-}^{32}P]$ATP. After labelling incubation the CPMV-infected (+) and mock-inoculated (-) extracts were subjected to immunoprecipitation with anti-VPg (lanes 2 and 3), anti-110K (lanes 5 and 6) or, preimmune serum (lanes 7 and 8). Positions of labelled immunoprecipitated products and marker proteins are indicated at the right and left side respectively.

Using ox-[α-^{32}P]GTP only viral proteins containing the core polymerase domain were labelled (Fig. 6). The 170K, 112K and 110K proteins were specifically immunoprecipitated with anti-VPg serum (Fig. 6, lane 2) and anti-24K (Fig. 6, lane 4), whereas labelled 60K and 84K proteins were not found in this case. Ox-GTP-labelled viral proteins were not isolated with preimmune serum (Fig. 6, lane 6), nor from extracts of uninfected cowpea leaves (Fig. 6, lanes 3, 5 and 7).

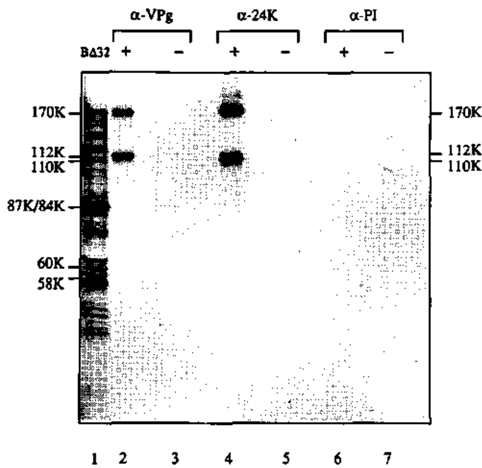


Fig. 6. SDS-polyacrylamide gel electrophoresis of labelled proteins from immunoprecipitates of extracts incubated with ox-[α-^{32}P]GTP. Labelled proteins from CPMV-infected (+) and mock-inoculated (-) extracts were immunoprecipitated with anti-VPg (lanes 2 and 3), anti-24K (lanes 4 and 5) or preimmune serum (lanes 6 and 7). Indexes on the right correspond to immunoprecipitated protein positions. On the left marker protein positions are also indicated.

The results make it clear that the 60K and 84K proteins can distinguish between ox-ATP and ox-GTP indicating that binding of ox-ATP occurs to a nucleotide specific site. The 60K and 84K (60K+24K) proteins have the 58K and VPg region in common and therefore the binding site is most likely located in this portion rather than in the 24K region. Moreover, labelling of the 24K protein with ox-ATP was not observed (Fig. 4, lane 5). Although the possibility that the ox-ATP binds to 60K and 84K proteins in the VPg region cannot be entirely excluded, it is much more likely that the binding is in the 58K region, which contains the NTBM. We have not been able to identify the 58K protein, because we do not have an antiserum at our disposal, but we speculate that the 58K protein will also bind ATP. Although the exact location of the ATP-binding site remains to be established, the conserved Lys500 residue with the ε-amino group located in the A-site of the NTBM in the 58K domain meets the needs for covalent linkage (Walker *et al.*, 1982; Dever *et al.*, 1987) (see Fig. 2). Furthermore, there are no indications for putative ribonucleotide binding sites other than the NTBM consensus sequence located in the 58K region. The NTBM lacks the conserved element that is found in GTP-binding proteins and in agreement with that the NTBM was found to be specific for ATP (Dever *et al.*, 1987).

A second ribonucleotide binding site is located in the 87K region, since the 87K core polymerase as well as the 110K, 112K, and 170K precursors, which contain the 87K domain, are active in binding of ATP and GTP. UV-crosslinking assays have shown that the 87K and the larger precursor proteins are also active in binding of CTP and UTP (data not shown). Therefore, the 87K region appears to be able to bind all four ribonucleotides as might be expected for a polymerase protein. The exact position of the ribonucleotide binding site in the 87K protein remains yet to be established. Furthermore, at present we cannot distinguish whether binding of ATP in the 170K binding occurs only at the 87K core polymerase domain or both at the 87K and the 58K domain at the same time.

Our finding that the 60K and 84K proteins have an active ATP-binding site raises the question if the binding has functional significance and which role can be attributed to the 60K and 84K proteins in viral RNA replication. Amino acid sequence homology studies show that the 60K protein of CPMV and the 2C protein of poliovirus (PV) share significant homology (Franssen *et al.*, 1984) and both proteins have been classified into the family of "picorna-like" NTBM-containing proteins (Gorbalenya *et al.*, 1989a, 1989b). Further similarities between PV and CPMV include their genome organization and expression (Argos *et al.*, 1984; Goldbach 1987) suggesting that these viruses employ a similar mechanism for viral RNA replication. For both viruses viral RNA replication takes place at cytoplasmic membrane vesicles. Indeed, both 60K and 2C proteins have been implicated in the induction and proliferation of vesicular membranes (Van Bokhoven *et al.*, 1992, Bienz *et al.*, 1983; Tershak, 1984), which have been shown to be the site of viral RNA replication (Assink *et al.*, 1973; De Zoeten *et al.*, 1983; Rezelman *et al.*, 1982). Recently, it was demonstrated that the NTBM of 2C is essential for viral RNA replication (Mirzayan and Wimmer, 1992) and that 2C possesses ATPase and GTPase activities (Rodriguez and Carrasco, 1993). No data are available on how the formation of these vesicles is induced, but 60K and 2C proteins might act by interference with the membrane synthesizing machinery of the host cell in a ribonucleotide-dependent fashion. Besides, the NTBM may be involved in multiple functions necessary for viral RNA replication. As for that, it has been suggested that NTBM-containing proteins encoded by RNA viruses are involved in RNA helicase activity during genome replication (Gorbalenya *et al.*, 1989a). Although experimental data are scarce, the NTBM-containing CI protein of plum pox virus (PPV, potyviridae), which shares some sequence homology with the CPMV 60K and polioviral 2C protein (Domier *et al.*, 1987), has been identified as an RNA-dependent RNA-helicase performing its activity at the expense of ATP (Lain *et al.*, 1990, 1991). Considering these possibilities, currently experiments are

in progress to further define a role for the B-RNA encoded NTBM in CPMV-RNA replication.

Acknowledgements

We thank Dr. O.C. Richards for suggesting the use of periodate-oxidized ribonucleotides, and Nicole ter Maten for photography. This work was supported by the Netherlands Foundation of Chemical Research (SON) with financial aid from the Netherlands Organization for Scientific Research (NWO).

References

Argos, P. (1988). A sequence motif in many polymerases. *Nucleic Acids Res.* **16**, 9909-9916.

Argos, P., Kamer, G., Nicklin, M.J.H., and Wimmer, E. (1984). Similarity in gene organization and homology between proteins of animal picorna viruses and a plant comovirus suggest common ancestry of these virus families. *Nucleic Acids Res.* **12**, 7251-7266.

Assink, A.M., Swaans, H., and 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**, 341-355.

Bienz, K., Egger, D., and Pasamontes, L. (1987). Association of polioviral proteins of the P2 genomic region with the viral replication complex and virus-induced membrane synthesis as visualized by electron microscopic immunocytochemistry and autoradiography. *Virology* **160**, 220-226.

Bienz, K., Egger, D., Rasser, Y., and Bossart, W. (1983). Intracellular distribution of poliovirus proteins and the induction of virus-specific cytoplasmic structures. *Virology* **131**, 39-48.

Bienz, K., Egger, D., Troxler, M., and Pasamontes, L. (1990). Structural organization of poliovirus RNA replication is mediated by viral proteins of the P2 genomic region. *J. Virol.* **64**, 1156-1163.

Blake, M.S., Johnston, K.H., Russell-Jones, G.J., and Gotschlich, E.C. (1984). A rapid, sensitive method for detection of alkaline phosphate-conjugated antibody on Western blots. *Anal. Biochem.* **136**, 175-179.

Bradley, M.K., Smith, T.F., Lathrop, F.H., and Livingston, D.M. (1987). Consensus topography in the ATP binding site of the SV40 and polyomavirus large tumor antigens. *Proc. Natl. Acad. Sci. USA* **84**, 4026-4036.

Clertant, P., and Cuzin, F. (1982). Covalent affinity labelling by periodate-oxidized [α-^{32}P] ATP of the large-T proteins of Polyoma and SV40 viruses. *J. Biol. Chem.* **257**, 6300-6305.

De Vos, A.M., Tong, L., Milburn, M.V., Matias, P.M., Jancarik, J., Noguchi, S., Nishimura, S., Miura, K., Ohtsuka, E., and Kim, S.H. (1988). Three-dimensional structure of an oncogene protein: catalytic domain of human c-H-*ras* p21. *Science* **239**, 888-893.

De Zoeten, G.A., Assink, A.M., Van Kammen, A., and Zabel, P. (1983). Association of cowpea mosaic virus-induced double-stranded RNA with a cytopathological structure in infected cells. *Virology* **59**, 341-355.

Dever, T.E., Glynias, M.J., and Merrick, W.C. (1987). GTP-binding domain: Three consensus sequence elements. *Proc. Natl. Acad. Sci. USA* **84**, 1814-1818.

Domier, L.L., Shaw, J.G., and Rhoades, R.E. (1987). Potyviral proteins share amino acid sequence homology with picorna-, como-, and cauliflower viral proteins. *Virology* **158**, 20-27.

Dorssers, L., Van der Krol, S., Van der Meer, J., Van Kammen, A., and Zabel, P. (1984). Purification of cowpea mosaic virus RNA replication complex: Identification of a virus-encoded 110,000-dalton polypeptide responsible for RNA chain elongation. *Proc. Natl. Acad. Sci. USA* **81**, 1951-1955.

Eggen, R. and Van Kammen, A., (1988). RNA replication in comoviruses. *In* "RNA Genetics" (P. Ahlquist, J. Holland, and E. Domingo, Eds.), Vol. 1, pp. 49-69. CRC Press, Boca Raton.

Eggen, R., Kaan, A., Goldbach, R., and Van Kammen, A. (1988). Cowpea mosaic virus RNA replication in crude membrane fractions from infected cowpea and *Chenopodium amaranticolor*. *J. Gen. Virol.* **69**, 2711-2720.

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, Leunissen J., Goldbach, R., Lomonossoff, G., and Zimmern, D. (1984). Homologous sequences in non-structural proteins from cowpea mosaic virus and picornaviruses. *EMBO J.* **3**, 855-861.

Fry, D.C., Kuby, S.A., and Mildvan, A.S. (1986). ATP-binding site of adenylate kinase: mechanistic implications of its homology with ras-encoded p21, F_1-ATPase, and other nucleotide-binding proteins. *Proc. Natl. Acad. Sci. USA* **83**, 907-911.

Garcia, J.A., Schrijvers, L., Tan, A., Vos, P., Wellink, J., and Goldbach, R. (1987). Proteolytic activity of cowpea mosaic virus encoded 24K protein synthesized in *Escherichia coli*. *Virology* **159**, 67-75.

Guinea, R., and Carrasco, L. (1990). Phospholipid biosynthesis and poliovirus genome replication, two coupled phenomena. *EMBO J.* **9**, 2011-2016.

Guinea, R., and Carrasco, L. (1991). Effects of fatty acids on lipid synthesis and viral RNA replication in poliovirus-infected cels. *Virology* **185**, 473-476.

Goldbach, R., Rezelman, G., and Van Kammen, A. (1980). Independent replication and expression of B-component RNA of cowpea mosaic virus. *Nature, (London)* **286**, 297-300.

Goldbach, R.W. (1987). Genome similarities between plant and animal viruses. *Microbiol. Sci.* **4**, 197-202.

Gorbalenya, A.E., and Koonin, E.V. (1989). Viral proteins containing the purine NTP-binding sequence pattern. *Nucleic Acids Res.* **17**, 8413-8440.

Gorbalenya, A.E., Blinov, V.M., Donchenko, A.P., and Koonin, E.V. (1989a). An NTP-binding motif is the most conserved sequence in a highly diverged monophyletic group of proteins involved in positive strand RNA viral replication *J. Mol. Evol.* **28**, 256-258.

Gorbalenya, A.E., Koonin, E.V., Donchenko, A.P., and Blinov, V.M. (1989b). Two related superfamilies of putative helicases involved in replication, recombination, repair and expression of DNA and RNA genomes. *Nucleic Acids Res.* **17**, 4713-4730.

Halliday, K.R. (1984). Regional homology in GTP-binding proto-oncogene products and elongation factors. *J. Cyclic Nucleotide Protein Phosphorylation Res.* **9**, 435-448.

Ishihama, A., and Nagata, K. (1988). Viral polymerases. Critical reviews in Biochemistry, Vol. 23, pp 27-76. CRC Press, Boca Raton.

Jurnak, F. (1988). The three-dimensional structure of c-H-ras p21: implications for oncogene and G protein studies. *Trends Biochem. Sci.* **13**, 195-198.

Kamer, G., and Argos, P. (1984). Primary structural comparison of RNA dependent polymerases from plant, animal and bacterial viruses. *Nucleic Acids Res.* **12**, 7269-7282.

Kuhn, R.J., and Wimmer, E. (1987). The replication of picorna viruses, p. 17-51. *In* "The molecular biology of positive strans RNA viruses". (D.J., Rowlands, B.W.J. Mahy, and M., Mayo, Eds.). Academic Press, Inc., New York & London.

La Cour, T.F.M., Nyborg, J., Thirup, S., and Clark, B.F.C. (1985). Structural details of the binding of guanosine diphosphate to elongation factor Tu from *Escherichia coli* as studied by X-ray crystallography. *EMBO J.* **4**, 2385-2388.

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

Lain, S., Martin, M.T., Riechmann, J.L. and Garcia, J.A. (1991). Novel catalytic activity associated with positive-strand RNA virus infection: nucleic acid-stimulated ATPase activity of the plum pox potyvirus helicaselike protein. *J. Virol.* **65**, 1-6.

Lain, S., Riechmann, J.L., and Garcia, J.A. (1990). RNA helicase: a novel activity associated with a protein encoded by a positive strand RNA virus. *Nucleic Acids Res.* **23**, 7003-7006.

Maynell, L.A., Kirkegaard, K., and Klymkowsky, M.W. (1992). Inhibition of polioviral RNA synthesis by brefeldin A. *J. Virol.* **66**, 1985-1994.

