MOLECULAR CLONING AND EXPRESSION OF FULL-LENGTH DNA COPIES OF THE GENOMIC RNAs OF COWPEA MOSAIC VIRUS

PIETER VOS

Cover shows nucleotide sequence analysis of cDNA clones modified by site-directed mutagenesis.

Cover design, Ingrid Hoornoedt

This work was supported by the Netherlands Foundation for Chemical Research (S.O.N.) with financial aid from the Netherlands Organisation for the Advance of Pure Research (Z.W.O.).

ISBN 90 - 71382 - 17 - 6

FEBO druk, Enschede

Promotoren: Dr. A. van Kammen, hoogleraar in de moleculaire biologie
Dr. R.W. Goldbach, hoogleraar in de virologie

Pieter Vos

Molecular cloning and expression of full-length DNA copies of the genomic RNAs of cowpea mosaic virus

Proefschrift
ter verkrijging van de graad van
doctor in de landbouwwetenschappen,
op gezag van de rector magnificus,
dr. C.C. Oosterlee,
in het openbaar te verdedigen
op vrijdag 4 september 1987
des namiddags te vier uur in de aula
van de landbouwuniversiteit te Wageningen

STELLINGEN

1. Het door cowpea mozaïek virus (CPMV) gecodeerde 32 kilodalton (KD) eiwit is geen proteïnase. (Dit proefschrift)

2. Uit het feit dat de proteolytische klieving van de door het CPMV M-RNA gecodeerde polyproteïnen niet geremd wordt door antilichamen gericht tegen het virale 170 kD eiwit, moet geconcludeerd worden dat deze antilichamen niet met het 24kD proteïnase van CPMV reageren. (Franssen et al., J.Virol.50,183-190,1984)

3. Bij de verdunningsproeven met het CPMV 170 kD eiwit, bedoeld om vast te stellen welke proteolytische splitsingen in dit eiwit inter-, dan wel intramoleculair verlopen, zijn niet de juiste controle-experimenten uitgevoerd. (Peng & Shih, J. Biol. Chem.259,3197-3201,1984)

4. De primaire klieving van het CPMV 200 kD eiwit vindt niet plaats door een cellulair proteïnase, maar door het CPMV 24 kD eiwit. (Tian & Shih,J.Virol.57,547-551,1986; Dit proefschrift)

5. De "Kozak"-regel voor initiatie van eiwitsynthese bij eukaryotische boodschapper-RNA's is te eenvoudig voor een groot aantal RNA's van virussen met een enkelstrengs RNA-genoom. (Kozak,Microbiol.Rev.47,1-45,1983)

6. Bij de proeven van Kang en Wu (1987) waarin het effect van basesubstituties in de SP6-promotor is onderzocht, is geen rekening gehouden met de aard van de betreffende mutaties. (Kang & Wu,Nucleic Acids Res.15,2279-2294,1987)

7. Plantevirussen zijn niet geschikt als vektoren voor het introduceren van soortvreemd DNA in plantecellen, maar leveren hooguit regulatiesignalen, die waardevol kunnen zijn bij het konstrueren van dergelijke vektoren.

8. Voor het verkrijgen van melkzuurstreptokokken met verhoogde proteolytische aktiviteit door middel van recombinant-DNA technieken met als mogelijke toepassing versnelde kaasrijping, is het van belang meer inzicht te verkrijgen in de regulatie van de genexpressie van het celwandproteinase van deze bacterien.

9. Het is rechtvaardiger om mensen, die door hun leefgewoontes hun gezondheid bewust in gevaar brengen, een hogere premie voor hun ziektekostenverzekering te laten betalen, dan mensen, die vanwege een aangeboren afwijking extra geneeskundige behandeling nodig hebben.

10. De wetenschapsjournalistiek zou aan kwaliteit winnen als een groter deel van de schrijvende pers kennis zou nemen van de elementaire beginselen van de exacte wetenschap en de didaktiek.

11. Een prikklok voor wetenschappers kan alleen maar verzonnen zijn door een burokraat, die meent dat wetenschap bedreven wordt gedurende vijf dagen per week van half negen tot half zes.

12. De APK-keuring voor personenauto's dient te worden uitgevoerd door instanties, die geen belang hebben bij het verrichten van reparaties.

13. Op wetenschappelijke vergaderingen en bijeenkomsten wordt te weinig rekening gehouden met mensen, die niet van koffie houden.

Stellingen behorende bij het proefschrift:
MOLECULAR CLONING AND EXPRESSION OF FULL-LENGTH DNA COPIES OF THE GENOMIC RNAs OF COWPEA MOSAIC VIRUS

Pieter Vos, Wageningen 4 september 1987

CONTENTS

Voorwoord

Een proefschrift kan meestal tot stand komen door de inzet van velen. Dit proefschrift vormt daarop geen uitzondering en bij deze gelegenheid wil ik iedereen bedanken die op welke wijze dan ook heeft bijgedragen aan de totstandkoming ervan.

In de eerste plaats moet ik iedereen bedanken die in de periode 1983-1987 deel heeft uitgemaakt van de "cowpea-groep". Niet alleen waren jullie altijd bereid tot discussie over experimenten of de uitkomst daarvan, maar bovendien waren jullie zonder uitzondering joviale collega's. In het bijzonder moet ik hierbij de namen van Jan Verver en Martine Jaegle noemen. Jan Verver heeft gedurende het grootste deel van mijn verblijf op de vakgroep direct meegewerkt aan de hier beschreven experimenten, waardoor er erg veel werk verzet kon worden. Martine Jaegle heeft een groot aandeel gehad in de constructie van infectieuze cDNA klonen. Ook Joan Wellink en Rik Eggen hebben zich bijzonder ingezet bij de infectieproeven met cDNA-klonen.

Kees Stam, Leon Frencken, Gerco Aangenent, Dick Smit en Michel Flipphi hebben in het kader van hun doctoraalonderzoek alleen een bijdrage geleverd aan dit proefschrift. Hoewel de ene wat meer succes had bij zijn experimenten dan de ander, hebben jullie zonder uitzondering een enorme inzet getoond.

Iedereen, die de afgelopen vier jaren het secretariaat heeft bemand, maar in het bijzonder Gré Heitkönig, wil ik bedanken voor het vele typewerk.

Piet Madern wil ik bedanken voor de vakkundige wijze waarop hij het vele teken- en fotowerk heeft verzorgd. De fotodienst van het NIZO wil ik bedanken voor het afdrukken van de omslagfoto.

Alle andere medewerkers van de vakgroep Moleculaire Biologie wil ik bedanken voor hun waardevolle tips en discussies, de prettige samenwerking en de gezellige sfeer op het lab.

Last but not least wil ik Ab van Kammen en Rob Goldbach bedanken. Ze hebben mij de afgelopen vier jaar vaak met raad en daad terzijde gestaan, in het bijzonder bij het verwerken van de vele gegevens tot prettig leesbare manuscripten. Op beiden is het woord promoter wat mij betreft zowel in de letterlijke als figuurlijke zin van toepassing.

CHAPTER 1

SCOPE OF THE INVESTIGATION

Scope of the investigation

Cowpea mosaic virus (CPMV) is a plant virus of which the genetic information is distributed among two messenger sense RNA molecules separately encapsidated into icosahedral particles (for reviews see e.g. Goldbach and Van Kammen, 1985 and Goldbach, 1987). These RNA molecules, denoted M and B RNA, have a small protein, VPg, covalently linked to their 5' ends and a poly(A) tail at their 3' ends. Both RNAs are translated into large polypeptides (polyproteins), which undergo proteolytical cleavages to yield the functional proteins. For M as well as B RNA the complete nucleotide sequence has been elucidated (Van Wezenbeek et al., 1983; Lomonossoff and Shanks, 1983).

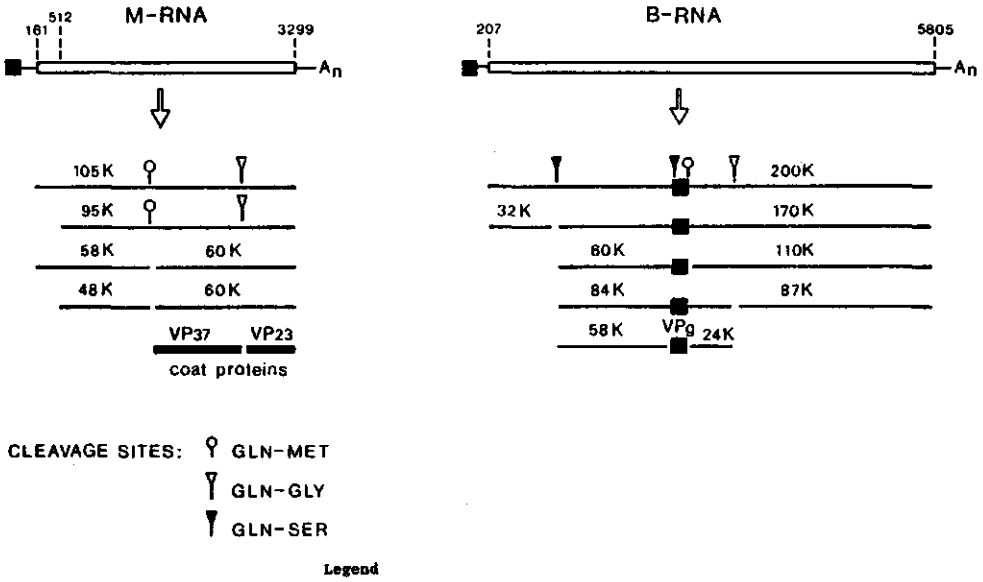


Figure 1. Model for the translation of the RNAs of CPMV and proteolytic processing of the CPMV polyproteins. The double-lined bars in the RNAs represent the long open reading frames, for which the start and stop codons are also indicated (Van Wezenbeek et al., 1983; Lomonossoff and Shanks, 1983). VPg is indicated by a black square, other protein sequences as single lines. The various cleavage sites are indicated on the primary translation products.
N.B. At the start of the investigation the cleavage sites on the B proteins were not yet determined.

At the start of the investigation described in this thesis it was known that B RNA was translated into a 200 kilodalton (kD) polyprotein rapidly

cleaved into polypeptides of 32 kD and 170 kD respectively. The other B RNA encoded polypeptides were shown to be all derived from the 170 kD polypeptide (see Figure 1) and for some of these proteins the function was known. B RNA was shown to be replicated and expressed independently in cowpea protoplasts indicating that the genetic information necessary for replication of the viral RNAs and proteolytic processing of the polyproteins is located on this RNA (Goldbach et al., 1980). Indeed the 110 kD polypeptide had been shown to represent the viral RNA dependent RNA polymerase (Dorssers et al., 1984; Franssen et al., 1984a) while the 60 kD polypeptide was found to be the direct precursor to VPg (Zabel et al., 1982). The observation that the 24 kD polypeptide showed significant sequence homology to the picornaviral P3c proteases suggested that this polypeptide represented a viral protease (Franssen et al., 1984a). Antibody inhibition studies had shown that the 32 kD polypeptide was involved in release of the 60 kD capsid protein precursor from the M RNA-encoded polyproteins, and therefore this polypeptide was proposed to be a second viral protease involved in cleavage of this site at a glutamine-methionine dipeptide sequence (Franssen et al., 1984b). The two capsid proteins VP37 and VP23 were known to be released from the 60 kD capsid protein precursor by cleavage at a glutamine-glycine site, but the proteolytic activity involved in this cleavage reaction was not yet identified.

The scope of the investigation described in this thesis was to gain further insight into the genetic organisation and gene expression of the CPMV RNAs. The approach was to construct full-length DNA copies of both genomic RNAs and try to express these copies in vitro and/or in vivo. If this could be achieved than the effect of specific mutations, introduced in these DNA copies using methods for site-directed mutagenesis, could be studied. The full-length DNA copies of M and B RNA appeared not to be expressed in vivo themselves, but in the course of our investigations efficient in vitro transcription systems became available, which allowed the synthesis of RNA molecules from the full-length DNA copies closely resembling the viral RNAs. These in vitro transcription systems proved to be extremely valuable for the expression of the full-length DNA copies, both in vitro and in vivo.

Chapter 2 gives an overview of the different methods, which can be employed

for the construction and molecular cloning of DNA copies of plant RNA viruses and viroids. In addition it is described how such copies have been used to elucidate the various properties of plant viral RNA genomes. The molecular cloning and expression in vitro of full-length DNA copies of M and B RNA is described in Chapters 3 and 4 respectively. Next in Chapter 5 we show how DNA clones of both RNAs of CPMV were constructed, which upon transcription in vitro produced RNA molecules able to be replicated and expressed in cowpea protoplasts. Finally in Chapter 6 the in vitro expression of specifically modified cDNA clones is described, which has resulted in the identification of the B RNA-encoded 24 kD protein as the protease responsible for all cleavages in the processing of the CPMV polyproteins and has led to a better understanding of the proteolytic processing reactions themselves.

CHAPTER 2

cDNA CLONING OF PLANT RNA VIRUSES AND VIROIDS

Chapter 3

cDNA Cloning of Plant RNA Viruses and Viroids

Pieter Vos

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

With 9 Figures

Contents

Pieter Vos

I. General Introduction

After the development of recombinant DNA technology and convenient techniques for rapid determination of nucleotide sequences, the construction of well-defined cloned DNA copies of plant viral RNA genomes and viroids is becoming more and more an integral part of the research on these pathogens. In this chapter I will attempt to show how the availability of such clones indeed has added a new dimension to plant virology.

In the first part of this chapter (section II) the technical aspects of synthesising cDNA will be described while, in addition, various cloning strategies will be compared and discussed. The following part (section III) will deal with applications of cDNA clones in experiments designed to unravel the multiplication cycle of plant RNA viruses. Hence, the use of these clones to study RNA replication, RNA recombination, gene expression and functions of viral proteins will be discussed. Section IV; in addition, will focus on the use of cDNA clones in viroid and satellite research whereas, finally, in section V some attention will be paid to diagnosis of plant virus diseases using cDNA clones.

II. Construction of Full-Length cDNA Clones

A. Introduction

Following the development of modern molecular cloning techniques, an ever growing number of plant virologists have been trying to construct full-length cDNA clones of plant RNA viruses and viroids, aiming to obtain specific probes for studying plant viruses, viroids and their diseases. While for several viroids full-size cDNA clones have been readily obtained (see e. g. Van Wezenbeek *et al.*, 1982; Cress *et al.*, 1983; Ohno *et al.*, 1983; Sano *et al.*, 1984; Tabler and Sanger, 1984), relatively few reports have been published on the construction of full-length cDNA clones of plant RNA viruses. This already indicates that synthesis of full-length double-stranded cDNA from relatively large RNA molecules, often containing extensive secondary structures, is far from easy. When writing this chapter full-size cDNA clones have been reported for the RNA of satellite tobacco necrosis virus (STNV) (Van Emmelo *et al.*, 1984), tobacco mosaic virus (TMV) (Meshi *et al.*, 1986; Dawson *et al.*, 1986), the three genomic RNAs of bromo mosaic virus (BMV) (Ahlquist and Janda, 1984) and the two genomic RNAs of cowpea mosaic virus CPMV (Vos *et al.*, 1984; Verver *et al.*, submitted). In the following paragraphs I will discuss in some detail how such clones can be prepared.

B. Synthesis of double-stranded cDNA

For first strand synthesis the reverse transcriptase of avian myeloblastosis virus (AMV) is routinely used. This enzyme, which is commercially

available since 1978, synthesises complementary DNA using RNA as template and requires like DNA polymerase a primer. Essential for successful first strand synthesis are both the quality of the reverse transcriptase used and the absolute purity and intactness of the template RNA. Potent RNase inhibitors are often included in the reaction mixture to prevent degradation of template RNA during the transcription reaction by ribonucleases present in most commercial reverse transcriptase preparations. Suitable reaction conditions have been described by several authors (see e. g. Efstratiadis *et al.*, 1976; Retzel *et al.*, 1980; Fields and Winters, 1982) and usually differ only on the monovalent cation concentration used. Actually, the optimal conditions for reverse transcription vary for different template RNAs and should therefore be determined independently for each RNA species. First strand synthesis can be followed by addition of labelled dNTPs to the reaction mixture and analysis of the reaction products on denaturing agarose gels (see Fig. 1).

The next step in the procedure is to synthesise the second strand using again reverse transcriptase or, as mostly preferred, *E. coli* DNA polymerase I. If the latter enzyme is used in combination with RNase H the first strand material synthesised can directly be used for second strand syn-

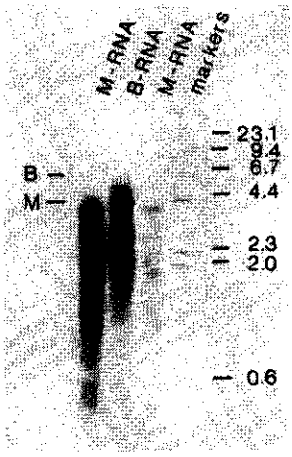


Fig. 1. Reaction products of first strand cDNA synthesis of cowpea mosaic virus M and B RNA analysed on a 1.2 % alkaline agarose gel (McDonnell *et al.*, 1977). On the left hand site the positions of full-size transcripts of M (3.5 kb) and B RNA (5.9 kb) are indicated. On the right hand site the positions of marker DNA fragments are shown.

thesis. The holo-enzyme of DNA polymerase I will perform "repair synthesis" on the RNA-DNA hybrid, replacing thereby the RNA strand attacked by the RNase H (Okayama and Berg, 1982). As a result a complementary strand will be obtained, containing several nicks that can be easily repaired by ligation. Although this approach is very efficient for synthesising the complementary strand and has the advantage that no primer is needed for second strand synthesis, it does not result in full-length double-stranded cDNA molecules comprising the ultimate 5'-terminal sequence of the RNA. To achieve the latter, second strand synthesis should be primed with an oligonucleotide homologous to the 5'-end of this RNA and extended using the Klenow fragment of DNA polymerase I. Since the Klenow enzyme cannot replace RNA hybridised to the template DNA, this RNA should be removed previously by alkali treatment of RNase treatment after heat denaturation. To improve the yield of full-length double-stranded cDNA it is advisable to seperate full-size single-stranded cDNA from smaller products prior to second strand synthesis. This can be accomplished in several ways, e. g. alkaline sucrose density gradient centrifugation (Meshi *et al.*, 1983), elution from denaturing polyacrylamide gels (Maxam and Gilbert, 1980) or elution from alkaline agarose gels (P. Vos, unpublished results).

C. Cloning Strategies

Once the double-stranded cDNA has been obtained, insertion in a suitable acceptor plasmid should follow. Blunt-end cloning of the double-stranded cDNA is most simple, but generally not very efficient. Furthermore, DNA fragments can be orientated in either direction, which in some cases may lead to a considerable number of unsuitable clones if one orientation, the wrong one, is preferred. For cowpea mosaic virus (CPMV) a strong orientation preference for full-length DNA copies of M RNA has been observed in some vectors (P. Vos, unpublished results). That, on the other hand, this approach can be very successful was demonstrated by Ahlquist and Janda (1984) by cloning in this way the three genomic RNAs of BMV.

Addition of linkers to the cDNA certainly enhances the efficiency of cloning, but knowledge of the nucleotide sequence of the RNA is then required to select suitable linkers. This disadvantage may be circumvented by methylation of the double-stranded cDNA (by specific methylases) prior to addition of linkers and subsequent digestion to generate sticky ends.

Alternatively first strand synthesis can be primed with an oligonucleotide extended at its 5'-end with the recognition sequence of a restriction enzyme for which no cleavage sites are present in the cDNA sequence. Upon second strand synthesis, initiated with an oligonucleotide homologous to the 5'-end of the RNA sequence, the double-stranded cDNA is digested with the restriction enzyme in question and ligated into a suitable vector using one sticky (3'-)end and one blunt (5'-)end. This protocol

ensures both efficient ligation and insertion in the desired orientation. For plant RNA viruses having a 3'-poly(A) tail (e. g. comoviruses, potyviruses, nepoviruses) first strand synthesis may be primed with a plasmid provided with an oligo(dT) tail as described by Okayama and Berg (1982). Second strand synthesis can next be primed with an oligonucleotide homologous to the 5'-end of the RNA and cloning involves only circularisation of the synthesised double-strand cDNA linked to the plasmid. This protocol serves again the important goal of efficient cloning in the desired orientation while, furthermore, the minimal length of the poly-d(A-T) track in the cDNA clone to be constructed can be previously determined by the length of the oligo (dT) tail linked to the plasmid primer.

Finally, when the RNA sequence is known, double-stranded cDNA may be cleaved with restriction enzymes yielding a number of fragments, which can be cloned separately and joined afterwards to a full-size copy (Meshi *et al.*, 1986; Dawson *et al.*, 1986; Vos *et al.*, 1984; Verver *et al.*, submitted). This approach is especially attractive for cloning genomic RNAs of considerable length.

D. Cloning in Transcription Vectors

Recently several *in vitro* transcription systems have been developed, which allow the synthesis of large quantities of RNA from cloned cDNA (Contreras *et al.*, 1982; Green *et al.*, 1983; Davanloo *et al.*, 1984). These systems are based on a homologous combination of a specific RNA polymerase and its promoter. RNA transcripts generated from such systems have been proven to be efficient messenger RNAs *in vitro* (Melton *et al.*, 1984; Ahlquist and Janda, 1984; Vos *et al.*, 1984), which in case of transcripts from viral DNA copies may express infectivity (Ahlquist *et al.*, 1984a; Mizutani and Colonno, 1985; Van der Werf *et al.*, 1986). Since DNA copies of most plant RNA viruses are not directly infectious (exception STNV, Van Emmelo, submitted), cloning downstream of a promoter sequence of an RNA polymerase, which allows efficient transcription *in vitro*, may be an attractive possibility to generate infectious RNA from cDNA clones (see for further discussion on this point section III. 2.). Efficient *in vitro* transcription systems have been described based on *E. coli* RNA polymerase on one hand or on the RNA polymerases of bacteriophages like SP6 or T7 on the other hand.

The *E. coli in vitro* transcription system is not as efficient as the transcription systems based on RNA polymerases of bacteriophages, but has the advantage that sequences downstream from the transcription initiation site are dispensable for efficient transcription *in vitro*. The *E. coli* promoter is defined by the "-35" and "-10" region, allowing cDNA inserts to be positioned directly at the transcription initiation site of a strong *E. coli* or *E. coli* phage promoter without affecting transcriptional efficieny. Hence Ahlquist and Janda (1984) designed a transcription vector, pPM1, based on the phage lambda PR promoter (see Fig. 2). Introduction of any DNA fragment with a 5'-proximal purine into the unique *Sma1* site in this vector

will result in transciptional start by *E. coli* RNA polymerase at the first nucleotide of the DNA fragment inserted. Moreover, addition of the cap analog m^7GpppN to the transcription mixture leads to the generation of capped transcripts (Contreras *et al.*, 1982; Ahlquist and Janda, 1984). Using this *E. coli* cloning/transcription system RNA transcripts from full-length cDNA clones of the three genomic RNAs of BMV and of TMV RNA have been synthesised, which perfectly resemble the natural viral RNAs with respect to their capped 5'-ends (Ahlquist *et al.*, 1984a; Meshi *et al.*, 1986; Dawson *et al.*, 1986).

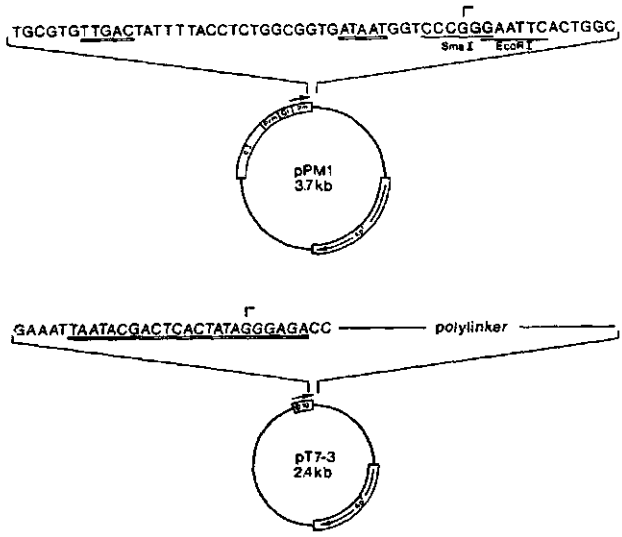


Fig. 2. Schematic representation of transcription vectors pPM1 and pT7-3 (see text) with in detail the sequences surrounding the transcription initiation sites (indicated by arrows). Sequences important for efficient transcription by either the *E. coli* or T7 RNA polymerase are indicated by double lines. In the plasmids the β-lactamase gene for ampiciline resistance is shown, pBR322 derived sequences are indicated by a single line and open boxes refer to phage λ or T7 derived sequences (Ahlquist and Janda, 1984; Queen, 1983; Dunn and Studier, 1983; Tabor and Richardson, 1985).

Similarly the RNA polymerase of bacteriophages like SP6 and T7 are very suitable to generate RNA transcripts *in vitro* from any DNA fragment positioned downstream of an SP6 or T7 promoter respectively (Green *et al.*, 1983; Vos *et al.*, 1984; Verver *et al.*, submitted, and Fig. 2). These RNA polymerases have the advantage over the *E. coli* RNA polymerase that they are very specific for their own promoter sequence and that they are at least

ten times more active *in vitro* compared to the *E. coli* enzyme. Yields of up to 20μg of RNA synthesised from 1μg of template DNA have been described (Tabler and Sänger, 1985; Verver *et al.*, submitted). Similar to *E. coli* RNA polymerase these phage-encoded polymerases are capable of incorporating the cap precursor m^7GpppN generating capped transcripts (Konarska *et al.*, 1984; P. Vos, unpublished results). A general disadvantage of the phage polymerases is, however, that sequences downstream of the transcription initiation site are important for efficient transcription and thus can not be removed. With both SP6 and T7 promoters the first five nucleotides downstream of the transcription initiation site form part of the promoter sequence (Dunn and Studier, 1983; P. Vos, unpublished results). Nucleotide changes in this region decreases the efficiency of transcription, although considerable amounts of RNA transcripts may still be produced. Investigation of this issue has indicated that as long as the transcription initiation site and the first downstream nucleotide remain unchanged, transcription efficiency is still 10—20 % compared to normal (Van der Werf *et al.*, 1986; P. Vos, unpublished results), which, in most cases, still gives sufficient yield. Vectors for the SP6 or T7 *in vitro* transcription system are commercially available and have multiple cloning sites downstream of the promoter sequence facilitating insertion of DNA fragments. To date construction of cDNA clones downstream of bacteriophage promoters has been reported for several RNA viruses like human rhino virus (Mizutani and Colonno, 1985), poliovirus (Van der Werf *et al.*, 1986), black beetle virus (Dasmahapatra *et al.*, 1986) and cowpea mosaic virus (CPMV) (Vos *et al.*, 1984; Verver *et al.*, submitted). Infectious transcripts from these cDNA clones were obtained in spite of non-viral extra sequences at the ends of the transcripts (see also III. 2).

III. DNA Copies as Tool to Study the Molecular Biology of Plant RNA Viruses

A. Introduction

This section will deal with the application of cDNA clones in studies on the molecular biology of plant RNA viruses. Only for a few plant RNA viruses full-length cDNA clones are currently available, although their number will certainly grow in the near future. When writing this chapter only for the genomic RNAs of cowpea mosaic virus (CPMV), bromo mosaic virus (BMV) and tobacco mosaic virus (TMV) full-size cDNA clones have been described. As a consequence the following pages will mainly focus on these viruses. Whenever necessary background information on these viruses will be given throughout the different paragraphs.

Pieter Vos

B. Infectivity of cDNA Clones

i) Infectivity of DNA Copies

Since plant RNA viruses have RNA as source of genetic information it is interesting to know if a double-strand DNA copy of such a genomic RNA is infectious upon inoculation of host cells. This phenomenon has indeed been found for some RNA viruses, like bacteriophage QB (Taniguchi *et al.*, 1978), and for human poliovirus (Racaniello and Baltimore, 1981). In all these cases the infectivity of the DNA copy was very low compared to that of viral RNA. This is not surprising, especially not for RNA viruses of eukaryotes, if one tries to imagine the pathway that such a copy must follow after introduction into the host cell to eventually produce an infectious RNA molecule. To achieve this the DNA copy, after entering the cytoplasm, has to reach the nucleus where transcription should take place. Transcripts having probably extra non-viral sequences at both ends should then be transported to the cytoplasm, where the multiplication process of the virus takes place. To get multiplication of these transcripts, they should be translated (to obtain viral replication proteins) and probably should be processed to RNA molecules that can be recognised by the viral replicase. One or more of these processes can be expected to be very inefficient and therefore may lead to a very low level of infectivity. While for viroids direct infectivity of double-stranded DNA copies have been demonstrated (see section IV. 2 for a more detailed discussion), for plant RNA viruses infectious DNA copies have so far not been reported. It could very well be that the infectivity of such copies is very low, as in the case of poliovirus, and therefore escaped detection. For CPMV, direct infectivity of double-stranded DNA copies has been tested extensively, but never resulted into a detectable level of infection (unpublished results). Even the use of (strong) eukaryotic promoter sequences, like the SV40 early promoter, the nopaline synthase promoter from *Agrobacterium tumefaciens* or the cauliflower mosaic virus (CaMV) 35 S promoter to enhance transient transcription after introduction in cowpea cells, was not successful.

An alternative way to express double-stranded DNA copies to give infectious particles again may be transfer of the cDNA to plants using *Agrobacterium tumefaciens*, which can also result in stable integration of the DNA copy into the plant chromosome (see chapter on agroinfection). Thus far this approach has been only successfully applied for plant DNA viruses, i. e. cauliflower mosaic virus (Grimsley *et al.*, 1986) and tomato golden mosaic virus (Rogers *et al.*, 1986).

With respect to RNA viruses partial DNA copies of TMV RNA have been integrated into tobacco (Bevan *et al.*, 1985; Powell Abel *et al.*, 1986). It was demonstrated that these DNA copies were transcribed in transformed plants and furthermore that also TMV-specific protein in casu coat protein could be detected. For CPMV cowpea cells have been transformed with a full-size DNA copy of the smaller of the two genomic RNAs, the M-RNA, provided either with the nopaline synthase (nos) promoter or the CaMV 35S promoter, using *Agrobacterium*-mediated gene transfer. This

DNA copy was efficiently expressed in transformed callus tissue obtained, yielding transcripts of the expected length of more than 3500 nucleotides long (Garcia *et al.*, 1986). These transcripts were remarkably stable although synthesis of CPMV RNAs occurs normally in the cytoplasm. Since the M-RNA of CPMV is dependent for its replication on the B-RNA encoded RNA dependent RNA polymerase, it will be of interest to determine if infection with B-RNA of transformed cowpea cells will result in replication of the M-RNA-like transcripts.

ii) Infectivity of in vitro Transcripts from DNA copies

The development of efficient *in vitro* transcription systems has offered an attractive alternative for the approaches described above to achieve infectivity from plant viral DNA copies. DNA copies preceded by a strong promoter specific for as certain RNA polymerase can be transcribed *in vitro* into their RNA counterparts. Using this approach infectious transcripts have been described for human rhinovirus and poliovirus (Mizutani and Collono, 1985; Van der Werf *et al.*, 1986), the insect virus black beetle virus (Dasmahapatra *et al.*, 1986) and the plant RNA viruses BMV (Ahlquist *et al.*, 1984a) and TMV (Dawson *et al.*, 1986; Meshi *et al.*, 1986).

The problem when synthesising RNA *in vitro* is to generate transcripts which resemble the original viral RNA with respect to its 5′ and 3′ end. It is obvious that extra terminal sequences may interfere with viral RNA replication, since this is generally a very specific process. From experiments done so far it is clear that the more the termini of *in vitro* RNA transcripts resemble those of the natural viral RNA the more infectious they are. For both the picornaviruses and viroids it was found that RNA transcripts were some orders of magnitude more infectious than the DNA copies themselves (Van der Werf *et al.*,, 1986; Tabler and Sänger, 1985; Mizutami and Collono, 1985). However the way in which extra sequences affect the infectivity of *in vitro* transcripts may differ from virus to virus depending on their replication mechanism.

As mentioned above for plant RNA viruses infectious transcripts have thus far only been obtained for BMV and TMV (Ahlquist *et al.*, 1984a; Dawson *et al.*, 1986; Meshi *et al.*, 1986). Both viruses have RNAs with a cap structure at their 5′ end, and for both capping of the *in vitro* transcripts was essential for infectivity. Extra nucleotides at the 5′ ends and/or the absence of a cap (almost) abolished the infectivity of the *in vitro* transcripts. On the other hand extra sequences at the 3′-end seem not to have such drastic effects. Such extra sequences have never been found in the progeny of *in vitro* transcripts, indicating that they are removed upon replication. Furthermore, for TMV reconstitution *in vitro* of the RNA transcripts with coat protein have been shown to further increase their infectivitiy (Meshi *et al.*, 1986). It is obvious that transcripts are protected in this way from nucleases and that the infection process occurs much more efficiently.

In various cases it has been observed that, without any obvious reason, some cDNA clones did not produce infectious transcripts. Although none

of these clones have been analysed in detail, changes at the nucleotide level are probably the reason for this lack of infectivity. Indeed it has been demonstrated that the RNA population of a virus may be rather heterogenous in sequence (Goelet *et al.*, 1982; Lomonossoff and Shanks, 1983), probably due to miscopying during RNA replication. The error rate of RNA-dependent RNA polymerases in general is 10^{-3} to 10^{-4}. Non-infectious clones can therefore have originated from RNA molecules which were not infectious. Alternatively, mistakes may be introduced by the reverse transcriptase during first strand synthesis. In this view the use of several independently obtained clones for infectivity assays is advisable. However, sequence analysis of non-infectious cDNA clones could still be very important to gain more insight in the process of virus infection and multiplication.

C. RNA Replication

Thus far little is known about the replication of eukaryotic RNA viruses, and this is certainly true for RNA viruses of plants. Only recently the idea is more and more emerging, that most, if not all, eukaryotic RNA viruses encode their own RNA-dependent RNA polymerase (replicase) which, as far as known, is very specific for its own RNA (for a review see Van Kammen, 1985). Most probably this specificity must have its origin in the primary and/or secondary structure at the termini of the RNA molecule, where RNA replication starts. Determination of the sequence signals involved in the initiation of replication would be importantly facilitated by the availability of appropriate systems to study this process *in vitro*. Such a system has thus far only been described for BMV (Bujarski *et al.*, 1982; Miller and Hall, 1983). From BMV-infected barley leaves viral RNA-synthesising extracts can be prepared, which are RNA dependent and capable of synthesising full-length complementary strands from added BMV RNA. Hence, since for BMV full-length cDNA clones have been constructed, a powerful system is available to study sequences involved in the initiation of replication of BMV RNA, by addition of RNA transcripts prepared from modified DNA templates to BMV- infected leaf extracts. Before discussing the results obtained so far in this system it is necessary to give a short description of the virus.