Mirzayan, C., and Wimmer, E. (1992). Genetic analysis of an NTP-binding motif in poliovirus polypeptide 2C. *Virology* **189**, 547-555.

Morrison, R.T., and Boyd, R.N. (1974). Organic chemistry. Third edition, pp 1258 Allyn and Bacon, Inc. Boston, U.S.A.

Möller, W., and Amons, R. (1985). Phosphate-binding sequences in nucleotide-binding proteins. FEBS *Lett.* **186**, 1-7.

Peters, S.A., Voorhorst, W.B.G., Wellink, J., and Van Kammen, A. (1992a). Processing of VPg-containing polyproteins encoded by the B-RNA from cowpea mosaic virus. *Virology* **191**, 90-97.

Peters, S.A., Voorhorst, W.G.B., Wery, J., Wellink, J., and Van Kammen, A. (1992b). A regulatory role for the 32K protein in proteolytic processing of cowpea mosaic virus polyproteins. *Virology* **191**, 81-89.

Rayford, R., Anthony, jr., D.D., O'Neill, jr., R.E., and Merrick, W.C. (1985). Reductive alkylation with oxidized nucleotides. *J. Biol. Chem.* **260**, 15708-15713.

Rodriguez, P., and Carrasco, L. (1993). Poliovirus protein 2C has ATPase and GTPase activities. *J. Biol. Chem.* (in press).

Rezelman, G., Franssen, H.J., Goldbach R.W., Ie, T.S., and Van Kammen A. (1982). Limits to the independence of bottom component RNA of cowpea mosaic virus. *J. Gen. Virol.* **60**, 335-342.

Tershak, D.R. (1984). Association of poliovirus proteins with the endoplasmatic reticulum. *J. Virol.* **52**, 777-783.

Van Bokhoven, H., Van Lent, J.W.M., Custers, R., Vlak, J.M., Wellink, J., and Van Kammen, A. (1992). Synthesis of the complete 200K polyprotein encoded by cowpea mosaic virus B-RNA in insect cells. *J. Gen. Virol.* **73**, 2775-2784.

Van Bokhoven, H., Verver, J., Wellink, J., and Van Kammen. (1993b). Protoplasts transiently expressing the 200K coding sequence of cowpea mosaic virus B-RNA support replication of M-RNA. *J. Gen. Virol.* in press.

Walker, J.E., Saraste, M., Runswick, M.J., and Gay, N.J. (1982). Distantly related sequences in the α- and β-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* **1**, 945-951.

Wellink, J., Jaegle, M., and Goldbach, R. (1987). Detection of a novel protein encoded by the bottom-component RNA of cowpea mosaic virus, using antibodies raised against a synthetic peptide. *J. Virol.* **61**, 236-238.

Wellink, J., Rezelman, G., Goldbach, R., and Beyreuther, K. (1986). Determination of the proteolytic processing sites in the polyprotein encoded by the bottom-component RNA of cowpea mosaic virus. *J. Virol.* **59**, 50-58.

Zabel, P., Jongen-Neven, I., and Van Kammen, A. (1976). *In vitro* replication of cowpea mosaic virus. II. Solubilization of membrane-bound replicase and partial purification of the solubilized enzyme. *J. Virol.* **17**, 679-685.

Chapter 6

The NTP-Binding Motif in the Cowpea Mosaic Virus B-polyprotein is Essential for Viral Replication

Sander A. Peters[1#], Ellen A.A. Nollen[1], Jan W.M. van Lent[2], Joan Wellink[1], and Ab van Kammen[1]

1. Department of Molecular Biology Agricultural University, Dreijenlaan 3, 6703 HA Wageningen, and 2) Department of Virology, Agricultural University, Binnenhaven 11, 6709 PD Wageningen, The Netherlands

Running title: The NTBM in the CPMV B-polyprotein

Present adress: Department of Virology, Institute of Medical Microbiology, State University Leiden, P.O. Box 320, 2300 AH Leiden, The Netherlands

Accepted for publication in J. Gen. Virology

The NTP-binding motif in the CPMV B-polyprotein is essential for viral replication

Abstract

We have assessed the functional importance of the ribonucleotide binding motif (NTBM) in the CPMV B-RNA encoded 58K domain by changing two conserved amino acids within the consensus A and B-sites (GKSRTGK500S and MDD545, respectively). Both Lys500Thr and Asp545Pro substitutions are lethal as mutant B-RNAs were no longer replicated in cowpea protoplasts. Transiently produced mutant proteins were not able to support *trans* replication of CPMV M-RNA in cowpea protoplasts in contrast to transiently produced wild-type B-proteins. Therefore loss of viral RNA synthesis was a result of a protein defect rather than due to an RNA template defect. Mutant B-polyproteins were correctly processed *in vitro* and *in vivo* and the regulatory function of the 32K protein on processing of B-proteins was not affected by these mutations. Since regulation of processing by the 32K protein depends on the interaction with the 58K domain, the mutations in the NTBM apparently do not interfere with this interaction. The Asp545Pro substitution left the binding properties of the 84K precursor of the 60K protein to ATP-agarose intact, whereas the Lys500Thr substitution decreased the binding capacity of the 84K protein, suggesting that the Lys500 residue is directly involved in ATP binding. The Lys500Thr substitution in the 58K domain resulted in an altered distribution of viral proteins, which failed to aggregate into large cytopathic structures that can be observed in wild-type B-RNA infected protoplasts. This might reflect the involvement of the NTBM in protein-protein interactions. Viral proteins containing the Asp545Pro substitution showed a normal distribution over the protoplasts.

Introduction

The genome of cowpea mosaic virus (CPMV) consists of two plus-strand RNAs, denoted B-RNA and M-RNA, that are translated into large polyproteins. The B-RNA encoded 200K polyprotein is proteolytically processed at specific Gln/Ser, Gln/Met and Gln/Gly sites by the viral encoded 24K proteinase to yield mature functional proteins (Wellink *et al.*, 1986; Garcia *et al.*, 1987; Vos *et al.*, 1988). A detailed processing scheme for the 200K polyprotein is shown in Fig. 1. Proteolytic

processing has been shown to be regulated by the B-RNA encoded 32K protein which forms a complex with the 170K protein (Peters *et al.*, 1992). The B-RNA and its encoded enzyme activities constitute an autonomous RNA replicon, since the B-RNA has been shown to replicate independently from M-RNA in isolated plant cells (Goldbach *et al.*, 1980). On the other hand B-RNA is dependent on M-RNA encoded activities for movement from cell-to-cell (Wellink and Van Kammen, 1989). The action of each individual protein in the replicative machinery is not fully understood. Sequence homology studies have shown a conservation between comoviral- and corresponding picorna-, nepo-, poty-, bymo-, and calicivirus nonstructural proteins (Franssen *et al.*, 1984a; Gorbalenya *et al.*, 1989a, 1990; Neill, 1990; King *et al.* 1991). These include proteins containing a consensus sequence that corresponds to the so-called nucleoside triphosphate binding motif (NTBM). This is a characteristic amino acid sequence motif for ribonucleotide utilizing proteins, consisting of two separate parts referred to as A and B-site (Walker *et al.*, 1982). Available experimental data suggests that NTBM-containing viral proteins might in fact be NTPases that are involved in processes such as duplex unwinding during RNA replication, trancription, mRNA translation, signal transduction, and membrane transport (Walker *et al.*, 1982; Gorbalenya and Koonin, 1989; Gorbalenya *et al.*, 1989b, and references therein).

In the CPMV B-polyprotein an NTBM is located in the N-terminal part of the 58K domain (Dever *et al.*, 1987; Gorbalenya and Koonin, 1989) and is contained within different processing products (see Fig. 1). Several lines of evidence support a role for the NTBM-containing 60K protein in viral RNA replication. Chemical cross-linking studies have shown that the 60K and 84K (60K+24K) proteins specifically bind ATP suggesting that the NTBM is functionally important (S. Peters, unpublished results). The B-RNA encoded 60K protein is involved in the induction of vesicular structures in insect cells (Van Bokhoven *et al.*, 1992). Such cytophatic structures are also observed in CPMV B-RNA infected cowpea mesophyll protoplasts and have been shown to be the site of viral RNA replication (Assink *et al.*, 1973; De Zoeten *et al.*, 1974; Rezelman *et al.*, 1982). Furthermore it has been suggested that the 60K protein has a role in anchoring the viral replication complex to these vesicular structures (Eggen and Van Kammen, 1988). Next to the vesicles amorphous structures are present that appear as electron-dense material upon electron microscopic analysis and were shown to contain B-RNA encoded proteins (Wellink *et al.*, 1988). The induction of these cytopathic structures is a property of B-RNA encoded proteins and has been shown to be independent of plant factors and viral replication (Van Bokhoven *et al.*, 1992).

To further investigate the role of the NTBM in viral proliferation we have introduced site specific mutations in the NTBM coding sequence and have tested

the effect on viral RNA replication, translation and polyprotein processing, and cellular localization of viral proteins. The results obtained show that the NTBM is essential for viral RNA replication.

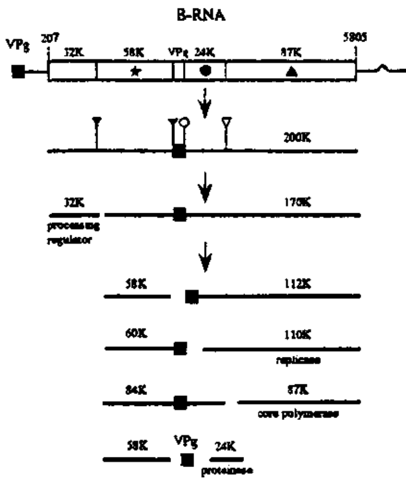


Fig. 1. Expression and genetic organization of B-RNA from CPMV. The RNA contains a long open reading frame (shown by open bar). The postions of the start and stop codon are indicated. Cleavage sites are indicated by (○) Gln/Met, (▽) Gln/Gly, and (▼) Gln/Ser. B-RNA encodes a 200K polyprotein which is processed into 32K and 170K proteins. The 170K is processed at three different sites into three sets of proteins of 58K+112K, 60K+110K, and 84K+87K. VPg can be generated via cleavage of the 112K protein or 84K protein. The 110K and 60K have been shown to be stable cleavage products. Symbols within the B-RNA represent conserved domains: ★ = nucleotide binding domain; ● = proteinase domain; ▲ = polymerase domain.

Materials and Methods

Oligo-directed site specific mutagenesis

To introduce mutations in the 58K coding region plasmid pTB1G (Eggen *et al.*, 1989) was used as starting material. Plasmid pTB1G contains the full-length cDNA of CPMV B-RNA, which is cloned behind a bacteriophage T7 RNA-polymerase promoter. RNA transcripts derived from pTB1G have been shown to be infectious in cowpea mesophyll protoplasts (Eggen *et al.*, 1989). The positions of restriction sites and nucleotides refer to the position in the B-RNA sequence determined by Lomonossoff and Shanks (1983). A NsiI-SstI fragment from pTB1G (positions 1393 to 2301) was inserted into the PstI-SstI site of M13mp19. The phosphorylated oligonucleotide (A) 5'-CCTGACTCATCAGCAAACTGGTACCAGTGCGTGACTTTCC-3', complementary to nts 1687-1726 of the 58K coding region, and (B) 5'-CGGCGGCAAAAGGATCCATCAGCAC-3', complementary to nts 1827-1851, were used for site-directed mutagenesis according to the method described by Kunkel (1985). The nucleotides that differ from the wild-type sequence are shown in bold. Using oligonucleotide A two nucleotide substitutions at position 1705 (A to C) and 1706 (A to C) create an additional KpnI site (underlined), whereas in oligonucleotide B two nucleotide substitutions at position 1839 (G to C) and 1840 (A to C) generate an

additional BamHI site (underlined). Oligodeoxynucleotides were synthesized with a cyclone DNA synthesizer (Biosearch). Standard recombinant DNA techniques were used for transformation in competent *E.coli* DH5aF' as described (Sambrook *et al.*, 1989). Recombinant clones M13mp19K500T and M13mp19D545P were analysed by restriction enzyme mapping and nucleotide sequence analysis as described (Sanger *et al.*, 1977; Korneluk *et al.*, 1985).

Construction of plasmids

The PvuII-SstI fragments from M13mp19K500T and M13mp19D545P were reinserted in the PvuII-SstI site (positions 1625 to 2301) of pTB1G and generated mutant transcription vectors pTBK500T and pTBD545P. To produce B-proteins in cowpea mesophyll protoplasts that contain amino acid substitutions in the NTBM, pMB200 (Van Bokhoven *et al.*, 1993) and pMB116 were used. pMB116 was generated by inserting a 3857-bp BglII/BamHI fragment from pMB200 into the BglII/BamHI digested vector pMON999 (Van Bokhoven *et al.*, 1993). pMB200 and pMB116 direct the transient expression of the 200K polyprotein and the 116K protein (32K+60K+24K) respectively in cowpea mesophyll protoplasts under the control of the CaMV 35S promoter. The PvuII-SstI fragments (positions 1625 to 2301) of these expression vectors were exchanged with the PvuII-SstI fragments from M13mp19K500T and M13mp19D545P and generated the mutant transient expression vectors pMB200K500T, pMB200D545P, pMB116K500T and pMB116D545P respectively.

In vitro transcription and translation

The procedures for *in vitro* transcription and translation of RNA templates have been described recently (Peters *et al.*, 1992). In the present experiments aliquots of transcription mixtures were translated in rabbit reticulocyte lysate from Promega Biotec.

Transfection of protoplasts, transient expression and preparation of subcellular fractions

Cowpea mesophyll protoplasts were transfected with CPMV-RNA, B-RNA or transient expression vectors or with both transient expression vectors and M-RNA as described by Van Bokhoven *et al.* (1993). Protoplasts were harvested 20 hr post transfection and subjected to fractionation by centrifugation at 30,000 × g for 30 min. Extracts were prepared as described by Franssen *et al.* (1982).

Northern blot analysis

Total nucleic acids from infected and uninfected cowpea mesophyll protoplasts were extracted 40 hr post transfection. Approximately 1 × 10^6 protoplasts were mixed with 200 µl 100 mM Tris-HCl pH7.5, 10 mM EDTA and 1% SDS and 200 µl phenol and vortexed for 2 min at room temperature. After addition of 200 µl chloroform the suspension was vortexed for 5 min followed by 10 min centrifugation at 16,000 × g. The aquous phase was transfered to a fresh tube and extracted with 1ml chloroform. The nucleic acid was recovered by precipitation with ethanol and sodium acetate and separated on a 1% denaturing agarose gel as described by McMaster and Carmichael (1977). The RNA was transfered from the gel to GeneScreene filter (NEN Research Products) and prehybridized as recommended by the manufacturer. Hybridization overnight was performed under the same conditions with 9 × 10^6 cpm of a B-RNA or M-RNA specific probe generated by random primer labelling as described (Feinberg and Vogelstein, 1983), using a SstI-BamHI fragment (positions 2301 to 3857) from pTB1G or a BglII-NcoI fragment (positions 189 to 3068) from pTM1G. After hybridization the blot was washed two times with 2 × SSC, 0.5% SDS at 42° for 30 min, followed by autoradiography.

Affinity chromatography on 5'ATP-agarose

A frozen pellet of approximately 1.5 × 10^6 protoplasts was mixed with 100 µl in binding buffer (20 mM Tris-HCl pH 7.5, 25 mM NaCl, 5 mM $MgCl_2$, 5 mM DTT, 0.5 mM PMSF, and 1 mM EDTA). Supernatant fractions prepared by centrifugation at 30,000 x g for 30 min were mixed with 50 µl 5'ATP-agarose (Sigma; the ligand is coupled via an 11 atom spacer to the agarose) in an eppendorf tube and incubated for 30 min at 4°. In addition, fractions were chromatographed in binding buffer in presence of 2.5 mM ATP to examine the binding specificity. The supernatant fraction was removed after the agarose was spun down for 30 sec at low speed. The agarose was washed three times with 1 ml bindingbuffer and then boiled in 50µl protein sample buffer for 3 min to release the bound proteins from the agarose. Subsequently, proteins were fractionated in an SDS-polyacrylamide gel and identified by immunoblotting.

Immunological detection of CPMV B-proteins and gel electrophoresis

Protein samples were heated in sample buffer (10% glycerol, 5% β-mercaptoethanol, 2% SDS, 0.01% bromophenol blue, 75 mM Tris-HCl pH 6.8) for 3

min at 100°. Proteins were subjected to fractionation in an 7.5% SDS-polyacrylamide gel according to Laemmli (1970). Gels were either dried and autoradiographed with Kodak X-Omat film or analysed by immunoblotting using anti-VPg (Eggen *et al.*, 1988) or anti-110K serum (Van Bokhoven *et al.*, 1992) as a primary antibody and anti-rabbit IgG alkaline phophatase (Promega Biotec) as a second antibody as described (Blake *et al.*, 1984). Immunofluorescent detection of CPMV B-proteins in infected cowpea mesophyll protoplasts was performed with the same antisera and fluorescin isothiocyanate (FITC)-conjungated goat anti-rabbit immunoglobulin (Nordic) as described by Van Lent *et al.* (1991).

Results

Mutations in the nucleoside triphosphate binding motif (NTBM) coding region are detrimental to viral RNA replication

The 58K domain of the B-polyprotein contains two highly conserved amino acid sequences, GKSRTGK500S and MDD545 respectively, that correspond to the consensus A-site and B-site found in nucleoside triphosphate utilizing proteins (Walker *et al.*, 1982; Dever *et al.*, 1987; Gorbalenya and Koonin, 1989; Gorbalenya *et al.*, 1990). To assess the functional importance of these two conserved elements, a single amino acid substitution was introduced in each element and the effect on virus viability was analysed. Translation of the BK500T-RNA will result in a polyprotein, in which the conserved Lys500 (AAG) residue in the A-site of the NTBM has been replaced by a Thr (ACC) residue. In BD545P-RNA the coding sequence for the conserved Asp545 (GAU) residue in the B-site of the NTBM has been changed into a Pro (CCU) residue. Upon transfection of cowpea mesophyll protoplasts with BK500T-RNA or BD545P-RNA, no accumulation of viral proteins was detected, using an immunofluoresence assay, whereas in B-RNA or CPMV-RNA transfected protoplasts viral proteins were produced (data not shown). This suggests that BK500T-RNA and BD545P-RNA cannot replicate in cowpea protoplasts. This was confirmed by a Northernblot analysis. The hybridization signal indicated an efficient replication of B-RNA and CPMV-RNA (Fig. 2, lanes 1 and 3), whereas hybridization could not be detected with RNA extracted from protoplasts transfected with BK500T-RNA, BD545P-RNA or mock-inoculated protoplasts (Fig. 2, lanes 2, 3 and 5). Apparently, either the NTBM or the coding sequence is crucial for viral replication.

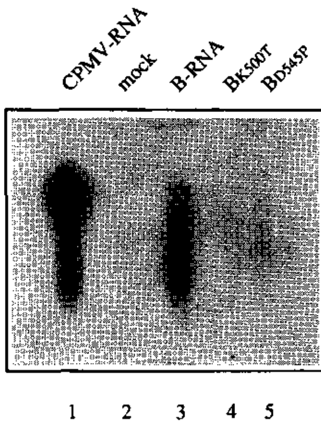


Fig. 2. Synthesis of CPMV B-RNA in transfected mesophyll protoplasts. Protoplasts were transfected with CPMV-RNA (lane 1), mock-inoculated (lane 2), wild-type B-RNA (lane 3), BK500T-RNA (lane 4) or BD545P-RNA (lane 5). After 40 hr of incubation total RNA was extracted from protoplasts. RNA products were separated, transfered onto membrane filter and hybridized with a B-RNA specific ^{32}P-labelled probe.

In vitro processing of mutant B-RNA translation products

To determine the integrity of mutant B-RNAs and to study the effect of the Lys500Thr and Asp545Pro substitutions on proteolytic processing *in vitro*, mutant B-transcripts were translated in rabbit reticulocyte lysate and the translation products were compared to those of wild-type B-RNA.

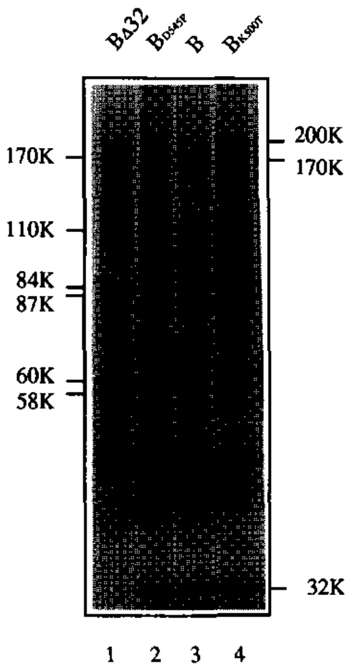


Fig. 3. Proteolytic processing of translation products containing amino acid substitutions in the NTBM consensus sequence located in the 58K domain of the CPMV B-polyprotein. BD545P-RNA (lane 2), B-RNA (lane 3) and BK500T-RNA (lane 4) were translated in presence of [^{35}S]methionine for 4 hr. The products were separated on a 7.5% SDS polyacrylamide gel and their positions are indicated at the right side. At the left side positions of marker proteins obtained from a BΔ32-RNA translation (Peters *et al.*, 1992) (lane 1) are indicated.

Similar to wild-type B-RNA, BK500T-RNA and BD545P-RNA were also translated into 200K polyproteins (Fig. 3 lanes 2 and 4). This indicates that the large open reading frame of both mutant RNAs was intact. Processing of the mutant 200K polyproteins generated the 170K and the 32K proteins, which comigrated with the wild-type primary processing products. This suggests that the mutant polyproteins are correctly processed. The primary processing of the BD545P-polyprotein, however, was less efficient after 4 hr of incubation resulting in smaller amounts of the 170K and 32K primary cleavage products, as compared to the amount of wild-type processing products (Fig. 3, lanes 2 and 3). Possibly as a result of the Asp545Pro substitution, the conformation of the nascent polypeptide chain has changed such that intramolecular cleavage at the Gln/Ser site is less favourable. Processing at secondary cleavage sites in the mutant 170K proteins was not observed (Fig. 3, lanes 2 and 4). Therefore the Lys500Thr and Asp545Pro substitutions appear to have no influence on the regulatory role of the 32K protein in processing of the 170K protein (Peters *et al.*, 1992).

Transient expression of mutant B-RNA coding sequences

To study the effect on proteolytic processing *in vivo*, CPMV B-RNA sequences carrying mutations in the NTBM coding region were transiently expressed in protoplasts using pMB200K500T and pMB200D545P. The processing products of the wild-type 200K protein produced by pMB200 and mutant 200K proteins from pMB200K500T and pMB200D545P were indistinguishable from the viral proteins found in CPMV-RNA inoculated protoplasts upon analysis with anti-110K serum (Fig. 4, lanes 1 to 4), or anti-VPg serum (Fig 4, lanes 5 to 14). This indicates faithful processing of the transiently produced wild-type and mutant proteins *in vivo*. *In vivo* processing of the mutant 200K-D545P proteins appeared to be equally efficient to wild-type 200K protein (Fig. 4, lanes 2, 3, 7 and 9), in contrast with *in vitro* processing, which showed a decrease in efficiency for the 200K-D545P-polyprotein processing (see Fig. 3, lane 2). Protoplasts transfected with pMB200K500T produced a lower amount of viral proteins than wild-type B-RNA or pMB200D545P transfected protoplasts. The lower level was most pronounced for the 170K, 110K and 87K proteins (see Fig. 4, lanes 2 and 4). Cleavage of the 116K protein (32K+60K+24K) produced by pMB116 generated 84K and 32K proteins (the latter protein not shown) (Fig. 4, lane 12), but further processing of the 84K protein into 60K and 24K was not observed. This is consistent with earlier *in vitro* results and again confirms our notion that the 32K protein can arrest the proteolytic processing of B-proteins (Peters *et al.*, 1992). With respect to cleavage patterns observed for the mutant 116K and the wild-type 116K proteins,

differences were not found either (Fig. 4, lanes 12 to 14). Apparently, the nonviability of the mutant transcripts BK500T and BD545P is not caused by interference of the Lys500Thr and Asp545Pro substitutions with translation or proteolytic processing of viral proteins.

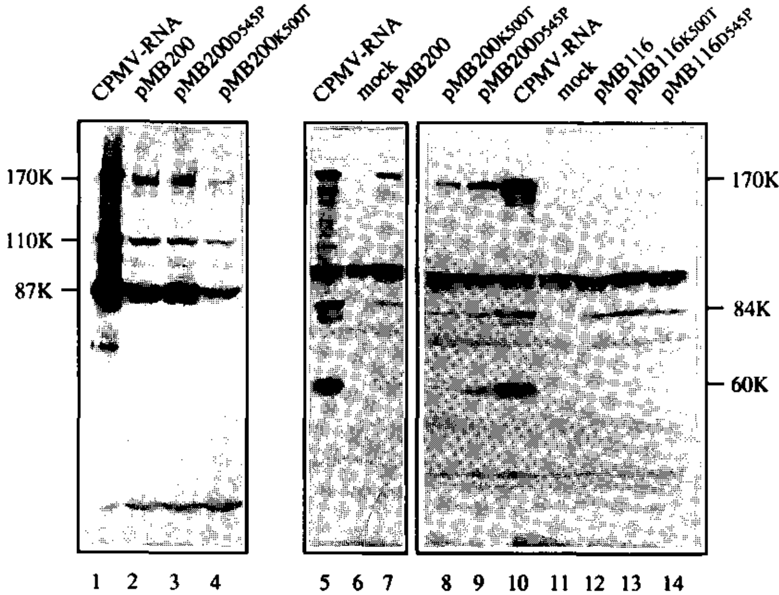


Fig. 4. Immunoblot analysis of proteins transiently produced in cowpea protoplasts. For detection of proteins anti-110K (lanes 1 to 4) and anti-VPg (lanes 5 to 14) sera were used as primary antibodies and anti-rabbit alkaline phosphatase conjungate as the second antibody. Viral proteins from CPMV-RNA inoculated protoplasts (lanes 1, 5 and 10) served as marker proteins and are indicated on the right and the left side. The nomenclature of the constructs that direct transient expression of B-RNA sequences, shown on top of the immunoblot, is described under Materials and Methods.

The CPMV B-proteins transiently produced by pMB200 are active in *trans* replication of M-RNA in protoplasts (Van Bokhoven *et al.*, 1993). In contrast to wild-type B-proteins the mutant proteins from pMB200K500T and pMB200D545P were not able to support the replication of M-RNA (see Fig. 5, lanes 2 and 3). This indicates that a protein function essential to viral RNA replication has been disrupted by the amino acid replacements.

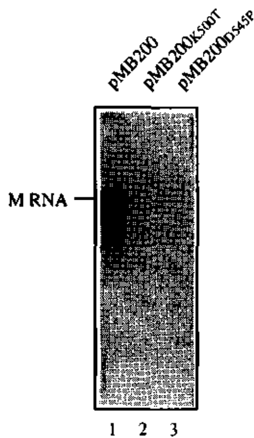


Fig. 5. Synthesis of CPMV M-RNA by B-proteins transiently expressed in protoplasts. Protoplasts were transfected with pMB200 (lane 1) pMB200K500T (lane 2) or pMB200D545P (lane 3) together with M-RNA. Total RNA extracted at 40 hr post transfection was separated in a 1% denaturing agarose gel followed by transfer of the RNAs onto GeneScreen. The blot was hybridized with a probe specific for M-RNA.

Chromatography of CPMV B-proteins on 5'ATP agarose

To investigate whether introduced amino acid substitutions in the NTBM disturbed a ribonucleotide binding activity, extracts prepared from transfected protoplasts were incubated with ATP-agarose. Bound viral proteins were subsequently eluted from the absorbent with protein sample buffer and subjected to an immunoblot analysis. In extracts prepared from CPMV infected protoplasts the 170K, 112K, 110K, 87K, 84K and 60K proteins, which were detected with anti-110K and anti-VPg serum, had bound to ATP-agarose (Fig. 6, lane 3 and Fig. 7, lane 6). The binding of 170K, 112K, 110K, 87K, 84K and 60K proteins was inhibited in presence of 2.5 mM ATP (Fig. 6, lane 5). The competition for binding by ATP indicates the specificity of binding of the afore-mentioned proteins. The ATP-binding sites are probably located in the 87K core-polymerase domain and in the 58K domain.