BMV is a virus of which the genetic information is divided among three positive-stranded RNA molecules, designated RNA 1, 2 and 3 respectively, with sizes of 3.2, 2.9 and 2.1 kilobases (kb) (Lane and Kaesberg, 1971; Lane, 1981; Ahlquist *et al.*, 1984 b). In infected plants an additionals RNA is abundantly present, RNA 4, which is as a subgenomic mRNA of RNA 3 and codes for the capsid protein. The viral RNAs all have a cap structure at their 5'-end and possess a tRNA-like structure at their 3'-end, including a CCA-terminus which can be aminoacetylated with tyrosine (Kiberstis and Hall, 1983; Loesch-Fries and Hall, 1982) (see Fig. 3). The sequence of the 3'-terminal 200 nucleotides is virtually identical among the three genomic BMV RNAs, indicating that this region may play an important

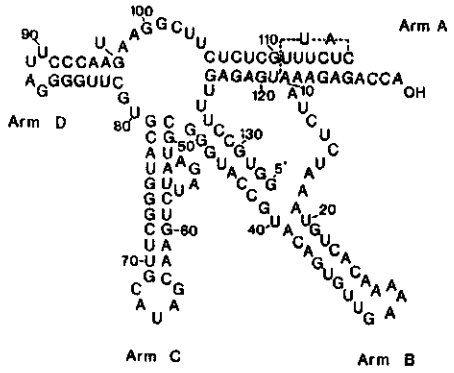


Fig. 3. t-RNA-like structure at the 3' end of the bromo mosaic virus RNAs as proposed by Rietveld *et al.* (1983). This stem structures arm A, B, C and D are indicated in the figure. The CCA-terminus is located in arm A, the anticodon AUA in arm C.

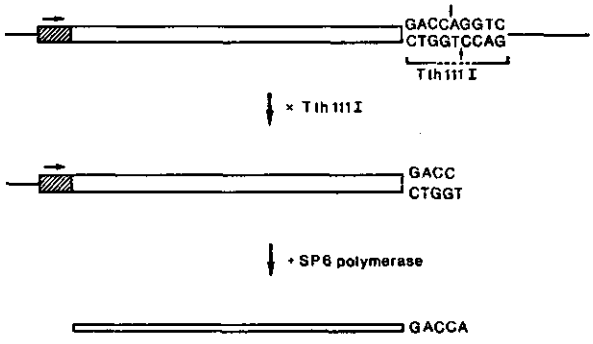


Fig. 4. Schematic representation of the cDNA clone containing the ultimate 201 nucleotides of RNA3 of bromo mosaic virus downstream of the SP_6 promotor. Single lines indicate vector sequences, the SP_6 promotor is indicated as a hatched bar, BMV sequences are indicated as an open bar. Note that the central three nucleotides in the nonameric recognition sequence of the restriction enzyme Tth111I are not important for cleavage, allowing site-directed mutagenesis at these three nucleotides without affecting restriction enzyme digestion.

role in initiation of (−) strand synthesis and aminoacetylation (Ahlquist *et al.*, 1984b). To investigate sequences involved in (−) strand synthesis and aminoacylation mutant viral RNAs were synthesised *in vitro* and offered as templates to either extracts of BMV-infected barley leaves (Miller and Hall, 1983) or partially purified preparation of aminoacetyl-tRNA synthetase from wheat germ (Kiberstis and Hall, 1983).

Since for aminoacetylation a correct 3′ CCA-terminus is an absolute requirement (Deutscher, 1983), it was important to generate *in vitro* BMV RNAs containing a correct CCA terminus. To meet this requirement a cDNA clone was constructed, downstream of an SP 6 promoter, which contained the 3′-terminal 201 nucleotides of BMV RNA3 terminated at a unique Tth111I restriction site. Upon transcription of Tth111I-linearised template DNA, RNA transcripts were synthesised ending with a correct CCA-terminus devoid of additional nucleotides (see Fig. 4). Furthermore the recognition sequence of Tth111I (GACNNNGTC) permitted changes in one or more of the three centrally located nucleotides of the nonameric sequence, allowing to study the effect of mutations at this end (see Fig. 4). The effect of several mutations in the RNA sequence on both template activity and amino acelytion were analysed, a. o. point mutations at the CCA terminus and anticodon loop (Dreher *et al.*, 1984), deletions in the stem and loop structures of the tRNA-like region (Bujarski *et al.*, 1985) and 5′-deletions and 3′-extensions of the terminal 201 nucleotides present in the cDNA clone (Miller *et al.*, 1986). The results obtained with the mutagenised transcripts are summarised below.

As expected both point mutations and extensions at the 3′-CCA-terminus dramatically affected aminoacetylation, but the effect on the template activity of the RNA was variable. Replication was not effected by limited extensions to the 3′ CCA terminus or nucleotide changes at the ultimate A, but longer extensions or changes of the penultimate C abolished replicase template activity.

Besides, mutations in the anticodon sequence had no effect on efficiency of aminoacylation, nor on the type of amino acid linked, but caused a drastic decrease of replicase template activity, at a rate independent of the nature of point mutation in the anticodon loop.

Finally, deletions in several regions of the stem and loop structures of the tRNA-like region had differential effects on both aminoacylation and template activity. Complete deletion of arm D (see Fig. 3) had no effect on aminoacetylation or template activity and even seemed to stimulate replication. On the other hand, deletions in arm C abolished replicase template activity and inhibited aminoacetylation partially. Deletion in arm B also affected both tyrosylation and (−) strand synthesis depending on the size of the deleted fragment. From deletion mutants at the 5′-end of the tRNA-like structure it became clear that all transcripts having the last 134 nucleotides of BMV RNA3 undergo normal aminoacylation and initiation of (−) strand synthesis.

One of the most important conclusions from these experiments is that the tRNA-like structure itself contains all signals for both aminoacylation

and initiation of (−) strand synthesis. At least *in vitro* the entire t-RNA-like region appears to be essential for replication, except for loop D. In this context it is noticeable that this latter loop is missing in the closely related broad bean mottle virus (BBMV). Furthermore initiation of (−) strand synthesis seems to start opposite the penultimate C of the CCA-terminus, since point mutations in the ultimate A or 3′ extensions did not affect replication. This observation has been confirmed by other experiments on BMV replication (Miller *et al.*, 1986).

The experiments furthermore confirm that for aminoacetylation a correct CCA-terminus is a prerequisite. Nucleotide substitutions in the anticodon sequence, on the other hand, do not affect this process, nor the identity of the linked amino acid. Besides, it is interesting to note that RNA templates that cannot be aminoacetylated are still very efficient templates *in vitro*, which makes any model demanding the use of esterified tyrosine for priming (−) strand synthesis very unlikely. The role aminoacetylation plays in the life cycle of BMV therefore remains obscure and still awaits to be elucidated.

The results described above indicate the attractive possibilities of *in vitro* transcripts in studying viral RNA sequences involved in replication. Using *in vitro* transcripts of RNA3 of (−) strand polarity it has recently been demonstrated that synthesis of the subgenomic RNA4 occurs by internal initiation of (−) strand RNA3 (Miller *et al.*, 1985). The results obtained with BMV are probably representative for other RNA viruses having genomic RNAs with similar structural features (cap, tRNA-like end). For other plant RNA viruses having a different genome structure, e. g. viruses with 3′-polyadenylated RNAs, a similar approach to study their RNA replication is evident. Although full-length clones suitable for the preparation or transcripts with well defined modifications are available for some of these viruses, the development of template-dependent *in vitro* replication systems, however, seems to impede those kind of experiments at this moment.

D. RNA Recombination

Until recently for plant RNA viruses evidence for RNA recombination was completely lacking. On the other hand, it has been known for years that for some animal RNA viruses RNA recombination is an important feature in their multiplication cycle. Synthesis of mRNAs of influenza virus for instance (Plotch *et al.*, 1981) is dependent upon certain RNA recombination events. RNA recombination has also been found and is regarded to be important in evolutionary variation of picornaviruses (King *et al.*, 1982; Emini *et al.*, 1984) The availability of cDNA clones, which are biologically active, creates the opportunity to generate well-defined, non-lethal RNA-mutants. Infection of plants with such mutants allows to investigate if new RNA molecules are generated through certain recombination events. The first example of studying recombination with a plant RNA virus, e. g. BMV, has been described by Bujarski and Kaesberg (1986) and will be dis-

cussed here. In the previous section (III. 3) it was demonstrated that BMV RNA3-transcripts derived from DNA-templates lacking arm D in the tRNA-like structure are still efficient templates for aminoacylation and replication *in vitro*. Since this arm D seems to be dispensable *in vitro*, Bujarski and Kaesberg inoculated barley plants with a mixture of wild type RNA 1 and 2 and a mutant RNA3, lacking this loop,and followed the fate of the altered RNA3 *in vivo*.

Although barley plants became infected with this combination of RNAs, less RNA3 and subgenomic RNA4 were found to accumulate than normal. The progeny RNA3 and RNA4 both possessed 3′ non-coding regions missing arm D in their tRNA-like structures and thus originated from the modified RNA3 inoculated. Since the deletion in the tRNA-like structure seemed to reduce the viability of RNA3, infected plants were incubated over a longer period to test if RNA variants more similar to wild type RNA3 would arise by recombination events. Indeed after prolonged infection revertant forms of RNA3 were observed in several plants, which eventually outcompeted the original RNA3 mutant and which had regained the deleted stem and loop region. These revertants were different from wild type RNA3 in having the t-RNA-like structure from RNA 1 or 2. Different plants had different revertants and sometimes more than one type of pseudo wild type RNA3 could be detected within the same plant. Analysis of the different RNA 3 variants showed that in most cases recombination had occurred between RNA3 and an homologous region in RNA 1 or 2, while also recombination outside homologous regions was demonstrated. Clearly the D-loop in the tRNA-like structure of BMV RNAs is not essential for infectivity, but its absence is somehow disadvantageous. As a result RNAs which have regained an intact t-RNA-like region by some recombination event have a selective advantage over the RNA3 mutant with the deleted stem and loop structure.

In summary it can be concluded that the experiments with the modified BMV RNA3 very elegantly demonstrate that recombination between plant viral RNAs is possible. It is probable that such a mechanism plays an important role in the evolution of plant RNA viruses. It certainly makes RNA viruses more flexible in adapting quickly to new environments or hosts.

E. Genetic Organisation and Gene Expression

Another major application of cDNA clones in plant virology has been to study the genetic organisation by expression of cDNA clones modified in specific regions. Since no methods for site-specific mutagenesis on RNA are available, the only approach is to generate RNA mutants indirectly by *in vitro* transcription of modified DNA clones. Relatively few reports have been published on the use of cDNA clones in identifying gene functions of RNA viruses *in vivo* for the simple reason that most mutations result in non-infectious genotypes. This means that in a number of cases the only outcome of a mutation in a gene of which the function is to be determined

will be that this gene is essential for infectivity. It is clear that there is a need for alternative ways to study functions of genes, for instance expression of these genes *in vitro*. However, this will only be possible in those cases where suitable assays for probing gene functions are available. An example of such a gene function that lends itself well for analysis *in vitro* is proteolytic processing, which is employed by various RNA viruses in their genetic expression, e. g. picornaviruses (Putnak and Phillips, 1981), comoviruses (Goldbach and Van Kammen, 1985) and potyviruses (Vance and Beachy, 1985; Dougherty *et al.*, 1985; Dougherty and Hiebert, 1985). This processing occurs faithfully upon translation *in vitro* of the viral RNAs (Pelham, 1978; Franssen *et al.*, 1984; Jackson, 1986). The *in vitro* translation and processing of mutant RNA transcripts of cowpea mosaic virus (CPMV) has resulted in important information on the proteolytic processing of this virus. Some of the most interesting results obtained will be discussed here.

The genome of cowpea mosaic virus (CPMV) is divided among two, separately encapsidated, single-stranded RNA molecules designated M and B RNA (for a recent review see Goldbach and Van Kammen, 1985). The viral RNAs are characterised by a small protein covalently linked to the 5'-end called VPg (Stanley *et al.*, 1978; Daubert *et al.*, 1978) and a poly(A) tail at their 3'-end (El Manna and Bruening, 1973). The translation of both RNAs into polyproteins and their subsequent proteolytic processing to yield the functional proteins have been described extensively (for a review see Goldbach and Van Kammen, 1985) (see Fig. 5). From nucleotide sequence analysis of the genomic RNAs (Van Wezenbeek *et al.*, 1983; Lomonossoff and Shanks, 1983) and from amino acid sequences of the N-terminal ends of various viral proteins (Wellink *et al.*, 1986) it is known that three types of cleavage sites are employed in proteolytic processing: glutamine-glycine, glutamine-serine and glutamine-methionine dipeptide sequences. There are several interesting aspects about this proteolytic processing which demand further analysis. What is the nature of the proteolytic activities; are these activities host or virus encoded and if they are virus encoded, which viral protein(s) is/are involved? What determines a cleavage site; is it only the dipeptide sequence or are surrounding amino acids or the three dimensional folding of the protein also important? In the following pages it will be demonstrated how the answers to these questions can be obtained by *in vitro* expression of cDNA clones modified by site-directed mutagenesis.

For this purpose full-length cDNA clones of both CPMV M and B RNA have been constructed downstream of T7 or SP6 promoters, which gave upon *in vitro* transcription and subsequent *in vitro* translation in rabbit reticulocyte lysate translation products indistinguishable from the *in vitro* translation products obtained with natural viral RNA (Vos *et al.*, 1984; Verver *et al.*, submitted). Based on time course translation studies and sequence homology to the picornaviral protease P-3c, the 24 K polypeptide encoded by B RNA (Fig. 5) has previously been proposed to be a viral protease involved in the proteolytic cleavage of the CPMV polyproteins

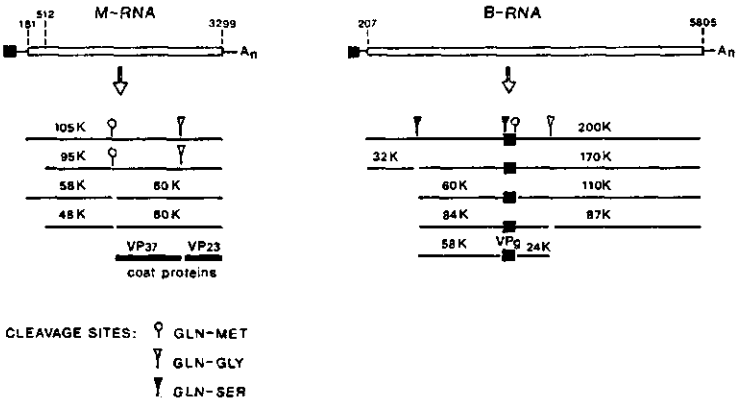


Fig. 5. Translation and processing of the genomic RNAs of cowpea mosaic virus. On the RNAs the long open reading is indicated by a double lined bar and the positions of start and stop condons are shown. Black squares refer to the genome linked protein VPg; other proteins are indicated as single lines with their estimated sizes in kilodaltons (K) presented above the protein.

(Franssen *et al.*, 1984a and 1984c). To test this a small deletion was introduced into the coding region for the 24 K polypeptide within a full-length B cDNA clone (Verver *et al.*, submitted). This deletion did not disturb the reading frame of the B RNA-encoded polyprotein in this clone, but resulted only in the loss of the coding sequence for 29 amino acids centrally located in the 24 K polypeptide (Fig. 5). Whereas translation of viral B RNA or *in vitro* transcripts of full-length cDNA clones of B RNA resulted in a 200 K polyprotein, which is rapidly processed into polypeptides of 170 K and 32 K (see Fig. 5), translation of the transcript from the deletion-containing B cDNA clone resulted in a slightly smaller primary translation product ($\pm$ 197 K), which was not proteolytically processed. This result demonstrates that cleavage at the glutamine-serine dipeptide between the 170 K and 32 K polypeptides is blocked by the small deletion in the 24 K polypeptide representing therefore a protease involved in this cleavage.

The action of the 24 K polypeptide was further investigated by constructing a hybrid cDNA clone containing sequences derived from both M and B RNA. In this hybrid clone a DNA fragment of a full-size B clone (pSPB21) containing the coding region for the 24 K polypeptide plus surrounding sequences, was inserted into a full-length M cDNA clone (pSPM5ΔBglII; see Fig. 6). The result of this construction is clone pMB210,

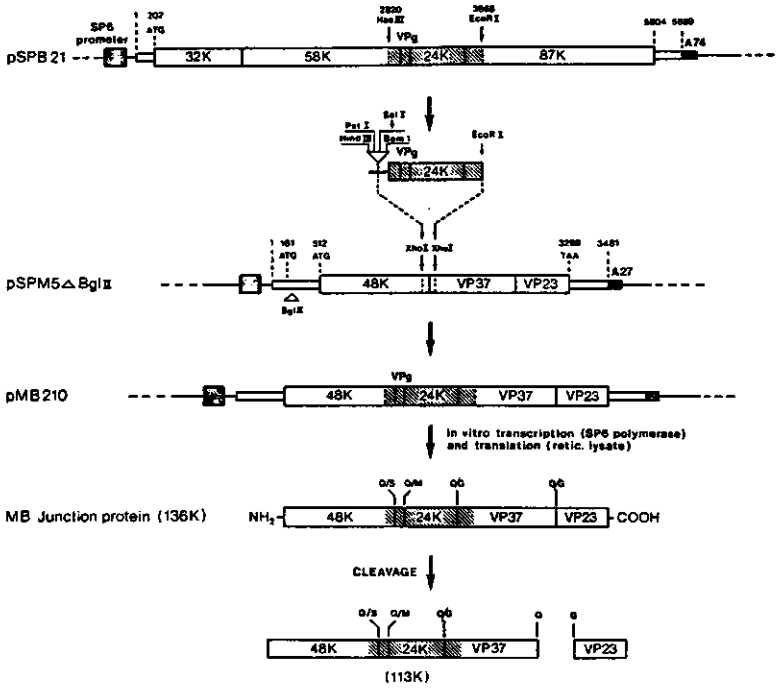


Fig. 6. Construction and schematic representation of a hybrid cDNA clone, containing sequences derived from M and B RNA of cowpea mosaic virus. A HaeIII-EcoRI fragment containing the coding region for the 24K polypeptide plus surrounding sequences was inserted into pUC9 to introduce a number of restriction sites at the left hand site of this DNA fragment. Subsequently this DNA fragment was released with *SalI* and *EcoRI* and inserted into the M cDNA clone pSPM5IBg1II (Vos *et al.*, 1984) digested with *XhoI*. The resulting clone pMB210 contained one contiguous open reading frame which could code for a protein of 136kD. After *in vitro* transcription with SP_6 polymerase and subsequent *in vitro* translation in rabbit reticulocyte lysates indeed a 136K polypeptide is produced, which is rapidly cleaved at the right most glutamine-glycine site (see text).
Open bars represent the open reading frame of polypeptides originating from M RNA (open region) or B RNA (hatched region). Small double-lined bars indicate the non-translated regions. The positions of the cleavage sites on the translation products are indicated by Q/S (glutamine-serine), Q/G (glutamine-glycine) and Q/M (glutamine-methionine).

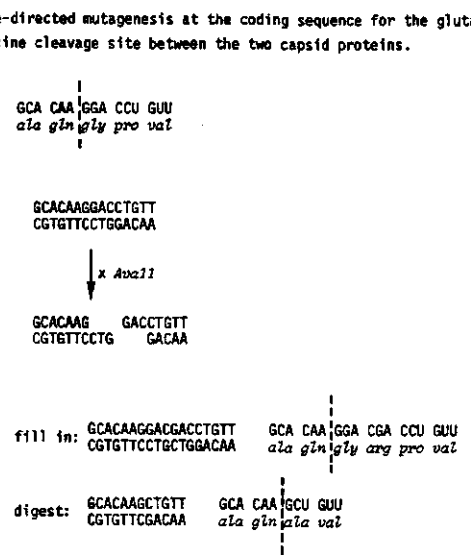


Fig. 7. Site-directed mutagenesis at the coding sequence for the glutamine-glycine cleavage site between the two capsid proteins. The nucleotide sequence surrounding the cleavage site is shown in capital letters, the resulting amino acid sequence in italics. As can be deduced from the sequence a recognition site for AvaII (GGACC) is present. After digestion this enzyme gives 5' protruding ends with an overhang of three nucleotides. "Filling in" with Klenow enzyme and religation results in three additional nucleotides, digestion sith S1-nuclease and religation results in the loss of three nucleotides. Mutagenesis is therefore possible without interrupting the open reading frame.

containing a long open reading frame of approximately 3800 nucleotides and which produced upon *in vitro* transcription and translation a protein with a molecular weight of 136 K (Fig. 6). In this protein, comprising sequences of both the B RNA- and M RNA-encoded polyproteins, four different proteolytic cleavage sites were present, one of which, the glutamine-glycine cleavage site between the two capsid protein (VP37 and VP23) sequences, was rapidly cleaved. This cleavage was not observed when the small deletion of 89 basepairs (see above) was introduced in the 24 K protein-coding region. This result demonstrates that the 24 K protein not only has proteolytic activity involved in cleavage of glutamine-serine sites (see above) but also recognizes glutamine-glycine cleavage sites.

Additionally, to analyse preliminary the properties of the proteolytic cleavage sites two derivatives were constructed from the hybrid cDNA

clone pMB210, which were modified at the coding sequence for the glutamine-glycine site between the two capsid proteins (Vos *et al.*, manuscript in preparation, and Fig. 7). The result of the first mutation was a deletion of one amino acid, changing the glutamine-glycine dipeptide sequence into a glutamine-alanine sequence. The result of the second mutation was an insertion of one amino acid residue at the right hand site of the glutamine-glycine dipeptide sequence, changing the sequence glutamine-glycine-proline into glutamine-glycine-arginine. Upon subsequent *in vitro* transcription and translation of these modified cDNA clones it was shown that removal of the glycine in the cleavage site abolished proteolytic processing completely, while introduction of the extra amino acid next to the glycine resulted in a considerable decrease of proteolytic cleavages at this site.

These and the other results discussed above nicely demonstrate how RNA transcripts derived from modified DNA templates can be utilised to identify and analyse the function and enzymatic activities of viral proteins. Using this approach for CPMV the 24 K polypeptide encoded by CPMV B RNA was definitely identified as a viral protease, recognizing both glutamine-serine and glutamine-glycine dipeptide sequences. Furthermore it was shown that not only a correct dipeptide sequence is an absolute requirement for proteolytic cleavage, but that also neighbouring amino acids affect the recognition of the cleavage site by viral proteases. While for CPMV this approach has resulted in an almost complete understanding of the proteolytic processing of this virus, this may also be expected for other viruses, which employ proteolytic processing for their genetic expression (e. g. nepoviruses, potyviruses).

IV. Viroids and Satellites

A. Introduction

This section will deal with the applications of cDNA clones in studies on viroids and satellites. Viroids and satellites are groups of small RNA molecules which differ from plant RNA viruses in many aspects. One of the major differences is that they are totally dependent for their replication on host enzymes (viroids) or the polymerase of a „helper virus" (satellites). In fact viroids and most satellites do not code for any protein at all, the former occuring as naked RNA molecules. An exception form the so-called „satellite viruses", i. e. satellite RNAs which encode their own capsid protein. The best studied example of these is Satellite Tobacco Necrosis Virus (STNV), a satellite virus of Tobacco Necrosis Virus (TNV), and part of this section will concern cDNA studies on this satellite virus. Most attention here will however be paid to viroids, since these pathogens have been extensively studied for many years while for several of them full-length cDNA clones have been constructed (Van Wezenbeek *et al.*, 1982; Cress *et al.*, 1983; Ohno *et al.*, 1983; Sano *et al.*, 1984; Tabler and Sänger, 1984).

Pieter Vos

Viroids are the smallest known agents of infectious diseases. They consist of a small circular RNA molecule of 270—360 nucleotides that is extensively basepaired to form a rod-like structure (for reviews see Diener, 1983 or Sänger, 1984) (see Fig. 8). Thus far, viroids have exclusively been found to be associated with diseases of higher plants and it has still remained obscure how such small molecules can cause diseases with sometimes dramatic consequences. Dickson *et al.* (1979) demonstrated that sequence heterogeneity among different isolates of the same viroid results in different virulence. Nucleotide sequencing of various isolates of potato spindle tuber viroid (PSTV) (Schölzer *et al.*, 1985) has indicated that mutations affecting pathogenicity are all centered in one region, denoted virulence modulation region. Recombinant DNA technology offers the opportunity to analyse these observations in more detail.

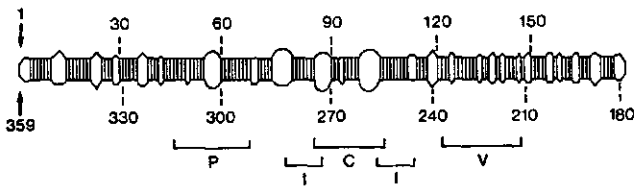


Fig. 8. Schematic representation of the secondary structure of PSTV, a similar structure can be drawn for other viroids as well. V refers to variable region, a region with relatively little sequence homology among the different viroids. P indicates pathogenicity region, a region in which most point mutations affecting virulence are centered. C refers to central conserved region, a region with a very high homology among the different viroids (Keese and Symons, 1985). I indicates the so-called hairpin I, one of the three hairpins, not present in the native configuration but formed transiently during thermal denaturation (Riesner *et al.*, 1979; Henco *et al.*, 1979; Diener, 1986).

Another interesting aspect of viroids concerns the replication mechanism. Replication of viroids takes place via oligomeric RNA intermediates of both positive and negative strand polarity, suggesting a rolling circle-type replication mechanism (Branch *et al.*, 1981; Owens and Diener, 1982; Branch and Robertson, 1984). The existence of oligomeric RNA intermediates requires specific cleavage and subsequent ligation to obtain covalently closed circular molecules (Branch and Robertson, 1984). Since isolation of oligomeric RNA intermediates from infected plants is very difficult, it was until recently not possible to follow the cleavage/ligation reaction in detail. However the availability of cloned multimeric viroid DNA copies has opened the way to study this process as will be demonstrated in this section.

In the first part of this section I will focus on the molecular cloning of viroids. For the construction of full-size cDNA clones of satellites and sat-

ellite viruses I would like to refer to paragraph 2, since cloning of satellite nucleic acid generally meets the same difficulties as cloning of the „helper" virus genome. In fact molecular cloning of satellites mostly is considerably less difficult than cloning of RNA viruses as their size is usually very small. On the other hand moleculear cloning of viroids is different from RNA viruses, due to the circularity and high level of secondary structure of the RNA. Additionally, in the second part of this section the use of DNA copies to improve our understanding of the molecular basis of replication and biological activities of viroids and satellite viruses will be discussed.

B. Molecular Cloning of Viroids

If the nucleotide sequence of the viroid in question is known, first strand synthesis may be primed with one of more specific oligonucleotides, complementary to sequences in the viroid RNA. When this approach is followed two aspects are very important. Fistly, since viroid preparations always contain circular and linear RNA molecules, the circular form of the RNA should be purified from preparative denaturing gels (Maxam and Gilbert, 1980; Van Wezenbeek *et al.*, 1982). Only in this way a high yield of full-length transcripts will be obtained, since linear template RNAs will not result in such transcripts. Secondly, the oligonucleotide primers used should be complementary to regions in the RNA sequence, in which internal basepairing of the viroid RNA is relatively weak, assuring efficient hybridisation of the primer to the RNA. To obtain of full-length double-stranded cDNA second strand synthesis must also be primed with an oligonucleotide.

An elegant approach for this has been developed by Haseloff and Zimmern (personal communication) in which a primer is employed that is partially complementary to the primer used for first strand synthesis (see Fig. 9). Upon annealing of such primer to full-size single-stranded cDNA a partially double-stranded circle wil be formed, which can be converted to a completely double-stranded molecule using the Klenow enzyme. The remaining single-stranded nicks are repaired by DNA ligase and finally the circular ds cDNA is linearised at a unique site and inserted into an appropriate vector.

An alternative approach is to use two primers for first strand synthesis that are complementary to different regions of the viroid RNA. Upon second strand synthesis by hairpin priming or by using the RNaseH/DNA-polymerase I method of Okayama and Berg (1982), cDNA clones will be obtained which together comprise the whole RNA sequence. After characterisation of the cDNA clones a full-size DNA copy can be constructed by joining of selected restriction fragments from these clones.

In addition, molecular cloning of viroid RNA can be achieved by polyadenylation of (linear) viroid molecules followed by priming with oligo(dT) (Sippel, 1973, Ohno *et al.*, 1983). Prior knowledge of the nucleotide sequence is not necessary in this case. Another advantage of this method is that generally cDNA clones will be obtained starting at a dif-

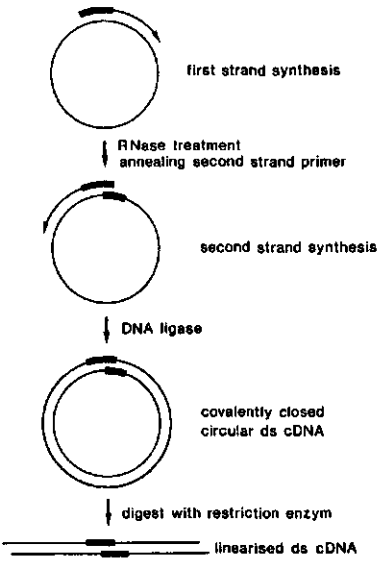


Fig. 9. Cloning strategy for circular RNA molecules, i. e. viroids, for which the nucleotide sequence is known, using an oligonucleotide primer for first strand synthesis and a primer for second strand synthesis partially complementary to the first strand primer. The result is a double-stranded circular DNA copy with a nick in both strands, which can be repaired by DNA ligase. Upon digestion with a unique restriction enzyme a lineair molecule is formed, which can be cloned in monomeric of multimeric forms in a suitable vector molecule.

ferent position of the viroid sequence, resulting in molecular cloning of the entire RNA sequence. This is due to the fact that in general viroid RNA will be linearised and thus polyadenylated at different position of the sequence. Finally, after polyadenylation, all cloning protocols originally developed for cellular mRNAs, like the highly efficient Okayama and Berg (1982) method, can be applied for viroid RNAs.

C. Application of cDNA Clones

i) Viroids

In contrast to the DNA copies of plant viruses available to date all DNA copies from viroids tested fo far have proven to be infectious. A prerequisite for this infectivity however is that the viroid-specific cDNA has to

be cloned in a multimeric form. Monomeric DNA copies have only very low infectivities or are not infectious at all (Diener, 1986). The same holds for RNA transcripts of viroid-specific cDNA clones. The rate of infectivity of multimeric transcripts of positive polarity is indistinguishable from that of natural viroid RNA. On the other hand, neither monomeric transcripts are infectiuos, nor multimeric transcripts of negative polarity (Diener, 1986). These observations suggest that mulitmeric transcripts of positive polarity represent true intermediates in viroid replication and that they are faithfully cleaved and ligated to circular monomers *in vivo* and therefore confer a high level of infectivity. It should be possible to identify sequences involved in this specific cleavage ligation reaction using recombinant DNA techniques.

Meshi *et al.*, (1985) have been trying to locate the region involved in cleavage of viroid multimers by testing the infectivity of cDNA clones containing more than one but less than two units of Hop Stunt Viroid (HSV) sequence. They found that duplication of the upper portion of the highly conserved central region and some neighbouring sequences (see Fig. 8)was sufficient for infectivity of monomeric DNA copies. Site-directed mutagenesis of cDNA clones of Citrus Exocortis Viroid (CEV) (Visvader *et al.*, 1985) has also identified this conserved central region as the possible processing site for viroid oligomers. Based on these and other observations Diener (1986) has recently proposed a model for the cleavage-ligation process of viroid multimers in which these multimers should form a highly basepaired, thermodynamically stable configuration to achieve the generation of monomers. In this model the upper portions of the central conserved region and hairpin I (see Fig. 8) of every two neighbouring monomeric sequences can form an extremely stable G-C-rich double-stranded structure exceeding the stability of native monomeric viroid RNA. Clearly additional experiments with modified cDNA clones will soon reveal the exact location of the cleavage ligation site.

Another promising application of cDNA clones in viroid research is to identify sequences involved in the virulence of the pathogen. These sequences, involved in symptom formation, are of great interest, because they might shed light on the interaction between viroid and host plant. Different isolates of the same viroid have been known for may years now, which replicate at a similar level but differ markedly in pathogenicity, varying from lethality to symptomless expression. Sequence analysis of various field isolates of both potato spindle tuber viroid (PSTV) and citrus exocortis viroid (CEV) have revealed a region in which most point mutations affecting virulence are centered (Schölzer *et al.*, 1985; Visvader and Symons, 1985). This region is called the virulence modulation region (Fig. 8). When writing this review no data were yet available on site-directed mutagenesis of DNA copies with the purpose to affect virulence. However, their availability in the near future may contribute to a better understanding of how viroids cause diseases.

ii) Satellite Viruses

Satellite viruses can be defined as encapsidated nucleic acids that encode their own capsid protein but are unable to multiply in cells without the assistance of a specific „helper virus" and that do not contribute in any way to the replication or infection of this helper virus nor have any sequence homology with their helper virus genome (for a review see Murant and Mayo, 1982). The best studied satellite virus is the satellite of tobacco necrosis virus (STNV). It was the first plant RNA „virus" of which the total nucleotide sequence was elucidated (Ysebaert *et al.*, 1980). The genome of STNV is 1239 residues long, the 5′ half of the RNA containing the coding region for the capsid protein and the 3′-terminal 622 nucleotides being untranslated. A full-sized cDNA clone of STNV has been constructed (Van Emmelo *et al.*, 1980; Van Emmelo *et al.*, submitted), which surprisingly turned out to be infectious. A plasmid containing a full-size DNA copy of STNV inserted by G-C-tailing in the PstI-site of a pBR322 derivative resulted in the production of STNV-particles when inoculated together with the TNV helper virus. A plasmid lacking the 5′-terminal 23 nucleotides of the STNV DNA copy was not infectious while a full-length cDNA clone lacking G-C-tails was not infectious as well. This indicates that the G-C-tail introduced for cloning could be of some importance for the generation of eukaryotic viral RNA from the prokaryotic plasmid DNA, although the mechanism remains obscure (Van Emmelo *et al.*, submitted).

Linker insertion at various places in the STNV DNA copy generally did not affect the infectivity, while the mutation in the STNV RNA was stably maintained in the progeny satellite virus. Insertions in the coat protein region, affecting the reading frame of this protein, did neither influence the infectivity although the ratio of free single-stranded STNV compared to the double-stranded replicative form was highly reduced (Van Emmelo *et al.*, submitted). Probably the STNV RNA is rapidly degraded when it is not encapsidated. Analysis of several deletion mutants revealed that in some cases short deletions were allowed without affecting infectivity.