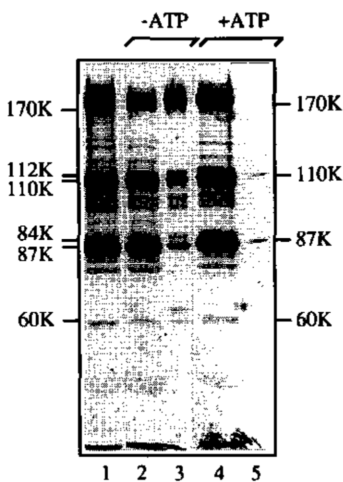


Fig. 6. Competition of binding of CPMV B-proteins to 5'ATP-agarose by ATP. Extracts prepared from 1.5×10^6 protoplasts infected with CPMV-RNA were incubated with ATP-agarose either in absence of ATP (lanes 2 and 3) or in presence of 2.5 mM ATP (lanes 4 and 5). After incubation unbound proteins (lanes 2 and 4) were separated from the ATP-agarose by centrifugation and immobilized proteins were then eluted from the absorbent by boiling in sample buffer (lanes 3 and 5). Proteins were fractionated in an SDS-polyacrylamide gel. Detection of viral proteins was performed with both anti-VPg and anti-110K serum. Marker proteins (lane 1) are indicated at the left side.

Since B-RNA encoded proteins are mainly associated in a complex (Franssen *et al.*, 1984b; Eggen and Van Kammen, 1988; Peters *et al.*, 1992), the effect of the mutations in the NTBM on ATP-binding should be studied in absence of the 87K domain, to exclude retention of B-proteins that are complexed with the core-polymerase but do not bind ATP. Indeed, binding was observed for the wild-type 84K from pMB116, although not all of the protein was retained by the ATP-agarose (Fig. 7, lanes 7 to 9). This is not due to limiting amounts of ATP-agarose, since a molar excess of B-proteins from the CPMV-RNA fraction over that of the 84K protein from pMB116 transfected protoplasts was bound (Fig. 7, lanes 6 and 9). Probably a small part of the 84K proteins is misfolded upon translation of MB116-RNA such that they are not able to bind ATP. The ATP-binding capacity of the 84K protein with the Lys500Thr replacement is clearly reduced, as the major part of this protein is present in the eluent fraction (Fig. 7, lanes 10 to 12). This difference in ratio of bound and unbound mutant 84K protein, compared to that observed for wild-type 84K protein, suggests that the Lys500 residue in the A-site of the NTBM is involved in the binding of ATP. However, the finding that a minor part of these mutant proteins was retained suggests that the binding capacity was not completely lost. The Asp545Pro mutation did not markedly decrease the ATP-binding capacity of the 84K protein (Fig. 7, lanes 13 to 15). Probably the Asp545 amino acid residue is not directly involved in the binding of ATP.

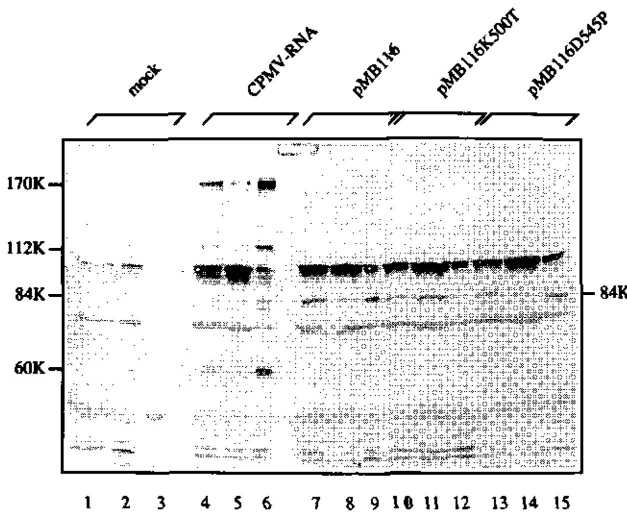


Fig. 7. Isolation of CPMV B-proteins by affinity chromatography on 5'ATP-agarose. Extracts prepared from 1.5 x 10^6 mock-inoculated protoplasts (lanes 1 to 3) or protoplasts transfected with CPMV-RNA (lanes 4 to 6), pMB116 (lanes 7 to 9), pMB116K500T (lanes 10 to 12), or pMB116D545P (lanes 13 to 15) were incubated with 5'ATP-agarose. After incubation unbound proteins in the supernatant (lanes 2, 5, 8, 11, and 14) were separated from the ATP-agarose and immobilized proteins were then eluted from the absorbent (lanes 3, 6, 9, 12 and 15). Lanes 1, 4, 7, 10, 13 contain untreated samples prepared from 0.5 x 10^6 protoplasts.

Proteins were fractionated in an SDS-polyacrylamide gel. To identify the viral proteins an immunoblot analysis, using anti-VPg serum was performed.

Distribution of CPMV B-polyproteins in cowpea protoplasts

To localize CPMV B-proteins in cowpea protoplasts in which 200K proteins have been transiently produced, an immunofluorescent staining with anti-VPg and anti-110K serum and FITC-conjugated goat-anti-rabbit IgG was performed. Using anti-VPg serum, the immunofluorescent label in protoplasts transfected with either CPMV-RNA, pMB200, or pMB200D545P is concentrated in distinct areas in the cytoplasm of the protoplast (Fig. 8, A, B and C). A similar labelling is obtained with anti-110K serum (data not shown), whereas such structures have not been detected in mock-inoculated protoplasts (data not shown). A striking difference in the immunofluorescent staining pattern was observed in protoplasts which were transfected with pMB200K500T. After treatment with anti-VPg and anti-110K serum, the immunofluorescent label appeared to be distributed over the entire protoplast (Fig. 8D and data not shown) and was not concentrated in specific areas of the cytoplasm, indicating that the Lys500Thr substitution has resulted in an altered localization of viral proteins.

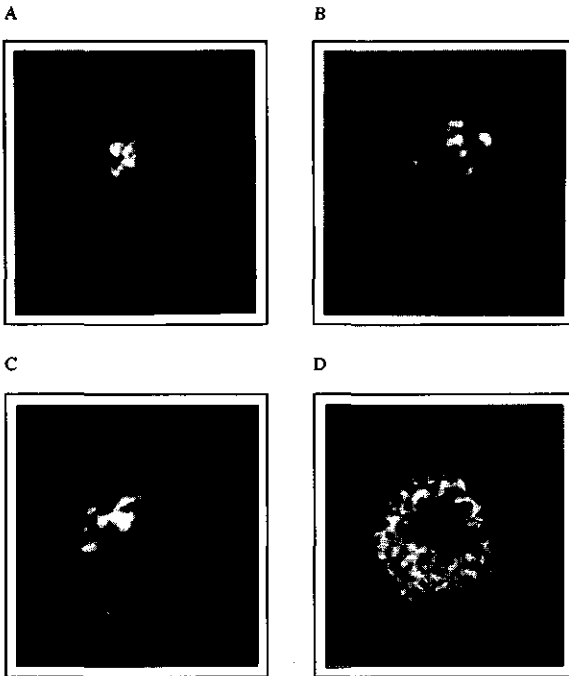


Fig. 8. Distribution of CPMV proteins in cowpea protoplasts. The immunofluorescent images represent typical examples of fluorescent staining of protoplasts transfected with CPMV-RNA (A), pMB200 (B), pMB200D545P (C) or pMB200K500T (D). Detection of viral proteins was performed with anti-VPg serum goat-anti-rabbit IgG conjugated with FITC.

Discussion

In this paper we have presented the first evidence that the B-RNA encoded NTBM plays an essential role in the CPMV RNA replication. Using site-directed mutagenesis, Lys500Thr and Asp545Pro substitutions were introduced in the consensus A and B-site of the NTBM. These mutations severely impair RNA replication as demonstrated by the absence of detectable RNA synthesis. The mutations may even be lethal to the virus, but we cannot entirely exclude the possibility of greatly reduced levels of RNA synthesis that are too low to be detected in a Northernblot analysis. Such a low level of replication might result in the occurrence of wild-type revertants. In case of the Thr500 and Pro545 mutants reversion to wild-type would require two nucleotide changes. Albeit at a very low frequency, restoration of two introduced nucleotide changes has occasionally been observed for poliovirus 2C proteins carrying a Lys135Ser or a Asp177Gly substitution in the A and B-site of the NTBM respectively (Teterina *et al.*, 1992), although this is dependent on the nature of the substitution, since Lys135Gln and Asp177Leu did not lead to reversion (Mirzayan and Wimmer, 1992). Nevertheless, our data do not indicate that reversion has occurred, supporting the notion that Lys500Thr and Asp545Pro substitutions eliminate CPMV RNA replication.
The introduced nucleotide changes might either result in a defective RNA template or a defective protein. To address this, we examined whether M-RNA replication in cowpea protoplasts could be supported by transiently produced proteins. Since the mutant proteins were not able to support the replication of M-RNA in contrast to CPMV B-proteins transient expressed from pMB200, it can be concluded that a protein defect underlies the nonviability of these mutants.
To investigate at which level the virus multiplication was blocked, we have studied the effect of these mutations on translation and polyprotein processing. Primary processing of the 32K protein occurs already on the nascent polypeptide chain by an *in cis* cleavage (Franssen *et al.*, 1984c) and probably depends on the folding of the growing polypeptide chain that brings the primary cleavage site in proximity to the 24K proteinase (Peters *et al.*, 1992). *In vitro* primary processing of the B-polyprotein with Asp545Pro replacement was delayed, but not inhibited. The Asp545Pro substitution probably affects the local flexibility of the polypeptide chain such that the required folding for *in cis* cleavage becomes more difficult. When the mutant B-RNA sequences were transiently expressed in protoplasts, 200K products were not detected. Apparently, incubation over a longer period is sufficient for complete primary processing, even for the 200K protein carrying the Asp545Pro mutation. Moreover, no abnormal differences in cleavage patterns compared to the expression of wild-type B-RNA were observed, suggesting the

precise processing of the mutant polyproteins. Previous results have pointed out the inhibitory function of the 32K protein on processing of the 84K protein *in vitro*, which depends on the interaction with the 58K domain (Peters *et al.*, 1992). Apparently, the 32K protein can also arrest the proteolytic processing of both wild-type and mutant 84K proteins in cowpea protoplasts, indicating that this function is not impeded by amino acid substitutions within the NTBM. Therefore the NTBM has no crucial role in the interdomain communication with the 32K protein, nor in translation and polyprotein processing. A similar conclusion was drawn from a genetic analysis on the polioviral NTBM (Mirzayan and Wimmer, 1992; Teterina *et al.*, 1992).

Experimental data from other characterized ATP/GTP utilizing proteins have shown that the conserved positively charged lysine residue in the A-site of the NTBM of these proteins interacts with the α-phosphate group of the ribonucleotide (Walker *et al.*, 1982; Möller and Amons, 1985; Bradley *et al.*, 1987; De Vos *et al.*, 1988). The negatively charged aspartic acid residues in the B-site form a salt bridge with the Mg^{2+} ion which, in turn, interacts with the negatively charged phosphate groups of the ribonucleotide and has been proposed to have a catalytic function in NTP-hydrolysis (La Cour *et al.*, 1985; Fry *et al.*, 1986; Jurnak, 1988; Weiner and Bradley, 1991). Recent data have indicated that the 60K and 84K proteins can directly bind ATP most likely to the NTBM in the 58K domain of these proteins (S. Peters, unpublished results). Both Lys500Thr and Asp545Pro substitutions might have interfered with an ATP consuming function that somehow is essential for successful replication, since the charge, hydrogen-bonding ability and hydrophobicity are altered by these replacements. To assess the functional importance of these amino acids in ribonucleotide binding, the mutant proteins were tested for their ability to bind ATP by affinity-chromatography. The results do not provide a clearcut answer, as the observed differences in ATP-binding are marginal. Nonetheless, the reduced ATP-binding capacity for the 84K protein with a Thr500 residue in the A-site of the NTBM was reproduced in several independent experiments. This effect could either be the result of an altered conformation of the ATP-binding pocket, or because the Lys500 is directly involved in binding. Apparently, the Asp545Pro substitution near the A-site can be tolerated without disturbing the ATP-binding properties. This supports the specific effect observed for the A-site mutant and points towards a direct involvement of the Lys500 residue in ATP-binding. Similar results have been reported for large T-antigen of SV40 that, along with the 'picorna-like' NTBM-containing proteins, has been classified in the same superfamily of NTBM-proteins (Gorbalenya *et al.*, 1990). Glu473-Asp474 to Ala-Ala amino acid substitutions in the Mg-binding B-site of large T-antigen do not influence the

ATP-binding properties, but render a protein that is defective in ATP-hydrolysis (Weiner and Bradley, 1991). By analogy, the Asp545Pro substitution might have produced a mutant 84K protein that has preserved its binding capacity but might be defective in ATP hydrolysis.

In protoplasts transfected with pMB200 the immunofluorescent label was localized in distinct areas in the cytoplasm which is similar to that observed in CPMV-RNA infected protoplasts. These structures probably represent the cytopathic structures that are seen as electron dense material upon electron microscopic analysis (Wellink *et al.*, 1988). Surprisingly, the viral proteins containing a single Lys500Thr amino acid substitution fail to aggregate into such structures and were found dispersed over the entire protoplast, whereas such an effect was not observed for the Asp545Pro substitution. The altered localization coincided with a lower amount of mainly 170K, 110K and 87K proteins in BK500T-RNA transfected protoplasts as compared to B-RNA or BD545P-RNA transfected protoplasts. Probably these unaggregated viral B-proteins are more susceptible to proteolytic attack, which might account for the lower expression level. The importance of the electron dense structures in the viral life cycle remains unclear, but it has been proposed that these structures represent sites where B-proteins are maintained in an active conformation for replication (Van Bokhoven *et al.*, 1993). The results described in this paper have indicated the NTBM is not involved in the assembly of the 32K/170K complex, but they may point towards the involvement in intermolecular contact that leads to multimerization of 32K/170K complexes. Whether this is promoted by binding of ATP, resulting in a conformation that allows the assembly into higher ordered complexes, remains unclear. Additional mutations in the NTBM will be required to further investigate the role of ATP-binding in protein-protein interactions.