A very interesting observation was the occurrence of a kind of cross protection by inoculating leaves with the STNV cDNA clone. It appeared that in leaves that were previously inoculated with TNV and a STNV DNA copy, modified in the coat protein region, secondary infection with TNV/STNV did not result in the appearance of STNV particles. The availability of infectious cDNA clones of satellite viruses offers therefore a good opportunity to study these viruses and their relation to the helper virus in great detail. Furthermore, since some satellite viruses probably replicate by the same mechanism as the helper virus, they could shed more light on the replication of the helper virus genome. Sequences involved in initiation of replication *in vivo* could be analysed by introducing specific mutations in the DNA copy of the satellite, without affecting the multiplication and expression of the helper virus itself.

V. Diagnosis of Plant Diseases Using DNA Copies of Plant Viruses and Viroids

A. Introduction

Plant viruses and viroids cause diseases in several economically important crops, e. g. potatoes, tomatoes, cucumbers and tobacco. Infection of plants often results in significant decrease in yield and in some cases total crop loss is possible. It is therefore very important to obtain seeds and seed plants absolutely free of viruses or viroids.

To test plant tissue for the presence of viruses or viroids, methods should be applied which allow rapid but sensitive screening of large amounts of samples routineously. For most plant viruses such methods exist based on immunochemical techniques. These methods, however, cannot be applied to viroids. Several years ago, a method was developed based on hybridisation of radioactively labelled cDNA to viroid RNA bound to nitrocellulose membranes to detect viroids in infected plants (Owens and Diener, 1981). This method, called „spot hybridisation", can also be applied to plant viruses for which immunochemical methods give poor results. In this section I will focus on the characteristics of this method in comparison with other methods used for diagnosis and detection of viruses and viroids.

B. Spot Hybridisation

Since viroids lack coat proteins, immunochemical methods cannot be applied for their detection in infected tissue. Until recently the only alternative detection methods were bioassays on suitable test plants or gel electrophoresis of extracted nucleic acids. In most cases the two methods were combined, since multiplication of viroids in intermediate hosts was necessary for detection on poly-acrylamide gels. Although gel electrophoretic detection techniques have improved a lot, recently culminating in the development of a bidirectional electrophoresis approach in combination with a highly sensitive silver staining technique (Schumacher *et al.*, 1983), they still are laborious and expensive and therefore not very suitable for routine indexing of large amounts of samples. Bioassays show often slow symptom development (Allen and Dale, 1981; Palukaitis *et al.*, 1981) and are not reliable since some viroids can replicate in test plants to normal levels without symptoms (Schölzner *et al.*, 1985).

Molecular hybridisation is a technique routinely used for detection of nucleic acids. Using radioactively labelled probes of high specific activity amounts as low as 0.1 pg of purified nucleic acid bound to a nitrocellulose membrane have been reported (Melton *et al.*, 1984). Therefore hybridisation should be a sensitive technique to detect viroid RNA in infected tissue and because of its simplicity should allow simultaneous screening of large numbers of samples. The last few years this method indeed has been applied successfully to a number of viroids (Owens and Diener, 1981 and

1984; Palukaitis *et al.*, 1985; Symons, 1985). Since methods for molecular hybridisation were already well established, molecular cloning of viroid sequences and methods for preparations of samples from infected plants were the only problems to be solved. Molecular cloning techniques for viroids have been described in the previous section (IV. 2). To date for mosts viroids cDNA clones are available which can be used for preparation or radioactively labelled cDNA. Furthermore methods have been developed for rapid and simple preparation of extracts of plant tissue, which can be performed in a single Eppendorf tube, giving samples suitable for direct „spotting" on membranes, that can bind nucleic acids (Owens and Diener, 1984; Palukaitis *et al.*, 1985). At this moment spot hybridisation is routinely used in several laboratories for the detection of potato spindle tuber viroid, cirtus exocortis viroid, coconut cadang cadang viroid and avocado sunblotch viroid.

Because of its simplicity and sensitivity the enzyme-linked immunosorbent assay (ELISA) will remain the principal diagnostic assay for detection of plant viruses. However, in some cases when ELISA is not satisfactory, molecular hybridisation might provide a solution. When viruses are present in a very low concentration in infected plants dot hybridisation, because of its higher sensitivity, is to be preferred. Other viruses are difficult to purify, are poor antigens or have variable antigenic properties. In these cases problems with ELISA can be expected and should be replaced by the spot hybridisation technique. Likewise, detection of non-encapsidated viruses, like the particle-deficient isolates of tobacco rattle virus (Harrison and Robinson, 1982) cannot be performed by ELISA and should give no problems using spot hybridisation.

Spot hybridisation is an attractive alternative for detection of plant viruses specially when immunochemical techniques fail to succeed. To date, for a number of plant RNA and DNA viruses this technique has been applied successfully and shown to be more sensitive than immunochemical techniques (Maule *et al.*, 1983; Sela *et al.*, 1984). The major disadvantage of the spot hybridisation is the use of radioactivity in the procedure, especially for diagnostic laboratories which are not well equipped for this kind of work. However, the development of non-radioactive probes (Leary *et al.*, 1983) can circumvent this problem and will certainly make this procedure accessible for more laboratories.

VI. Conclusions and Future Aspects

In this chapter we have seen how recombinant DNA technology has enabled experiments with plant RNA viruses and viroids, which until recently were considered to be impossible. Of course it is very difficult to predict how plant virology will develop in the near future, but we certainly may anticipate that the application of cDNA clones will become more and more crucial in experiments which are intended to unravel the properties of plant viruses and viroids. It is obvious that the number of plant RNA viruses for which full-length DNA copies are available will grow steadily.

Apart from studying viral replication and gene expression such clones may be used to identify viral genes involved in specific virus-host plant interactions, hence determining pathogenicity, host range and symptom development. Another application of plant RNA viruses may be the use as gene vectors. Since generally viral genes are expressed very efficiently (e. g. genes for coat and inclusion body proteins), viruses may be altered in suitable gene vectors. Recently French *et al.* (1986) indeed detected chloramphenicol acetyltransferase (CAT) activity in barley protoplasts when infected with BMV RNA 1 and 2 and a modified BMV RNA 3 containing a bacterial CAT gene. A possible advantage of plant viruses as gene vector is that difficult transformation procedures can be circumvented since viral RNAs containing desired genes can simply be inoculated on host plants. Viruses have, on the other hand, various serious disadvantages and limitations as gene vectors. Important disadvantages could be the limited host range of most viruses and the limitations of the size of the genes to be introduced. Furthermore, the error rate of RNA polymerases is known to be relatively high, which may cause rapid introduction of non-desired mutations in inserted genes. Finally viral genes involved in pathogenicity should be altered in such a way that the virus is tolerated by the plant and does not cause crop losses. In summary, it may therefore be concluded that the application of plant viruses as gene vectors is only a distant prospect.

Another application of plant viral cDNAs (and this may sound contradictory) is their use as antiviral agents. As early as in 1929 it was described that infection of a plant with a (mild) virus strain may result in resistance of this plant to other (severe) strains of the same virus (McKinney, 1929). This phenomenon, called "cross protection", also occurs with viroids (Fernow, 1967). This property has for some plant viruses and viroids been used in practice. There are, however, several disadvantages about this procedure. For example, a mild strain may change into a highly pathogenic strain by only one or a few point mutations (Palukaitis and Zaitlin, 1984), while a virus strain causing mild symptoms in one host plant can be very virulent in others. These problems could be circumvented if the viral genes responsible for cross protection could be transferred to host plants instead of using whole viruses. Powell Abel *et al.* (1986) recently demonstrated that tobacco plants transformed with a partial DNA copy of TMV, encoding and expressing the coat protein, indeed showed symptom delay upon infection with the corresponding whole virus. Although the molecular basis of this form of induced resistance is still obscure, these first experiments promise good possibilities for protecting economical important crops against virus infections.

VII. References

Ahlquist, P., Janda, H., 1984: cDNA cloning and *in vitro* transcription of the complete bromo mosaic virus genome. Mol. Cell. Biol. **4**, 2876—2882.

Ahlquist, P., French, R., Janda, M., Loesch-Fries, L. S., 1984a: Multicomponent

RNA plant virus infection derived from cloned viral cDNA. Proc. Natl. Acad. Schi. U.S.A. **81,** 7066—7070.

Ahlquist, P., Dasgupta, R., Kaesberg, P., 1984b: Nucleotide sequence of the bromo mosaic virus genome and its implications for viral replication. J. Mol. Biol. **172,** 369—383.

Allen, R. N., Dale, J. L., 1981: application of rapid biochemical methods for detecting avocado sunblotch disease. Ann. Appl. Biol. **98,** 451—461.

Bevan, M. W., Mason, S. E., Goelet, P., 1985: Expression ot tobacco mosaic virus coat protein by a cauliflower mosaic virus promoter in plants transformed by *Agrobacterium.* EMBO J. **4,** 1921—1926.

Branch, A. D., Robertson, H. D., Dickson, E., 1981: Longer-than- unit-length viroid minus strands are present in RNA from infected plants. Proc. Natl. Acad. Sci. U.S.A. **78,** 6381—6385.

Branch, A. D., Robertson, H. D., 1984: A replication cycle for viroids and other small infectious RNAs. Science **223,** 450—455.

Bujarski, J. J., Hardy, S. F., Miller, W. A., Hall, T. C., 1982: Use of dodecyl-i-D-maltoside in the purification and stabilisation of RNA polymerase from bromo mosaic virus-infected barley. Virology, **11,** 465—473.

Bujarski, J. J., Dreher, T. W., Hall, T. C., 1985: Deletions in the 3′-terminal tRNA-like structure of bromo mosaic virus RNA differentially affect aminoacylation and replication *in vitro.* Proc. Natl. Acad. Sci. U.S.A., **82,** 5636—5640.

Bujarski, J. J., Kaesberg, P., 1986: Genetic recombination between RNA component of a multipartite plant virus. Nature, **321,** 528—531.

Contreras, R., Cheroute, H., Degrave, W., Fiers, W., 1982: Simple, efficient *in vitro* synthesis of capped RNA useful for direct expression of cloned eukaryotic genes. Nucleic Acids Res., **10,** 3133—3142.

Cress, D. E., Kiefer, M. C., Owens, R. A., 1983: Construction of infectious potato spindle tuber viroid cDNA clones. Nucleic Acids Res., **11,** 6821—6835.

Dasmahapatra, B., Dasgupta, R., Saunders, K., Selling, B., Gallagher, T., Kaesberg, P., 1986: Infectious RNA derived by transcription from cloned cDNA copies of the genomic RNA of an insect virus. Proc. Natl. Acad. Sci. U.S.A., **83,** 63—66.

Daubert, S. D., Bruening, G., Najarian, R. C., 1978: Protein bound to the genome RNAs of cowpea mosaic virus. Eur. J. Biochem., **92,** 45—51.

Davanloo, P., Rosenberg, A. H., Dunn, J. J., Studier, F. W., 1984: Cloning and expression of the gene for bacteriophage T7 RNA polymerase. Proc. Natl. Acad. Sci. U.S.A., **81,** 2035—2039.

Dawson, W. O., Beck, D. L., Knorr, D. A., Grantham, G. L., 1986: cDNA cloning of the complete genome of tobacco mosaic virus and production of infectious transcripts. Proc. Natl. Acad. Sci. U.S.A., **83,** 1832—1836.

Deutscher, H. P., 1983: Enzymes of nucleic acid synthesis and modification, **vol. 2** (Jacob, S. T., ed.) 159—183. CRC Press, Boca Raton, Florida.

Dickson, E., Robertson, H. D., Niblett, C. L., Horst, R. K., Zaitlin, M., 1979: Minor differences between nucleotide sequences of mild and severe strains of potato spindle tuber viroid. Nature, **277,** 60—62.

Diener, T. O., 1983: Viroids. Adv. Virus. Res., **28,** 214—283.

Diener, T. O., 1986: Viroid processing: A model involving the central conserved region and hairpin I. Proc. Natl. Acad. Sci. U.S.A., **83,** 58—62.

Dougherty, W. G., Hiebert, E., 1985: Molecular Plant Virology, **Vol. II** (Davies, J. W., ed.) 39—42. CRC Press, Bocas Raton, Florida.

Dougherty, W. G., Allison, R. F., Parks, T. D., Johnston, R., E., Feild, M. J., Armstrong, F. B., 1985: Nucleotide sequence at the 3′ terminus of pepper mottle

virus genomic RNA: evidence for an alternative mode of potyvirus capsid protein gene organization. Virology, **146**, 282—291.

Dreher, T. W., Bujarski, J. J., Hall, T. C., 1984: Mutant viral RNAs synthesised *in vitro* show altered aminoacylation and replicase template activities. Nature, **311**, 171—175.

Dunn, J. J., Studier, F. W., 1983: Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. J. Mol. Biol., **166**, 477—435.

Efstratiadis, A., Kafatos, F. C., Maxam, A. M., Maniatis, T., 1976: Enzymatic *in vitro* synthesis of globin genes. Cell, **7**, 279—288.

El Manna, H., Bruening, G., 1973: Polyadenylate sequences in the ribonucleic acids of cowpea mosaic virus. Virology, **56**, 198—206.

Emini, E. A., Leibowitz, J., Diamond, D. C., Bonin, Wimmer, E., 1984: Recombinants of mahoney and sabin strain poliovirus type 1: analysis if *in vitro* phenotypic markers and evidence that resistance to guadine maps in the non-structural proteins. Virology, **137**, 74—85.

Van Emmelo, H., Devos, R., Ysebaert, H., Fiers, W., 1982: Construction and characterisation of a plasmid containing a nearly full-size DNA copy of satellite tobacco necrosis virus RNA. J. Mol. Bio., **143**, 259—270.

Van Emmelo, J., Ameloot, P., Plaetinck, G., Fiers, W., 1984: Controlled synthesis of the coat protein of satellite tobacco necrosis virus in *Escherichia coli*. Virology, **136**, 32—40.

Van Emmelo, J., Ameloot, P., Fiers, W. (submitted for publication): Linker insertion mutagenesis of satellite tobacco necrosis virus genome and its expression in plants.

Fernow, K. H., 1967: Tomato as a test plant for detecting mild strains of potato spindle tuber viroid. Phytopathology, **57**, 1347—1352.

Fields, S., Winter, G., 1982: Nucleotide sequences of influenza virus segments 1 and 3 reveal mosaic structure of a small viral RNA segment. Cell, **28**, 303—313.

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

Franssen, H., Goldbach, R., 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 Research, **1**, 39—49.

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

Franssen, H., Leunissen, J., Goldbach, R., Lomonossoff, G., Zimmern, D., 1984c: Homologous sequences in non-structural proteins from cowpea mosaic virus and picorna viruses. EMBO J., **3**, 855—861.

French, R., Janda, M., Ahlquist, P., 1986: Bacterial gene inserted in an engineered RNA virus: Efficient expression in monocotyledonous plant cells. Science, **231**, 1294—1297.

Garcia, J. A., Hille, J., Vos, P., Goldbach, R., 1986: Transformation of cowpea *Vigna unguiculata* cells with a full-length DNA copy of cowpea mosaic virus M RNA. Plant Science, in press.

Goelet, P., Lomonossoff, G. P., Butler, P. J. G., Akam, M. E., Gait, N. J., Karn, J., 1982: Nucleotide sequence of tobacco mosaic virus RNA. Proc. Natl. Acad. Sci. U.S.A., **79**, 5818—5823.

Goldbach, R. W., Schilthuis, J. G., Van Kammen, A., 1981: Comparison of *in vitro*

translation of cowpea mosaic virus RNAs. Biochem. Biophys. Res. Commun. **99**, 89—94.

Goldbach, R. W., Van Kammen, A., 1985: Molecular Plant Virology, **Vol. II** (Davies, J. W., ed.) 83—120. CRC Press, Boca Raton, Florida.

Green, M., Maniatis, T., and Melton, D., 1983: Human i-globin pre-mRNA synthesised *in vitro* is accurately spliced in Xenopus oocyte nuclei. Cell, **32**, 681—694.

Grimsley, N., Hohn, B., Hohn, Th., Walden, R., 1986: "Agroinfection", an alterative route for viral infection of plants by using the Ti-plasmid. Proc. Natl. Acad. Sci. U.S.A., **83**, 3282—3286.

Harrison, B. D., Robinson, D. J., 1982: Genomic reconstitution and nucleic acid hybridisation as methods of identifying particle-deficient isolates of tobacco rattle virus in potato plan with stem-mottle disease. H. of Virol. Methods, **5**, 255—266.

Henco, K., Sänger, H. L., Riesner, D., 1979: Fine structure melting of viroids as studied by kinetic methods. Nucleic Acids Res., **6**, 3041—3059.

Van Kammen, A., 1985: The replication of plant virus RNA. Microbiological Sciences, **2**, 170—174.

Jackson, R. J., 1986: A detailed kinetic analysis of the *in vitro* synthesis and processing of encephalomyocarditis virus products. Virology, **149**, 114—127.

Keese, P., Symons, R. H., 1985: Domains in viroids: Evidence of intermolecular RNA rearrangements and their contribution to viroid evolution. Proc. Natl. Acad. Sci. U.S.A., **82**, 4582—4586.

Kiberstis, P. A., Hall, T. C., 1983: N-acetyl-tyrosine at the 3' end of bromo mosaic virus RNA has little effect on infectivity. J. Gen. Virol, **64**, 2073—2077.

King, A. M. Q., McCahon, D., Slade, W. R., Newman, J. W. I., 1982: Biochemical evidence of recombination within the unsegmented RNA genome of aphtovirus. J. Virol., **41**, 66—77.

Lane, L., Kaesberg, P., 1971: Biophysically homogenous bromegrass mosaic virus contains multiple genetic components. Nature, **232**, 40—43.

Lane, L., 1981: Handbook of plant virus infections and comparative diagnosis (Kurstak, E., ed.) 333—376. Elsevier North Holland Biomedical Press, Amsterdam.

Leary, J. L., Brigati D. J., Word, D. C., 1981: Rapid and sensitive colorimetric method for visualising biotin-labeled DNA probes hybridised to DNA or RNA immobilised on nitrocellulose: Bio-blots. Proc. Natl. Acad. Sci. U.S.A., **80**, 4045—4059.

Loesch-Fries, L. S., Hall, T. C., 1982: *In vivo* aminoacylation of bromo mosaic and barley strip mosaic virus RNAs. Nature, **298**, 771—773.

Lomonossoff, G., Shanks, M., 1983: The nucleotide sequence of cowpea mosaic virus B RNA. EMBO J., **2**, 2253—2258.

Maule, A. J., Hull, R., Donson, J., 1983: The application of spot hybridisation to the detection of DNA and RNA viruses in plant tissues. J. of Virol. Methods, **6**, 215—224.

Maxam, A. M., Gilbert, W., 1980: In: Methods in Enzymology, **65**, 499—560 (Grossman, L., Moldave, K., eds.) Academic Press, Inc., New York.

McDonnell, M. W., Simon, H. N., Studier, F. W., 1977: Analysis of restriction fragments of T7 DNA and determination of molecular weights in neutral and alkaline gels. J. Mol. Biol., **110**, 119—146.

McKinney, H. H., 1929: Mosaic diseases in the Canary Islands, West Africa and Gibraltar. J. Agric. Res., **39**, 557—566.

Melton, D. A., Krieg, P. A., Rebagliati, M. R., Maniatis, T., Zinn, K., Green, M. R., 1984: Efficient *in vitro* synthesis of biologically active RNA and RNA hybrid-

isation probes from plasmids containing a bacteriophage SP_6 promoter. Nucleic Acids Res., **12**, 7035—7056.
Meshi, T., Kiyanna, R., Ohno, T., Okada, Y., 1983: Nucleotide sequence of the coat protein cistron and the 3' non-coding region of cucumber green mottle mosaic virus (watermelon strain) RNA. Virology, **127**, 54—64.
Meshi, T., Ishikama, H., Watamabe, Y., Yamaya, H., Okada, Y., Sano, T., Shikata, E., 1985: The sequence necessary for the infectivity of hop stunt viroid cDNA clones. Mol. Gen. Genet., **200**, 199—206.
Meshi, T., Ishikawa, M., Motoyoshi, F., Semba, K., Okada, Y., 1986: *In vitro* transcription of infectious RNAs from full-length cDNAs of tobacco mosaic virus. Proc. Natl. Acad. Sci. U.S.A., **83**, 5043—5047.
Miller, W. A., Bujarski, J. J., Dreher, T. W., Hall, T. C., 1986: Minusstrand initiation by bromo mosaic virus replicase within the 3' tRNA-like structure of native and modified RNA templates. J. Mol. Biol., **187**, 537—546.
Miller, W. A., Hall, T. C., 1983: Use of micrococcal nuclease in the purification of highly template dependent RNA-dependent RNA polymerase from brome mosaic virus-infected barley. Virology, **125**, 236—241.
Miller, W. A., Dreher, T. A., Hall, T. C., (1985): Synthesis of bromo mosaic virus subgenomic RNA *in vitro* initiation on (−) sense genomic RNA. Nature, **313**, 68—70.
Mizutani, S., Colonno, R. J., 1985: *In vitro* synthesis of an infectious RNA from cDNC clones of human rhinovirus type 14. J. Virol., **56**, 628—632.
Murant, A. F., Mayo, M. A., 1982: Satellites of plant viruses. Ann. Rev. Phytopathol., **20**, 49—70.
Okayama, H., Berg, P., 1982: High-efficiency cloning of full-length cDNA. Mol. Cell. Biol., **2**, 161—170.
Ohno, T., Takamatsu, N., Meshi, T., Okado, Y., 1982: Hop shunt viroid: Molecular cloning and nucleotide sequence of the complementary cDNA copy. Nucleic Acids Res., **11**, 6180—6197.
Owens, R. A., Diener, T. O., 1981: Sensitive and rapid diagnosis of potato spindle tuber viroid disease by nucleic acid hybridization. Science, **213**, 670—672.
Owens, R. A., Diener, T. O., 1982: RNA intermediates in potato spindle tuber in viroid replication. Proc. Natl. Acad. Sci. U.S.A. **79**, 113—117.
Owens, R. A., Diener, T. O., 1984: Spot hybridisation for detection of viroids and viruses. In: Methods in Virology, vol. VII (Maramorosch, K., Koprowski, H., eds.) Academic Press Inc., New York.
Palukaitis, P., Rakowski, A. G., Alexander, D., McE, Symons, R. H., 1981: Rapid indexing of the sunblotch disease of avocados using a complementary DNA probe to avocado sunblotch viroid. Ann. Appl. Biol., **98**, 439—449.
Palukaitis, P., Cotts, S., Zaitlin, M., 1985: Detection and identification of viroids and viral nucleic acids by "dot-blot" hybridisation. Acta Horticulturae, **164**, 109—118.
Pelham, H. R. B., 1978: Translation of encephalomyocarditis virus RNA *in vitro* yields an active proteolytic processing enzyme. Eur. J. Biochem., **85**, 457—462.
Plotch, S., J., Bouloy, M., Ulmanen, I., Krug, R. H., 1981: A unique cap ($m^7GpppXm$)-dependent influenza virion endonuclease cleaves capped RNAs to generate the primer that initiate viral RNA transcription. Cell, **23**, 847—858.
Putnak, J. R., Philips, B. A., 1981: Picornaviral structure and assembly. Microbiol. Rev., **45**, 287—315.
Queen, C., 1983: A vector that uses phage signals for efficient synthesis of proteins in *Escherichia coli*. J. Mol. Appl. Genet., **2**, 1—10.

Racaniello, V. R., Baltimore, D., 1981: Cloned poliovirus complementary DNA is infectious in mammalian cells. Science, **214**, 916—919.

Retzel, E. F., Collet, M. S., Faras, A. J., 1980: Enzymatic synthesis of deoxyribonucleic acid by the avian retrovirus revers transcriptase *in vitro*: optimum conditions required for transcription of large ribonucleic acid templates. Biochemistry, **19**, 513—518.

Rezelman, G., Goldbach, R., Van Kammen, A., 1980: Expression of bottom component RNA of cowpea mosaic virus in cowpea protoplasts. J. Virol, **36**, 366—373.

Rogers, S. G., Bisaro, D. H., Horsch, R. B., Fraley, R. T., Hoffmann, N. L., Brand, L., Elmer, J. S., Lloyd, A. H., 1986: Tomato golden mosaic virus A component DNA replicates autonomously in transgenic plants. Cell, **45**, 593—600.

Sano, T., Uyeda, I., Shikata, E., Ohno, T., Okada, Y., 1984: Nucleotide sequence of cucumber pale fruit viroid: homology to hop stunt viroid. Nucleic Acids Res., **12**, 3427—3434.

Riesner, D., Henco, K., Rokol, U., Klotz, G., Kleinschmidt, A. K., Gross, H. J., Domdey, H., Sänger, H. L., 1979: Structure and structure formation of viroids. J. Mol. Biol., **133**, 85—15.

Sänger, H. L., 1984: In: The Microbe 1984: Part I Viruses, 281—334 (Mahy, B. W. J., Pattison, J. R., eds.) Cambridge University Press, Cambridge, U. K.

Schuhmacher, H., Randles, J. W., Riesner, d., 1983: a two-dimensional electrophoretic technique for the detection of circular viroids and virusoids. Anal. Biochem., **135**, 288—295.

Schnölzer, M., Haas, B., Ramm, K., Hoffman, H. and Sänger, H. L., 1985: Correlation between structure and pathogenicity of potato spindle tuber viroid (PSTV). EMBO J., **3**, 2181—2190.

Sela, I., Riechman, H., Weissbach, A., 1982: Comparison of dot molecular hybridisation and enzyme linked immunosorbent assay for detecting tobacco mosaic virus in plant tissue and protoplasts. Phytopathology, **74**, 385—389.

Sippel, A. E., 1973: Pruification and characterisation of adenosine triphosphate: ribonucleic acid adenyltransferase from *Escherichia coli*. Eur. J. Biochem., **37**, 31.

Stanley, J., Rottier, R., Davies, J. W., Zabel, P., 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.

Symons, R. H., 1985: Diagnostic approaches for the rapid and specific detection of plant viruses and viroids. In: Plant Microbe Interactions, Molecular and Genetic Perspectives , vol. I, 93—124 (Kosuge, T. and Nester, E. W., eds.). McMillan Publishing Co., New York.

Tabler, M., Sänger, H. L., 1984: Cloned single- and double-stranded DNA copies of potato spindle tuber viroid (PSTV) RNA and co-inoculated subgenomic DNA fragments are infectious. EMBO J., **3**, 3055—3062.

Tabor, S., Richardson, C. C., 1985, A bacteriophage T7 RNA polymerase/promotor system for controlled exclusive expression of specific genes. Proc. Natl. Acad. Sci. U.S.A., **82**, 1074—1078.

Taniguchi, T., Palmieri, M., Weissmann, C., 1978: QB DNA-containing hybrid plasmids giving rise to QB phage formation in the bacterial host. Nature, **274**, 223—228.

Vance, V. B., Beachy, R. N., 1984: Translation of soybean mosaic virus RNA *in vitro*: evidence of protein processing. Virology, **132**, 271—281.

Verver, J., Goldbach, R., Garcia, J., Vos, P., submitted for publication. *In vitro*

expression of a full-length DNA copy of cowpea mosaic virus B RNA: Identification of the B RNA-encoded 24-kilodalton protein as a viral protease.

Visvader, J. E., Symons, R. N., 1985: Eleven new sequence variants of citrus exocortis viroid and the correlation of sequence with pathogenicity. Nucleic Acids Res., **13**, 2907—2920.

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

Wellink, J., Rezelman, G., Goldbach, R., Beyreuther, K., 1986: Determination of the proteolytic processing sites in the polyprotein encoded by the bottom-component RNA of cowpea mosaic virus. J. Virol., **60**, 50—58.

Van der Werf, S., Bradley, J., Wimmer, E., Studier, F. W., Dunn, J. J., 1986: Synthesis of infectious poliovirus RNA by purified T7 RNA polymerase, Proc. Natl. Acad. Sci. U.S.A. **83**, 2330—2334.

Van Wezenbeek, P., Vos, P., Van Boom, J., Van Kammen, A., 1982: Molecular cloning and characterization of a complete DNA copy of potato spindle tuber viroid RNA. Nucleic Acids Res., **10**, 7947—7957.

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

Ysebaert, N., Van Emmelo, J., Fiers, W., 1980: Total nucleotide sequence of a nearly full-size DNA copy of satellite tobacco necrosis virus RNA. J. Mol. Biol. **143**, 273—287.

CHAPTER 3

STUDY OF THE GENETIC ORGANIZATION OF A PLANT VIRAL RNA GENOME BY IN VITRO EXPRESSION OF A FULL-LENGTH DNA COPY

Pieter Vos, Jan Verver, Peter van Wezenbeek,
Ab van Kammen and Rob Goldbach

(EMBO J. 3: 3049-3053, 1984)

The EMBO Journal vol.3 no.13 pp.3049 - 3053, 1984

Study of the genetic organisation of a plant viral RNA genome by *in vitro* expression of a full-length DNA copy

Pieter Vos, Jan Verver, Peter van Wezenbeek[1], Ab van Kammen and Rob Goldbach

Department of Molecular Biology, Agricultural University, De Dreijen 11, 6703 BC Wageningen, and [1]Department of Biochemistry, University of Nijmegen, Geert Grooteplein Noord 21, 6526 EZ Nijmegen, The Netherlands

Communicated by H.L. Sänger

The genetic approach for elucidating functions encoded by RNA plant viruses has been hampered by the lack of methods to select desired mutants following random mutagenesis. An alternative might be to copy RNA genomes into DNA and use methods for site-directed mutagenesis to modify specific regions of the DNA copy. Transcription of the DNA copy will subsequently produce viral RNA with desired mutations. We have constructed a full-length DNA copy of the smaller of the two cowpea mosaic virus (CPMV) RNAs, referred to as M RNA. The DNA copy was positioned downstream from the promoter of bacteriophage SP6 and using SP6 RNA polymerase, this copy and two derivatives of it containing a specific deletion and insertion, respectively, have been transcribed into RNA molecules which are efficiently translated in rabbit reticulocyte lysates. The results obtained show that the subsequent *in vitro* transcription and translation of DNA copies may be a powerful tool to unravel the genetic properties of viral RNA genomes.

Key words: cowpea mosaic virus/cDNA/site-directed mutagenesis/*in vitro* transcription/*in vitro* translation

Introduction

Cowpea mosaic virus (CPMV) is a plant virus with a bipartite, positive-stranded RNA genome that is translated into large polypeptides from which the functional proteins are derived by proteolytic cleavages (Rezelman *et al.*, 1980; Goldbach *et al.*, 1981; Franssen *et al.*, 1982, 1984a). Both *in vivo* and *in vitro* translation studies have resulted in detailed knowledge about the expression of the larger of the two genomic RNAs, the B RNA (5889 nucleotides) (Rezelman *et al.*, 1980; Goldbach *et al.*, 1980; Zabel *et al.*, 1982; Goldbach and Rezelman, 1983; Lomonossoff and Shanks, 1983; Franssen *et al.*, 1984a, 1984b). On the other hand, the expression of the smaller of the two RNAs, the M RNA, (3481 nucleotides) (Van Wezenbeek *et al.*, 1983) is poorly understood. In infected cowpea cells (*Vigna unguiculata*) the only M RNA-specific proteins identified are the two capsid proteins VP37 (37 000 daltons) and VP23 (23 000 daltons). Translation of M RNA *in vitro* yields two overlapping polypeptides of ~105 000 (105 K) and 95 000 (95 K) daltons (Figure 1), which have never been observed *in vivo*, presumably because they are rapidly processed into functional proteins (Pelham, 1979; Franssen *et al.*, 1982, 1984b; Huez *et al.*, 1983). The information available seems to indicate that these polypeptides are produced from two different initiation sites on M RNA (Franssen *et al.*, 1982; Van Wezenbeek *et al.*, 1983). They are recognised and cleaved by a B RNA-encoded protease resulting in cleavage products of 60 000 daltons (60 K), 58 000 daltons (58 K) and 48 000 daltons (48 K) (Figure 1) (Franssen *et al.*, 1982). Whereas the 60-K cleavage product has been shown to represent a common precursor to both capsid proteins (Franssen *et al.*, 1982), the functions of the 58-K and 48-K polypeptides are still unknown. To gain further insight into the expression of M RNA, we decided to construct a full-length cDNA copy of M RNA. This would enable us to study the genetic organisation of M RNA by expression of DNA copies modified by site-directed mutagenesis. Here we report the successfull *in vitro* expression of DNA copies by subsequent *in vitro* transcription and translation. Application of this approach to two mutagenised DNA copies definitely reveals the presence of two translation initiation sites on M RNA and has led to a further mapping of the proteolytic cleavage site used by the B RNA-encoded protease.

Results

Molecular cloning

For the construction of a full-length DNA copy we used a number of cDNA clones described previously (Van Wezenbeek *et al.*, 1983). Since various rearrangements have been shown to occur during cloning of cDNA from M RNA (Van Wezenbeek *et al.*, 1983), this approach was preferred to synthesising new full-length cDNA. Only two new M13 clones, denoted CM131RF and CM414RF were constructed, which contained the 3′ and 5′ terminal sequences of M RNA, respectively (see Figure 2). For the synthesis of double-

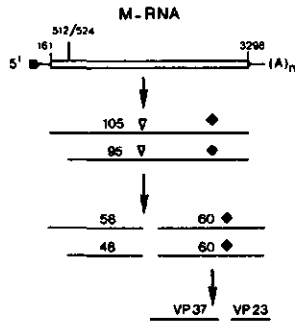


Fig. 1. Model for the translation of the M RNA of CPMV. The open reading frame is indicated with a double-lined bar. The positions of potential start codons (162, 512, 524) and stop codon (3298) are indicated. Open triangles correspond to the putative glutamine-methionine cleavage site recognised by the 32-K B RNA-encoded protease (Franssen *et al.*, 1982; Van Wezenbeek *et al.*, 1983). Black lozenges refer to the glutamine-glycine cleavage site in the 60-K capsid protein precursor, by which VP37 and VP23 are released (Van Wezenbeek *et al.*, 1983). VPg is indicated by a black square.

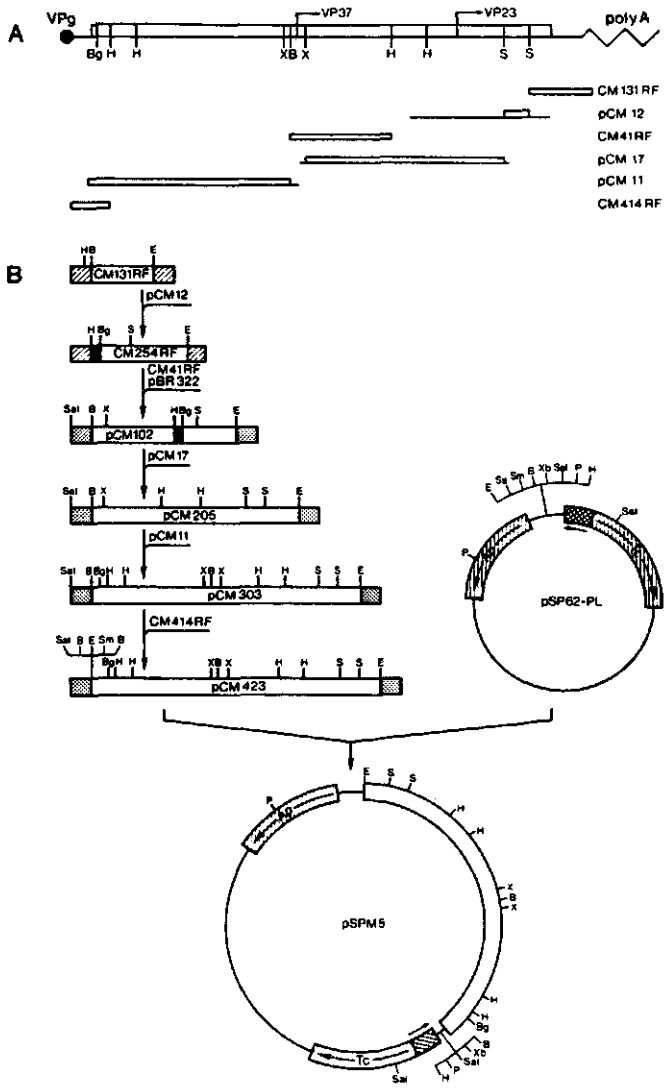


Fig. 2. Construction of a full-length DNA copy of CPMV M RNA. **Panel A:** schematic representation of the various cDNA clones used for the construction of the full-size DNA copy. The upper line represents the complete M RNA sequence with the double-lined bar indicating the position of the long open reading frame. The positions of the N-terminal ends of VP37 and VP23 are indicated. Prefix p refers to pBR322-derived clones, RF refers to M13mp9-derived clones. The positions of relevant recognition sites of *Bam*HI (B), *Bgl*I (Bg), *Eco*RI (E), *Hind*III (H), *Pst*I (P), *Sal*I (Sal), *Sau*3A (S), *Sma* (Sm), *Sst*I (Ss), *Xba*I (Xb) and *Xho*I (X) are shown. The regions of the M RNA-derived inserts in various clones are shown, with the segments used for the construction indicated by double-lined bars. M RNA-derived cDNA clones pCM11, pCM12, pCM17 and CM41RF have been described previously (Van Wezenbeek *et al.*, 1983). **Panel B:** construction of a full-size DNA copy from various cDNA clones. Sources of DNA fragments are indicated as follows: open boxes, M RNA-specific sequences; dotted boxes, adjacent pBR322 sequences; dashed boxes, adjacent M13 vector sequences; black box, *Bgl*I-*Hind*III adaptor fragment (see below); chequered box, SP6 promoter fragment. The figures are not drawn to scale. For explanation of symbols see **panel A**. For further details see text.

stranded cDNA containing the ultimate 5′ end, an oligonucleotide primer for second strand DNA synthesis homologous to the first 16 nucleotides of M RNA was used. Double-stranded cDNA obtained in this way was digested with *Hind*III and ligated into M13mp9 digested with *Hind*III and *Hind*II, producing clone CM327RF. The polylinker sequence of M13mp7 was subsequently inserted into the *Eco*RI site of CM327RF, resulting in clone CM414RF (Figure 2). In this way a number of restriction sites were positioned upstream of the 5′ end of the CPMV M cDNA facilitating manipulations with clone CM414RF. The DNA clone containing the 3′ end (CM131RF) was constructed by insertion

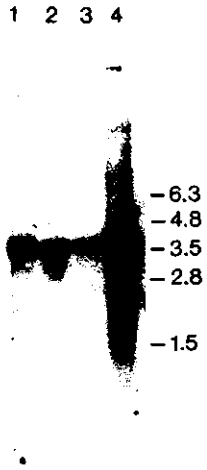


Fig. 3. Northern blot analysis of *in vitro* transcripts, obtained with phage SP6 RNA polymerase. RNA samples were glyoxylated and electrophoresed in a 1% agarose gel according to McMaster and Carmichael (1977) and transferred to nitrocellulose (Thomas, 1980). The blot was hybridised to nick-translated pCM423, washed as described (Thomas, 1980) and exposed to X-ray film (Sakura) for 16 h. **Lanes 1, 2** and **3** contain *in vitro* transcripts obtained from *Eco*RI-digested 0.05 μg pSPM5, pSPM5△*Xho*I and pSPM5△*Bgl*II respectively; **Lane 4**, 0.2 μg M RNA extracted from CPMV M components (Klootwijk *et al.*, 1977; Davies *et al.*, 1978). The numbers at the right side of the figure indicate the sizes (in kilobases) of glyoxylated marker DNA fragments derived from plasmid pCM423.

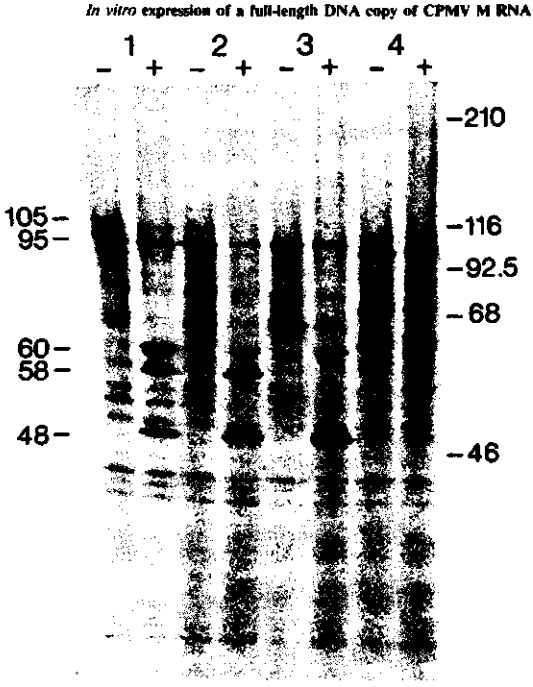


Fig. 4. *In vitro* translation and processing of transcripts obtained with SP6 polymerase. In each assay 0.5 μg of RNA was translated *in vitro* in 10 μl rabbit reticulocyte lysate using [^{35}S]methionine as radioactively labeled amino acid (Goldbach *et al.*, 1981). The *in vitro* translation products were incubated with an equal volume of 30 000 *g* supernatant fraction from B components inoculated protoplasts (+), or left untreated (−) and analysed on a 12.5% SDS-polyacrylamide gel. **Lane 1**, translation products from natural M-RNA; **lanes 2, 3** and **4** translation products from *in vitro* transcripts obtained with SP6 RNA polymerase from pSPM5, pSPM5△*Bgl*II and pSPM5△*Xho*I, respectively. The numbers at the right side of the gel indicate the sizes (in kd) of the marker proteins used: myosin, 210; *E. coli* β-galactosidase, 116; phosphorylase b, 92.5; bovine serum albumin, 68; ovalbumin,46. The numbers at the left side indicate the positions of the 105 and 95 primary translation products of M RNA and their proteolytic processing products (60, 58 and 48). The actual mol. masses of 105 and 95 are 116 and 102 kd, respectively (Van Wezenbeek *et al.*, 1983).

of a 360-bp DNA fragment, obtained upon digestion of double-stranded cDNA with *Sau*3A, into *Bam*HI and *Sma*I digested M13mp9. CM131RF contained 27 residues of the poly(A) tail and since a C residue was present adjacent to the *Sau*3A recognition sequence, the *Bam*HI site of M13mp9 was restored in this clone. The cDNA clones CM131RF, CM414RF and four previously described clones were used for the construction of a full-size cDNA clone of CPMV M RNA.

Construction of a full-length cDNA copy

The construction of a full-size DNA copy of M RNA is outlined in Figure 2. The first step in the construction was a three-point ligation of the large *Hind*III-*Bam*HI fragment of CM131RF, a 169-bp *Sau*3A fragment of pCM12 and an 83-bp *Hind*III-*Bgl*II adaptor fragment, derived from clone pCM11, resulting in clone CM254RF. As the *Sau*3A recognition sequence of the 169-bp fragment was flanked by an A residue the adaptor fragment obtained from pCM11 introduced a *Bgl*II site allowing unique cleavage next to the inserted *Sau*3A fragment in a later stage of the construction. Clone pCM102 was constructed by insertion of the 606-bp *Eco*RI-*Hind*III fragment of CM254RF and the 707-bp *Hind*III-*Bam*HI fragment of CM414RF into pBR322 digested with *Eco*RI and *Bam*HI. Insertion of the 1366-bp *Xho*I-*Sau*3A fragment of pCM17 into *Xho*I and *Bgl*II digested pCM102 resulted in clone pCM205, which now contained an M RNA-derived sequence from the *Bam*HI site at position 1504 downstream to the 3′ end. Next, the 1557-bp *Bam*HI fragment of pCM11 was ligated to pCM205, linearised with *Bam*HI, resulting in clone pCM303, which contained almost the entire cDNA except for the ultimate 5′ end. To construct a full-size clone, the *Sal*I-*Bgl*II fragment of CM414RF, containing the 5′ end, was ligated to *Sal*I and *Bgl*II digested pCM303, resulting in clone pCM423. Both restriction enzyme analysis and nucleotide sequence analysis revealed that clone pCM423 contained the complete sequence of M RNA including 27 residues of the 3′-terminal poly(A) tail.

Since bacteriophage SP6 RNA polymerase has been shown to be very useful for *in vitro* transcription (Kassavetis *et al.*, 1982; Green *et al.*, 1983), the full-length copy was positioned downstream of the SP6 promoter sequence in plasmid pSP62-PL. For this purpose the cDNA insert of pCM423, as obtained upon total digestion with *Eco*RI and partial digestion with *Bam*HI, was ligated to *Bam*HI, and *Eco*RI-digested pSP62-PL (a gift of J. Robertson). The resulting clone pSPM5 (Figure 2) and its derivatives were used for *in vitro* transcription experiments.

In vitro expression

For *in vitro* transcription plasmid pSPM5 was linearised with *Eco*RI to obtain run-off transcripts. Using SP6 RNA polymerase discrete transcripts of 3.6 kb were produced, which represented the complete plus-strand of M RNA but containing 78 additional nucleotides at the 5′ end, mainly derived from the leader sequence of the SP6 promoter fragment, and seven nucleotides added to the poly(A) tail, originating from the *Eco*RI linker (Figure 3). It should be noted that a further difference between the *in vitro* transcripts and virion M RNA is the absence of the genome-linked protein VPg at the 5′ terminus (Stanley *et al.*, 1978; Daubert *et al.*, 1978). Routinely 10 μg of RNA was obtained from 5 μg of linearised pSPM5, corresponding to an efficiency of 1000% since only 40% of the template DNA was available for the RNA polymerase and only one strand is used for transcription.

As shown in Figure 4 these molecules are efficiently translated in rabbit reticulocyte lysates, resulting in efficient production of the 105-K and 95-K precursor proteins also synthesised from natural virion RNA. The 105-K and 95-K polypeptides obtained from *in vitro* synthesised RNA were, moreover, faithfully cleaved by the B RNA-encoded protease into 60-K, 58-K and 48-K polypeptides (Franssen *et al.*, 1984b), which further underlined the complete co-linearity of the constructed DNA copy with M RNA (see Figure 4).

Site-directed mutagenesis

Having verified the integrity of the full-size DNA copy, the approach of successive *in vitro* transcription and translation was subsequently applied to DNA copies modified by site-directed mutagenesis. Two different mutant copies were constructed. One to test whether indeed two translation initiation sites are present on the M RNA, another to confirm the previous localisation of the recognition sequence for the B RNA encoded protease in the primary translation products of M RNA. The unique *Bgl*II site present in pSPM5 (at position 189 of M RNA, see Figure 2) was exploited to construct the first mutant. After digestion of pSPM5, the protruding ends of the linearised plasmid were filled in and re-ligated resulting in plasmid pSPM5△*Bgl*II. The result of the addition of four nucleotides in this mutant is a shift in the reading frame of M RNA by which a translational start at the first AUG codon at position 161 of the M RNA sequence is expected to be terminated 22 codons downstream at an UAA stop codon. The second mutant was made by deleting the 174-bp *Xho*I fragment, corresponding to the M RNA sequence between positions 1446 and 1620 (see Figure 2). This deletion does not alter the reading frame, but results in the loss of the coding sequence for 58 amino acids surrounding the glutamine-methione dipeptide sequence, proposed to be the cleavage site used for the release of the 60-K coat protein precursor (Van Wezenbeek *et al.*, 1983). Both modified DNA copies, linearised with *Eco*RI, were transcribed using SP6 polymerase and the transcripts obtained were translated in rabbit reticulocyte lysates. Compared with pSPM5 transcripts, RNA transcribed from pSPM5△*Bgl*II had the same length whereas, as expected, transcripts obtained from pSPM5△*Xho*I were slightly smaller (Figure 3). Upon translation of RNA obtained from plasmid pSPM5△*Bgl*II only the 95-K polypeptide was produced. The absence of synthesis of the 105-K polypeptide indicates that the AUG codon of position 161 represents the initiation codon for this polypeptide. Since the synthesis of the 95-K polypeptide remained unaffected translation must have initiated at the AUG codon at position 512 or 524 (Figure 1). Processing of the 95-K polypeptide produced from pSPM5△*Bgl*II, gave the 60-K coat protein precursor and the 48-K polypeptide, but not the 58-K polypeptide (Figure 4). The results obtained with mutant pSPM5△*Bgl*II definitely demonstrate that the 105-K and 95-K polypeptides are synthesised from two different start codons, and that the 95-K polypeptide is not derived from the 105-K polypeptide by proteolytic cleavage.

Translation of the RNA transcript from the second mutant clone, pSPM5△*Xho*I, resulted in the synthesis of two polypeptides with significantly lower mol. wt. compared with the products of pSPM5 and natural M RNA (Figure 4). This reduction in mol. wt. of ~7 kd conforms to the length of the deleted *Xho*I fragment. Neither of the two translation products of pSPM5*Xho*I were cleaved by the protease specified by B RNA, indicating that the amino acid sequence required for the release of the 60-K capsid protein precursor was encoded by the *Xho*I fragment. This fits the previous suggestion that the proteolytic cleavage site in the primary translation products is located between the glutamine and methionine residues corresponding to positions 1536–1541 of the M RNA.

Discussion

Here we report the successful *in vitro* expression of a full-size DNA copy of CPMV M RNA. Application of this method to DNA copies modified by site-directed mutagenesis has contributed to a further understanding of the genetic organisation of the M RNA. First, it has now been definitely demonstrated that the 105-K and 95-K polypeptides obtained upon *in vitro* translation of M RNA are synthesized from two different AUG codons. Secondly, the position of the cleavage site used by the B RNA-encoded protease to release the 60-K coat protein precursor from the 105-K and 95-K polypeptides has been determined.

As proposed by the scanning model of Kozak (1978, 1981) ribosomal subunits attach to the 5′ end of mRNAs and move downstream the leader sequence. AUG codons are preferentially recognised as initiator codons when a purine, A rather than G, is present at position −3 and a guanine is present at position +4. The AUG codon at position 161 of the M RNA, representing the initiator codon for the 105-K polypeptide, has an adenine residue at position −3, but a uracil residue at position +4 and only partially matches the consensus sequence for initiator codons. This may account for the fact that many ribosomes do not recognise this AUG triplet as initiator codon and start translation at a second 'in phase' initiator codon, presumably the AUG codon at position 512. This codon matches the consensus sequence very well, since guanine residues are present at positions −3 and +4. It may also explain why, upon *in vitro* translation of M RNA, synthesis of 95-K polypeptide is favoured over the synthesis of the 105-K polypeptide (see Figure 4 and Goldbach *et al.*, 1981; Franssen *et al.*, 1982). The occurrence of two different translation initiation sites on M RNA might be a general feature of comoviruses, since the M RNAs of all comoviruses tested so far are translated *in vitro* into two large polypeptides (Goldbach and Krijt, 1982). Furthermore, it is interesting to note that in the case of CPMV B RNA (Lomonossoff and Shanks, 1983), which specifies only one unique precursor polypeptide, the sequence surrounding the initiator AUG conforms perfectly to the 'Kozak rule'.

The large capsid protein VP37 has been shown to possess a N-acetylated methionine residue at its N terminus (Bruening, 1981). From previous work (Van Wezenbeek *et al.*, 1983) it has been proposed that the coding region of VP37 starts with the AUG codon of position 1538. This would lead however to a protein with a mol. mass of 41.2 kd, instead of 37 kd. Therefore the AUG triplet of position 1643 remained an attractive alternative, since it would give a protein with a mol. mass of ~37.2 kd. Our results with cDNA clone pSPM5△*Xho*I, in which the AUG codon of position 1538 has been removed and not that of position 1643, demonstrate however that the N-terminal amino acid of the 37-K protein is encoded by the AUG codon at position 1538. Recently analysis of the C-terminal amino acids of the 48-K polypeptide has confirmed the position of the cleavage site.

The results presented in this paper show that certain aspects of the genetic organisation and expression of viral RNA genomes can be studied by site-directed mutagenesis of full-length DNA copies and their subsequent transcription and translation *in vitro*. Finally, the genetic studies on the functions of the CPMV genome may even be further extended *in vivo* by transfection of cowpea cells with mutagenised RNA molecules obtained by the procedure described here.

Materials and methods

Nucleic acids and enzymes

CPMV was grown in *Vigna unguiculata L.*, 'California Blackeye' and M components were purified as described previously (Klootwijk *et al.*, 1977). CPMV RNA molecules were extracted from separated components as described by Davies *et al.* (1978). Plasmid DNA was isolated as described by Godson and Vapnek (1973) or by the method of Klein *et al.* (1980). Sources of enzymes were as described previously (Van Wezenbeek *et al.*, 1983). The 16-mer d(TATTAAAATCTTAATA) was synthesized by the chemical phosphotriester method (Van der Marel *et al.*, 1982).

Molecular cloning

Double-stranded cDNA was synthesised as described previously (Van Wezenbeek *et al.*, 1983). Conditions used for ligation of DNA fragments and subsequent transformation of competent *Escherichia coli* cells have been described (Van Wezenbeek *et al.*, 1983). DNA fragments used for ligation were purified either by extraction from low-melting agarose (Weislander, 1979) or by elution from polyacrylamide gel slices (Maxam and Gilbert, 1977). Recombinant DNA clones were characterised by restriction enzyme analysis and/or nucleotide sequencing (Sanger *et al.*, 1977, 1980).

In vitro transcription

In vitro transcriptions were carried out in a volume of 100 μl containing 40 mM Tris-HCl pH 7.5, 6 mM $MgCl_2$, 10 mM dithiothreitol, 4 mM spermidine, 0.5 mM of each of four NTPs, 5 μg template DNA, 15 units of SP6 polymerase (New England Nuclear Corp.) and 50 units of RNasin (Biotec Inc.). After a 30 min incubation at 37°C, 5 μg of DNase I (Worthington), freed of contaminating RNase activity by absorption to macaloid (Maniatis *et al.*, 1982), was added and incubation was continued for a further 15 min at 37°C. The reaction mixture was extracted once with phenol/chloroform, passed through Sephadex G-50 and precipitated with ethanol.

In vitro translation

M RNA extracted from CPMV M components and *in vitro* transcripts were translated in rabbit reticulocyte lysate using [^{35}S]methionine as radioactive amino acid as described previously (Goldbach *et al.*, 1981). Translation mixtures were incubated with an equal volume of 30 000 *g* supernatant fraction from cowpea protoplasts inoculated with B components as described by Franssen *et al.* (1982). *In vitro* translation products were analysed on a 12.5% SDS-polyacrylamide gel (Franssen *et al.*, 1982).

Acknowledgements

We thank Jacques van Boom for synthesis of the oligodeoxynucleotide, Peter Little and Jim Robertson for advise in the use of SP6 RNA polymerase and gifts of SP6 promoter fragments, Pim Zabel for critical reading of the text, Piet Madern for art work and photography and Mona van de Peppel for typing the manuscript. This work was supported by the Netherlands Foundation for Chemical Research (S.O.N.) with financial aid from the Netherlands Organisation for the Advancement of Pure Research (Z.W.O.).

References

Bruening,G. (1981) in Marcus,E. (ed.) *The Biochemistry of Plants*, Vol. 6, Academic Press, NY, pp.571-631.

Daubert,S.D., Bruening,G. and Najarian,R.C. (1978) *Eur. J. Biochem.*, **92**, 45-51.

Davies,J.W., Verver,J.W.G., Goldbach,R.W. and Van Kammen,A. (1978) *Nucleic Acid Res.*, **5**, 4643-4661.

Franssen,H., Goldbach,R., Broekhuijsen,M., Moerman,M. and Van Kammen,A. (1982) *J. Virol.*, **41**, 8-17.

Franssen,H., Goldbach,R. and Van Kammen,A. (1984a) *Virus Res.*, **1**, 39-49.

Franssen,H.J., Moerman,M., Rezelman,G. and Goldbach,R. (1984b) *J. Virol.*, **50**, 183-190.

Godson,G.N. and Vapnek,P. (1973) *Biochim. Biophys. Acta*, **299**, 516-520.

Goldbach,R., Rezelman,G. and Van Kammen,A. (1980) *Nature*, **286**, 297-300.

Goldbach,R.W., Schilthuis,J.G. and Rezelman,G. (1981) *Biochem. Biophys. Res. Commun.*, **99**, 89-94.

Goldbach,R. and Krijt,J. (1982) *J. Virol.*, **43**, 1151-1154.

Goldbach,R. and Rezelman,G. (1983) *J. Virol.*, **46**, 614-619.

Green,M.R., Maniatis,T. and Melton,D.A. (1983) *Cell*, **32**, 681-694.

Huez,G., Cleuter,Y., Bruck,C., Van Vloten-Doting,L., Goldbach,R. and Verduin,B. (1983) *Eur. J. Biochem.*, **130**, 205-209.

Kassavetis,G.A., Butler,E.T., Roulland,D. and Chamberlin,M.J. (1982) *J. Biol. Chem.*, **257**, 5779-5788.

Klein,R.D., Selsing,E. and Wells,R.D. (1980) *Plasmid*, **3**, 88-91.

Klootwijk,J., Klein,I., Zabel,P. and Van Kammen,A. (1977) *Cell*, **11**, 73-82.

Kozak,M. (1978) *Cell*, **15**, 1109-1123.

Kozak,M. (1981) *Nucleic Acids Res.*, **9**, 5233-5252.

Lomonossoff,G.P. and Shanks,M. (1983) *EMBO J.*, **2**, 2253-2258.

Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning, A Laboratory Manual*, published by Cold Spring Harbor Laboratory Press, NY.

Maxam,A.M. and Gilbert,W. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 560-564.

McMaster,G.K. and Carmichael,G.G. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 4835-4838.

Pelham,H.R.B. (1979) *Virology*, **96**, 463-477.

Rezelman,G., Goldbach,R. and Van Kammen,A. (1980) *J. Virol.*, **36**, 366-373.

Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463-5467.

Sanger,F., Coulson,A.R., Barrell,B.G., Smith,A.J.H. and Roe,B.A. (1980) *J. Mol. Biol.*, **143**, 161-178.

Stanley,J., Rottier,P., Davies,J.W., Zabel,P. and Van Kammen,A. (1978) *Nucleic Acids Res.*, **5**, 4505-4522.

Thomas,P.S. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 5201-5205.

Van der Marel,C.A., Wille,G., Marugg,J.E., De Vroom,E., Tromp,M., Van Boekel,C.A.A. and Van Boom,J.H. (1982) *Rec. Trav. Chim. Pays-Bas*, **10**, 239-241.

Van Wezenbeek,P., Verver,J., Harmsen,J., Vos,P. and Van Kammen,A. (1983) *EMBO J.*, **2**, 941-946.

Weislander,L. (1979) *Anal. Biochem.*, **98**, 305-309.

Zabel,P., Moerman,M., Van Straaten,F., Goldbach,R. and Van Kammen,A. *J. Virol.*, **41**, 1083-1088.

Received on 24 August 1984; revised on 2 October 1984

CHAPTER 4

IN VITRO EXPRESSION OF A FULL- LENGTH DNA COPY OF COWPEA MOSAIC VIRUS B RNA: IDENTIFICATION OF THE B RNA-ENCODED 24-KILODALTON PROTEIN AS A VIRAL PROTEASE

Jan Verver, Rob Goldbach, Juan A.Garcia and Pieter Vos

(EMBO Journal, 6, 549 - 554, 1987)

The EMBO Journal vol.6 no.3 pp.549–554, 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

Jan Verver, Rob Goldbach[1], Juan A.Garcia[2] and Pieter Vos

Department of Molecular Biology, Agricultural University, De Dreijen 11, 6703 BC Wageningen, [1]Department of Virology, Agricultural University, Binnenhaven 11, 6709 PD Wageningen, The Netherlands, and [2]Centro de Biologia Molecular, Universidad Autonoma de Madrid, 28049 Madrid, Spain

Communicated by A.-L.Haenni

Double-stranded cDNA was synthesized from B component RNA of cowpea mosaic virus and cloned into appropriate vectors. Using four clones, together representing the entire B RNA sequence, a full-length DNA copy was constructed and subsequently positioned downstream of a phage SP6 or T7 promoter. RNA molecules transcribed from this full-size DNA copy using SP6 or T7 RNA polymerase were efficiently translated in rabbit reticulocyte lysates into a 200-kd polypeptide similar to RNA isolated from viral B components. Moreover this polypeptide was rapidly cleaved into 32-kd and 170-kd polypeptides, exactly like the 200-kd polypeptide encoded by viral B RNA. *In vitro* transcription and translation of a DNA copy in which an 87-bp-long deletion in the coding sequence for the 24-kd polypeptide was introduced revealed that the 24-kd polypeptide bears the proteolytic activity involved in the primary cleavage of the B RNA-encoded polyprotein.

Key words: cowpea mosaic virus/cDNA/*in vitro* transcription/*in vitro* translation/viral protease

Introduction

Cowpea mosaic virus (CPMV) is a plant virus with a genome consisting of two separately encapsidated, positive-sense RNA molecules designated M and B RNA (for a recent review see Goldbach and Van Kammen, 1985). The viral RNAs are characterized by a small protein, denoted VPg, covalently linked to their 5'-end (Stanley *et al.*, 1978; Daubert *et al.*, 1978) and a poly(A)-tail at their 3'-end (El Manna and Bruening, 1973). They are both translated into large polypeptides ('polyproteins') from which the functional proteins are derived by proteolytic cleavages (Rezelman *et al.*, 1980; Goldbach *et al.*, 1981). In all these respects CPMV resembles the animal picornaviruses. The marked similarity in genome organization and the sequence homology among their non-structural proteins, moreover, suggest a genetic relationship (Franssen *et al.*, 1984a).

B RNA is translated both *in vivo* and *in vitro* into a 200-kd polyprotein, which is rapidly cleaved into a 170-kd and a 32-kd polypeptide. The 170-kd polypeptide is then further processed into polypeptides with sizes of 110, 87, 84, 60, 58, 24 and 4 kd (VPg) (Rezelman *et al.*, 1980; Goldbach and Rezelman, 1983; Franssen *et al.*, 1984b and Figure 1). These cleavage products have been mapped on the 200-kd polyprotein by sequencing of their amino-terminal ends (Wellink *et al.*, 1986; Zabel *et al.*, 1984). The results revealed that three different types of cleavage sites are employed during the processing of the B RNA-encoded polyprotein: glutamine–methionine (1x), glutamine–serine (2x) and glutamine–glycine (1x) dipeptide sequences.

Inhibition studies using antisera directed against various viral proteins strongly suggested that the 32-kd protein (see Figure 1) represents a protease involved in the processing of the unique glutamine–methionine cleavage site in the M RNA-encoded polyproteins (Franssen *et al.*, 1984c). Hence involvement of this protease in cleavage at the glutamine–methionine site of the B polyprotein is also very probable. Time course translations have provided indications that the 24-kd polypeptide encoded by a B RNA represents a second viral protease (Franssen *et al.*, 1984b). This suggestion was further supported by the sequence homology of this polypeptide to the picornaviral protease P-3C (Franssen *et al.*, 1984a). We have now constructed and cloned a full-length DNA copy of B RNA in order to introduce mutations on specific sites in this copy and to express mutagenized copies *in vitro* according to the approach described previously for M RNA (Vos *et al.*, 1984). Following this approach we have been able to demonstrate that the 24-kd protein encoded by B RNA indeed represents a protease involved in the primary cleavage of the 200-kd polyprotein.

Results

Construction and molecular cloning of a full-length DNA copy of B RNA

Double-stranded cDNA was synthesized from B RNA (Van Wezenbeek *et al.*, 1983) using an oligonucleotide homologous to the first 16 nucleotides of B RNA to prime second-strand synthesis. The DNA obtained was digested with suitable restriction enzymes at positions which could be deduced from the nucleotide sequence of B RNA (Lomonossoff and Shanks, 1983). In this

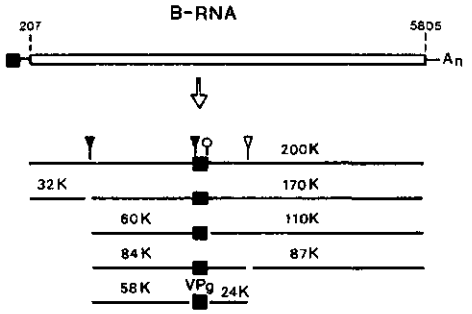


Fig. 1. Model for the translation of the B RNA of CPMV. The double-lined bar in the RNA refers to the long open reading frame running from start codon at position 207 to a stop codon at position 5805 (Lomonossoff and Shanks, 1983). VPg is indicated as a black square, other protein sequences as single lines. Cleavage sites are indicated by open circles (glutamine–methionine), open triangles (glutamine–glycine) or closed triangles (glutamine–serine) (Wellink *et al.*, 1986).

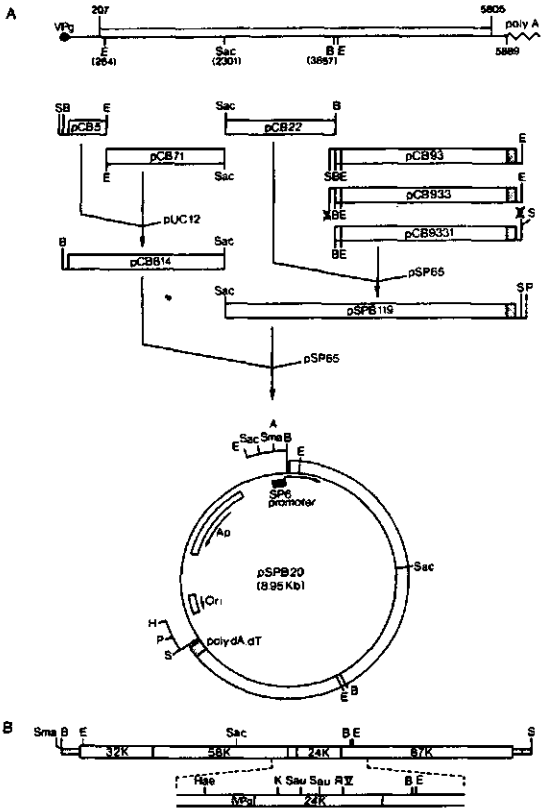


Fig. 2. Construction of a full-length DNA copy of CPMV B RNA. **Panel A:** cloning strategy for construction of a full-size DNA copy from four smaller cDNA clones. The upper line represents the complete B RNA sequence with the double-lined bar indicating the long open reading frame. The positions of relevant recognition sites of *BamHI* (B), *EcoRI* (E), *HindIII* (H), *PstI* (P), *SalI* (S), *SacI* (Sac) and *SmaI* (Sma) are shown. In the various cDNA clones obtained vector-derived sequences are indicated by a single line, B RNA-derived sequences by open boxes and the poly d(A-T) track of 55 residues as a dashed box. Double-stranded cDNA was synthesized using oligo(dT) as primer for first strand synthesis and an oligonucleotide homologous to the first 16 nucleotides of B RNA as primer for second strand synthesis. Digestion of the double-stranded cDNA with *BamHI* and *SacI* and subsequent ligation into pUC8 linearized with *SmaI* and *BamHI* and pUC12 digested with *SacI* and *BamHI* resulted in cDNA clones pCB93 and pCB22. Clone pCB93 contained a B RNA-specific insert from the *BamHI* site at position 3857 to the 3'-end including a poly d(A-T) track, cDNA clone pCB22 contained a viral cDNA sequence from the *SacI* recognition sequence at position 2301 to the *BamHI* site at position 3857. In addition, double-stranded cDNA was digested with *SacI* and *EcoRI* and ligated into pUC12 linearized with *SacI* and *EcoRI*. One of the cDNA clones obtained, pCB71, contained the *EcoRI*–*SacI* cDNA fragment, corresponding to positions 264–2301 of the B RNA sequence. To obtain a cDNA clone containing the ultimate 5'-end of the B RNA sequence, first strand synthesis was primed with a 388-bp *KpnI* fragment (positions 2746–3034), derived from cDNA clone pCB22. Double-stranded cDNA synthesized in this way was digested with *EcoRI* and subsequently ligated into *EcoRI*- and *SmaI*-digested pUC8 resulting in clone pCB5, which contained a cDNA insert corresponding to positions 1–264 of the B RNA sequence. cDNA clones pCB5, pCB71, pCB22 and pCB93, together representing the entire B RNA-sequence, were used to construct a full-size DNA copy. For this purpose, first a cDNA clone containing the first 2301 nucleotides of B RNA was constructed (pCB814) by joining the cDNA inserts of clone pCB5, obtained upon digestion with *SalI* and *EcoRI*, and clone pCB71, obtained upon digestion with *EcoRI* and *SacI*, via a three-point ligation into pUC12 linearized with *SalI* and *SacI*. A unique *SalI* site was introduced in clone pCB93 downstream of the poly d(A-T) track. Firstly, the *SalI* site already present in this construct was removed by digesting the DNA with *SalI* and subsequent religation after filling-in the recessed ends by the Klenow fragment of *E. coli* DNA polymerase I. Secondly, the resulting clone pCB933 was partially digested with *EcoRI*, made blunt-end with S1 nuclease and the linearized DNA, obtained by digestion of one of the two *EcoRI* sites eluted from the gel (Tautz and Renz, 1984). Finally *SalI* linkers were coupled to the DNA, cohesive ends were generated by digestion with *SalI* and the DNA was religated. One of the clones obtained in this way, pCB9331, contained a unique *SalI* site downstream of the poly d(A-T) track as intended. The *BamHI*–*SalI* fragment from this clone and the *SacI*–*BamHI* fragment from clone pCB22 were joined via a three-point ligation with *SacI* and *SalI* digested pSP65 to produce clone pSPB119, containing a cDNA copy corresponding to the 3'-end of B RNA from position 2301 downstream. Finally a full-size cDNA clone was constructed, denoted pSPB20, by joining of the *BamHI*–*SacI* fragment of pCB814 with the *SacI*–*PstI* fragment of pSPB119 in a three-point ligation with *BamHI*- and *PstI*-digested pSP65. pSPB20 has a total length of 8.95 kb and the B RNA-specific cDNA can be excised from this plasmid with *SmaI* and *SalI* (see also panel B). This 5.95-kb *SmaI*–*SalI* fragment, containing the full-length DNA copy, was inserted into plasmid pT7-1 digested with *SmaI* and *SalI*, resulting in clone pTB3 (not shown). Using this clone B RNA-like transcripts could be synthesized initiated at the T7 promoter present in the construct. **Panel B:** schematic representation of the full-length DNA copy. The long open reading frame with the positions of various B RNA-specific proteins is indicated by a large open bar, other B RNA-specific sequences are indicated by double lines. The region in the DNA copy used in the construction of a mutant with a deletion in the coding region for the 24-kd polypeptide is presented enlarged (see Materials and methods). Relevant recognition sites for restriction enzymes are indicated: Hae (*HaeIII*), K (*KpnI*), Sau (*Sau3A*), RV (*EcoRV*), for other enzymes see panel A.