Acknowledgements

We thank Nicole ter Maten for photography. This work was supported by the Netherlands foundation for Chemical Research (SON) with financial aid from the Netherlands Foundation for Scientific Research (NWO).

References

Assink, A.M., Swaans, H., and 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**, 341-355.

Blake, M.S., Johnston, K.H., Russel-Jones, G.J., and Gotschlich, E.C. (1984). A rapid, sensitive method for detection of alkaline phosphate-conjugated antibody on Western blots. *Anal. Biochem.* **136**, 175-179.

Bradley, M.K., Smith, T.F., Lathrop, R.H., Livingston, D.M., and Webster, T.A. (1987). Consensus topography in the ATP binding site of the simian virus 40 and polyoma virus large tumor antigens. *Proc. Natl. Acad. Sci. USA* **84**, 4026-4030.

Dever, T.E., Glynias, M.J., and Merrick, W.C. (1987). GTP-binding domain: Three consensus sequence elements. *Proc. Natl. Acad. Sci. USA* **84**, 1814-1818.

De Vos, A.M., Tong, L., Milburn, M.V., Matias, P.M., Jancarik, J., Noguchi, S., Nishimura, S., Miura, K., Ohtsuka, E., and Kim, S.H. (1988). Three-dimensional structure of an oncogene protein: catalytic domain of human c-H-*ras* p21. *Science* **239**, 888-893.

De Zoeten, G.A., Assink, A.M., Van Kammen, A., and Zabel, P. (1974). Association of cowpea mosaic virus-induced double-stranded RNA with a cytopathological structure in infected cells. *Virology* **59**, 341-355.

Eggen, R., Kaan, A., Goldbach, R., and Van Kammen, A. (1988). Cowpea mosaic virus RNA replication in crude membrane fractions from infected cowpea and *Chenopodium amaranticolor*. *J. Gen. Virol.* **69**, 2711-2720.

Eggen, R., and Van Kammen, A. (1988). RNA replication in comoviruses. *In* "RNA Genetics" (P. Ahlquist, J. Holland, and E. Domingo, Eds.), Vol. 1, pp. 49-69. CRC Press, Boca Raton.

Eggen, R., Verver, J., Wellink, J., de Jong, A., Goldbach, R., and Van Kammen, A. (1989a). Improvements of the infectivity of *in vitro* transcripts from cloned cowpea mosaic virus cDNA: impact of terminal nucleotide sequences. *Virology* **173**, 447-455.

Feinberg, A.P., and Vogelstein, B. (1983). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **132**, 6-13.

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., Leunissen, J., Goldbach, R., Lomonossoff, G., and Zimmern, D. (1984a). Homologous sequences in non-structural proteins from cowpea mosaic virus and picornavirus. *EMBO J.* **3**, 855-861.

Franssen, H., Moerman, M., Rezelman, G., and Goldbach, R. (1984b). Evidence that the 32,000-Dalton protein encoded by bottom component RNA of cowpea mosaic virus is a proteolytic processing enzyme. *J. Virol.* **50**, 183-190.

Franssen, H., Goldbach, R., and Van Kammen, A. (1984c). Translation of bottom component RNA of cowpea mosaic virus in reticulocyte lysate: faithful proteolytic processing of the primary translation product. *Virus Res.* **1**, 39-49.

Fry, D.C., Kuby, S.A., and Mildvan, A.S. (1986). ATP-binding site of adenylate kinase; mechanistic implications of its homology with ras-encoded p21, F_1-ATPase, and other nucleotide-binding proteins. *Proc. Natl. Acad. Sci. USA* **83**, 907-911.

Garcia, J.A., Schrijvers, L., Tan, A., Vos, P., Wellink, J., and Goldbach, R. (1987). Proteolytic activity of cowpea mosaic virus encoded 24K protein synthesized in *Escherichia coli*. *Virology* **159**, 67-75.

Goldbach, R., Rezelman, G., and Van Kammen, A. (1980). Independent replication and expression of B-component RNA of cowpea mosaic virus. *Nature (London)* **286**, 297-300.

Gorbalenya, A.E., and Koonin, E.V. (1989). Viral proteins containing the purine NTP-binding sequence pattern. *Nucleic Acids Res.* **17,** 8413-8440.

Gorbalenya, E.G., Blinov, V.M., Donchenko, A.P., and Koonin, E.V. (1989a). An NTP-binding motif is the most conserved sequence in a highly diverged monophyletic group of proteins involved in positive strand RNA viral replication. *J. Mol. Evol.* **28,** 256-268.

Gorbalenya, A.E., Koonin, E.V., Donchenko, A.P., and Blinov, V.M. (1989b). Two related superfamilies of putative helicases involved in replication, recombination, repair and expression of DNA and RNA genomes. *Nucleic Acids Res.* **17,** 4713-4730.

Gorbalenya, A.E., Koonin, E.V., and Wolf, Y.I. (1990). A new superfamily of putative NTP-binding domains encoded by genomes of small DNA and RNA viruses. FEBS *Lett.* **262,** 145-148.

Jurnak, F. (1988). The three-dimensional structure of c-H-ras p21: implications for oncogene and G protein studies. *Trends Biochem. Sci.* **13**, 195-198.

King, A.M.Q., Lomonossoff, G.P., and Ryan, M.D. (1991). Picornaviruses and their relatives in the plant kingdom. *Sem. Virol.* **2**, 11-17.

Korneluk, R.G., Quan, G., and Gravel, R.A. (1985). Rapid and reliable dideoxy sequencing of double-stranded DNA. *Gene* **40,** 317-323.

Kunkel, T.A. (1985). Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl Acad. Sci. USA* **82,** 488-492.

La Cour, T.F.M., Nyborg, J., Thirup, S., and Clark, B.F.C. (1985). Structural details of the binding of guanosine diphosphate to elongation factor Tu from *Escherichia coli* as studied by X-ray crystallography. *EMBO J.* **4**, 2385-2388.

Laemmli, U.K. (1970). Cleavage of structural proteins during assembly of the head of the bacteriophage T_4. *Nature (London)* **227,** 680-685.

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

McMaster, G.K., and Carmicheal, G.G. (1977). Analysis of single- and double-stranded nucleic acid on polyacrylamide and agarose gels by using glyoxal and acridine orange. *Proc. Natl. Acad. Sci. USA* **74,** 4835-4838.

Mirzayan, C., and Wimmer, E. (1992). Genetic analysis of an NTP-binding motif in poliovirus polypeptide 2C. *Virology* **189,** 547-555.

Möller, W., and Amons, R. (1985). Phosphate-binding sequences in nucleotide-binding proteins. FEBS *Lett.* **186**, 1-7.

Neill, J.D. (1990). Nucleotide sequence of a region of the feline calicivirus genome which encodes picorna-like RNA-dependent RNA-polymerase, cysteine protease and 2C polypeptide. *Virus Res.* **17**, 145-160.

Peters, S.A., Voorhorst, W.G.B., Wery, J., Wellink, J., and Van Kammen, A. (1992). A regulatory role for the 32K protein in proteolytic processing of cowpea mosaic virus polyproteins. *Virology* **191,** 81-89.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). "Molecular Cloning: A laboratory manual", 2nd ed. Cold Spring Harbour Laboratory, Cold Spring Harbor, New York.

Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467.

Teterina, N.L., Kean, K.M., Gorbalenya, A.E., Agol, V.I., and Girard, M. (1992). Analysis of the functional significance of amino acid residues in the putative NTP-binding pattern of the poliovirus 2C protein. *J. Gen. Virol.* **73**, 1977-1986.

Van Bokhoven, H., Van Lent, J.W.M., Custers, R., Vlak, J.M., Wellink, J., and Van Kammen, A. (1992). Synthesis of the complete 200K polyprotein encoded by cowpea mosaic virus B-RNA in insect cells. *J. Gen. Virol.* **73**, 2775-2784.

Van Bokhoven, H., Verver, J., Wellink, J., and Van Kammen, A. (1993). Protoplasts transiently expressing the 200K coding sequence of cowpea mosaic virus B-RNA support replication of M-RNA. *J. Gen. Virol.* **74**, 2233-2241

Van Lent, J., Storms, M., Van der Meer, F., Wellink, J., and Goldbach, R. (1991). Tubular structures involved in movement of cowpea mosaic virus are also formed in infected cowpea protoplasts. *J. Gen. Virol.* **72**, 2615-2623.

Vos, P., Laegle, M., Wellink, J., Van Kammen, A., and Goldbach, R. (1988). Two viral proteins involved in the proteolytic processing of the cowpea mosaic virus polyproteins. *Nucleic Acids Res.* **16**, 1967-1985.

Walker, J.E., Saraste, M., Runswick, M.J., and Gay, N.J. (1982). Distantly related sequences in the α- and β-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* **1**, 945-951.

Weiner, B.M., and Bradley, M.K. (1991). Specific mutation of a regulatory site within the ATP-binding region of simian virus 40 large T antigen. *J. Virol.* **65**, 4973-4984.

Wellink, J., Rezelman, G., Goldbach, R., and Beyreuther, K. (1986). Determination of the proteolytic processing sites in the polyprotein encoded by the bottom-component RNA of cowpea mosaic virus. *J. Virol.* **59**, 50-58.

Wellink, J., Van Lent, J., and Goldbach, R. (1988). Detection of viral proteins in cytophatic structures in cowpea protoplasts infected with cowpea mosaic virus. *J. Gen. Virol.* **66**, 751-755.

Wellink, J., and Van Kammen, A. (1989). Cell-to-cell transport of cowpea mosaic virus requires both 58K/48K proteins and the capsid proteins. *J. Gen. Virol.* **70**, 2279-2286.

Chapter 7

Discussion

7.1 The CPMV B-RNA encoded 60K protein is a multi-functional protein

The study described in this thesis was undertaken to elucidate the functioning of CPMV B-RNA encoded 32K and 60K proteins in CPMV-RNA replication. At first it was attempted to synthesize wild-type and mutant viral proteins in *E.coli* employing an inducible expression system which would permit large scale production of the protein and, subsequently, characterization of their biochemical activities. Previous studies had shown high level synthesis of the B-RNA encoded 110K and 87K proteins in *E.coli* (Richards *et al.*, 1989) and therefore such a prokaryotic expression system seemed appropriate to produce the 60K protein as well. However, production of the 60K protein in *E.coli* met with great difficulty. The system suffered from internal initiation, frequent DNA rearrangements, and the 60K protein appeared to be unstable in *E.coli*. The attempts were further obstructed because the protein was found to be highly toxic even when produced in small amounts, inhibiting cell growth and, eventually, causing cell death (data not shown). A frame shift mutation in the N-terminal part of the 60K coding sequence only produced truncated products, resulting from internal initiation downstream from the frame shift mutation, of which the largest protein had a molecular weight of approximately 50K. These products were not detrimental to *E. coli*. Furthermore, the toxic effects caused by 60K amino acid sequences could be overcome by producing them as 92K (32K+60K) or 116K (32K+60K+24K) precursor proteins. Probably the 32K protein imposes a conformation such that toxic sequences within the N-terminal part of the 60K protein become unavailable for interaction with the cellular structures. The majority of these proteins accumulated in pellet fractions. Although the 116K protein was further processed this mainly resulted into 92K and 24K proteins, whereas cleavage into the 32K and 60K proteins hardly occurred. Usually the primary cleavage is very rapid and efficient, but in this case is presumably impeded because of an aberrant folding of the proteins in *E. coli*. Considering these results, the *E. coli* expression system was found to be inappropriate to produce the 60K protein.

As CPMV and poliovirus have a comparable genomic organization, the polioviral 2BC3AB processing intermediate (Takegami *et al.*, 1983) can be considered to be the CPMV 60K homologue, although sequence homology is confined mainly to the N-terminal domain of the 60K protein and the 2C protein (Argos *et al.*, 1984). Remarkably, polioviral 2BC, 3AB, 2B and 3A processing products, if synthesized in *E.coli*, inhibited cell growth and caused cell lysis, accompanied by drastic permeability of the inner bacterial membrane (Lama *et al.*, 1990; Lama and

Carrasco, 1992). These polioviral proteins have been implicated in the induction of newly synthesized membrane vesicles to which the viral replication complex becomes attached (Bienz *et al.*, 1987, 1990). Poliovirus infection has been shown to stimulate phospholipid biosynthesis, which apparently is essential for viral replication (Vance *et al.*, 1980; Guinea and Carrasco, 1990). The full significance of these observations remains yet to be established. In this regard the role of phospholipids in plus-strand RNA synthesis by Flockhouse replicase (FHV, nodaviruses) is worthwhile mentioning. An RNA-dependent RNA polymerase extracted from FHV-infected cells has been shown to produce exclusively minus-strand progeny RNA from a plus-strand template resulting in dsRNA products. However, in presence of phospholipids the system carries out complete replication producing mainly plus-strand RNA predominantly in a single-stranded form. Moreover, if membranes to which FHV RNA-replicase is associated are disrupted by detergent, the ability to synthesize plus-strand RNA in presence of phospholipids is lost, while minus-strand RNA synthesis is not affected (Wu *et al.*, 1992). Although the mechanism by which phospholipids participate in the activation of plus-strand RNA synthesis remains unclear, these results suggest that phospholipds directly interact with one or more components of the replicase and activate an enzymatic function vital for plus-strand synthesis. The activation of the dnaA protein involved in *E.coli* chromosome replication by phospholipid binding may be illustrative for such a stimulation. The dnaA protein forms a part of the *E. coli* DNA replication complex that appears to be tightly coupled to special phospholipid membranous structures (mesosomes) via specific interaction with the diacidic head group of phospholipid cardiolipin. Binding of cardiolipin to the inactive ADP-bound form of dnaA protein subsequently promotes replacement of ADP by ATP and activates the dnaA protein to function as an initiator replication protein (Sekimizu and Kornberg, 1988; Sekimizu *et al.*, 1988).

Based on amino acid sequence homology and the high affinity for membranes, the similar effects caused by CPMV 60K and polioviral homologous proteins in *E.coli* might result from interaction with phospholipids. Although the mechanism of the interaction and modification of membranes remains obscure, it has been reported that hydrophobic domains contained within the 2B and 3A proteins, which presumably form basic amphipathic α-helices, might be involved (Lama and Carrasco, 1992). Indeed such structural elements have been implicated in their interaction with phospholipid membranes (Agawa *et al.*, 1991). These amphipathic segments have the potential to seek surfaces between membrane and aqeous phase (Eisenberg *et al.*, 1984). Probably their hydrophobic surfaces are buried within the hydrophobic interior of a phospholipid bilayer, where they interact with fatty acyl chains, whereas the positively charged residues are probably positioned

at the polar non-polar interface to interact with the ionized phosphates (Segrest *et al.*, 1992; Jones *et al.*, 1992), thereby disrupting lipid bilayers and promoting cell lysis (Eisenberg *et al.*, 1984). Secondary structure analysis according to Chou and Fasman (1978), Garnier *et al.* (1978) and Sander and Schneider (1991) predict that amino acid segments in the N-terminal part of the 60K protein of CPMV and two other comoviruses red clover mottle virus (RCM and cowpea severe mosaic virus (CPSMV) are arranged in an α-helical conformation. These peptide segments share a relatively low mean hydrophobicity, a high mean hydrophobic moment according to the hydrophobicity plot of Eisenberg *et al.* (1984) and positively charged amino acids at the periphery of the polar face (see Fig. 1). By analogy it may be this amphipatic element in the CPMV 60K protein that is toxic to *E.coli*. Somehow this may be related to the induction of membranous vesicles that appear in CPMV infected plant cells (Assink *et al.*, 1973). Indeed, the observation that membranous vesicles were induced in insect cells in which the 60K protein was synthesized provides further evidence for the involvement of the 60K protein in membrane interaction and modification (Van Bokhoven *et al.*, 1992). Furthermore, the 60K protein has been located in the proximity of the vesicles, supporting the notion that the 60K serves to anchor the viral replication complex to the membranes. Following these considerations, it is tempting to speculate that interaction with membranes, in particular with the phospholipid-tails, might prove to be essential for the activation of the 60K protein in replication of CPMV RNA.

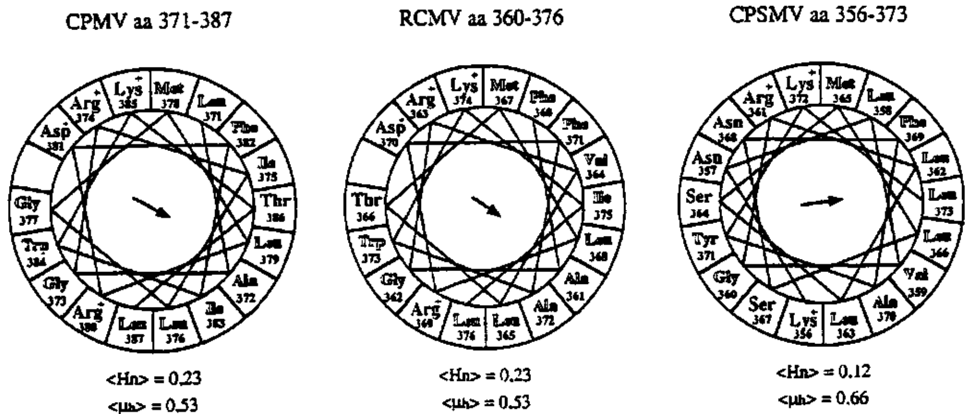


Fig. 1. Analysis of the potential α-helices in the N-terminal sequences of the 60K proteins of CPMV, RCMV and CPSMV. According to the helical wheel program of Schiffer and Edmundson (1967) amino acids are arranged as an ideal α-helix (100° rotation per residue) seen down the long axis from the amino terminal end. Each residue is shown with charge and sequence

number corresponding to the position from the amino terminus of the B-polyprotein. The normalized consensus scale values according to Eisenberg *et al.* (1984) for each type of amino acid are used to calculate the mean hydrophobicity <Hn> and the mean hydrophobic moment <μh>. Amphipatic α-helices were considered to be potential membrane-surface seeking when meeting Eisenberg's criterion $<\mu h> + 0.4<Hn> \geq 0.6$. The arrow represents the mean hydrophobic moment vector. Hydrophobic residues are outlined and hydrophylic residues are shown in bold.

The occurrence of a conserved ribonucleotide binding motif (NTBM) in the 58K domain has been noticed, but actual evidence for the involvement in binding of ribonucleotides by NTBM-containing B-proteins was lacking. In chapter 6 the functional importance of the putative ribonucleotide binding site has been examined. Covalent binding of chemically modified ATP to 60K and 84K proteins was observed which appeared to be specific, since binding of chemically modified GTP was not observed. Independent from these experiments, specific ATP-binding was observed using affinity chromatography (chapter 7). Moreover, a Lys→Thr amino acid substitution in the A-site of the NTBM, supposed to be involved in binding of the phosphate-moiety of the ribonucleotide, resulted in a decreased ATP-binding capacity of the 84K protein, whereas an Asp→Pro amino acid replacement in the Mg^{2+} binding B-site of the NTBM did not affect the ATP-binding properties. In addition both mutations appeared to be lethal to the virus. It therefore appears that the NTBM in the 58K domain is involved in an ATP-consuming function essential for viral RNA replication.
The CPMV B-RNA encoded 60K protein belongs to a group of homologous NTBM-containing proteins including the polioviral 2C protein and the CI protein from plum pox virus (PPV). The latter protein has been shown to function as an ATP-dependent RNA helicase (Lain *et al.*, 1990). Whether CPMV would require a helicase activity during replication to remove secondary structures in the template RNA and/or to unwind template and the nascent daughter strand RNAs is not known. However, the NTBM-containing 60K and 84K proteins are potential candidates to accomplish such a task. This has been the incentive to test fractions from CPMV-infected cowpea leaves enriched in the NTBM-containing 60K, 84K and 170K proteins for unwinding activity. In the unwinding assay, essentially performed according to Lain *et al.* (1990), a double stranded RNA substrate with 5' and 3' overhanging ends was used, as helicases are known to require single stranded RNA regions for binding prior to unwinding. Yet, no indications for a possible helicase activity were found (data not shown). Similar results have been obtained for the polioviral 2C protein. The NTBM in the 2C protein plays an essential role in the polioviral RNA replication (Mirzayan and Wimmer, 1992a;

Teterina *et al.*, 1992) and both RNA binding capacity and ATP/GTPase activity were shown to be intrinsic properties of the 2C protein (Mirzayan and Wimmer, 1992b; Rodriguez and Carrasco, 1993). However, biochemical evidence for an ATP/GTP dependent RNA unwinding activity has not been obtained for the 2C protein (Rodriguez and Carrasco, 1993). Moreover, purified polioviral RdRp has been demonstrated to unwind an RNA duplex during chain elongation (Cho *et al.*, 1993). Furthermore, the CI protein of PPV is more closely related to the flavi- and pestiviral supergroup of NTBM-containing proteins, which form a different supergroup. Thus it is possible that 2C-like proteins do not have helicase activity.
In addition to the functional properties described above the 60K domain of the 170K protein appears to participate in the processing of both M and B-polyproteins. Its role in polyprotein processing will be discussed in the next section.

7.2 32K protein assisted folding of the 170K protein

The processing cascade of both M and B-polyproteins is regulated by the 32K protein. *In vitro* translation of B-RNA and transient expression of B-cDNA sequences in cowpea protoplasts indicate that *cis* cleavages in the 170K proteins by the 24K proteinase are arrested by the 32K protein, whereas in absence of the 32K protein rapid and faithful processing of the 170K protein occurs (chapter 3 and 7). Previous results have pointed towards the 58K domain as a target for interaction with the 32K protein (Franssen *et al.*, 1984) and have shown that the 32K protein is required as a cofactor in M-polyprotein processing (Vos *et al.*, 1988). Consistent with these observations are the results described in chapter 2, which indicate that a 32K/170K complex cleaves the M-polyprotein *in trans*. The part of the 58K region to which the 32K protein binds, is probably only exposed in the nascent polypeptide chain. This is suggested by the observation that the M-polyprotein processing defect of the 32K deletion mutant could not be complemented by the 32K protein when supplied *in trans*. However, when the 32K and 170K proteins were jointly synthesized *trans* processing activity was rescued, indicating that the interaction occurs already before synthesis of the 170K protein has been completed.
Comparison of the amino acid sequences encoded by CPMV, CPSMV and RCMV B-RNAs reveals an extensive homology and suggests that polyproteins will probably be processed by a corresponding mechanism (Shanks and Lomonossoff, 1992; Chen and Bruening, 1992). Previous results have shown the inability of CPMV proteins to cleave the M-polyproteins of RCMV and CPSMV, indicating a substrate specificity (Goldbach and Krijt, 1982; Chen and Bruening, 1992 and references

therein). The 24K proteinase sequences of these comoviruses are very similar, whereas the 32K proteins of CPMV, CPSMV and RCMV is the least conserved protein among the B-encoded proteins. Although nothing is known about the interactions between the 32K/170K complex and the M-polyprotein substrate, it may be the 32K protein that determines substrate specificity, perhaps via interaction with the M-polyprotein.

Experiments with insertion and deletion mutants indicate that a peptide segment in the C-terminal part of the 58K portion is involved in the interaction with the 32K protein (chapter 3). Amino acid sequence analysis reveals a large hydrophobic domain of approximately 40 amino acids just upstream of the VPg region, found conserved among other members of the comovirus group (e.g. CPSMV, RCMV), which is a likely candidate for interaction with the 32K protein. It can be expected that such a large hydrophobic segment will not be exposed to the cytosol but rather would be located in the interior of the 170K protein. This would render the peptide segment inaccessible for the 32K protein, which might explain the observed failure to reconstitute an active 32K/170K complex with separately produced proteins (chapter 2).

It is unclear by what mechanism the 32K protein arrests the *cis* cleavage activity in the 170K protein. Obviously, the information which specifies the conformation for correct processing of the 170K protein resides solely within the amino acid sequence of the 170K polypeptide (chapter 3). However, following binding, the 32K protein might occupy peptide segments that otherwise would participate in the folding of the emerging 170K polypeptide chain. Thus it seems reasonable to suppose that transient C-terminal hydrophobic sequences in the 58K part, which are read out by the 32K protein, endow the nascent B-polyprotein with the necessary folding signals. Thereby the 170K protein presumably obtains a functional intermediate conformation, such that cleavage sites within the 170K protein become inaccessible to the 24K proteinase.

Quite separate considerations lead to the idea that release and folding of the 170K protein in the host cell needs to be assisted. Release and solubilization of the 60K protein from host membranes of CPMV infected plant cells have only succeeded in the presence of detergent. Upon removal of detergent the 60K protein becomes insoluble and aggregates (data not shown), presumably because of reexposure of hydrophobic surfaces. It can be imagined that such interactive surfaces are also exposed during protein synthesis. Thus, the occurrence of non-specific intra- and intermolecular interactions would be an inherent risk for the growing polypeptide chain. Such a risk might be reduced by binding of the 32K proteins to the interactive regions within the 58K domain. Obviously, a protein binding to a polypeptide chain during translation will have the advantage of binding the

proper peptide segment without having to seek it among a crowd of other competing molecules.
The 32K protein enables the 24K protein to function as a 'double duty' proteinase. In the 32K/170K complex *trans* cleavage activity is promoted to produce the structural proteins translated from M-RNA. As pointed out above the dissociation of the 32K protein from the complex is an irreversible event. This implies that the binding of the 32K protein should persist long enough to ensure the production of sufficient capsid precursor proteins. Indeed, in CPMV infected plant cells the 170K protein appears to be the most abundant protein (Rezelman *et al.*, 1980) and is probably slowly processed. This delayed processing might reflect the fine-tuning of the 24K proteinase activity in favour of particle formation. On the other hand, the 32K protein has to dissociate from the complex to switch on *cis* cleavage activity of the 24K proteinase domain in the 170K protein to produce the individual the B-proteins involved in replication. Furthermore, dissociation of the 32K protein must occur at such a rate as required to keep pace for the progress of viral RNA replication. Therefore the question arises what triggers the disassembly of the complex? When M and B-RNA are simultaneously translated *in vitro* the M-polyproteins are cleaved into the 58K/48K proteins and the capsid precursor by the 32K/170K proteinase, but further cleavage of the 170K protein was not observed (data not shown). Apparently, M-polyprotein processing itself does not trigger dissociation of the complex. Perhaps the answer should be sought in the localization of the replicative proteins. Somehow the B-proteins end up attached to viral-induced membranous vesicles, known to be the site at which viral replication occurs. The 32K/170K protein complex may be in a conformation competent to become attached to membranes. Once associated with the membranes such a hydrophobic environment may promote the dissociation of the 32K protein thereby triggering processing of the 170K protein and enabling the formation of a functional replication complex that is anchored via the 58K part to the membranes. Such a mechanism is speculative and needs further prove, but it proposes an essential role for the 32K protein in the formation of a functional membrane bound replication complex, although the 32K protein itself will probably not be a component of the final structure. The role of the 32K protein can be interpreted as to prevent premature cleavages and nonspecific interactions until the B-polyprotein properly associates with the membranes. Increasingly, examples are found of proteins destined to become membrane bound, which assemble into functional structures only if their folding is assisted by the antifolding capacity of other proteins known as molecular chaperones. Chaperones form a family of several unrelated classes of proteins that mediate the correct assembly of other polypeptides but are not themselves components of the

final structures (Ellis and Hemmingsen, 1989). Thus, the 32K protein can be considered a molecular chaperone in building a functional membrane bound replication complex.