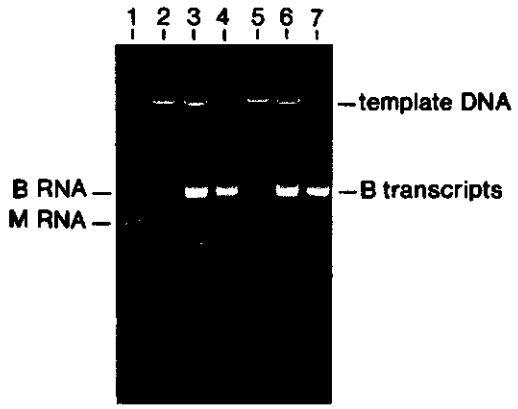


Fig. 3. Native 1.2% agrarose gel with reaction products of several *in vitro* translation assays using T7 RNA polymerase. On the right-hand side the positions of template DNA and RNA transcripts are shown. On the left-hand side the positions of M and B RNA isolated from virus particles are indicated. **Lane 1** shows M and B RNA isolated from virus particles. **Lanes 2** and **5** show linearized template DNA of pTB3 and pTB3Δ1 respectively. Reaction products of *in vitro* transcriptions with pTB3 and pTB3Δ1 are shown without DNase treatment (**lanes 3** and **6** respectively) or DNase treated (**lanes 4** and **7** respectively). Reaction conditions were as described in Materials and methods using 20 units T7 RNA polymerase per μg of template DNA. Template DNA was removed by DNase treatment as described previously (Vos *et al.*, 1984).

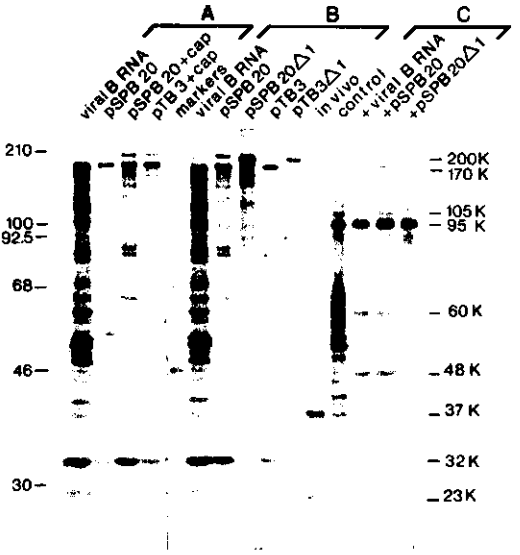


Fig. 4. *In vitro* translation products of transcripts obtained from various CPMV-specific cDNA clones using SP6 or T7 polymerase. Numbers at the left-hand side indicate the sizes (in kd) of marker proteins used: myosin, 210; phosphorylase b, 100 and 92.5; bovine serum albumin, 68; ovalbumin 46; carbonic anhydrase, 30. Numbers at the right-hand side correspond to the positions of relevant viral proteins encoded by natural M or B RNA. **Panel A:** *in vitro* translation products of viral B RNA and of transcripts synthesized by SP6 RNA polymerase from pSPB20 or by T7 RNA polymerase from pTB3, with (+ cap) or without a cap structure at the 5′ end. **Panel B:** *in vitro* translation products of viral B RNA and of transcripts synthesized by SP6 or T7 RNA polymerase using 'wild type' DNA copies as template DNA (pSPB20, pTB3) or DNA copies in which a deletion in the coding region for the 24-kd polypeptide was introduced (pSPB20Δ1 and pTB3Δ1). **Panel C:** *in vitro* translation products of viral M RNA, untreated (control) or incubated with unlabelled translation products of viral B RNA, or of transcripts from pSPB20 or pSB20Δ1. M RNA was translated in the presence of [^{35}S]methionine as described previously (Goldbach *et al.*, 1981) and after 1 h a 3 times excess of unlabelled translation products of viral B RNA or B RNA transcript was added and incubation continued at 30°C for another 5 h. In the case of the control sample, incubation was also continued for 5 h. The lane between panels B and C shows the [^{35}S]methionine-labelled proteins extracted from CPMV-infected protoplasts.

way a set of restriction fragments was generated from the cDNA synthesized which, after cloning in suitable vectors, represented the entire B RNA sequence. After verifying their integrity by restriction enzyme and/or sequence analysis the cDNA inserts of four different clones were selected and ligated to generate a full-length cDNA clone of B RNA. The strategy of this multi-step cloning procedure is outlined in Figure 2A.

The resulting clone pSPB20 had a total length of 8.95 kb with a unique *SalI*-site downstream of a poly d(A-T) track of 55 bp that could be utilized to synthesize run-off transcripts initiated at the SP6-promoter by phage SP6 RNA polymerase (Green *et al.*, 1983; Vos *et al.*, 1984). Alternatively the viral insert of clone pSPB20 was positioned downstream of a phage T7 promoter, resulting in clone pTB3, to allow *in vitro* transcription by RNA polymerase from phage T7 instead of SP6 (Figure 2).

In vitro expression

To generate run-off transcripts plasmids pSPB20 and pTB3 were linearized downstream of their poly d (A-T) track by digestion with *SalI* (Figure 2) and trancribed using SP6 or T7 RNA polymerase, repectively (Vos *et al.*, 1984; Davanloo *et al.*, 1984) (see Figure 3). Discrete transcripts of 6.0 kb were produced containing, in addition to B RNA-specific sequences, ~35 extra nucleotides to the 5′ end, derived from the SP6 or T7 promoter and multiple cloning site, and nine nucleotides at the 3′ end of the 55 residue-long poly(A) tail. A further difference with B RNA from virions was the absence at the 5′-end of the genome-linked protein VPg. Depending on the amount of SP6 or T7 polymerase added the efficiency of the *in vitro* transcription reactions could be increased, yielding up to 20 μg of RNA from 1 μg of template DNA. Occasionally transcripts were capped by addition of the cap precursor m^{7}GpppG to the transcription reaction (Konarska *et al.*, 1984). The efficiency of capping depended on the ratio of m^{7}GpppG to GTP in the reaction mixture. In the case of SP6 polymerase capping was very efficient giving 50% capped transcripts with equal amounts of GTP and m^{7}GpppG. For the T7 polymerase a four times molar excess of m^{7}GpppG to GTP had to be used to achieve 50% capping (results not shown).

The *in vitro* transcripts obtained were translated in rabbit reticulocyte lysates resulting in efficient production of a 200-kd protein, co-migrating with the 200-kd polyprotein produced from natural B RNA (Figure 4A). Moreover this 200-kd polypeptide was cleaved into two polypeptides with sizes of 170 and 32 kd, similar to the primary cleavage of B RNA-encoded 200-kd polyprotein (Figure 4A). These results show that the DNA copy contained a single long open reading frame similar to viral B RNA and that the translation product contains the proteolytic activity for cleavage of the polyprotein at the right position. No differences were observed between the *in vitro* translation products of *in vitro* RNAs synthesized by T7 or SP6 RNA polymerase. The presence of the cap structure at the 5′ end of

the *in vitro* transcripts did not significantly increase the translational activity of the transcripts, which is in contrast to results obtained with several *in vitro* synthesized eukaryotic mRNAs (Krieg and Melton, 1984). The minor differences in the amounts of translation products obtained (Figure 4A) are probably due to variations in RNA concentration in the *in vitro* translation reactions. The relatively large amounts of smaller translation products obtained with viral B RNA probably result from translation of degraded RNA molecules.

Assignment of proteolytic activity to the 24-kd protein encoded by B RNA

The availability of a full-length DNA copy of B RNA enabled us to study the genetic organization of B RNA by *in vitro* expression of DNA copies, modified by site-directed mutagenesis. One of the features of CPMV expression that lends itself well for *in vitro* analysis is the proteolytic processing of the primary translation products. This process occurs very reliably upon *in vitro* translation in rabbit reticulocyte lysates (Pelham, 1979; Franssen *et al.*, 1984b,c; Peng and Shih, 1984). The 24-kd polypeptide encoded by B RNA may have an important role in the proteolytic processing reactions since this protein shows sequence homology to the picornaviral protease P-3c (Franssen *et al.*, 1984a).

To test whether the 24-kd polypeptide indeed possesses proteolytic activity we introduced a small deletion in the coding region for this protein within the DNA copy of B RNA and determined possible defects in proteolytic cleavage of the translation product of this deletion mutant. For this purpose an 87-bp *Sau*3A fragment, corresponding to positions 3240–3327 of the B RNA sequence, was excised from the full-size DNA copy (Figure 2B and Materials and methods). This deletion did not change the reading frame, but resulted in the loss of the coding sequence for 29 amino acids of the 24-kd polypeptide. The resulting deletion mutant, denoted pSPB20Δ1 or pTB3Δ1 depending on the choice of the promoter, was subjected to subsequent *in vitro* transcription and translation after linearization with *Sal*I. Compared to the primary translation products of both viral B RNA and transcripts of clones pSPB20 or pTB3, the translation product of the deletion mutant was, as expected, slightly smaller (see Figure 4B). Furthermore this product did not undergo the faithful proteolytic processing as observed with the translation products of pSPB20 and pTB3 (figure 4B). This finding indicates that the 24-kd polypeptide indeed represents a proteolytic enzyme, involved in the primary cleavage of the 200-kd viral polyprotein, which occurs at a glutamine–serine pair (Wellink *et al.*, 1986).

The 32-kd protein released by the primary cleavage has been reported to represent also a viral protease, involved in the cleavage of the M RNA encoded polyproteins at a glutamine–methionine site (Franssen *et al.*, 1982, 1984c). We were interested in the effect of the deletion in the 24-kd protease, i.e. blocking of the release of the 32-kd protease, on the activity of this second protease. For this purpose *in vitro* translation products of viral B RNA or of transcripts from the cDNA clones pSPB20 and pSPB20Δ1, respectively, were mixed with the M RNA-encoded polyproteins and incubated for 5 h at 30°C (Figure 4C). Significant processing of the M RNA-encoded 95-kd polyprotein to polypeptides of 60 and 48 kd (Franssen *et al.*, 1982) was observed when *in vitro* translation products of viral B RNA or of transcript from clone pSPB20 were added, but upon addition of translation product from the mutant clone pSPB20Δ1 processing could not be detected (Figure 4C). A possible explanation for this effect of the B RNA-encoded 24-kd polypeptide on the processing of the primary translation products of M RNA at a glutamine–methionine site is discussed below.

Discussion

Various animal and plant RNA viruses use proteolytic processing to synthesize their proteins. Most of the current knowledge on viral protein cleavage reactions is derived from studies with picornaviruses (for recent review see Toyoda *et al.*, 1986; Nicklin *et al.*, 1986; Jackson, 1986) and CPMV (Goldbach and Van Kammen, 1985). Until recently it was generally believed that in the proteolytic processing of the picornaviral polyproteins, in addition to the viral protease 3C, a cellular enzyme was involved (Korant *et al.*, 1980; Burroughs *et al.*, 1984). By analogy to the picornaviruses a similar involvement of a host-encoded protease in the cleavage of the CPMV polyproteins may be surmized. In this paper we report the construction of a full-length cDNA clone of CPMV B RNA and its subsequent exploitation for studying the proteolytic processing of the CPMV translation products. The *in vitro* transcription and translation experiments described here demonstrate that the B RNA-encoded 24-kd polypeptide represents a viral protease, being involved in the primary cleavage of the 200-kd polyprotein, thereby excluding the possible involvement of a cellular enzyme in this process.

Bacteriophage SP6 RNA polymerase has been demonstrated to be very suitable for *in vitro* transcription of DNA templates (e.g. Green *et al.*, 1983; Melton *et al.*, 1984; Vos *et al.*, 1984; Tabler and Sänger, 1985). In addition to SP6 RNA polymerase we now also have used T7 RNA polymerase to synthesize RNA from templates provided with a T7 promoter fragment. RNA synthesis by T7 RNA polymerase also appeared to be very efficient, giving yields comparable with those of SP6 polymerase-dependent reactions and, furthermore, efficient capping of the RNA transcripts could be achieved similar to the SP6 polymerase (Konarska *et al.*, 1984). The advantage of the use of T7 instead of SP6 RNA polymerase in *in vitro* transcription assays may be the thorough understanding of the genetics of phage T7. Transcription of T7 genes under the control of T7 promoters has been extensively studied (see for instance Dunn and Studier, 1983). Moreover, plasmids are available containing the T7 RNA polymerase gene under the control of the λP_L promoter, permitting simple, large-scale purification of the enzyme from induced *Escherichia coli* cells (Tabor and Richardson, 1985; this paper).

Recently the proteolytic cleavage sites in the 200-kd polyprotein from B RNA have been identified (Wellink *et al.*, 1986). Three types of cleavage sites are used for processing, i.e. glutamine–serine, glutamine–glycine and glutamine–methionine dipeptide sequences. For processing of the M RNA-encoded polyproteins only two of these three different cleavage sites are used: a glutamine–methionine dipeptide sequence to release the 60-kd capsid protein precursor, and a glutamine–glycine dipeptide sequence to release the two capsid proteins from this precursor (Van Wezenbeek *et al.*, 1983; Vos *et al.*, 1984). Inhibition studies using antisera raised against various viral proteins and *in vitro* processing studies using partially purified proteins from B component-inoculated protoplasts strongly suggested that the 32-kd protein encoded by B RNA (see Figure 1) is involved in the cleavage of the M RNA-encoded polyproteins at the glutamine–methionine site (Franssen *et al.*, 1984b,c). If this protein indeed represents a protease recognizing this site then it is most probably also involved in cleavage of the single glutamine–

methionine site in the B polyprotein, which is situated at the left border of the 24-kd polypeptide sequence (Figure 1).

Since we have now demonstrated that the 24-kd polypeptide also possesses proteolytic activity, responsible for the primary cleavage of the 200-kd polyprotein at a glutamine–serine dipeptide sequence, the question arises which of the two proteases recognizes the glutamine–glycine cleavage sites. Comparison of the three types of cleavage sites used in the processing of the CPMV polyproteins shows that they fall into two categories. Although all cleavage sites share a glutamine residue at the first position of the dipeptide sequence the nature of the residue at the second position is quite variable, being either small and polar (glycine and serine) or large and non-polar (methionine). It is therefore tempting to assume that the 32-kd polypeptide recognizes only glutamine–methionine cleavage sites, while the 24-kd polypeptide may be involved in cleavage of both the glutamine–serine and glutamine–glycine cleavage sites. Indeed, preliminary experiments seem to indicate that the 24-kd polypeptide is also involved in the cleavage at a glutamine–glycine site in the M RNA-encoded polyproteins (data not shown).

Surprizingly the small deletion in the coding region of the 24-kd polypeptide also affected the processing of the M RNA-encoded polyproteins at a glutamine–methionine site, which is supposed to be recognized by the 32-kd protease (Franssen *et al.*, 1984c). While such processing was observed upon addition of translation products from clone pSPB20 this was clearly not the case with translation products from clone pSPB20Δ1. The most plausible explanation for this inhibitory effect is that release of the 32-kd polypeptide from the 200-kd polyprotein (by the 24-kd protease) is a prerequisite for the 32-kd polypeptide to become also proteolytically active. Clearly further experiments are necessary to confirm this latter hypothesis and to understand fully how the two different B RNA-encoded proteases co-operate in the release of all mature proteins from the viral polyproteins.

Materials and methods

Nucleic acids and enzymes

CPMV was propagated in *Vigna unguiculata* L., 'California Blackeye' and B components were purified as described previously (Kloetwijk *et al.*, 1977). CPMV B RNA was extracted from purified B components as described by Zimmern (1975) or Davies *et al.* (1978). Plasmid DNA isolations on large or small scale were performed by the alkaline lysis method (Birnboim and Doly, 1979). All enzymes used were purchased from Boehringer Mannheim B.V. or Anglian Biotechnology Corp. and were used according to the manufacturer's indications. T7 RNA polymerase was isolated as described below. Plasmid pT7-1 was purchased from United States Biochemical Corp. The 16-mer d(TATTAAAATCAATACA) was synthesized by the chemical phosphotriester method (Van der Marel *et al.*, 1982).

Molecular cloning

Double-stranded cDNA was synthesized from CPMV RNA as described previously (Van Wezenbeek *et al.*, 1983). DNA fragments were joined by T4 DNA ligase and transformed into competent *E. coli* HB101 cells as described by Van Wezenbeek *et al.* (1983). DNA fragments used for ligation were generally purified either by extraction from agarose gels (Tautz and Renz, 1984; Weislander, 1979) or by elution from polyacrylamide gel slices (Maxam and Gilbert, 1977). Recombinant DNA clones were analysed by restriction enzyme mapping and/or nucleotide sequencing (Sanger *et al.*, 1977, 1980).

Introduction of a deletion in the full-length DNA copy of B RNA

A *HaeIII–EcoRI* fragment (corresponding to positions 2815–3865 of the B RNA sequence) (see Figure 2B) was isolated from a full-length cDNA clone of B RNA (pSPB20) and inserted in plasmid pUC9 previously digested with *SmaI* and *EcoRI*. The resulting clone, pLB281, was cut with *KpnI* and *EcoRV*, both having unique recognition sites in this plasmid, yielding a small DNA fragment of 301 bp corresponding to positions 3134 (*KpnI*) to 3435 (*EcoRV*) of the B RNA sequence (see Figure 2B), together with a large fragment. The 301-bp fragment was subsequently digested with *Sau*3A resulting in three fragments of 106 bp (*KpnI*, 3134, to *Sau*3A, 3240), 87 bp (*Sau*3A, 3240, to *Sau*3A, 3326) and 108 bp (*Sau*3A, 3327, to *EcoRV*, 3435). In a three-point ligation the 106-bp *KpnI–Sau*3A fragment and 108-bp *Sau*3A–*EcoRV* fragment were joined together to the large *KpnI–EcoRV* fragment of pLB281. In this way a clone was constructed, pLB281ΔSau87, that contained a *HaeIII–EcoRI* fragment (corresponding to positions 2815–3865 of the B RNA sequence), missing the 87-bp *Sau*3A fragment (positions 3240–3327). To construct a full-length cDNA clone of B RNA containing this deletion, the *KpnI–EcoRI* fragment (positions 3134–3865) of clone pLB281ΔSau87 was exchanged for the corresponding fragment of pSPB20, yielding clone pSPB20Δ1. The *SmaI–SalI* fragment of pSPB20Δ1, containing the full-size DNA copy of B RNA with the desired deletion was also positioned downstream of the T7 promoter in plasmid pT7-1 in a similar way as decribed for clone pTB3 (see Results and Figure 2) yielding clone pTB3Δ1.

Purification of T7 RNA polymerase

As source for the purification of T7 RNA polymerase plasmid pT7P2 was used in *E. coli* strain NF1 (Bernard *et al.*, 1979). Isolation was according to Tabor and Richardson (1985) with some modifications. Plasmid pT7P2 contained a *BamHI–EcoRI* fragment from the low copy number plasmid pGP1 (a kind gift of Stanley Tabor; described in Tabor and Richardson, 1985) cloned into the high copy number plasmid pUC8. This fragment contained the T7 RNA polymerase gene under control of the λP_L promoter. Due to a temperature-sensitive cI repressor gene in bacterial strain NF1 (Bernard *et al.*, 1979) synthesis of T7 polymerase can be triggered by raising the temperature from 28 to 42°C.

Cells were grown at 28°C until an OD_{600} of 0.4 was reached and then incubated for 90 min at a temperature of 42°C. Cells were collected by centrifugation and lysed according to the procedure of Butler and Chamberlin (1982). The lysate was centrifuged for 30 min at 20 000 *g* and $(NH_4)_2SO_4$ was added to the supernatant to a saturation level of 30%. The $(NH_4)_2SO_4$ suspension was centrifuged for 30 min at 20 000 *g* and the pellet fraction containing most of the T7 RNA polymerase was further purified as described by Tabor and Richardson (1985).

In vitro transcription and translation

In vitro transcription using SP6 polymerase was performed as described previously (Vos *et al.*, 1984). *In vitro* transcription with T7 RNA polymerase was carried out in 100 μl containing 40 mM Tris–HCl pH 8.0, 20 mM $MgCl_2$, 10 mM DTT, 100 μl/ml BSA, 1 mM of each four NTPs, 5 μg linearized template DNA, 5–50 U of T7 polymerase and 50 U of RNasin. To synthesize capped transcripts the GTP concentration was lowered to 0.2 mM and the cap precursor m^7GpppG was added to a final concentration of 2 mM.

Aliquots of the transcription mixture were translated directly, without any purification, in rabbit reticulocyte lysates using [^{35}S]methionine as radioactive amino acid as described previously (Goldbach *et al.*, 1981). *In vitro* translation products were analysed on a 12.5% SDS–polyacrylamide gel (Franssen *et al.*, 1982).

Acknowledgements

We thank Jacques van Boom for synthesis of the oligodeoxynucleotide, Richard Jackson for the gift of the rabbit reticulocyte lysates, Ab van Kammen for critical reading of the text, Piet Madern for artwork and photography and Gré Heitkönig for typing the manuscript. This work was supported by The Netherlands Foundation for Chemical Research (SON) with financial aid from The Netherlands Organization for the Advancement of Pure Research (ZWO).

References

Bernard,H.U., Remaut,E., Vickers Hersfield,M.V., Das,H.K., Helinski,D.R., Yanofsky,C. and Franklin,N. (1979) *Gene*, **5**, 59–76.
Birnboim,H.C. and Doly,J. (1979) *Nucleic Acids Res.*, **7**, 1513–1523.
Burroughs,J.N., Sanger,D.V., Clarke,B.E., Rowlands,D.J., Billian,A. and Collen,D. (1984) *J. Virol.*, **50**, 878–883.
Butler,E.T. and Chamberlin,M.J. (1982) *J. Biol. Chem.*, **257**, 5772–5778.
Daubert,S.D., Bruening,G. and Najarian,R.C. (1978) *Eur. J. Biochem.*, **92**, 45–51.
Davanloo,P., Rosenberg,A.H., Dunn,J.J. and Studier,F.W. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 2035–2039.
Davies,J.W., Verver,J.W.G., Goldbach,R.W. and Van Kammen,A. (1978) *Nucleic Acids Res.*, **5**, 4643–4661.
Dunn,J.J. and Studier,F.W. (1983) *J. Mol. Biol.*, **166**, 477–535.
El Manna,M.M. and Bruening,G. (1973) *Virology*, **56**, 198–206.
Franssen,H., Goldbach,R., Broekhuijsen,M., Moerman,M. and Van Kammen,A. (1982) *J. Virol.*, **41**, 8–17.
Franssen,H., Leunissen,J., Goldbach,R., Lomonossoff,G. and Zimmern,D. (1984a) *EMBO J.*, **3**, 855–861.
Franssen,H., Goldbach,R. and Van Kammen,A. (1984b) *Virus Res.*, **1**, 39–49.
Franssen,H., Moerman,M., Rezelman,G. and Goldbach,R. (1984c) *J. Virol.*, **50**, 183–190.

Goldbach,R. and Rezelman,G. (1983) *J. Virol.*, **46**, 614–619.
Goldbach,R.W. and Van Kammen,A. (1985) In Davies,J.W. (ed.) *Molecular Plant Virology*. CRC Press, Boca Raton, FL, Vol. II, pp. 83–120.
Goldbach,R.W., Schilthuis,J.G. and Rezelman,G. (1981) *Biochem. Biophys. Res. Commun.*, **99**, 89–94.
Green,M.R., Maniatis,T. and Melton,D.A. (1983) *Cell*, **432**, 681–694.
Jackson,R.J. (1986) *Virology*, **149**, 114–127.
Klootwijk,J., Klein,I., Zabel,P. and Van Kammen,A. (1977) *Cell*, **11**, 73–82.
Konarska,M.M., Padgett,R.A. and Sharp,P.A. (1984) *Cell.*, **38**, 731–736.
Korant,B.D., Chow,N.L., Lively,M.O. and Powers,J.C. (1980) *Ann. NY Acad. Sci.*, **343**, 304–318.
Krieg,P.A. and Melton,D.A. (1984) *Nucleic Acids Res.*, **11**, 7057–7070.
Lomonossoff,G.P. and Shanks,M. (1983) *EMBO J.*, **2**, 2253–2258.
Van der Marel,C.A., Wilke,G., Maagg,J.E., De Vroom,E., Tromp,M., Van Boekel,C.A.A. and Van Boom,J.H. (1982) *Rec. Trav. Chim. Pays-Bas*, **10**, 239–241.
Maxam,A.M. and Gilbert,W. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 560–564.
Melton,D.A., Krieg,P.A., Rebagliati,M.R., Maniatis,T., Zinn,K. and Green,M.R. (1984) *Nucleic Acids Res.*, **12**, 7035–7056.
Nicklin,M.J.H., Toyoda,H., Murray,M.G. and Wimmer,E. (1986) *Biotechnology*, **4**, 33–42.
Pelham,H.R.B. (1979) *Virology*, **96**, 463–477.
Peng,X.X. and Shih,D.S. (1984) *J. Biol. Chem.*, **259**, 3197–3201.
Rezelman,G., Goldbach,R. and Van Kammen,A. (1980) *J. Virol.*, **36**, 366–373.
Sanger,F., Nickler,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
Sanger,F., Coulson,A.R., Barrel,B.G., Smith,A.J.H. and Roe,B.A. (1980) *J. Mol. Biol.*, **143**, 161–178.
Stanley,J., Rottier,P., Davies,J.W., Zabel,P. and Van Kammen,A. (1978) *Nucleic Acids Res.*, **5**, 4505–4522.
Tabler,M. and Sänger,H.L. (1985) *EMBO J.*, **4**, 2191–2199.
Tabor,S. and Richardson,C.C. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 1074–1078.
Tautz,D. and Renz,M. (1983) *Anal. Biochem.*, **132**, 14–19.
Toyoda,H., Nicklin,M.J.H., Murray,M.G. and Wimmer,E. (1986) In Inouye,M. and Sarma,M. (eds), *Protein Engineering: Application in Science, Medicine and Industry*. Academic Press, New York, in press.
Vos,P., Verver,J., Van Wezenbeek,P., Van Kammen,A. and Goldbach,R. (1984) *EMBO J.*, **3**, 3049–3053.
Weislander,L. (1979) *Anal. Biochem.*, **98**, 305–309.
Wellink,J., Rezelman,G., Goldbach,R. and Beyreuther,K. (1986) *J. Virol.*, **59**, 50–58.
Van Wezenbeek,P., Verver,J., Harmsen,J., Vos,P. and Van Kammen,A. (1983) *EMBO J.*, **2**, 941–946.
Zabel,P., Moerman,M., Lomonossoff,G., Shanks,M. and Beyreuther,K. (1984) *EMBO J.*, **3**, 1629–1634.
Zimmern,D. (1975) *Nucleic Acids Res.*, **2**, 1189–1201.

Received on November 6, 1986; revised on December 18, 1986

CHAPTER 5

INFECTIOUS RNA TRANSCRIPTS DERIVED FROM FULL- LENGTH DNA COPIES OF THE GENOMIC RNAs OF COWPEA MOSAIC VIRUS

Pieter Vos, Martine Jaegle, Joan Wellink, Jan Verver, Rik Eggen,
Ab van Kammen and Rob Goldbach

(submitted for publication)

ABSTRACT

A set of full-length DNA copies of both M and B RNA of cowpea mosaic virus (CPMV) were cloned downstream of a phage T7 promoter. Upon *in vitro* transcription using T7 RNA polymerase M or B RNA-like transcripts were obtained from these DNA copies with only two additional non-viral nucleotides at the 5' end and five non-viral nucleotides to the 3' end. In cowpea protoplasts the transcripts of several cDNA clones of B RNA were able to replicate leading to detectable synthesis of viral RNA and proteins. Transcripts of M cDNA clones inoculated together with these B RNA transcripts were also expressed, although the number of protoplasts, in which both transcripts were expressed, was very low. Preliminary infectivity tests with mutagenized RNA transcripts indicate essential roles of the B RNA encoded 24K and 32K polypeptides in viral RNA replication.

INTRODUCTION

Cowpea mosaic virus has a genome consisting of two messenger sense RNA molecules referred to as B RNA and M RNA (for recent reviews see e.g. Goldbach and Van Kammen, 1985 and Goldbach, 1987). Both RNAs are required for infectivity in plants but B RNA can replicate independently in cowpea protoplasts (Goldbach et al., 1980). The RNAs have a small protein, VPg, covalently linked to their 5' ends and a poly(A) tail at their 3' ends. M and B RNA are translated into large polyproteins, which are subsequently cleaved by a virus encoded proteinase into smaller functional polypeptides (Vos et al., submitted). In replication and expression as well as in genome structure and organisation CPMV resembles the animal picornaviruses. There is also significant amino acid sequence homology among several proteins encoded by picornaviruses and CPMV (Franssen et al., 1984a), suggesting genetic relationship.

The lack of well-defined mutants of RNA viruses has impeded studies on the molecular biology of these pathogens. However the recent development of powerful in vitro transcription systems (Contreras et al., 1982; Green et al., 1983; Davanloo et al., 1984) now enables generation of such mutants indirectly by transcription in vitro of DNA copies modified by site-directed mutagenesis. Such RNA transcripts are particularly useful if they could be expressed in vivo, i.e. are replicated upon inoculation of host cells. For several animal, insect and plant RNA viruses, the construction of cDNA clones from which infectious transcripts could be synthesized has been reported (Ahlquist et al., 1984; Mizutani and Colonno, 1985; Dasmahapatra et al., 1986; Dawson et al., 1986; Meshi et al., 1986; Van der Werf et al., 1986). For CPMV, full-length DNA copies of both genomic RNAs have been constructed downstream of phage SP6 and T7 promoters (Vos et al., 1984; Verver et al., 1987) allowing in vitro synthesis of RNA transcripts resembling the viral RNAs. Although these transcripts directed the synthesis of CPMV-specific proteins in vitro (Vos et al., 1984; Verver et al., 1987), they were not able to infect cowpea plants or protoplasts. In this paper we report the construction and molecular cloning of a series of full-length DNA copies of both M and B RNA of CPMV in transcription vector pX1. We show that transcripts synthesized from these new cDNA clones, which have

only few additional non-viral nucleotides at both the 5' and 3' end are replicated and expressed in cowpea protoplasts.

RESULTS

Molecular cloning

For molecular cloning of full-length DNA copies we employed vector pX1, which was specially designed for this purpose (see Figure 1a). This vector is derived from the T7-based transcription vector pT7-5 (a kind gift of Stanley Tabor) and has an unique StuI restriction site at the transcription initiation site in the T7 promoter sequence. Any DNA fragment inserted into this StuI site by blunt end ligation can be transcribed using T7 RNA polymerase into RNA molecules with only two additional G-residues at the 5' end.

Initially the approach used by Ahlquist and Janda (1984) for cloning of the genomic RNAs of BMV was followed for the construction of full-size DNA copies of M and B RNA in vector pX1, i.e. synthesis of double-stranded cDNA with specific primers for first and second strand synthesis and blunt end ligation of the full-size double-stranded cDNA, isolated from preparative gels, to specific vectors. This approach did not result in full-length clones of B RNA, while for M RNA several full-size clones were obtained but these were all orientated in the (-) sense direction with respect to the T7 promoter sequence presumably due to a strong orientation preference for M cDNA in plasmid pX1.

In view of these results a different method for cloning of the viral RNAs was employed based on the procedure described by Okayama and Berg (1982). Plasmid pX1 was linearised with HindIII, blunt-ended with S1-nuclease and provided with oligo d(T) tails of 50-100 residues (Okayama and Berg, 1982). The tailed plasmid DNA was digested with StuI to remove one of the d(T) tails and subsequently used as plasmid primer for first strand synthesis.

The second strand was synthesized using either Klenow polymerase and a specific primer after removal of the template RNA (Vos et al., 1984) or a

combination of E. coli DNA polymerase I and RNase H (Okayama and Berg, 1982). After second strand synthesis the linear double-stranded cDNA linked to the plasmid DNA was circularised with T4 DNA ligase and transformed into competent E. coli HB101 cells. The advantage of this procedure is that the cDNA is forced into the desired orientation and that the ligation step involves only circularization of a linear blunt ended molecule which warrants efficient cloning. An extra advantage in case of CPMV was that the minimal length of the poly d(A-T) track in the clones could be determined by the length of the oligo d(T) tail linked to the plasmid-primer, which made it possible to construct cDNA clones with long poly d(A-T) tracks.

Using this cloning procedure we were indeed able to construct very long cDNA clones of both M and B RNA in the desired orientation with poly d(A-T) tracks ranging from 60 to more than 200 residues. For M RNA several clones with the complete sequence except for a few nucleotides at the ultimate 5' end were obtained but all cDNA clones of B RNA lacked approximately 1600 nucleotides of the 5' end probably due to a "strong stop" for the reverse transcriptase at this position. To supply the missing sequences of the cDNA of M RNA we constructed a set of 5' end clones containing the first 272 nucleotides of M RNA (see figure 1B). For B RNA we constructed two series of overlapping cDNA clones using primers for first strand synthesis homologous to regions located more to the 5' end of the RNA (see figure 1b). Together with the long 3' end clones we obtained in this way three different sets of overlapping cDNA clones representing the entire B RNA sequence.