7.3 Alternate processing of the 170K protein

The biological functions required for replication of the genome of CPMV are provided by B-RNA. B-RNA encodes a 200K polyprotein that undergoes several processing steps to release the replicative proteins. The primary cleavage is very rapid and yields the 32K and 170K proteins. VPg is one of the components of the 170K protein. Although the precise role of the VPg is not known for certain, its occurrence at the 5'-end of each progeny RNA points towards a function as a primer in the viral RNA synthesis. The processing studies described in chapter 4 demonstrate that the 24K proteinase can cleave the 170K protein at three sites in the first instance. This results in three different sets of proteins, 60K+110K, 84K+87K and 58K+112K, of which the 60K, 84K and 112K (VPg+110K) have the potential to serve as a VPg precursor. The low level of *in trans* 24K proteinase activity of B-proteins indicates the preference of processing B-polyproteins via a *cis*-like mechanism (chapter 4) and argues against the 60K and in favour of the 84K and 112K proteins to function as a VPg precursor. Processing of the 84K protein occurs via two pathways, either into 60K+24K or into 58K+26K(VPg+24K) proteins. Whereas the 60K appears to be stable, the 26K protein is probably rapidly further processed into VPg and 24K proteins. Cleavage of the 112K protein probably also occurs via two pathways, either into VPg+110K or 26K+87K proteins (chapter 5). The significance of the cleavage pathway via the 26K intermediate protein remains yet to be established, as the 26K product has not so far been found in CPMV infected cells. The 112K has the conserved polymerase domain in common with the 87K and 110K protein and each of these proteins appears to be functional in binding of ribonucleotides (chapter 6 and 7). Apparently, different sets of functional proteins can be generated from the 170K protein. Thus, it can be conceived that the alternate cleavage of the 170K protein provides CPMV with a mechanism to build different replication complexes to regulate viral RNA synthesis, perhaps at the level of template selection. Plus and minus-strand RNAs are probably synthesized throughout the replication cycle, but plus-strand RNA is produced in an excess over minus-strand RNA. The molecular basis of the asymetric RNA synthesis remains obscure but must somehow be regulated. For a number of positive-strand RNA viruses including CPMV and poliovirus different lines of evidence show that viral induced membrane vesicles are the site of plus-strand RNA synthesis, whereas minus-strand synthesis can be achieved in

absence of membranes (Dorssers *et al.*, 1984; Semler *et al.*, 1988). If this holds for CPMV, different processing pathways may provide CPMV the opportunity to produce soluble replication complexes for minus-strand synthesis and membrane bound replication complexes for positive-strand synthesis. In either case the initiation of RNA synthesis will start with the release of VPg for primary viral RNA synthesis. It now appears attractive to assume that the 112K protein functions as VPg precursor in minus-strand RNA synthesis. In the 84K and 60K protein VPg is part of precursors that have in the 58K domain a hydrophobic region for anchoring the replication complexes in membranes (Goladbach *et al.*, 1982). In the 112K VPg-precursor this domain is absent. Furthermore, time course studies in the appearance of the different processing products upon *in vivo* translation of B-RNA (chapter 3) have shown strong preference for the 170K protein to be cleaved into 110K+60K and 84K+87K proteins over 58K+112K proteins. Such preference is also reflected in the low amount of the 112K protein observed *in vivo* compared to 60K and 84K proteins. The different proportion in which the 112K and 84K and 60K protein produced are in agreement with the difference in negative - and positive RNA strand synthesis during viral RNA replication. Of course, this circumstantial evidence for regulation of negative and positive RNA strand synthesis needs to be consolidated by further experiments.

References

Agawa, Y., Lee, S., Ono, S., Aoyagi, H., Ohno, M., Taniguchi, T., Anzai, K., and Kirino, Y. (1991). Interaction with phospholipid bilayers, ion channel formation, and antimicrobial activity of basic amphipatic α-helical model peptides of various chain lengths. *J. Biol. Chem.* **266**, 20218-20222.

Argos, P., Kamer, G., Nicklin, M.J.H., and Wimmer, E. (1984). Similarity in gene organization and homology between proteins of animal picornaviruses and a plant comovirus suggest common ancestry of these virus families. *Nucleic Acids Res.* **12**, 7251-7267.

Assink, A.M., Swaans, H., and 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.

Bienz, K., Egger, D. and Pasamontes, L. (1987). Association of polioviral proteins of the P2 genomic region with the viral replication complex and virus-induced membrane synthesis as visualized by electron microscopic immunocytochemistry and autoradiography. *Virology* **160**, 220-226.

Bienz, K., Egger, D., Troxler, M., and Pasamontes, L. (1990). Structural organization of poliovirus RNA replication is mediated by viral proteins of the P2 genomic region. *J. Virol.* **64**, 1156-1163.

Chen, X., and Bruening, G. (1992). Cloned dna copies of cowpea severe mosaic virus genomic RNAs: infectious transcripts and complete nucleotide sequence of RNA1. *Virology* **191**, 607-618.

Cho, M.W., Richards, O.C., Dmitrieva, T.M., Agol, V., and Ehrenfeld, E. (1993). RNA duplex unwinding activity of poliovirus RNA-dependent RNA polymerase $3D^{pol}$. *J.Virol.* **67**, 3010-3018.

Chou, P.Y., and Fasman, G.D. (1978). Prediction of the secondary structure of proteins from their amino acid sequence. *Adv. in Enzym.* **47**, 45-178.

Dorssers, L., Van der Krol, S., Van der Meer, J., Van Kammen, A., and Zabel, P. (1984). Purification of cowpea mosaic virus RNA replication complex: Identification of a virus-encoded 110;000-dalton polypeptide responsible for RNA chain elongation. *Proc. Natl. Acad. Sci. USA.* **81**, 1951-1955.

Eisenberg, D., Schwarz, E., Komaromy, M., and Wall, R. (1984). Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J. Mol. Biol.* **179**, 125-142.

Ellis, J.R., and Hemmingsen, M. (1989). Molecular chaperones: proteins essential for the biogenesis of some macromolecular structures. *Trends Biochem. Sci.* **14**, 339-342.

Franssen, H., Moerman, M., Rezelman, G., and Goldbach, R. (1984). Evidence that the 32,000-Dalton protein encoded by bottom component RNA of Cowpea mosaic virus is a proteolytic processing enzyme. *J. Virol.* **50**, 183-190.

Garnier, J., Osguthorpe, D.J., and Robson, B. (1978). Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J. Mol. Biol.* **120**, 97-120.

Goldbach, R., and Krijt, J. (1982). Cowpea mosaic virus-encoded protease does not recognize primary translation products of M RNA from other comoviruses. *J. Virol.* **43**, 1151-1154.

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.

Guinea, R., and Carrasco, L. (1990). Phospholipid biosynthesis and poliovirus genome replication, two coupled phenomena. *EMBO J.* **9**, 2011-2016.

Jones, M.K., Anantharamaiah, G.M., and Segrest, J.P. (1992). Computer programs to identify and classify amphipatic α helical domains. *J. Lipid Res.* **33**, 287-296.

Lain, S., Riechmann, J.L., and Garcia, J.A. (1990). RNA helicase: a novel activity associated with a protein encoded by a positive strand RNA virus. *Nucleic Acids Res.* **18**, 7003-7006.

Lama, J., and Carrasco, L. (1992). Expression of poliovirus nonstructural proteins in *Escherichia coli* cells; modification of membrane permeability induced by 2B and 3A. *J. Mol. Chem.* **267**, 15932-15937.

Lama, J., Martinez-Abarca, F., and Carrasco, L. (1990). Search for poliovirus proteins responsible for membrane modifications. Abstract B15, VIIth meeting of the European study group on picornaviruses, Canterbury, U.K.

Mirzayan, C., and Wimmer, E. (1992a). Genetic analysis of an NTP-binding motif in poliovirus polypeptide 2C. *Virology* **189**, 547-555.

Mirzayan, C., and Wimmer, E. (1992b). Genetic analysis of an NTP-binding motif in poliovirus polypeptide 2C. Abstract P4-28, Third international positive strand RNA symposium. Clearwater, Florida, U.S.A.

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.

Richards, O.C., Eggen, R., Goldbach, R., and Van Kammen, A. (1989). High level synthesis of cowpea mosaic virus RNA polymerase and protease in *Escherichia coli. Gene* **78**, 135-146.

Rodriguez, P., and Carrasco, L. (1993). Poliovirus protein 2C has ATPase and GTPase activities. *J. Biol. Chem.* **268**, 8105-8110.

Sander, C., and Schneider, R. (1991). Database of homology-derived protein structures. *Proteins* 9, 56-68.

Schiffer, M., and Edmundson, A.B. (1967). Use of helical wheels to represent the structures of proteins and to identify segments with helical potential. *Biophys. J.* **7**, 121-135.

Segrest, J.P., Jones, M.K., De Loof, H., Brouillette, C.G., Venkatachalapathi, Y.V., and Anantharamaiah, G.M. (1992) The amphipatic helix in the exchangeable apolipoproteins: a review of secondairy structure and function. *J. Lipid Res.* **33**, 141-166.

Sekimizu, K., and Kornberg, A. (1988). Cardiolipin activation of dnaA protein, the initiation protein of replication in *Escherichia coli. J. Biol. Chem.* **263**, 7131-7135.

Sekimizu, K., Yung, B., and Kornberg, A. (1988). The dnaA protein of *Escherichia coli. J. Biol. Chem.* **263**, 7136-7140.

Semler, B.L., Kuhn, R.J., and Wimmer, E. (1988). Replication of the poliovirus genome. *In* "RNA Genetics" (P. Ahlquist, J. Holland, and E. Domoingo, Eds.), vol 1, pp 23-58. CRC Press, Boca Raton.

Shanks, M., and Lomonossoff, G.P. (1992). The nucleotide sequence of red clover mottle virus bottom component RNA. *J. Gen. Virol.* **73**, 2473-2477.

Takegami, T., Kuhn, R.J., Anderson, C.W., and Wimmer, E. (1983). Membrane-dependent uridylylation of the genome-linked protein VPg of poliovirus. *Proc. Natl. Acad. Sci. USA* **80**, 7447-7451.

Teterina, N.L., Kean, K.M., Gorbalenya, A.E., Agol, V.I., and Girard, M. (1992). Analysis of the functional significance of amino acid residues in the putative NTP-binding pattern of the poliovirus 2C protein. *J. Gen. Virol.* **73**, 1977-1986.

Van Bokhoven, H., Van Lent, J.W.M., Custers, R., Vlak, J.M., Wellink, J., and Van Kammen, A. (1992). Synthesis of the complete 200K polyprotein encoded by the cowpea mosaic virus B-RNA in insect cells. *J. Gen. Virol.* **73**, 2775-2784.

Vance, D.E., Trip, E.M., and Paddy, H.B. (1980). Poliovirus increases phosphatidyl choline biosynthesis in Hela cells by stimulation of the rate-limiting reaction catalyzed by CTP:phosphocholine cytidyltransferase. *J. Biol. Chem.* **255**, 1064-1069.

Vos, P., Jaegle, M., Wellink, J., Van Kammen, A., and Goldbach, R. (1988). Two viral proteins involved in the proteolytic processing of the cowpea mosaic virus polyproteins. *Nucleic Acids Res.* **16**, 1967-1985.

Wu, S-X., Ahlquist, P., and Kaesberg, P. (1992). Active complete in vitro replication of nodavirus RNA requires glycerophospholipid. *Proc. Natl. Acad. Sci. USA* **89**, 11136-11140.

Samenvatting

In dit proefschrift wordt onderzoek beschreven dat als doel heeft meer inzicht te verkrijgen in het mechanisme van de replicatie van cowpea mosaic virus (CPMV). Het onderzoek heeft zich met name gericht op de analyse van de activiteit van enkele virale eiwitten, waarvan de functie in de replicatie van het virus nog zeer onduidelijk was.

De genetische informatie van CPMV is verdeeld over twee RNA moleculen, die M- en B-RNA worden genoemd. In het B-RNA is informatie gecodeerd voor de replicatie van het virale RNA, terwijl M-RNA codeert voor de virus manteleiwitten en eiwitten die betrokken zijn bij het transport van het virus van cel naar cel. In plantecel protoplasten, geïnfecteerd met CPMV B-RNA, kan B-RNA onafhankelijk van M-RNA repliceren. M-RNA daarentegen is voor de replicatie volledig afhankelijk van B-RNA.

De genomische RNA moleculen zijn enkelstrengs en bevatten een klein viraal eiwit, het VPg, aan het 5'-uiteinde. Aan het 3'-uiteinde bevindt zich een poly(A)-staart. De beide RNA moleculen hebben een positieve polariteit en functioneren als boodschapper RNA in de gastheercel, waar ze worden vertaald (translatie) in grote eiwitten, zogenaamde polyproteïnen. De polyproteïnen bevatten verschillende functionele domeinen en door specifieke proteolytische klievingen worden eiwitten uit het polyproteïn vrijgemaakt (processing) die elk een eigen functie in de replicatie en verspreiding van het virus hebben. Het proteïnase, dat verantwoordelijk is voor het vrijmaken van functionele eiwitten, is een 24K eiwit, dat zelf deel uit maakt van het door het B-RNA gecodeerd 200K polyprotein. Voordat de synthese van het 200K polyproteïn is voltooid, wordt al het N-terminaal gelegen 32K eiwit afgesplitst, dat zich onmiddelijk aan de groeiende aminozuurketen van het 170K eiwit bindt tot een 32K/170K eiwitcomplex. Vervolgens kan het 170K eiwit, via 60K, 84K, 110K en 112K tussenproducten, uiteindelijk in 58K, VPg, 24K en 87K eindproducten worden gekliefd. Van de B-eiwitten, die allen betrokken zijn bij de CPMV-RNA replicatie, functioneert het 110K eiwit, dat uit het 24K proteinase en het 87K RNA-afhankelijke RNA-polymerase bestaat, als het actieve replicase. De functie van het 32K eiwit, het 58K eiwit, alsook die van de tussen producten in de replicatie, was aan het begin van dit onderzoek onduidelijk.

Het kleinere M-RNA wordt getransleerd in overlappende 105K of 95K polyeiwitten. De eerste klieving in deze eiwitten resulteert in het vrijkomen van een 60K eiwit en 48K/58K eiwitten. Na klieving van het 60K eiwit ontstaan de beide manteleiwitten VP37 en VP23. Hoewel het 24K proteïnase in staat is alle bovengenoemde klievingen uit te voeren, blijkt voor de *trans*-klieving van het M-polyproteïn tevens het 32K eiwit nodig te zijn. Het 32K eiwit bezit zelf geen proteolytische activiteit, maar functioneert als cofactor van het 24K proteïnase.