Construction and selection of full-size cDNA clones

Full-length DNA copies were constructed from the partial cDNA clones in a single ligation step as outlined in Figure 1B. To minimize the possibility of introducing errors in the sequence DNA fragments used for the construction of full-size DNA copies were in all cases a mixture of molecules obtained from different cDNA clones. Finally 8 clones with full-length cDNA of M RNA (pTM201-208) and 14 clones with full-length cDNA of B RNA were selected (pTB101-114).

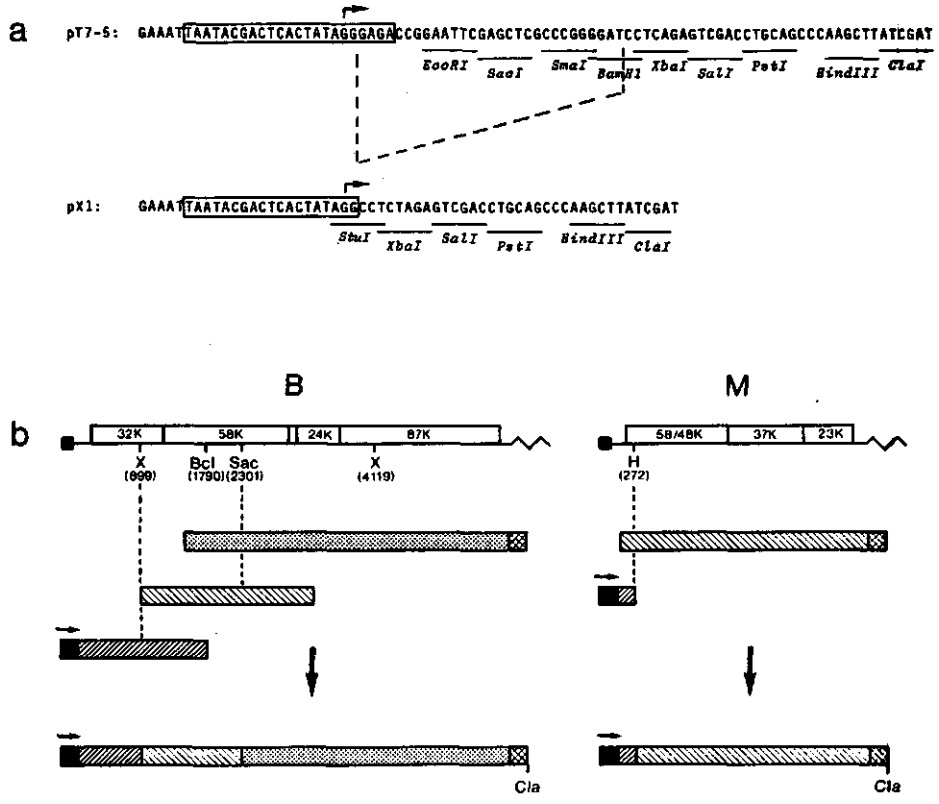


FIGURE LEGENDS

Fig. 1. (a) Sequences surrounding the transcription initiation sites in plasmids pT7-5 and pX1. The T7 promoter sequence (Dunn and Studier, 1983) is boxed with the transcription initiation site indicated by a bent arrow. Restriction sites in the polylinker sequence downstream of the T7 promoter are indicated. Dashed lines indicate the region in pT7-5 removed to construct plasmid pX1. A 312-bp RsaI-HindIII fragment of plasmid T7-5 (nucleotides -253 to +59 with respect to the transcription initiation site of the T7 promoter) was inserted into M13mp8 digested with SmaI and HindIII resulting in M13 clone, M13EG2. Single-stranded phage DNA was isolated from M13EG2 (Sanger et al., 1980) and annealed to 50 times excess of a 20-mer with the sequence 5'-ACTCTAGAGGCCTATAGTGA-3'. The complementary strand was synthesised according to Zoller and Smith (1982) and the reaction mixture was transfected into competent E.coli JM103 cells. Recombinant phages with the desired deletion were selected by hybridization with the oligonucleotide, used as primer for complementary strand synthesis (Zoller and Smith, 1982). Finally the desired deletion was transferred to plasmid pT7-5 resulting in plasmid pX1 by exchanging the normal 94-bp PvuII-HindIII-fragment (positions -35 to +59 with respect to the transcription initiation site of the T7 promoter) for the corresponding DNA fragment obtained from M13 clone M13EK2 containing this deletion. (b) Schematic representation of the construction of full-length cDNA clones of CPMV M and B RNA. On top M and B RNA are presented with the large open reading frames as open bars and the non-translated regions as single lines. VPg is indicated by a black square and at the 3' end of the RNAs the poly(A) tail is indicated. In the large open bars the positions of the coding regions for the various viral proteins are shown. Relevant positions of recognition sequences for restriction endonucleases used in the construction of full-size copies are shown: X = XbaI; Bcl = BclI; Sac = SacI;

H = HindIII; Cla = ClaI. The hatched and dotted bars indicate cDNA sequences in the various cDNA clones, with the cross hatched areas indicating the poly d(A-T) tracks. The black box represents the T7 promoter sequence with the direction of transcription shown by an arrow.
The cDNA clones of M and B RNA containing the poly d(A-T) tracks were constructed using a d(T) tailed plasmid primer as described in the results section. For M RNA cDNA clones containing the ultimate 5' end were obtained using oligo dT as a primer for first strand synthesis and an oligonucleotide homologous to nucleotides 1-16 of M RNA for second strand synthesis. Double stranded cDNA synthesized in this way was digested with HindIII and ligated into StuI and HindIII digested pX1. For B RNA cDNA clones having the ultimate 5' end were obtained using an oligonucleotide complementary to positions 2269 to 2306 of B RNA as primer for first strand synthesis and an oligonucleotide homologous to the first 16 nucleotides of B RNA as primer for second strand synthesis. Double-stranded cDNA obtained in this way was digested with BclI and ligated into StuI and BamHI digested pX2 (note: pX2 is identical to pX1, but contains a larger multiple cloning site with a unique BamHI site). Since the overlap between the 5' and 3' end clones of B RNA was rather small making it difficult to construct full-size copies, cDNA clones were constructed which overlapped with a large part of both the 5' and 3' end clones. Again cDNA was synthesized now using an oligonucleotide complementary to positions 3311-3330 of the B RNA as a primer for first strand synthesis. After second strand synthesis the double-stranded cDNA was digested with XbaI and ligated to XbaI and SmaIlinearized pUC12 resulting in cDNA clones containing the middle part of the B RNA sequence. To construct full-size cDNA clones of M RNA the 3.3 kb BglII (position 189)-ClaI (downstream the poly d(A-T) track) fragment of the 3' end clones was inserted into BglII and ClaI digested plasmid DNA of the clones containing the ultimate 5' end resulting in several cDNA clones with the entire M RNA sequence.
Full-size cDNA clones of B RNA were constructed via a three-point ligation using the XbaI-site (position 899), SacI-site (position 2301) and ClaI-site (downstream the poly d(A-T) track) as indicated in the figure, resulting in several cDNA clones having the complete sequence of B RNA.

The selected clones all had the following features. Firstly they contained a complete viral sequence linked to the T7 promoter in a way that upon transcription run-off RNA transcripts were obtained with only two extra nucleotides at the 5' end compared to the viral RNAs. It is noteworthy that in these clones the sequence downstream the transcription initiation site (positions +3 to +6) differs from the T7 promoter sequence reported to be essential for efficient transcription *in vivo* (i.e.: GGGAGA→GGTATT; Dunn and Studier, 1983). *In vitro* transcription by T7 RNA polymerase however was still very efficient. Secondly the poly d(A-T) track in the cDNA sequence of all clones was at least 60 residues long and therefore transcription resulted in RNA molecules with poly(A) tails of 60 nucleotides or more. At the 3' end of the poly(A) tail the RNA transcripts contained 4 to 5 non-viral nucleotides originating from the ClaI-site used to linearize the template DNA prior to transcription. The RNAs isolated from virions contain poly(A)tails of 50-400 residues with an average of 90 for B RNA and 150 for M RNA respectively (Ahlquist and Kaesberg, 1979). Finally as will be published elsewhere the RNA transcripts gave the same *in vitro* translation products as the viral RNA and these translation products were proteolytically processed exactly like the translation products of the viral RNAs (Vos *et al.*, submitted).

Transfection of protoplasts

RNA transcripts were synthesized from the full-length cDNA clones of B RNA using T7 RNA polymerase. 100 µg of RNA transcript of each cDNA clone was inoculated into cowpea protoplasts which were subsequently labelled with ^{35}S-methionine 15 hr post-infection and harvested at 40 hr post-infection. 30,000 x g supernatant fractions were prepared and analysed for the presence of B RNA encoded proteins by immunoprecipitation with antiserum directed against the B RNA encoded 24K polypeptide (Wellink et al., 1987).

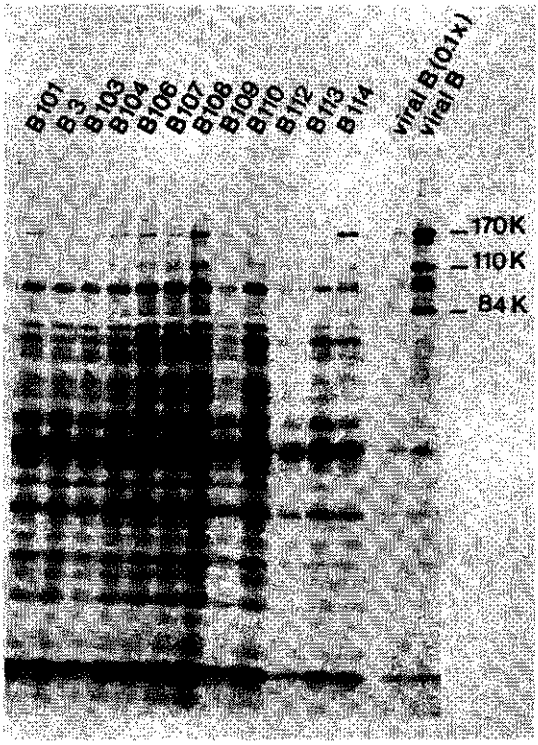


Fig. 2. Immunoprecipitation of proteins extracted from protoplasts transfected with RNA-transcripts of B cDNA clones or viral B RNA. 30.000 x g supernatant fractions of transfected protoplasts were prepared immunoprecipitated with antiserum directed against the B RNA encoded 24K polypeptide (Wellink et al., 1987) and analysed on a 12.5% SDS-polyacrylamide gel. Numbers at the top indicate the RNA transcripts used to transfect protoplasts, e.g. B101 corresponds to RNA transcripts derived from cDNA clone pTB101. For each experiment approximately 100 µg of transcripts was used to transfect 5 x 10^6 protoplast. Lanes at the right hand site of the gel correspond to protoplasts infected with 10 µg of viral B RNA, for which in one case the 30.000 x g supernatant fraction was diluted ten times with a 30.000 x g supernatant fraction of uninfected protoplast before immunoprecipitation. Numbers at the right hand site of the gel indicate the positions of the virus specific 170K, 110K and 84K polypeptides.

The 170K, 110K and 84K polypeptides, all containing the 24K polypeptide domain (Goldbach and Rezelman, 1983; Wellink et al., 1987), which are characteristically detected in protoplasts infected with viral B RNA were also found in protoplasts inoculated with transcripts of most of the B cDNA clones. The amount of these proteins in protoplasts inoculated with B cDNA transcripts, however, was much lower than in protoplasts infected with RNA isolated from B components (Figure 2). The 24K polypeptide itself was not detected, because it is only present in small amounts as a free protein (Wellink et al., 1987). The amounts of B RNA-specific proteins detected in the protoplasts varied significantly for transcripts from individual clones. The transcripts of some cDNA clones did not result in detectable amounts of viral proteins including the previously described cDNA clone pTB3 (Verver et al., 1987). Inoculation with the transcripts of the other B cDNA clones led to detectable amounts of B RNA specific proteins. This indicated that these transcripts were infectious since it has been shown that replication of viral RNA has to occur before viral protein synthesis can be detected (Goldbach et al., 1980).

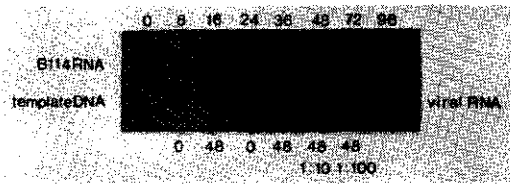


Fig. 3. Dot blot of nucleic acids extracted from transfected protoplasts. Each spot represents the total nucleic acids isolated from 10^5 protoplasts as described in Material and Methods, which was hybridized to a DNA probe specific for positive-stranded B RNA prepared from a recombinant M13 phage (Brown et al., 1982). Numbers at the top and bottom indicate times in hours after transfection. The top lane shows spots at different times after transfection of protoplasts transfected with RNA transcripts of cDNA clone pTB114. The bottom lane shows spots of protoplasts with template DNA or viral M and B RNA at 0 and 48 hours post transfection, the latter also ten and hundred times diluted.

The infectivity of the RNA transcripts of the B cDNA clones was further proven by analyzing protoplasts for the presence of viral RNA at different times after inoculation. Protoplasts were inoculated with RNA transcripts of B cDNA clone pTB114 and at various times post infection examined for the presence of B RNA (see Figure 3). For comparison protoplasts inoculated with total viral RNA and template DNA were similarly analysed. At time zero about 10 ng of the initial 100 µg transcript RNA could be detected, which corresponds to about 1000 molecules per cell. The amount of B specific RNAs decreased rapidly in the first 24 hours and then increased again, indicating that after a certain lag period replication of the transcripts started. These results demonstrate that the RNA transcripts of the B cDNA are indeed replicated in cowpea protoplasts.

Having demonstrated that transcripts of the B cDNA clones are expressed in cowpea protoplasts, transcripts of the M cDNA clones were also analyzed for their capability to infect protoplasts. Unlike B RNA M RNA is not replicated and translated independently in cowpea protoplasts and therefore transcripts of M cDNA clones need to be co-transfected with B RNA. As viral B RNA is always contaminated with trace amounts of M RNA irrespective of the isolation procedure (Goldbach et al., 1980; Goldbach, 1987) infectious transcripts of one of the cDNA clones of B RNA had to be used for co-transfection. In view of the relatively low infectivity of these transcripts the sensitivity of detection of M RNA encoded proteins was expected to be significantly lower than that of B RNA encoded proteins. Indeed, when protoplasts were infected with RNA transcripts of pTB114 together with transcripts of any of the selected M cDNA clones capsid protein synthesis could not be detected by immunoprecipitation with antibodies directed against the capsid proteins.

Therefore as an alternative method for detection of viral proteins in protoplasts the indirect immuno fluorescent antibody technique (Hibi, 1975; Maule et al., 1980) was applied. With this method fluorescent cells could be clearly detected among protoplasts infected with RNA transcripts of B cDNA clone pTB114 using antibodies directed against the 24K polypeptide and in addition the number of infected cells could be determined. Subsequently RNA transcripts of cDNA clone pTM203 with or without transcripts of pTB114

RNA transcripts	µg RNA per 5 x 10^6 protoplasts	% fluorescent cells with α-24K serum	% fluorescent cells with α-capsid serum
B114	100	0.8	0
B114*	100	2.0	0
M203*	100	0	0
B114 + M203*	200	1.5	0.1
B114 A1*	100	0	0
B114 A2*	100	0	0
CPMV-RNA			
B RNA + M RNA	2	1.5	1.4
"	4	4.0	3.9
"	20	20	20
"	50	50	50

Table 1. Detection of viral RNA expression in cowpea protoplast with immunofluorescent labelling. Asteriks refers to the presence of the cap structure analog ^{7}meGpppG during *in vitro* transcription, which resulted under the conditions used in approximately 50% capped transcripts (Verver *et al.*, 1987). In all cases about 20.000 protoplasts were examined by the indirect immuno fluorescence technique (Maule *et al.*, 1980). Incubations were carried out with a mixture of the anti-capsid serum and the 24K antiserum and the percentage of protoplasts stained with the anti-24K serum (α-24K) or both the anti-24K serum and anti-capsid serum (α-capsid) are indicated. The percentages shown represent an average of several independent experiments.

were examined by immuno fluorescence for the presence of M or B RNA-encoded proteins using a mixture of antibodies directed against the 24K polypeptide and the capsid proteins (see Table 1). Following this approach it was also possible to discriminate between protoplasts infected with only B RNA transcripts and protoplasts infected with both M and B RNA transcripts, since the immuno fluorescence caused by anti-24K serum on one hand and the antiserum against the capsid proteins on the other hand was quite different. Staining with the anti 24K serum resulted in a bright, localized immuno-fluorescence, whereas the anti capsid serum produced a more or less dim fluorescent staining over the whole cell. Infection with 100 µg of both

RNA transcripts of cDNA clone pTB114 and pTM203 resulted in the detection of B RNA specific proteins in about 1% of the protoplasts and of M RNA specific proteins in about 0.05% of the protoplast (see Table 1). This immuno fluorescent was never detected in uninfected cells. This indicates that transcripts of the M cDNA clone pTM203 were indeed infectious and that the lack of success with the detection of M RNA specific proteins by immuno precipitation of ^{35}S-methionine labelled cells must be due to the limited number of infected cells. The low level of infectivity of the cDNA transcripts compared to viral RNA found in the other experiments seems to be caused by the limited number of successfully infected cells and not by a low level of expression in a higher number of cells. If this was taken into account the infectivity of the transcripts of M cDNA was similar to the infectivity of the transcripts of B cDNA.

One of the reasons of the low infectivity of the transcripts could be rapid degradation of the RNA upon infection of the protoplasts (Verver et al., 1987). The presence of a cap structure at the 5' end is known to increase the stability of RNA molecules and therefore capped transcripts were also tested for infectivity (Furuichi et al., 1977; Shimotohno et al., 1977; Green et al., 1983). The number of infected protoplasts was indeed 2 to 3 times higher when capped RNA transcripts were used, although still low compared to viral RNA (see table 1).

It was also tried to infect cowpea plants with transcripts from both M and B cDNA clones, which should result in systemic infection. Under the conditions used, so far symptoms in plants infected with transcripts have not been detected, indicating that detection of infectivity in plants is even more difficult than in protoplasts.

Effect of deletions on the infectivity of transcripts from the B cDNA clones

Two deletion mutants, previously studied in vitro (Vos et al., submitted) were tested for their capability to be expressed in vivo.

The first mutant, pTB114Δ1, contained a 87 bp deletion in the coding sequence for the 24K protease (Verver et al., 1987). The effect of this "in frame" deletion is a complete blocking of proteolytic processing of the

viral polyproteins (Vos et al., submitted). Because the functional proteins can no longer be formed this mutant was expected to be defective in replication. Indeed we were not able to detect expression of the RNA transcripts of this mutants, neither by immunoprecipitation of proteins from ^{35}S-methionine labelled protoplasts nor by immuno fluorescent staining (see Table 1).

The second deletion mutant, pTB114Δ2, contained a 360 bp in frame deletion in the coding sequence for the 32K polypeptide (Vos et al., submitted). In vitro expression studies on this mutant revealed that the proteolytic processing of this mutant is similar to wild type B RNA. The only difference is the generation of a truncated 32K polypeptide with a molecular weight of 18 kilodalton (kD). Since there are no indications for a function of the 32K polypeptide in the expression of B RNA from in vitro translation studies (Vos et al., submitted), the effect of this mutation could not be anticipated. Similar to the first deletion mutant we were not able to detect expression of the RNA transcripts of clone pTB114Δ2 in cowpea protoplasts. The lack of infectivity of the RNA transcripts of this cDNA clone suggests that the 32K polypeptide somehow plays a role in replication of the viral RNAs. Furthermore these experiments again demonstrate and confirm that for detection of viral proteins replictation of the RNA transcripts is an absolute requirement.

DISCUSSION

In this paper we have described the construction of a series of full-length cDNA clones of both M and B RNA of CPMV downstream from the T7 promoter in plasmid pX1. Upon transcription of ClaI-linearised template DNAs of these clones using T7 RNA polymerase RNA transcripts were obtained closely resembling the viral RNAs. It was demonstrated in several ways that transcripts of the cDNA clones of B RNA are replicated and expressed in cowpea protoplasts and that transcripts of the M cDNA clones are also infectious.

Infectious RNA transcripts synthesized by transcription in vitro of full-length DNA copies of RNA viruses have been reported for human rhino

virus and poliovirus (Mizutani and Colonno, 1985; Van der Werf et al., 1986), the insect virus black beetle virus (Dasmahapatra et al., 1986) and two plant viruses, bromo mosaic virus (BMV) (Ahlquist et al., 1984) and tobacco mosaic virus (TMV) (Dawson et al., 1986; Meshi et al., 1986). The degree of infectivity of these transcripts varies from virus to virus but in general was much lower than that of the corresponding viral RNA, in agreement with our present results with RNA transcripts of cDNA clones of CPMV. The RNA transcripts used in our studies had two additional G-residues at the 5' end and 4 to 5 extra nucleotides to the 3' poly(A) tail. These extra non-viral sequences, which cannot be eliminated using one of the currently known in vitro transcription systems, presumably interfere with replication of the RNAs. In particular extra sequences at the 5' end may dramatically decrease the infectivity of in vitro transcripts of cDNA clones, as has been found for TMV (Dawson et al., 1986), poliovirus (Van der Werf et al., 1986) and BMV (Paul Ahlquist, personal communication). In all these cases a limited number of extra sequences at the 3' end did not significantly influence infectivity. A further reason for the relatively low infectivity of the in vitro transcripts from the cDNA clones of CPMV may be the absence of VPg in such transcripts. Although the function of VPg is not yet known, one of its functions might be protection of the RNA against ribonucleases. Transcripts having a triphosphate at the 5' end as used in this paper may be very susceptible to degradation. This may explain the higher infectivity of capped transcripts (although CPMV RNAs isolated from virions are not capped). On the other hand the higher infectivity of capped transcripts may be caused by a more efficient translation in the cowpea protoplasts during the first hours after infection (Shih et al., 1976), which may indirectly increase the infectivity of the RNA.

Another interesting observation is the difference in infectivity of transcripts from each of the full-size cDNA clones of B RNA tested. Such variation has also been observed among transcripts from cDNA clones of BMV and TMV (Ahlquist et al., 1984; Dawson et al., 1986; Meshi et al., 1986). Since sofar further information on the properties of these deviant cDNA clones is lacking, the reasons for the lower or complete lack of infectivity remains unclear. Possibly these cDNA clones are derived from non-infectious RNA molecules present in the viral RNA population used in the

preparation of the cDNA. Alternatively errors introduced during synthesis of the double-stranded cDNA may have led to reduced or even complete loss of infectivity.

Infection experiments with *in vitro* transcripts of two different deletion mutants, pTB114Δ1 and pTB114Δ2, show that both mutations destroy the infectivity of B RNA. While this was anticipated for mutant pTB114Δ1, the lack of infectivity of transcripts from pTB114Δ2 is less clear. *In vitro* translation studies of RNA transcripts of the latter deletion mutant show normal generation of the 170K polypeptide and further processing by cleavage into 110K and 60K or 87K and 84K proteins (Vos *et al.*, submitted). For replication of the RNAs the 110K polypeptide, the viral RNA dependent RNA polymerase and the 60K polypeptide (Dorssers *et al.*, 1983; Eggen and Van Kammen, in press) are considered to be essential. The processing of the B RNA-encoded polyprotein is achieved by the 24K polypeptide encoded by B RNA. At this moment no specific function of the 32K protein in the expression of B RNA is known, but it has been demonstrated that the 32K polypeptide plays an important role in proteolytic processing of the M RNA encoded polyproteins (Franssen *et al.*, 1984b; Vos *et al.*, submitted). The results obtained in this paper suggest that the 32K polypeptide directly or indirectly also plays a role in replication of the viral RNAs. Alternatively the change in the RNA structure due to this deletion may affect the replication of B RNA.

This is the first report of the expression *in vivo* of cDNA clones of a plant virus genome with a genome-linked protein and a 3' poly(A) tail. The infectivity of the *in vitro* transcripts which lack a VPg demonstrate that the genome-linked protein is not necessary for the infectivity of the CPMV RNAs. This finding confirms the conclusion drawn from experiments which show that the infectivity of CPMV RNA is not lost if VPg is removed by protease degradation. The availability of infectious RNA transcripts opens the way to study the expression and genetic organization of CPMV *in vivo* by expression of RNA transcripts obtained from DNA copies engineered on specific sites.

MATERIALS AND METHODS

Nucleic acids and enzymes

CPMV was propagated in Vigna unguiculata L., "California Blackeye", and M or B components were purified as described previously (Klootwijk et al., 1977). Viral RNA was extracted from purified M or B components as described by Zimmern (1975). Large or small scale plasmid or M13RF DNA isolations were performed according to Birnboim and Doly (1979). T7 RNA polymerase was isolated as described previously (Verver et al., 1987) or purchased from PL Biochemicals. All other enzymes were purchased from Boehringer Mannheim B.V. or PL Biochemicals. All oligonucleotides used were synthesized by the chemical phosphotriester method (Van der Marel et al., 1982).

Molecular cloning

Double-stranded cDNA was synthesized from CPMV RNA as described previously (Van Wezenbeek et al., 1983) or by the method of Okayama and Berg (1982). Occasionally single-stranded cDNA was eluted from alkaline agarose gels by a modified freeze squeeze method (Tautz and Renz, 1983). DNA ligations were performed as described previously (Verver et al., 1987). Recombinant DNA clones were analysed by restriction enzyme mapping and/or nucleotide sequencing (Sanger et al., 1977, 1980; Guo et al., 1983).

Transfection of cowpea protoplasts and detection of viral proteins

Template DNAs were linearized with ClaI and run-off transcripts were synthesized using T7 RNA polymerase as described previously (Verver et al., 1987). Occasionally transcripts were capped by addition of the cap structure analogue ^{7}meGpppG to the reaction mixture (Verver et al., 1987). After transcription template DNA was removed by DNase treatment (Vos et al., 1984) and in vitro transcripts were recovered from the reaction mixture by precipitation overnight with 2 M LiCl. Cowpea mesophyl protoplasts were prepared essentially as described by Hibi et al. (1975). Generally 5 x 10^{6} protoplasts were infected with 50-200 µg of RNA transcripts or viral RNA using the polyethylene glycol (PEG) method (Dawson et al., 1978; Maule et al., 1980). Washed protoplasts were incubated at 25°C under continuous illumation (Rottier et al., 1979). Protoplasts were stained with antiserum against the B RNA encoded 24K polypeptide and/or anti-capsid protein for detection of these proteins in infected protoplasts by the indirect fluorescent antibody technique (Hibi et al., 1975; Maule et al., 1980). Immunoprecipitations of 30,000 x g supernatant fractions of protoplasts labelled with ^{35}S-methionine were performed according to Wellink et al. (1986).

Analysis of RNA

At suitable periods after transfection protoplasts were collected by centrifugation and disrupted by freezing at -80°C followed by resuspension in 100 mM Tris 8.0, 10 mM EDTA, 2% SDS. This mixture was extracted with an equal volume of phenol:chloroform 1:1 and nucleic acids were collected from the waterphase by ethanol precipitation. Pellets were directed dissolved in 50% formamide, 2.2 M formaldehyde, 20 mM MOPS pH 7.0, 5 mM NaAc, 0.1 mM EDTA and incubated for 15 minutes at 55°C. For spot hybridisations aliquots of the RNA samples in the formaldehyde buffer were spotted directly on Gene Screen + hybridization (New England Nuclear Corp.) membranes and baked for two hours at 80°C. Gene Screen membranes were hybridized to strand specific M13 probes according to Thomas et al. (1980).

ACKNOWLEDGEMENTS

We thank Prof.Dr. Jacques van Boom for synthesis of the oligo-deoxynucleotides, Dr. Stanley Tabor for the gift of T7 promoter plasmids, Piet Madern for artwork and Gré Heitkönig for typing the manuscript. This work was supported by the Netherlands Foundation for Chemical Research (S.O.N.) with financial aid from the Netherlands Organisation for the Advancement of Pure Research (Z.W.O.).

REFERENCES

Ahlquist, P. and Kaesberg, P. (1979) Nucleic Acids Res., 7, 1195-1204.
Ahlquist, P. and Janda, M. (1984) Moll. Cell. Biol., 4, 2876-2882.
Ahlquist, P., French, R., Janda, M. and Loesch-Fries, L.S. (1984) Proc.Natl. Acad. Sci. USA, 81, 7066-7070.
Birnboim, C. and Doly, J. (1979) Nucleic Acids Res., 7, 1513-1523.
Brown, D.M., Frampton, J., Goelet, P. and Kam, J. (1982) Gene, 20, 139-144.
Contreras, R., Cheroute, H., Degrave, W. and Fiers, W. (1982) Nucleic Acids Res., 10, 3133-3142.
Dasmahapatra, B., Dasgupta, R., Saunders, K., Selling, B., Gallagher, T. and Kaesberg, P. (1986) Proc. Natl. Acad. Sci. USA, 83, 63-66.
Davanloo, P., Rosenberg, A.H., Dunn, J.J. and Studier, F.W. (1984) Proc.Natl. Acad. Sci. USA, 81, 2035-2039.
Dawson, J.R.O., Dickerson, P.E., King, J.H., Sakai, F., Trim, A.R.H. and Watts, J.W. (1978) Z. Naturforsch., 33c, 548-551.
Dawson, W.O., Beck, D.L., Knorr, D.A. and Grantham, G.L. (1986) Proc. Natl. Acad. Sci. USA, 83, 1832-1836.
Dunn, J.J. and Studier, F.W. (1983) J. Mol. Biol., 166, 477-435.
Franssen, H., Leunissen, J., Goldbach, R., Lomonossoff, G. and Zimmern, D. (1984a) EMBO J. 3, 855-861.
Franssen, H., Moerman, M., Rezelman, G. and Goldbach, R. (1984) J. Virol. 50, 183-190.
Furuichi, Y., La Fiandra, A. and Shatkin, A.J. (1977) Nature, 266, 235-239.
Goldbach, R., Rezelman, G. and Van Kammen, A. (1980) Nature, 286, 297-300.
Goldbach, R. and Rezelman, G. (1983) J. Virol., 46, 614-619.
Goldbach, R. and Van Kammen, A. (1985) In Davies, J.W. (ed), Molecular Plant Virology. CRC Press, Boca Raton, FL, Vol. II, pp 83-120.
Goldbach, R. (1987) In Harrison, B.D. and Murant, A.F. (eds), The Plant Virus: viruses with bipartite RNA genomes and isometric particles. Plenum Press, New York.
Green, M., Maniatis, T. and Melton, D. (1983) Cell, 32, 681-694.
Guo, L.H., Yang, R.C.A. and Wu, R. (1983) Nucleic Acids Res., 11, 5521-5539.
Hibi, R., Rezelman, G. and Van Kammen, A. (1975) Virology, 64, 308-318.
Klootwijk, J, Kleine, I., Zabel, P. and Van Kammen, A. (1977) Cell, 11, 73-82.
Van der Marel, C.A., Wille, G., Marugg, J.E., De Vroom, E., Tromp, M., Van Boekel, C.A.A. and Van Boom, J.H. (1982) Rec. Trav. Chim. Pays Bas, 10, 239-241.
Maule, A.J., Boulton, M.I. and Wood, K.R. (1980) Phytopathol. Z., 97, 118-126.
Meshi, T., Ishikawa, M., Motoyoshi, F., Semba, K. and Okada, Y. (1986) Proc. Natl. Acad. Sci. USA 83, 5043-5047.
Mizutani, S. and Colonno, R.J. (1985) J. Virol., 56, 626-632.
Okayama, H. and Berg, P. (1982) Moll. Cell. Biol., 2, 161-170.
Rottier, P.J.M., Rezelman, G. and Van Kammen, A. (1979) Virology, 92, 299-309.
Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA, 74, 5463-5467.
Sanger, F., Coulson, A.R., Barrel, B.G., Smith, A.J.H., Roe, B.A. (1980) J. Mol. Biol., 143, 161-178.
Shih, D.S., Dasgupta, R. and Kaesberg, P. (1976) J. Virol., 19, 736-742.
Shimotohno, K., Kodama, Y., Hashimoto, J. and Miura, K.I. (1977) Proc. Natl. Acad. Sci. USA, 74, 2734-2738.
Tautz, D. and Renz, M. (1983) Anal. Biochem., 132, 14-19.
Thomas, P.S. (1980) Proc. Natl. Acad. Sci. USA, 77, 5201-5205.
Verver, J., Goldbach, R., Garcia, J. and Vos, P. (1987) EMBO J., 6, 549-554.
Vos, P., Verver, J., Van Wezenbeek, P., Van Kammen, A. and Goldbach, R. (1984) EMBO J., 3, 3049-3053.
Vos, P., Verver, J., Wellink, J., Jaegle, M., Van Kammen, A. and Goldbach, R. Role of the cowpea mosaic virus B RNA-encoded 24K and 32K polypeptides in the proteolytic processing of the viral proteins, submitted for publication.
Wellink, J., Rezelman, G., Goldbach, R. and Beyreuther, K. (1986) J. Virol., 59, 50-58.
Wellink, J., Jaegle, M. and Goldbach, R. (1987) J. Virol., 61, 236-238.
Van der Werf, S., Bradley, J., Wimmer, E., Studier, F.W. and Dunn, J.J. (1986) Proc. Natl. Acad. Sci. USA, 83, 2330-2334.
Van Wezenbeek, P., Verver, J., Harmsen, J., Vos, P. and Van Kammen, A. (1983) EMBO J., 2, 941-946.
Zimmern, D. (1975) Nucleic Acids Res., 2, 1189-1201.
Zoller, M.J. and Smith, M. (1982) Nucleic Acids Res., 10, 6487-6500.

CHAPTER 6

ROLE OF THE COWPEA MOSAIC VIRUS B RNA-ENCODED 24K AND 32K POLYPEPTIDES IN THE PROTEOLYTIC PROCESSING OF THE VIRAL PROTEINS

Pieter Vos, Jan Verver, Martine Jaegle, Joan Wellink, Ab van Kammen and Rob Goldbach

(submitted for publication)

ABSTRACT

A series of specific deletion mutants derived from a full-length cDNA clone of cowpea mosaic virus (CPMV) B RNA was constructed with the aim to study the role of viral proteins in the proteolytic processing of the primary translation products. For the same purpose cDNA clones were constructed having sequences derived from both M and B RNA of CPMV. In vitro transcripts prepared from these clones with T7 RNA polymerase, were efficiently translated in rabbit reticulocyte lysates. The translation products obtained were processed in the lysate by specific proteolytic cleavages into smaller products, which made it possible to study subsequently the effect of various mutations on this process. The results obtained indicate that the B RNA encoded 24K polypeptide represents a protease responsible for all cleavages in the polyproteins produced by both CPMV B and M RNA. For efficient cleavage of the glutamine-methionine site in the M RNA encoded polyprotein the presence of a second B RNA encoded protein, the 32K polypeptide, is essential, although the 32K polypeptide itself does not have protease activity. Some cleavage-site mutants were constructed in which the coding sequence for the glutamine-glycine cleavage site between the two capsid proteins was changed. Subsequent in vitro transcription and translation of these cleavage site mutants show that a correct dipeptide sequence is a prerequisite for efficient cleavage but that the folding of the polypeptide chain also plays an important role in the determination of a cleavage site.

1. Introduction

The genetic information of cowpea mosaic virus (CPMV) is divided among two, separately encapsidated, positive-stranded RNA molecules, designated M and B RNA (for recent reviews see Goldbach and Van Kammen, 1985 and Goldbach, 1987). The viral RNAs have a small protein, denoted VPg (Daubert et al., 1978; Stanley et al., 1978), covalently linked to their 5' end, and a poly(A) tail at their 3' end (El Manna and Bruening, 1973). Both RNAs are translated into large polyproteins, from which the functional proteins are derived by proteolytic cleavages (Rezelman et al., 1980; Goldbach and Rezelman, 1983).

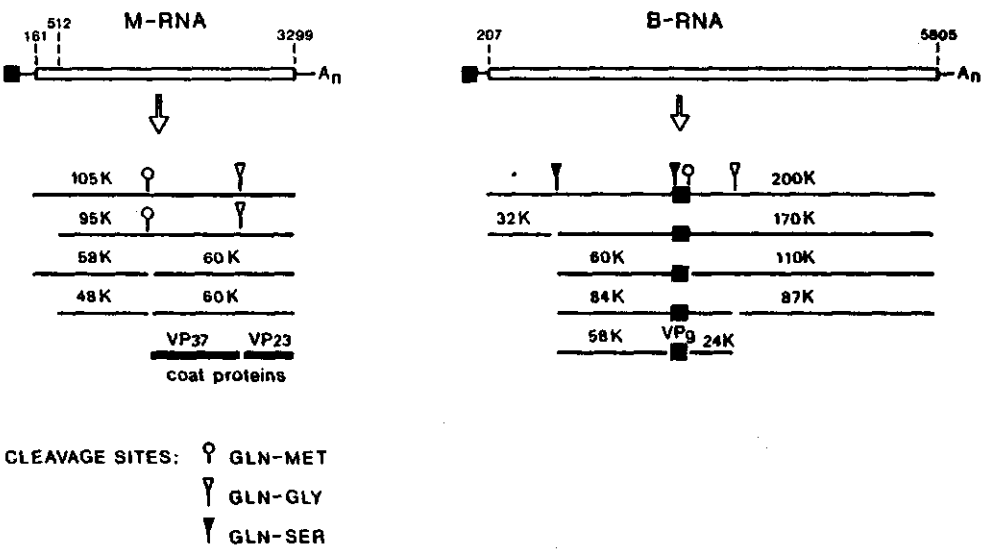


Figure 1. Model for the translation of the RNAs of CPMV and proteolytic processing of the CPMV polyproteins. The double-lined bars in the RNAs represent the long open reading frames, for which the start and stop codons are also indicated (Van Wezenbeek et al., 1983; Lomonossoff and Shanks, 1983). VPg is indicated by a black square, other protein sequences as single lines. The various cleavage sites are indicated on the primary translation products.
N.B. At the start of the investigation the cleavage sites on the B proteins were not yet determined.

CPMV B RNA is translated into a 200 kilodalton (kD) polyprotein which is rapidly cleaved into 32 kD and 170 kD polypeptides (see Figure 1). This primary cleavage is achieved intramolecularly by the B RNA encoded 24K polypeptide at a glutamine-serine dipeptide sequence (Franssen et al., 1984a; Verver et al., 1987). The 170K polypeptide is then further cleaved either at a glutamine-methionine site giving the 110K and 60K polypeptide or at a glutamine-glycine site giving the 87K and 84K polypeptides (Wellink et al., 1986). While cleavage of the glutamine-glycine site is performed by the 24K polypeptide (Garcia et al., 1987) the proteolytic activity responsible for cleavage of the glutamine-methionine site has not yet been identified. VPg is released from the 60K polypeptide at a glutamine-serine dipeptide sequence and the 24K polypeptide can be released either from the 84K polypeptide by cleavage of the glutamine-methionine site or from the 110K polypeptide by cleavage of the glutamine-glycine site.

CPMV M RNA is translated into two C-terminal overlapping polypeptides of 105 and 95 kD respectively (see Figure 1). These polypeptides are in vivo immediately cleaved at a glutamine-methionine site releasing the 60 kD capsid protein precursor, which in turn is rapidly cleaved at a glutamine-glycine site giving the two capsid proteins VP37 and VP23. Cleavage at the glutamine-methionine pair releasing the capsid protein precursor can be blocked in vitro by antibodies raised against the B RNA encoded 32K polypeptide (Franssen et al., 1984b). Therefore the 32K polypeptide has been proposed to be the proteinase responsible for this cleavage reaction. Expression studies in E. coli of cDNA clones containing sequences derived from both M and B RNA have shown that proteolysis of the cleavage site between the two capsid protein at a glutamine-glycine dipeptide sequence is accomplished by the 24K protease encoded by B RNA (Garcia et al., 1987).

In summary at least two proteinases, the B RNA-encoded 24K and 32K polypeptides, seem to be involved in the proteolytic processing of the CPMV polyproteins, at three different types of cleavage sites. In this paper we have investigated in more detail the role of the 24K and 32K proteins in the proteolytic processing reactions and have also studied the structural requirements for cleavage-sites. For this purpose the in vitro translation of RNA-transcripts derived from specifically engineered cDNA clones was investigated. The results obtained reveal that a single virus-encoded pro-

teinase, the 24K polypeptide, can recognize all three types of cleavage sites in the CPMV polyproteins, but that the 32K polypeptide also plays a role in the proteolytic processing.

2. Results

Construction of mutagenized cDNA clones

As starting material for the analyses described in this paper a full-length cDNA clone of B RNA, pTB114, was used, which recently has been shown to express infectivity in cowpea protoplasts (Vos et al., 1987). To study the function of the 32K and 24K proteins encoded by B RNA five different deletion mutants were constructed: pT114Δ1, pTB114Δ2, pTB114Δ3, pTB114Δ4 and pTB32*. For the same purpose two cDNA clones were constructed pTMB110 and pTMB120, containing sequences of both M and B RNA.

cDNA clone pTB114Δ1 was constructed by exchange of a small DNA fragment of pTB114 for the corresponding fragment of a previously constructed cDNA clone pTB3Δ1 (Verver et al., 1987) (see figure 2). The result is an 87 bp deletion (positions 3240 to 3327) in the coding sequence for the 24K polypeptide leading to deletion of 29 amino acids centrally located in this polypeptide. In vitro transcription and translation of pTB3Δ1 already showed that this clone encodes a polyprotein, which does not longer achieve the primary cleavage of the glutamine-serine dipeptide sequence and for that reason does not release the 32K polypeptide from the polyprotein (Verver et al., 1987).

Mutant pTB114Δ2, with a deletion in the coding sequence for the 32K polypeptide, was constructed by removing an 375 bp AccI fragment (positions 360-735) from the full-size cDNA clone pTB114 (see Figure 2). Clone pTB114Δ2 has lost the coding sequence for 125 amino acids of the 32K polypeptide and this large deletion is likely to affect the properties of the 32K polypeptide.

A third deletion mutant, pTB114Δ3, contains a very large deletion (positions 346 to 2815 of B RNA) removing almost the complete coding sequence for both the 32K and 58K polypeptides including the glutamine-serine cleaving site between these two proteins (see Figure 2). This deletion does not affect the reading frame and the <u>in vitro</u> translation products of RNA transcripts of this clone are expected to have a molecular weight of approximately 120 kD. This 120 kD product contains only the complete amino acid sequence of VPg, the 24K polypeptide and the 87K polypeptide.

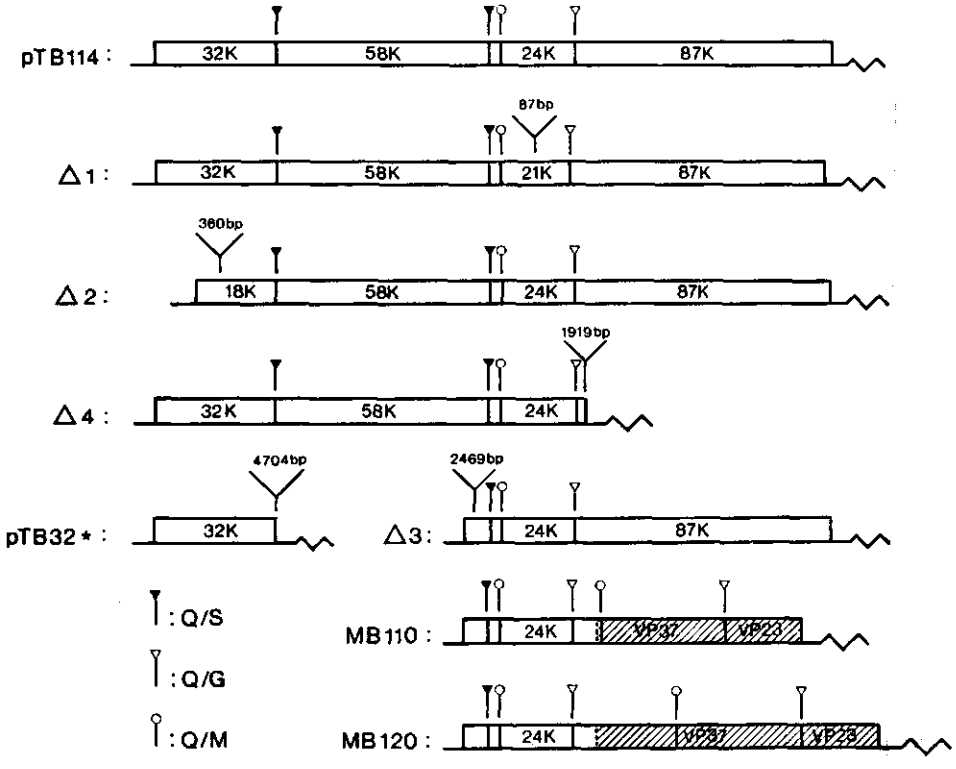


Fig. 2. Schematic representation of the various mutagenized cDNA clones. The open reading frames are indicated by double-lined bars, non-coding sequences as single lines. The positions of the viral proteins in the long open reading frames are shown. The size of the deletions in B cDNA clones pTB114Δ1, pTB114Δ2, pTB114Δ3, pTB114Δ4 and pTB32* are indicated. Δ1, Δ2, Δ3 and Δ4 refer to DNA clone pTB114Δ1, pTB114Δ2, pTB114Δ3 and pTB114Δ4 respectively. MB110 and MB120 refer to the hybrid DNA clones pTMB110 and pTMB120. The positions of the potential cleavage sites in all DNA clones are indicated. For details on the construction of the clones see Materials and Methods.

In cDNA clone pTB114Δ4 a 1961 bp PvuII-HpaI-fragment (positions 3700 to 5661 of B RNA) is removed from the full-length DNA copy in clone pTB114 (see Figure 2). The result of this deletion is a frame shift leading to an expected primary translation product of about 110 kD terminating in the sequence of the 87K polypeptide 18 amino acids downstream of the glutamine-glycine cleavage site between the 24K and 87K polypeptides. This mutant therefore contained the coding sequence for the 24K polypeptide and all B RNA encoded proteins located N-terminally from this protein in contrast to clone pTB114Δ3, which contained the coding sequence for the 24K polypeptide and the 87K polypeptide located at the C-terminal side of the 24K polypeptide.

For a further analysis of the functions of the 32K polypeptide a cDNA clone pTB32* was constructed, which allows the synthesis in vitro of the 32K polypeptide as a free protein. Using methods for site-directed mutagenesis (see Materials and Methods) the AGT codon for the serine residue at the glutamine-serine cleavage site between the 32K and 58K polypeptides was changed into TGA giving an amber stop codon at this position. Sequences downstream from this stop codon were removed and exchanged for part of the non-translated region and poly d(A-T) track of a cDNA clone of M RNA. Upon in vitro transcription of pTB32*, RNA molecules are generated with the exact coding sequence for the 32K polypeptide flanked by the 5' leader sequence of B RNA and the 3' non-translated region and poly(A) tail of M RNA (see Figure 2).

To study the proteolytic activity of the 24K protease with respect to the cleavage sites in the M polyproteins cDNA clones were constructed having sequences of both M and B RNA. In clone pTB114Δ3 the DNA fragment from the BamH1-site at position 3865 of B RNA to the 3' end was removed and replaced by coding sequences for the C-terminal part of the M RNA encoded polyproteins in such a way that the reading frame was not affected. As a result almost the complete coding sequence for the 87K polypeptide was exchanged for the coding sequences for the two capsid proteins (see Figure 2). Two such hybrid cDNA clones were constructed, pTMB110 and pTMB120, of which the latter contained a larger part of the coding sequence for the M RNA encoded polyproteins. In these clones large open reading frame are present expected to give rise to translation products of 108 kD (pTMB110) and

140 kD (pTMB120) respectively. In the expected translation products to two cleavage sites of the M polyproteins, glutamine-methionine and glutamine-glycine, are present as well as three cleavage sites of the B RNA encoded polyprotein (see Figure 2).

In vitro expression of cDNA clones with deletions in the coding sequence of B RNA

The various cDNA clones were transcribed by T7 RNA polymerase and the RNA transcripts were translated in rabbit reticulocyte lysates (Verver et al., 1987). Such transcripts have been shown to be translated very efficiently in vitro and to undergo faithful proteolytic processing (Vos et al., 1984; Verver et al., 1987). Figure 3A shows the PAAGE patterns of the translation products of RNA transcripts obtained from the various cDNA clones.

RNA transcripts of the full-length cDNA clone pTB114 were translated into a 200 kD polyprotein similar to the primary translation product of viral B RNA, which was rapidly cleaved into 170K and 32K polypeptides. Upon prolonged incubation in the lysate the 170K polypeptide was processed either (at a glutamine-methionine site) into 110K and 60K polypeptides or (at a glutamine-glycine site) into 87K and 84K polypeptides which comigrated in this gel (Figure 3A and Franssen et al., 1984a; Wellink et al., 1986). The 60K polypeptide was further processed into the 58K polypeptide and VPg. In comparison with viral B RNA the in vitro synthesized B RNA was translated more efficient and gave less secondary translation products. This is probably due to partial breakdown of the viral RNA as a consequence of the isolation procedure. As a result an equal amount of viral RNA produces a lower amount of the complete polyprotein and more smaller translation products (Kozak, 1983).

The small deletion in the 24K polypeptide in cDNA clone pTB114Δ1 has already been shown to block the primary cleavage of the 200K polyprotein at a glutamine-serine site by which the 32K polypeptide is released (Verver et al., 1987). In vitro translation of RNA transcripts of this clone resulted in accumulation of a translation product of 200 kD migrating slightly

faster than the 200K polyprotein. Even after prolonged incubation of this translation product remained stable in the reticulocyte lysate and processing into smaller polypeptides could not be detected. This indicates that the 24K protease is involved in all cleavages in the B RNA encoded polyprotein or that at least generation of the 170K polypeptide is a prerequisite for the secondary cleavages to occur. In the latter case another proteolytic activity may still be involved.

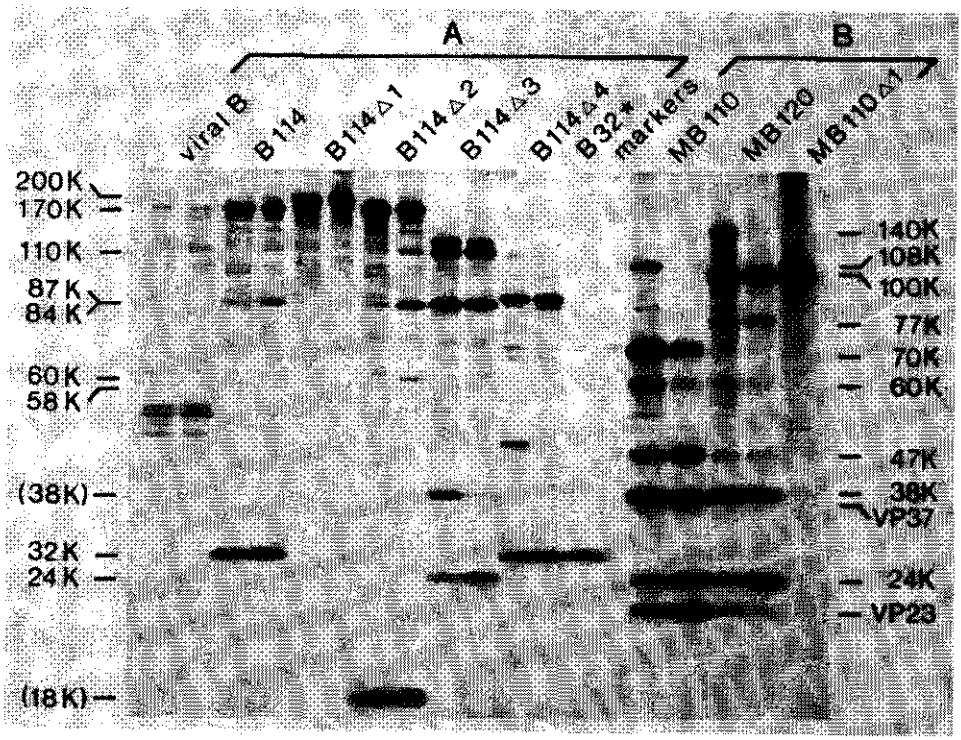


Fig. 3. In vitro translation of RNA transcripts of the various specifically modified DNA clones. In most cases (except clone pTB32* and pTMB110Δ1) the translation products are shown after one hour and twenty hours incubation in the reticulocyte lysate, left hand lane and right hand lane respectively. Marker proteins were from top to bottom: myosin, 210 kD; phosphorylase b, 100 kD and 92.5 kD; bovine serum albumine, 68 kD; ovalbumin, 46 kD and carbonic anhydrase, 30 kD.
Panel A: In vitro translation products of RNA transcripts of deletion mutants derived from the full-size B cDNA clone pTB114. At the left the positions of the various B RNA specific proteins are indicated and in addition the 38 kD protein (38K, see text) and the truncated 32 kD protein (18K). cDNA clone pTB32* was translated for one hour only.
Panel B: In vitro translation products of RNA transcripts of the hybrid DNA clones, containing sequences from both M and B RNA. To the right the positions of the most abundant translation and processing products are indicated, VP37 and VP23 refer to the 37 kD and 23 kD capsid proteins. For DNA clone pTMB110Δ1 only the 20 hours translation experiment is shown.

Translation of RNA transcripts of pTB114Δ2, the cDNA clone with the 375 bp deletion in the coding sequence for the 32K polypeptide, was very similar to translation of RNA transcripts of pTB114. The only difference was the generation of a primary translation product of approximately 190 kD, which was rapidly cleaved into the 170K polypeptide and a truncated 32K polypeptide with a molecular weight of approximately 18 kD. Further processing of the 170K polypeptide of pTB114Δ2 occurred precisely as in the 170K polypeptide encoded by the wild type B RNA. This suggests that the 32K polypeptide is not involved in any of the cleavages in the 170K polypeptide but that these cleavages are all catalyzed by the 24K protease. Hence the 24K protease seems to be capable of cleaving all three types of cleavage sites used in the CPMV polyproteins, which makes involvement of a second viral protease, i.e. the 32K polypeptide unnecessary.

This conclusion was confirmed by translation of RNA transcripts of cDNA clone pTB114Δ3. The primary translation product of this clone had a molecular weight of about 125 kD and was efficiently cleaved into smaller proteins. Cleavage of the glutamine-methionine site at the left hand site of the 24K protease results in the 110K polypeptide and a protein with a size of 15 kD (not to be seen in Figure 3A). Cleavage of the glutamine-glycine site at the right hand site of the 24K protease gives the 87K polypeptide and a polypeptide with a molecular weight of 38 kD. Cleavages at both sites of the 24K protease released this polypeptide from the polyprotein. These results confirm that apart from glutamine-serine sites the 24K protease can also cleave glutamine-methionine and glutamine-glycine dipeptide sequences. Surprisingly with this mutant relatively large amounts of free 24K protease were generated. One of the reasons for this observation may be that the 110K polypeptide, a precursor of the 24K protease, is generated in the reticulosyte lysate much more efficient from the 125 kD translation product of pTB114Δ3 than from the viral 170K polypeptide.

Translation of the RNA transcripts of cDNA clone pTB114Δ4, lacking almost the complete coding sequence of the 87K polypeptide, resulted as expected in a primary translation product with a size of about 115 kD. The 32K polypeptide is rapidly released from this primary translated product leaving a polypeptide slightly large than the 84K polypeptide since a few N-terminal amino acids of the 87K polypeptide are still linked to the 84K

polypeptide (see figure 2). This extra sequence at the C-terminus of the 84K polypeptide is cleaved off after prolonged incubation in the lysate by cleavage at the glutamine-glycine dipeptide sequence at the right hand site of the 24K protease. Translation of RNA transcripts of pTB114Δ4 did not result in detectable amounts of the 24K polypeptide in contrast to translation of RNA transcripts of pTB114Δ3. This may indicate that the 84K polypeptide is not the direct precursor of this polypeptide while the 110K polypeptide is.

In conclusion the results obtained by *in vitro* expression of these mutant cDNA clones of B RNA demonstrate that the 24K protease performs all four cleavages at three different dipeptide sequences in the B polyproteins and indicate that involvement of a second proteolytic activity is not necessary.

In vitro expression of DNA clones containing cDNA sequences of both M and B RNA

The proteolytic processing of the M RNA encoded polyproteins was investigated with two cDNA clones containing sequences of both M and B cDNA (see Figure 2). The clones, pTMB110 and pTMB120, have a single long open reading frame and both contain the complete coding sequence of the 24K protease and the two M RNA-encoded capsid proteins including the two cleavage sites to release the structural proteins. The only difference between pTMB110 and pTMB120 is that the latter encodes a larger primary translation product in which the two cleavage sites originating from the M polyproteins are situated at a larger distance from the 24K protease (see Figure 2).

Upon translation of the RNA transcripts of clone pTMB110 a polypeptide with a molecular weight of about 108 KD was synthesized, which was rapidly processed into smaller proteins (Fig. 3B). The identity of these processing products was determined by immunoprecipitation using antisera against the 24K polypeptide, VP37 and VP23 (data not shown).

The two glutamine-glycine sites in the 108K polyprotein of pTMB110 were efficiently cleaved. Cleavage of the first glutamine-glycine site gave a N-terminal polypeptide of 38 kD and a C-terminal polypeptide of 70 kD.

Cleavage of the second glutamine-glycine site resulted in a C-terminal polypeptide of 85 kD, which was rapidly processed into smaller products and the small capsid protein VP23. Cleavage at both glutamine-glycine sites produced a 47 kD polypeptide. The two glutamine-methionine sites in the translation products of pTMB110 were also cleaved by the 24K protease. Predominant proteins resulting from cleavage at these glutamine-methionine sites were the 24K protease itself, resulting from processing at both the glutamine-methionine and glutamine-glycine cleavage sites bordering this polypeptide, and the 60K capsid protein precursor. Cleavage of the glutamine-glycine site in this capsid protein precursor released both capsid proteins VP37 and VP23.

The results obtained with cDNA clone pTMB120 were very similar to those obtained with clone pTMB110. The difference between pTMB110 and pTM120 is that the distance between the glutamine-glycine site at the right hand site of the 24K protease and the glutamine-methionine site at the left hand site of VP37 is much larger in clone pTMB120. Therefore all translation products containing this region are about 30 kD larger than the corresponding translation products of pTMB110. Accordingly, the primary translation product of the RNA transcripts of pTMB120 had a molecular weight of about 140 kD compared to 108 kD for pTMB110. The major processing products of this cDNA clone were polypeptides of 100 kD (70 kD in pTMB110), 77 kD (47 kD in pTMB110) and further the 60 kD capsid protein precursor, the 38 kD polypeptide, the 24K protease and VP23. The 60K capsid protein precursor and both capsid proteins VP37 and VP23 were present in lower amounts in case of pTMB120 compared to pTMB110. Processing of the cleavage sites originating from the M RNA encoded polyproteins was therefore less efficient in the translation products of pTMB120, possibly because in the translation products of pTMB120 these cleavage sites are located further away from the 24K protease region. Translation of the RNA transcripts of clone pTMB110Δ1, a derivative of pTMB110, containing the 87 bp deletion in the coding sequence of the 24K protease, showed that all cleavages in this hybrid polyprotein indeed depended on the presence and activity of the intact 24K polypeptide (see Figure 3B).

The results of the translation and subsequent processing of the RNA transcripts of these hybrid cDNA clones indicate that all cleavage sites

present in these chimaeric polyproteins can be cleaved by the 24K protease, including the glutamine-methionine and glutamine-glycine dipeptide sequences originating from the M RNA-encoded polyprotein.

Processing of the M RNA encoded polyproteins by translation products of the various deletion mutants of B RNA

Although we now have demonstrated that the 24K protease is able to recognize both cleavage sites of the M polyprotein located on a chimaeric polyprotein, these cleavage sites are *in vivo* not located on the same protein molecule as the 24K protease. To investigate the natural processing at these cleavage sites (in trans) unlabeled translation products of RNA transcripts of wild type and mutated B cDNA clones were added to the M RNA encoded polyproteins labeled with ^{35}S-methionine (see Figure 4).

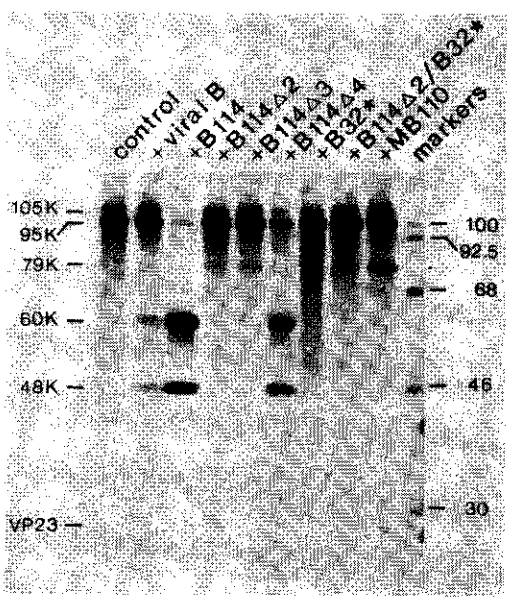


Fig. 4. Proteolytic processing of the M RNA encoded primary translation products by translation products of the various specifically modified DNA clones. RNA transcripts of M cDNA clone pTM203 were translated in the presence of ^{35}S-methionine and after one hour a 3 times excess of unlabeled translation products of one of the DNA clones as indicated on top of the gel was added and the incubation was continued at 30°C for another 20 hours. Control refers to addition of a 3 times excess of reticulocyte lysate to which no RNA was added. B114Δ2/B32* refer to the addition of a 1:1 mixture of the translation products of DNA clones pTB114Δ2 and pTB32*. To the left the positions of M RNA encoded proteins are indicated (Franssen *et al.*, 1982). To the right the positions of marker proteins phosphorylase b (100 kD and 85 kD), bovine serum albumin (68 kD), ovalbumin (46 kD) and carbonic anhydrase (30 kD) are indicated.

RNA transcripts of cDNA clone pTM203 (Vos et al., submitted) were translated in rabbit reticulocyte lysate resulting in efficient production of the 105K and 95K polyproteins (see figure 4). Upon addition of the translation products of viral B RNA or in vitro synthesized B RNA from cDNA clone pTB114 the M RNA encoded 95K polyprotein was processed into the 48K polypeptide and the 60K coat protein precursor. The 105K protein is processed into the 60K and 58K polypeptides but since the 105K polypeptide is present in only small amounts the 58K polypeptide is not easily detectable. Translation products of the deletion mutant in the 24K protease, pTB114Δ1, did not give any processing (data not shown). Addition of the translation products of the RNA transcripts of DNA clones pTB114Δ2, pTB114Δ3 or pTMB110, all lacking an intact 32K polypeptide but having an intact 24K protease, did not result in processing of the 95K polypeptide at the glutamine-methionine site but resulted in cleavage of the glutamine-glycine site giving the small capsid protein VP23 and a 79 kD polypeptide. Addition of translation products of RNA transcripts of pTB114Δ4, having an intact 24K and 32K polypeptide resulted in release of the capsid protein precursor indicating that the 32K polypeptide is important for efficient processing in trans of the glutamine-methionine site in the M polyproteins.

To investigate in more detail the role of the 32K polypeptide in processing of the M polyproteins, the translation products of clone pTB32* were analysed with respect to their activity on the M RNA encoded polyproteins. RNA transcripts of this cDNA clone are efficiently translated into a free 32K polypeptide (see Figure 3A). Addition of the translation products of the RNA transcripts of pTB32* to the M polyproteins did not result in proteolytic cleavage of the 105K and 95K polypeptides. Furthermore the translation products of pTB32* were not able to complement the activity of a defective 32K polypeptide in trans. When the M polyproteins were incubated with translation products of pTB114Δ2 addition of the translation product of pTB32* (i.e. the 32K polypeptide) to this mixture did not result in cleavage of the 105K and 95K primary translation products. The results from the processing of the M RNA encoded polyproteins by the translation products of the various cDNA clones confirm that the 24K protease is responsible for the proteolytic cleavages in these polyproteins, but indicate that 32K polypeptide is somehow also involved. In the absence of the

32K polypeptide the 24K protease cleaves the glutamine-glycine site between the two capsid proteins. When the 24K polypeptide and 32K polypeptide ar both present proteolytic cleavage is directed to the glutamine-methionine dipeptide sequence. This activity of the 32K cannot be complemented *in trans*, probably because for the latter cleavage the 32K polypeptide must be released from the same polyprotein as the 24K protease.

Proteolytic processing of specifically modified cleavage sites

To analyse the effect of amino acid substitutions at or around cleavage sites several derivates of cDNA clone pTMB110 were constructed modified at the glutamine-glycine dipeptide sequence between the two capsid proteins. This cleavage site was selected because in the coding sequence an *AvaII* restriction site is present at this site facilitating modification of the sequence. The restriction enzyme *AvaII* gives a 5' extension of 3 nucleotides. Hence, digestion with *AvaII* and subsequent "filling-in" and religation results in insertion of 3 nucleotides, while digestion with *AvaII*, followed by removal of the protruding ends with S1-nuclease and religation results in deletion of 3 nucleotides. Employing the *AvaII*-site for site-directed mutagenesis as indicated does therefore not alter the reading frame of the modified sequence.

Four different cleavage site mutants of pTMB110 were constructed, two employing the *AvaII* restriction site and two by oligonucleotide-directed mutagenesis (see Materials and Methods). In the translation products of pTMB110 the amino acid sequence at the cleavage sites between the two capsid proteins is glutamine/glycine-proline-valine. In pTMB113 and pTMB114 this amino acid sequence is changed into glutamine/alanine-valine and glutamine/ glycine-arginine-proline-valine respectively. In clones pTMB115 and pTMB116 the glutamine-glycine cleavage sites is changed into glutamine-methionine and glutamine-serine respectively. The latter two cleavage sites are also used at other positions in the CPMV polyproteins.

RNA transcripts of pTMB110 and the four mutants were translated in rabbit reticulocyte lysates (see Figure 5). The translation and processing products of pTMB110 have already been described earlier in this paper (see

Figure 3B). In Figure 5 in addition the 85kD polypeptide and 48 kD polypeptide are indicated, the former resulting from cleavage at only the glutamine-glycine site between the two capsid proteins and the latter resulting from cleavage at only the glutamine-methionine site to the left of VP37. Normally these polypeptides are rapidly processed in smaller products.

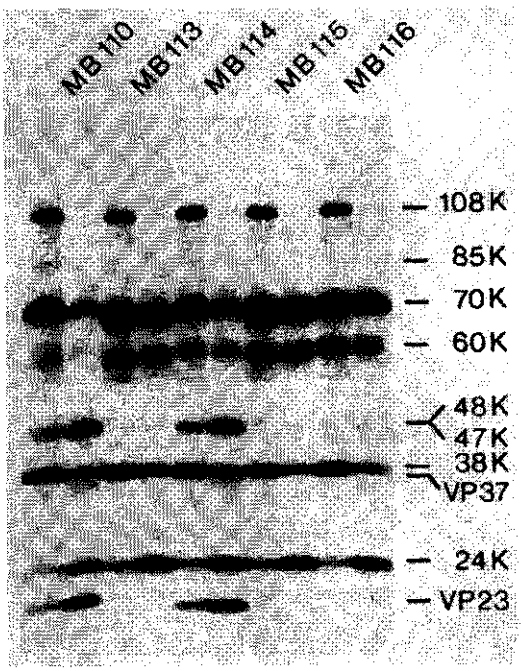


Fig. 5. In vitro translation of RNA transcripts of cDNA clone pTMB110 and its derivatives modified at the glutamine-glycine cleavage sites between the two capsid proteins. RNA transcripts were incubated for one hour (left hand lane) and 20 hours (right hand lane) in rabbit reticulocyte lysates. To the left the positions of various translation and processing products are indicated (see also figure 3B). In clones pTMB113, pTMB115 and pTMB116 the coding sequence for the glutamine-glycine dipeptide sequence is changed to glutamine-alanine, glutamine-methionine and glutamine-serine respectively. In clone pTMB114 the amino acid to the right of the glutamine-glycine dipeptide sequence is changed from proline into arginine. For details on the construction of these mutants see Materials and Methods.

The results obtained with the modified glutamine-glycine cleavage site can be summarized as follows. Processing at the glutamine-glycine site was not significantly affected by insertion of an extra amino acid i.e. an arginine next to the glycine in pTMB114. Changing of the glutamine-glycine dipeptide sequence in all cases decreased the cleavage efficiency of this site irrespective of nature of the change. This means that changing of the glutamine-glycine into glutamine-serine or glutamine-methionine, the two other cleavage sites used in the proteolytic processing of the CPMV polyproteins, also affected the cleavage efficiency at this position. Decrease of the cleavage efficiency at this specific site was easily detected by lower amounts of the proteins which are released by this cleavage (i.e. VP23, VP37, the 47 kD polypeptide and the 85 kD polypeptide) and the greater stability of precursor proteins having this cleavage site (i.e. the 70 kD polypeptide and the 60 kD capsid protein precursor) (see Figure 5).

3. Discussion

In this paper we have investigated the proteolytic processing reactions in the CPMV polyproteins by *in vitro* expression of specifically modified cDNA clones. The results obtained clearly demonstrate that the B RNA encoded 24K polypeptide catalyses all cleavages in the CPMV polyproteins. However, proteolytic cleavage of glutamine-methionine site in the M RNA encoded polyproteins by the 24K proteinase requires the presence of the B RNA encoded 32K polypeptide, indicating an essential role of the latter polypeptide in the cleavage reaction at this site. This confirms the earlier finding that antibodies raised against the 32K polypeptide can block the primary cleavage reaction of the M RNA encoded polyproteins (Franssen *et al.*, 1986).