De genoomstructuur, genoomorganisatie en genexpressie strategie van CPMV vertoont overeenkomsten met die van een aantal positief-enkelstrengs RNA virussen, waaronder poliovirus, die gezamelijk de groep van de 'picorna-virus like' superfamilie vormen. De analogie weerspiegelt waarschijnlijk eenzelfde replicatie mechanisme. Een aantal belangrijke aspecten van RNA-replicatie, van virussen die tot deze superfamilie behoren, worden in hoofdstuk 1 besproken, omdat deze aspecten een rol spelen bij de interpretatie en discussie van de experimentele resultaten, die zijn beschreven in de daarna volgende hoofdstukken.

In hoofdstuk 2 worden experimenten beschreven die zijn gericht op het ophelderen van de functie van het door het B-RNA gecodeerde 32K eiwit in de processing van polyproteïnen. In een volledige lengte DNA copie van het B-RNA zijn in het, voor het 32K en 58K coderende gebied, inserties en deleties aangebracht. Vervolgens zijn *in vitro* transcripten van deze insertie- en deletie-mutanten *in vitro* getransleerd en is het effect van de mutaties op zowel de M- als B-polyproteïn processing bestudeerd. De resultaten geven aan dat de primaire klieving van het B-polyproteïn in 32K en 170K eiwitten een zeer efficiënt proces is dat via een intramoleculair klievingsmechanisme verloopt. De proteolytische klieving is niet alleen afhankelijk van de Gln/Ser klievingsplaats, maar ook van een juiste vouwing van het 200K polyproteïn. De juiste vouwing wordt mede bepaald door een hydrofobe interactie tussen het 32K eiwit en het aan het 32K grenzende 58K domein van het B-polyproteïn. Na klieving blijft het 32K eiwit door deze interactie aan het 170K eiwit gebonden en is de vouwing van het 170K eiwit in dit complex zodanig dat de verdere proteolytische klievingen in het 170K eiwit niet, of slechts zeer langzaam, verlopen. Eveneens werd gevonden dat, wanneer 32K en 170K eiwitten apart worden gesynthetiseerd, ze alleen dan een complex kunnen vormen, indien synthese van beide eiwitten tegelijkertijd en in elkaars aanwezigheid geschiedt. Dit geeft aan dat de interactie tussen het 32K eiwit en het 58K domein al tijdens de synthese van het B-polyeiwit moet plaatsvinden wil er een functioneel 32K/170K complex gevormd kunnen worden. Het 32K/170K complex dat aldus ontstaat, is in staat de M-polyproteïnen op de Gln/Met klievingsplaats in 60K en 58K/48K eiwitten te klieven. Het 170K eiwit dat *in vitro* in afwezigheid van het 32K eiwit wordt gesynthetiseerd, wordt onmiddelijk verder gekliefd en dit leidt tot het vrijkomen van de verschillende eerder genoemde B-eiwitten. De functie van het 32K eiwit in polyproteïn processing is dus tweeledig. Enerzijds remt het de *cis*-klievingsactiviteit van het 24K proteïnase in het 170K eiwit en anderzijds maakt het de *trans*-klievingsactiviteit van dit eiwit, nodig in de processing van M-polyproteïnen, mogelijk.

Aangenomen wordt dat VPg een functie in de initiatie van CPMV-RNA replicatie heeft, maar men heeft dat nog niet kunnen aantonen. Om het vrijkomen van VPg uit VPg-precursor eiwitten nader te analyseren, is in hoofdstuk 3 de klieving van het 170K eiwit en de 60K, 84K en 112K processing tussenproducten met behulp van *in vitro* expressie van specifieke mutanten van een volledige B-cDNA kloon bestudeerd. Afhankelijk van de positie van de eerste klieving kunnen er drie paar eiwitten uit het 170K eiwit vrijkomen, namelijk 60K+110K, 84K+87K en 58K+112K eiwitten. Uit het 112K (VPg+110K) en het 84K eiwit kan VPg vrijkomen, maar het 60K eiwit blijkt *in vitro* stabiel te zijn. Het is echter niet ondenkbaar dat het 60K eiwit in membraangebonden toestand wel *in trans* wel kan worden gekliefd in 58K en VPg. Door gebruik te maken van cowpea protoplasten die werden getransfecteerd met expressievectoren kon de processing van verschillende B-eiwitten ook *in vivo* worden onderzocht (hoofdstuk 4). Er werd gevonden dat het 112K eiwit op twee posities wordt gekliefd, resulterend of in VPg+110K, of in 26K+87K eiwitten. Dit suggereert dat het 112K eiwit als directe voorloper van VPg in de replicatie kan functioneren. In tegenstelling tot het 110K eiwit, dat stabiel is in protoplasten, wordt het 26K eiwit, dat zowel uit 84K als uit 112K eiwitten ontstaat, in VPg+24K eiwitten verder gesplitst. Het is echter niet duidelijk of de klievingen via het 26K intermediaire product *in vivo* een rol spelen, daar dit eiwit in CPMV-geïnfecteerde cowpea-cellen niet kon worden aangetoond.
Het 58K domein in de B-RNA gecodeerde eiwitten bevat een nucleoside trifosfaat bindingsmotief (NTBM) dat sequentie homologie heeft met NTBM-bevattende eiwitten die ATPase activiteit hebben. Op grond hiervan wordt het NTBM in het 58K domein verondersteld specifiek te zijn voor de binding van ATP. Het NTBM bestaat uit een A-element en een B-element, waarbij de aminozuren in het A-element een rol in de binding van fosfaatgroepen van het ribonucleotide hebben en de aminozuren in het B-element essentiëel zijn voor de binding van een Mg^{2+}-cation. De B-eiwitten zijn op ribonucleotide bindende eigenschappen onderzocht door gedeeltelijk gezuiverde eiwitfracties van CPMV-geïnfecteerde cowpea bladeren, die zijn verrijkt in B-eiwitten, te incuberen met perjodaat-geoxideerde ribonucleotiden (hoofdstuk 5). Vervolgens werden de afzonderlijke B-eiwitten door middel van immunoprecipitaties geïsoleerd en geanalyseerd. Er werd gevonden dat 60K eiwit en 84K eiwitten, die beide het 58K domein bevatten, specifiek ATP binden, maar bindingsaffiniteit voor GTP is niet waargenomen. Daarnaast bleken het 87K, 110K, 112K en het 170K eiwit, die allen het 87K polymerase bevatten, behalve ATP ook GTP, UTP en CTP te binden en zal overeenkomen met het voorkomen van een functioneel polymerase domein in deze eiwitten.

Door plaats specifieke mutaties in het NTBM van het 58K domein aan te brengen en vervolgens het effect van de mutaties op enerzijds de vermenigvuldiging van het mutante B-RNA in cowpea protoplasten en anderzijds op de vertaling van het gemuteerde B-RNA en de processing van de producten *in vitro* te onderzoeken, werd de functie van het 58K domein nader geanalyseerd (hoofdstuk 6). Een Lys→Thr aminozuursubstitutie in het A-element alsook een Asp→Pro aminozuursubstitutie in het B-element van het NTBM zijn lethaal voor de replicatie van CPMV. Het defect in de replicatie is te wijten aan een defect eiwit en niet het gevolg van een defecte RNA-matrijs en daarom kan aan het NTBM een essentiële functie in de CPMV-RNA replicatie worden toegeschreven. Vooralsnog blijft echter de exacte functie van het NTBM in de replicatie onduidelijk. De Lys→Thr en Asp→Pro substituties in het NTBM hebben geen effect op de translatie van B-RNA, of de klieving van het B-polyproteïn, noch op de interactie van het 58K domein met het 32K eiwit. Wel bindt het 84K eiwit met de Lys→Thr substitutie in het NTBM minder goed ATP, terwijl de ATP-binding voor het 84K eiwit met de Asp→Pro substitutie onveranderd blijft. Dit duidt erop dat het NTBM in het 58K domein daadwerkelijk bij de ATP-binding is betrokken. Het lijkt daarom waarschijnlijk dat het 58K eiwit een ATP-verbruikende functie heeft die essentiëel is voor de virus RNA replicatie. Er is verondersteld dat het 58K domein een helicase activiteit zou bevatten, maar deze activiteit werd voor B-eiwitten die het 58K domein bevatten niet gevonden. Wel heeft de Lys→Thr substitutie een opvallend effect op de localizatie van de B-gecodeerde eiwitten in geïnfecteerde cellen. In cellen geïnfecteerd met wild-type B-RNA zijn B-eiwitten samengepakt tot wat electronendichte structuren lijken te zijn. De functie van deze structuren in de virus replicatie is niet duidelijk. De Lys→Thr substitutie leidde ertoe dat B-eiwitten werden verspreid in het cytoplasma van de protoplasten, terwijl de Asp→Pro substitutie niet resulteerde in een dergelijke verspreiding van B-eiwitten, maar identiek aan wild-type B-eiwitten zijn samengepakt. Dit resultaat suggereert een mogelijke relatie tussen ATP-binding en complexvorming van B-eiwitten. Wat daarvan de betekenis zou kunnen zijn is niet duidelijk en zal door verder onderzoek moeten worden vastgesteld.

In hoofdstuk 7 wordt besproken welke betekenis de resultaten, die zijn verkregen met het onderzoek naar de eigenschappen van de 32K en 60K eiwitten, hebben met betrekking tot de functie van deze eiwitten in de replicatie van CPMV-RNA. De replicatie van CPMV-RNA vindt plaats in replicatie complexen die zijn gebonden aan membraanblaasjes in het cytoplasma. De binding van het replicatie complex vindt plaats via het 58K domein van de door het B-RNA gecodeerde 170K precursor. Een mogelijke functie van het 32K eiwit zou kunnen zijn er voor te zorgen dat het 170K eiwit de juiste conformatie heeft die binding aan membranen mogelijk

maakt, terwijl tevens voortijdige klievingen en niet specifieke interacties voorkomen worden, tot het 170K eiwit op de juiste manier geassocieerd is met de membranen. Het is niet ondenkbaar dat de interactie van het 32K/170K complex met membranen, de hydrofobe interactie tussen het 32K eiwit en het 58K domein van het 170K eiwit doet verbreken en dat, volgend op de dissociatie van het 32K eiwit, klieving van het 170K eiwit wordt geactiveerd. Tevens lijkt het niet uitgesloten dat door de interactie met membranen ook een ATP-verbruikende functie van het 58K domein wordt geactiveerd, die mogelijk een rol speelt bij het tot stand komen van een functioneel replicatie complex, waarna de replicatie van het CPMV RNA kan beginnen.

Curriculum Vitae

Sander Antonius Peters is geboren op 14 mei 1960 te 's Gravenhage. Op 30 mei 1980 behaalde hij het Atheneum-B diploma aan de Thorbecke Scholengemeenschap te Arnhem. Na zijn militaire dienstplicht te hebben vervuld, begon hij op 1 oktober 1981 de studie planteziektenkunde aan de Landbouwuniversiteit te Wageningen. In het kader van de doctoraalstudie werd onderzoek bij de vakgroepen Virologie (prof. dr. R.W. Goldbach), Biochemie (prof. dr. C. Veeger) en Moleculaire Biologie (prof. dr. A. van Kammen) gedaan. Bij T.N.O./M.B.L. heeft hij gedurende 6 maanden als stageaire bij de afdeling Biochemie, sectie recombinant DNA, onderzoek verricht (prof. dr. P.H. Pouwels). Vervolgens werd op 19 september de universitaire studie met het behalen van het ingenieursdiploma afgesloten. Van 1 november 1988 tot en met 31 oktober 1992 was hij als onderzoeker in opleiding in dienst bij de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) en werkzaam bij de vakgroep Moleculaire Biologie van de Landbouwuniversiteit Wageningen (prof. dr. A. van Kammen). Het onderzoek dat hij verrichtte aan het cowpea mozaïek virus is in dit proefschrift beschreven. Sinds 1 januari 1993 is hij als wetenschappelijk medewerker in dienst bij de Rijksuniversiteit Leiden en werkzaam bij de afdeling Virologie (prof. dr. W.J.M. Spaan) van het Academisch Ziekenhuis te Leiden.

Nawoord

Na de koele feiten een woord van dank voor de mensen die zich mede hebben ingezet bij het tot stand komen van dit proefschrift. Allereerst wil ik jou, Ab van Kammen, bedanken voor je ondersteuning. Ik kon met problemen altijd bij je terecht. Je enthousiasme en inspiratie werkten aanstekelijk en je was een onmisbare stuwende kracht in het onderzoek. Joan Wellink, jou wil ik bedanken voor de vrijheid die je me liet in het onderzoek en je kritische begeleiding waarmee je de lijn en de continuïteit van het onderzoek hebt helpen waarborgen. De overige leden van de CPMV familie, Jan Verver, Titia Sijen, Hans van Bokhoven, Heleen Frings, Geertje Rezelman, de gastmedewerkers Jean-Michel Mesnard, Olivier Le Gall, Jean-Christophe Boyer en de toenmalige studenten Diederik de Bruijn, Maaike Hamel, Wilfried Voorhorst, Jan Wery, Richard Milek, Jan-Willem Donkers, Gijs van Tol, Hilda Kuin, Ingeborg Kooter en Ellen Nollen wil ik bedanken voor hun inzet en toewijding, de compu's, trafo's, protoplasten en andere zaken als ik weer eens zonder zat, de gemoedelijke sfeer waarin het prettig werken was en de gezellige diners en snookeravonden. Van de vakgroep Virologie wil ik Rob Goldbach voor zijn belangstelling en zijn bijdrage in de werkbesprekingen en Jan van Lent voor het microscopie werk bedanken. Peter van Druten, Piet Madern en fotolocatie de Dreyen voor het teken- en fotowerk, vooral wanneer dat weer eens een haastklus was. En natuurlijk alle Molbiollers die me met hun blijvende belangstelling, ook toen ik de vakgroep verliet, hebben gesteund, een vakgroep waarin ik me mede dankzij jullie steeds heb thuisgevoeld. Buiten de vakgroep heeft de promotie duidelijk een stempel gedrukt op het privé-leven. Daarom wil ik ook mijn familie en vrienden voor hun niet aflatende interesse, vooral in de minder goede tijden van het onderzoek, bedanken. Lieve Jenny, jij als geen ander bent hierbij betrokken geweest. Je zorgde altijd voor de broodnodige ontspanning, in de overtuiging dat er meer dan wetenschap alleen is. Je warmte en vertrouwen waren en blijven onmisbaar.