Proteolytic cleavage of the M RNA encoded polyproteins differs from cleavage of the B RNA-encoded polyprotein because the proteinase responsible for the cleavage reactions is located on a different protein molecule and therefore the proteolytic cleavage reactions are always bimolecular. It

is interesting to note that the glutamine-methionine cleavage site of the M polyproteins is efficiently cleaved by the 24K protease without involvement of the 32K polypeptide when this sites is part of a hybrid MB polyprotein (cDNA clones pTMB110 and pTMB120) in which the 24K protease is also present. This is not due to a different conformation of this site in the hybrid polyprotein, because for cleavage in trans of the hybrid protein, e.g. the translation product of pTMB110Δ1, the 32K polypeptide is again required (data not shown). The stimulatory effect of the 32K polypeptide on cleavage of this glutamine-methionine site therefore seems to be limited to cleavage in trans and plays no role in intramolecular cleavage reactions.

It is not clear how the 32K polypeptide is involved in this cleavage reaction. A preprequisite for efficient cleavage of the glutamine-methionine site in the M polyproteins is that both the 24K proteinase and the 32K polypeptide originate from the same precursor protein. The specific effect of the 32K polypeptide was absent if this polypeptide was supplied in trans e.g. by the translation products of clone pTB32*. An explanation for this may be that the 32K polypeptide produced from clone pTB32* has a different conformation compared to a 32K polypeptide that is released from the B RNA encoded 200K polyprotein. This possibility was investigated by addition of the translation products of a cDNA clone producing an active 24K protease to the translation products of cDNA clone pTB114Δ1, which gives a non-processed 200 kD polyprotein due to a deletion in the 24K coding region. The 32K polypeptide was freed from this aberrant 200 kD polyprotein but did not stimulate the cleavage of the glutamine-methionine site in the M polyproteins (data not shown). Probably the 32K polypeptide has to remain associated with the 170K polyprotein after its release from the 200K polyprotein and enhances the glutamine-methionine cleavage in the M polyprotein only in this conformation. This conforms to previous data indicating that the 32K polypeptide occurs in infected cells in complexes with the B RNA encoded 170K and 84K polypeptides (Franssen et al., 1984b).

When the M RNA-encoded polyproteins are incubated with the translations products of cDNA clones producing the 24K protease but lacking the 32K polypeptide cleavage takes place predominantly at the glutamine-glycine site between the two capsid proteins. When the 32K polypeptide is also present cleavage takes place predominantly at the glutamine-methionine site

and even after long incubation free capsid proteins can hardly be detected. Again when the M polyproteins are incubated with extracts from B infected cowpea protoplasts only the glutamine-methionine site seems to be attacked (Franssen et al., 1982). Probably the 32K polypeptide has a regulatory function in processing of the M polyproteins and directs in its presence cleavage to the glutamine-methionine site while in its absence the glutamine-glycine site is cleaved.

While the cleavages in the M RNA encoded polyproteins must occur intermolecular, the cleavages in B polyprotein may occur either intramolecular or intermolecular. Time course translations have demonstrated that the glutamine-serine dipeptide sequence releasing the 32K polypeptide occurs intramolecular (Franssen et al., 1984a). We have studied this issue by following the proteolysis of the translation product of clone pTB114Δ1, having a deletion in the 24K domain, by the translation products of several other cDNA clones (data not shown). The results obtained indicate that the glutamine-serine site releasing the 32K polypeptide and glutamine-glycine site releasing the 87K polypeptide can also be efficiently cleaved in trans. The glutamine-methionine site in the translation product of pTB114Δ1 was not cleaved in trans. This suggests that this cleavage only occurs in cis, although it cannot be excluded that the deletion in the abberant 200 kD polyprotein of pTB114Δ1 influences this cleavage reaction. The 32K polypeptide does not enhance any of the cleavages of the B polyproteins in trans and thus its stimulatory effect seems be limited to the glutamine-methionine cleavage in the M polyproteins.

The question arises in which conformation the 24K polypeptide carries out the cleavage reactions: as a free protein or as part of a larger polypeptide. Alternatively different polypeptides containing the 24K domain may have different cleavage specificities. The results presented here do not give a clear answer to these questions, however some observations are noticable in this respect. Firstly, the DNA clones producing the most free 24K polypeptide i.e. pTB114Δ3 and pTMB110 give the best cleavage in trans of the glutamine-glycine sites in both the M and B polyprotein (P. Vos, unpublished results and Figure 4). Secondly the translation products of DNA clone pTB114Δ4 can efficiently cleave the glutamine-methionine site in the M polyprotein indicating that the 84K polypeptide together with the 32K

polypeptide can perform this cleavage reaction. Finally the 84K polypeptide itself is not further cleaved into the 60K and 24K polypeptides by cleavage of the glutamine-methionine site at the left border of the 24K protease, while the 110K polypeptide is efficiently cleaved at the glutamine-glycine site to the right of the 24K protease into this polypeptide and the 87K polypeptide.

The cleavage sites in the CPMV polyproteins are characterized by a specific dipeptide sequence: glutamine-glycine, glutamine-serine or glutamine-methionine (Wellink et al., 1986). Apart from this dipeptide sequence the folding of the protein at the position of the dipeptide sequence is probably very important and also surrounding amino acids could play a role. For CPMV there is a weak consensus sequence for the amino acids surrounding a cleavage site having an alanine at position -4 and an alanine or proline at position -2 (Wellink et al., 1986).

In this paper we have investigated the effect of several amino acid changes at the glutamine-glycine dipeptide sequence between the two capsid proteins. The results obtained are very interesting since changing of the glutamine-glycine dipeptide sequence into either of the two other employed dipeptide sequences, glutamine-serine or glutamine-methionine, resulted in a dramatic decrease of cleavage efficiency. On the other hand the insertion of a basic amino acid i.e. an arginine next to the cleavage site did not have a significant effect on the cleavage efficiency. Furthermore changing the glutamine-glycine into glutamine-alanine did decrease the cleavage efficiency but did not completely block the cleavage reaction. The results obtained are very interesting but demonstrate that a detailed biochemical analysis of the cleavage reactions and investigation of the three-dimensional structure of the polypeptide domains surrounding the cleavage sites must supplement molecular genetic studies for complete understanding of the proteolytic processing reactions.

4. Material and Methods

Molecular cloning

Plasmid DNA isolations on large and small scale were performed by the alkaline lysis method (Birnboim and Doly, 1979). Alle enzymes were purchased from Boehringer Mannheim B.V. or New England Biolabs Corp. and used according to the manufacturers indications. DNA fragments were joined by T4 DNA ligase and transformed into competent E. coli HB101 as described previously (Van Wezenbeek et al., 1983). DNA fragments used for ligation were generally purified by extraction from agarose gels by a modified freeze-squeeze method (Tautz and Renz, 1983) or by elution from polyacrylamide gel slices (Maxam and Gilbert, 1977). Recombinant DNA clones were analyzed by restriction enzym mapping and if necessary nucleotide sequence analysis (Sanger et al., 1977, 1980). Oligonucleotide-directed mutagenesis using single-stranded recombinant M13 DNA was performed essentially as described by Zoller and Smith (1984).

In vitro transcription and translation

In vitro transcriptions using T7 RNA polymerase were performed as described previously (Verver et al., 1987) and the RNA was recovered from the reaction mixture by precipitation overnight with 2M LiCl at 4°C. Rabbit reticulocyte lysates were purchased from Green Hectares (Oregon, Wisconsin) and made messenger RNA dependent by treatment with micrococcal nuclease as described by Jackson (1983). 0.5 µg of RNA transcript was translated in a volume of 10 µl using ^{35}S-methionine (New England Nuclear Corp.) as radioactive amino acid (Goldbach et al., 1981). In vitro translation products were analysed on 12.5% SDS-polyacrylamide gels (Franssen et al., 1982).

Construction of modified cDNA clones

cDNA clone pTB114Δ1 was constructed by exchanging the KpnI-BamH1-fragment (positions 3134 to 3857 of B RNA) in clone pTB114 for the corresponding fragment of clone pTB3Δ1 (Verver et al., 1987).

For construction of cDNA clone pTB114Δ2 a PstI-XbaI fragment (positions 346 to 899 of B RNA) was inserted into pUC19 linearized with PstI and XbaI, thereby removing the unique AccI-site in this vector. From the recombinant clone obtained a 375 bp AccI fragment (corresponding to position 360-735 of B RNA) was removed by digestion with AccI and subsequent religation. The PstI-XbaI fragment with de deleted AccI-fragment was isolated from this clone and exchanged for the corresponding fragment in the full-size cDNA clone of B RNA pTB114, generating pTB114Δ2.

For construction of clone pTB114Δ3 first a HaeIII-EcoRI fragment (positions 2815 to 3865 of B RNA) was inserted into pUC9 digested with SmaI and EcoRI, introducing several unique restriction sites in front of the HaeIII-site. The DNA fragment was then recovered from this clone by digestion with PstI (polylinker region) and BamH1 (position 3857 of B RNA) and inserted into PstI (position 346 of B RNA) and BamH1 digested pTB114. Due to the construction nucleotides 346 to 2815 of B RNA were removed and replaced by 14 nucleotides of the multible cloning site of pUC9, however, the open reading frame was not disturbed.

cDNA clone pTB114Δ4 was constructed by deletion of a PvuII-HpaI fragment (positions 3700 to 5661 of B RNA) from pTB114. pTB114 DNA was partially digested with PvuII and completely digested with HpaI. A 6.4 kb DNA fragment containing the desired sequences was isolated and the DNA religated generating clone pTB114Δ4. Due to the construction the reading frame in this mutant downstream the PvuII-site was disturbed resulting in a translation stop 18 amino acids downstream of this site.

For constructing of cDNA clones pTMB110 and pTMB120 the BamH1-ClaI fragment of pTB114Δ3 (postions 3857 to poly(A) tail of B RNA) was replaced in case of pTMB110 for a BamH1-ClaI fragment of pTM203 (positions 1504 to 3' end tail of M RNA) and in case of pTMB120 for a PvuII-ClaI fragment of pTM203 (positions 699 to 3' end of M RNA) (Vos et al., submitted). In the latter case the BamH1-site of pTB114Δ3 was "filled in" prior to ligation to the PvuII-ClaI fragment of the cDNA of M RNA. The constructions were designed in such a way that long open reading frames were formed expected to result upon translation in chimaeric proteins containing amino acid sequences of M and B polypeptides.

cDNA clone pTB32* was constructed using a specific oligonucleotide designed to change the glutamine-serine dipeptide sequence in a glutamine-stop codon sequence. This oligonucleotide, a 22-mer with the sequence 5-CAAUGCACAGTGAGCTCAGAGA-3 is homologous for the first 10 nucleotides to positions 1175 to 1187 of B RNA and for the last 10 nucleotides to positions 1340 to 1350 of B RNA. Nucleotides 11, 12 and 13 of the oligonucleotide do not hybridize to B specific sequences and are necessary to change the AGT serine codon into a TGA stop codon. Apart from creating a stop codon the oligonucleotide should create a specific deletion in the B RNA specific cDNA in the M13 clone. As a result immediately downstream the TGA stop codon a GAGCTC SstI restriction site is created.

A 726 bp XbaI-PvuII fragment (positions 899 to 1625 of B RNA) was inserted into M13mp10 previously digested with XbaI and SmaI. The recombinant single-stranded DNA was annealed to the oligonucleotide and the complementary strand was synthesized (Zoller and Smith, 1982). The reaction mixture was transformed into competent E. coli JM103 cells and desired mutants were selected by selective hybridization to the oligonucleotide (Zoller and Smith, 1982). Finally clone pTB32* was constructed via a three-point ligation employing a 3.3 kb ClaI-XbaI fragment from pTB114 containing vector sequences the T7 promoter and nucleotides 1 to 899 of B cDNA, the 292 bp XbaI-SstI fragment from the mutagenized M13 clone and a SstI-ClaI fragment of approximately 200 bp containing the 3' non-translated region and poly d(A-T) track of the M cDNA clone pTM203.

Constructing of derivates of clone pTMB110 modified at the glutamine-glycine dipeptide sequence between the two capsid proteins

For manipulations with the coding sequence of the glutamine-glycine dipeptide sequence a DNA fragment from a full-length cDNA clone of M RNA was inserted into M13mp10. This recombinant M13 clone, called M13EC1, contained a HindIII-SstI fragment corresponding to positions 2576 to 3428 of M RNA. After mutagenesis RF DNA of the engineered clone was isolated and the mutant cDNA fragment exchanged for the wild type DNA fragment in clone pTMB110.

For construction of clones pTMB113 and pTMB114 RF DNA was isolated of M13 clone M13EC1 and partially digested with AvaII. M13EC1 contains two AvaII-sites, one in the vector sequence and another in the cDNA fragment, the latter one at the position of the glutamine-glycine dipeptide sequence. Linearized M13EC1 was isolated from gel and the protruding ends were either filled in with the Klenow fragment of E. coli DNA polymerase I or digested with S1 nuclease. Subsequently the DNA obtained in this way was religated and transformed into competent E. coli JM103 cells. Recombinant phages were selected by dideoxy sequencing. Filling-in of the AvaII-site alters the glutamine-glycine-proline sequence into glutamin-glycine-arginine-proline i.e. CAA-GGA-CCT into CAA-GGA-CGA-CCT. Digestion with S1-nuclease of the AvaII-site alters the glutamine-glycine-proline sequence into glutamine-alanine i.e. CAA-GGA-CCT into CAA-GCT. The DNA fragment containing 3 additional nucleotides was transferred to pTMB110 resulting in pTMB114, the DNA fragment missing 3 nucleotides was transferred to pTMB110 resulting in pTMB113.

For construction of clones pTMB115 and TMB116 single-stranded DNA was isolated of M13EC1 and annealed to specific oligonucleotides. The complementary strand was synthetized using the Klenow fragment of E. coli DNA polymerase I in the presence of T4 DNA ligase to repair the single-stranded gap in the full-length complementary DNA annealed to the circular template DNA. The reaction mixture was directly transformed into competent E. coli JM103 cells and mutagenized phages were selected by hybridization with the oligonucleotides. The oligonucleotide for construction of pTMB115 was a 24-mer with the sequence 5'-CTATAGCACAAATGCCTGTTTGTG-3', changing the glutamine-glycine sequence into glutamine-methionine, the oligonucleotide for construction of pTMB116 was a 24-mer with the sequence 5'-CTATAGCACAATCACCTGTTTGTG-3' changing the glutamine-glycine sequence into glutamine-serine.

5. Acknowledgements

We thank Jacques van Boom and Unilever Research Laboratories for synthesis of oligodeoxynucleotides, Henk Franssen for critically reading the manuscript, Richard Jackson for advise in in vitro translation and processing studies, Piet Madern for artwork and Gré Heitkönig for typing the manuscript. This work was supported by the Netherlands Foundation for Chemical Research (S.O.N.) with financial aid from the Netherlands Organisation for the Advancement of Pure Research (Z.W.O.).

6. References

Birnboim, C. and Doly, J. (1979). Nucleic Acids Res. 7, 1513-1523.
Daubert, S.D., Bruening, G. and Najarian, R.C. (1978). Eur. J. Biochem. 92, 45-51.
El Manna, M. and Bruening, G. (1973). Virology 56, 198-206.
Franssen, H., Goldbach, R., Broekhuijsen, M., Moerman, M. and Van Kammen, A. (1982). J. Virol. 41, 8-17.
Franssen, H., Goldbach, R. and Van Kammen, A. (1984a). Virus Research 1, 39-49.
Franssen, H.J., Moerman, M., Rezelman, G. and Goldbach, R. (1984b). J. Virol. 50, 183-190.
Garcia, J.A., Schrijvers, L., Tan, A., Vos, P., Wellink, J. and Goldbach, R. (1987) Virology, in press.
Goldbach, R.W., Schilthuis, J.G. and Van Kammen, A. (1981). Biophys. Res. Commun. 99, 89-94.
Goldbach, R. and Krijt, J. (1982). J. Virol. 43, 1151-1154
Goldbach, R. and Rezelman, G. (1983). J. Virol. 46, 614-619.
Goldbach, R.W. and Van Kammen, A. (1985). CRC Press, Boca Raton, Florida.
Goldbach, R.W. (1987). Comoviruses: molecular biology and replication, in: The Plant Virus: viruses with bipartite RNA genomes and isometric particles (Harrison, B.D. and Murant A.F., eds) Plenum Press, New York.
Kozak, M. (1983). Microbiol. Rev. 47, 1-45
Lomonossoff, G. and Shanks, M. (1983). EMBO J. 2, 2253-2258.
Maxam, A.M. and Gilbert, W. (1980). Academic Press, Inc., New York.
Rezelman, G., Goldbach, R. and Van Kammen, A. (1980). J. Virol. 36, 366-373.
Sanger, F., Nicklen, S. and Coulson, A.R. (1977). Proc. Natl. Acad. Sci. USA 74, 5463-5467.
Sanger, F., Coulson, A.R., Barrel, B.G., Smith, A.J.H., Roe, B.A. (1980). J. Mol. Biol. 143, 161-178.
Stanley, J., Rottier, P., Davies, J.W., Zabel, P. and Van Kammen, A. (1978). Nucleic Acids Res. 5, 4505-4522.
Tautz, D. and Renz, M. (1983). Anal. Biochem. 132, 14-19.
Verver, J., Goldbach, R., Garcia, J. and Vos, P. (1987). EMBO J. 6, 549-554.
Vos, P., Verver, J., Van Wezenbeek, P., Van Kammen, A. and Goldbach, R. (1984). EMBO J. 3, 3049-3053.
Vos, P., Jaegle, M., Wellink, J., Verver, J., Eggen, R., Van Kammen, A. and Goldbach, R. Infectious RNA transcripts derived from full-length DNA copies of the genomic RNAs of cowpea mosaic virus, submitted for publication.
Wellink, J., Rezelman, G., Goldbach, R. and Beyreuther, K. (1986). J. Virol. 59, 50-58.
Van Wezenbeek, P., Verver, J., Harmsen, J., Vos, P. and Van Kammen, A. (1983). EMBO J. 2, 941-946.
Zoller, M.J. and Smith, M. (1982). Nucleic Acids Res. 10, 6487-6500.

CHAPTER 7

CONCLUDING REMARKS

Concluding remarks

The experiments described in this thesis were designed to unravel various aspects of the mechanism of gene expression of cowpea mosaic virus (CPMV). For this purpose full-length DNA copies of both genomic RNAs of CPMV were constructed. Using powerful *in vitro* transcription systems RNA transcripts closely resembling the viral RNAs were prepared from these clones, which were efficiently translated *in vitro* and were able to infect cowpea protoplasts. Mutations were introduced in specific regions in the cDNA clones using methods for site-direced mutagenesis. Investigation of the effect of these mutations on viral gene expression has increased our insight in the multiplication cycle of CPMV.

Most of the mutagenized cDNA clones described in this thesis were designed to study the proteolytic processing of the CPMV polyproteins. The proteolytic processing of CPMV lends itself very well to be studied *in vitro* because it reliably mimicks the processing *in vivo* upon translation of the viral RNAs in rabbit reticulocyte lysates (Pelham, 1978; Franssen *et al*., 1984). The results obtained here clearly indicate that the B RNA encoded 24K polypeptides is the proteinase involved in all cleavage reactions in the CPMV polyproteins. The question remains however how to processing of the CPMV polyproteins is regulated *in vivo*. Some cleavage reactions which occur slowly *in vitro* occur so rapidly *in vivo*, that the precursor polypeptides cannot be detected. In the proteolytic processing of the M RNA encoded polyproteins the 32K polypeptide also plays a role, because in its absence the glutamine-methionine site is cleaved very oorly by the 24K proteinase (chapter 6). Furthermore it is not clear whether the 24K polypeptide is proteolytically active in infected cells in the form one or mroe of the precursor polypeptides: the 170K, 110K and 84K polypeptides or as a free protein. It may therefore be expected that the proteoytic cleavages *in vivo* are strictly regulated depending on the different processes in the life cycle of the virus, like RNA multiplication and virus assembly.

The only way to gain more insight in the regulation of the proteolytic cleavage reactions is to study the expression of specifically modified cDNA

clones *in vivo*. In view of that the expression of RNA transcripts of full-length cDNA clones in cowpea protoplasts is very promising. The expression of such transcripts in protoplasts was successful, but their infectivity in comparison to viral RNA was rather low and therefore the usefulness of these transcripts at this moment is limited. Besides the infection of whole cowpea plants could not be achieved. It is obvious that the infectivity of the RNA transcripts needs to be improved to increase their applicability. The difference in infectivity between viral RNA and the *in vitro* sunthesized RNA msut originate from the different properties of the RNA molecules. Compared to viral RNA the *in vitro* synthesized RNAs lack the genome-linked protein VPg and have two additional nucleotides at the 5' end and 4 to 5 extra nucleotides to the 3' end of the RNA. The *in vitro* synthesized RNA transcripts of full-length cDNA clones of tobacco mosaic virus (TMV) and bromo mosaic virus (BMV) have a relatively high level of infectivity (Dawson *et al.*, 1986; Meshi *et al.*, 1986; Ahlquist *et al.*, 1984), but these transcripts have exactly the same 5' end as the viral RNAs; i.e. a cap structure. Using one of the currently known *in vitro* transcription systems it seems difficult to synthesize *in vitro* transcripts from DNA copies of CPMV RNA, which have exactly the same 5' end as the viral RNAs. It may be hoped that the infectivity of the transcripts will increase if the number of extra nucleotides at the 5' end is reduced, but the synthesis of such transcripts will most likely be less efficient (Kang and Wu, 1987; Dunn and Studier, 1983). It is also possible that the extra nucleotides at the 3' end affect the infectivity, although in case of other viruses for which the *in vitro* transcription of cDNA clones has been reported sofar a limited number of extra nucleotides added to the 3' end did not significantly influence infectivity (Ahlquist *et al.*, 1984; Van der Werf *et al.*, 1986; Dasmahapatra *et al.*, 1986; Meshi *et al.*, 1986; Dawson *et al.*, 1986). It should be possible to construct cDNA clones, which allow the *in vitro* synthesis of RNA transcripts terminating in a poly(A) tail without extra nucleotides at the 3' end using a restriction enzym for linearization of the template DNA, which restricts the DNA upstream of its recogntion sequence like e.g. BspM1.

The possibilities for investigating the various aspects of the multiplication cycle of CPMV and of virus-host interaction will increase signifi-

cantly when RNA transcripts can be synthesized which are much more infectious than the present ones. Since B RNA replicates independently in cowpea protoplasts because it encodes all proteins necessary for replication of the viral RNAs and proteolytic processing of the polyproteins, various mutations can be introduced in M RNA transcripts without affecting the multiplication and expression of B RNA. The infection of cowpea protoplasts with B RNA transcripts together with specifically modified M RNA transcripts may lead to a better understanding of the proteolytic processing of the M RNA encoded polyproteins *in vivo* and to the elucidation of sequences involved in viral RNA replication.

References

Ahlquist, P., French, R., Janda, M. and Loesch-Fries, L.S. (1984) Proc. Natl. Acad. Sci. USA **81**, 7066-7070.

Dasmahapatra, B., Dasgupta, R., Saunders, K., Selling, B., Gallagher, T. and Kaesberg, P. (1986) Proc. Natl. Acad. Sci USA **83**, 63-66.

Dawson, W.O., Beck, D.L., Knorr, D.A. and Grantham, G.L. (1986) Proc. Natl. Acad. Sci. USA **83**, 1832-1836.

Dunn, J.J. and Studier, F.W. (1983) J. Mol. Biol. **166**, 477-535.

Franssen, H., Goldbach, R. and Van Kammen, A. (1984) Virus Research **1**, 39-49.

Kang, C. and Wu, C-W. (1987) Nucleic Acids Res. **15**, 2279-2295.

Meshi, T., Ishikawa, M., Motoyoshi, F., Semba, K. and Okada, Y. (1986) Proc. Natl. Acad. Sci. USA **83**, 5043-5047.

Pelham, H.R.B. (1978) Eur. J. Biochem. **85**, 457-462.

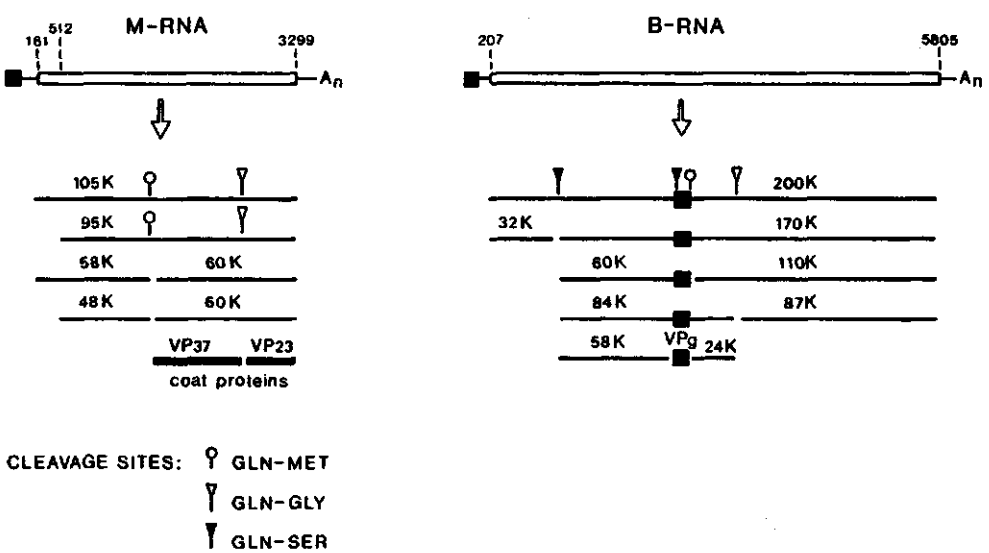
Van der Werf, S., Bradley, J., Wimmer, E., Studier, F.W. and Dunn, J.J. (1986) Proc. Natl. Acad. Sci. USA **83**, 2330-2334.

CHAPTER 8

SAMENVATTING

Samenvatting

Onderwerp van het hier beschreven onderzoek was het cowpea mozaïek virus (CPMV), een plantevirus, waarvan de genetische informatie verdeeld is over twee enkelstrengs RNA-moleculen, M- en B-RNA. Deze RNAs bezitten een klein eiwit, VPg genoemd, covalent gebonden aan het 5'-uiteinde en een poly(A) staart aan het 3'-uiteinde en kunnen direct als boodschapper-RNAs fungeren. De beide RNAs worden vertaald in grote polyproteinen, die vervolgens worden gesplitst in kleinere functionele eiwitten door proteolyse van specifieke dipeptide sequenties (zie Figuur 1).



Figuur 1. Schema voor de vertaling van de RNAs van CPMV en proteolytische processing van de CPMV polyproteinen. De dubbele lijnen in het RNA geven de lange open leesramen aan; stop- en startcodons zijn ook aangegeven. VPg is aangeduid met een zwart vierkantje, andere eiwitten door enkele lijnen. De verschillende klievingsplaatsen op de polyproteïnen zijn aangegeven.

Het B RNA wordt vertaald in een 200 kilodalton (kD) polyproteine, dat vrijwel onmiddellijk wordt gesplitst in twee eiwitten van respectievelijk 32 kD en 170 kD. De overige B RNA specifieke eiwitten ontstaan allen uit het 170 kD eiwit (zie Figuur 1). Het M RNA wordt vertaald in twee elkaar

overlappende polyproteinen van respectievelijk 105 en 95 kD met een zelfde C- terminus maar een verschillende N- terminus. Van deze grote polyproteinen wordt eerst de 60 kD manteleiwitprecursor vrijgemaakt, die vervolgens wordt gesplitst in de capsideëiwitten VP37 en VP23.

Het B RNA is in staat onafhankelijk van het M RNA te repliceren en zijn genetische informatie tot expressie te brengen in cowpea protoplasten. Voor systemische infectie van intacte planten, waarbij het transport van cel tot cel essentieel is, moet het M RNA ook aanwezig zijn. Er is dus een duidelijke verdeling van gegetische functies tussen de beide RNAs: het B RNA bevat de informatie voor de virale RNA replicatie en voor de proteolytische splitsing van de virale polyproteinen en het M RNA codeert voor de beide manteleiwitten en mogelijk voor een transportfunctie.

De bijzondere manier van genexpressie en de duidelijke verdeling van functies tussen de beide RNAs maakt het CPMV een interessant onderzoeksobject. Daarnaast is CPMV ook evolutionair gezien zeer interessant, aangezien het sterk verwant is aan de dierlijke picornavirussen (poliovirus, mond- en klauwzeer en het verkoudheidsvirus zijn hiervan bekende vertegenwoordigers). Mogelijk zijn al deze virussen uit een gemeenschappelijk voorouder ontstaan.

Het doel van de experimenten, die in dit proefschrift zijn beschreven, was het inzicht in de levenscyclus van CPMV te vergroten. Het idee was om de genetische expressie van virale cDNA klonen te bestuderen, die op specifieke plaatsen veranderd waren. Voor DNA zijn verschillende methodes beschreven voor het introduceren van specifieke mutaties, terwijl het niet mogelijk is om gericht mutaties in RNA aan te brengen. Voorwaarde was echter wel dat het mogelijk zou zijn om op de één of andere manier expressie van virale DNA-kopiën te bewerkstelligen.

In eerste instantie werd begonnen met de konstruktie van een volledige DNA-kopie van de kleinste van de twee virale RNA's, het M-RNA, omdat bij de aanvang van het onderzoek hier een groot aantal onvolledige cDNA-klonen van aanwezig was, die gezamenlijk de complete cDNA-sequentie bevatten (Hoofdstuk 3). Stap voor stap werd van deze subklonen een volledige cDNA-kloon van het M-RNA gekonstrueerd in plasmidevector pBR322. Omdat gebleken was dat voor het verwante diervirus poliovirus dubbelstrengs DNA-kopieën infectie van gastheercellen konden bewerkstelligen, is dit ook

voor de cDNA-kloon van het M-RNA van CPMV geprobeerd. Net zoals bij andere planten-RNA-virussen echter, waren de DNA kopieën van CPMV zelf niet direct "infectieus", dat wil zeggen dat bij inoculatie van planten of van suspensies van bladcelprotoplasten met de DNA-kopieën geen virus vermenigvuldiging plaatsvond.

Tijdens het onderzoek kwamen efficiënte in vitro transcriptiesystemen beschikbaar, die het mogelijk maakten om met behulp van bacteriofaag polymerases in vitro µg hoeveelheden enkelstrengs RNA over te schrijven van gekloneerd dubbelstrengs DNA. Een dergelijke in vitro transcriptie systeem, gebaseerd op de promoter en het RNA-polymerase van bacteriofaag SP6, is gebruikt om RNA-transcripten te maken van een volledige cDNA-kloon van M-RNA. De aldus verkregen transcripten bevatten de volledige sequentie van M-RNA en bovendien nog een gering aantal extra nucleotiden aan zowel het 5'- als 3'-uiteinde. Daarnaast misten deze transcripten het VPg, het kleine eiwit dat covalent gebonden is aan het 5'-uiteinde van de virale RNA's. Infectieproeven met deze RNA-transcripten waren helaas niet succesvol, maar de transcripten bleken wel efficiënte boodschapper-RNAs in een in vitro translatiesysteem en gaven daarin dezelfde translatieproducten als viraal M RNA. Daardoor werd het mogelijk om bepaalde aspecten van de expressie van de virale RNA's, die zich leenden om in vitro te worden bestudeerd, nader te onderzoeken. Aldus werd aangetoond dat de 105 kD en 95 kD polyproteinen ontstaan door initiatie van eiwitsynthese op verschillende startcodons (hoofdstuk 3).

Net zoals van M-RNA werd vervolgens ook van B-RNA een volledige DNA-kopie gekonstrueerd en gekloneerd in een plasmide achter een SP6 promoter sequentie. Bovendien werd van deze kloon een volledige cDNA-kloon afgeleid met een bacteriofaag T7 promoter in plaats van SP6 promoter. Met behulp van SP6 of T7 RNA polymerase konden van deze kloons RNA-transcripten worden gesynthetiseerd die een volledige B-RNA sequentie bevatten. Ook deze RNA-transcripten werden in vitro efficiënt vertaald in eiwitten, die identiek waren aan de translatieprodukten van viraal B-RNA (Hoofdstuk 4). Door de in vitro expressie van specifieke deletiemutanten van de volledige B cDNA-kloon werd veel nieuwe informatie verkregen over de proteolytische processing van de CPMV polyproteinen. Al eerder was aangetoond dat de proteolytische processing van de virale polyproteinen in een celvrij transla-

tiesysteem vergelijkbaar is met de processing in vivo, zodat deze processing bij uitstek geschikt is om in vitro te bestuderen. Deze specifieke proteolyse vindt plaats op drie verschillende soorten knipplaatsen. Nader onderzoek leerde dat het B-RNA gecodeerde 24 kD eiwit verantwoordelijk is voor alle splitsingen in het B polyproteine (hoofdstuk 6). Dit roept vragen op betreffende de regulatie van de processing. Het is niet uitgesloten dat het 24 kD eiwit in verschillende configuraties (vrij eiwit, 84 kD, 110 kD) verschillende specifieke activiteiten bezit. Door de in vitro expressie van een hybride cDNA-kloon met de genetische informatie van het B-RNA gecodeerde 24K protease en de M-RNA gecodeerde manteleiwitten kon worden aangetoond dat de beide splitsingen in het M polyproteine op respectievelijk een glutamine-methionine en glutamine-glycine plaats ook door het 24K eiwit worden gekatalyseerd (hoofdstuk 6). Bij de klieving van de M RNA gecodeerde polyproteinen door de translatieprodukten van verschillende mutante B cDNA-klonen bleek echter dat voor efficiënte splitsing van de glutamine-methionine knipplaats "in trans" de aanwezigheid van het B RNA gecodeerde 32 kD polypeptide essentieel is. Dit eiwit werd tot voor kort aangezien voor het protease dat deze glutamine-methionine dipeptide sequentie splitste, gebaseerd op remmingsproeven met antilichamen tegen verschillende B-RNA gecodeerde eiwitten. Het is nu gebleken dat het 32 kD eiwit wel bij deze splitsing betrokken is, maar dat het 24 kD eiwit toch verantwoordelijk is voor de proteolytische splitsing op de plaats van deze dipeptide sequentie. Het 32 kD eiwit blijft waarschijnlijk nadat het 200 kD polyproteine gesplitst is in het 170 kD en 32 kD eiwit gecomplexeerd aan het 170 kD of na verdere splitsing aan het 84 kD eiwit. Deze eiwitcomplexen voeren de glutamine-methionine splitsing veel efficiënter uit dan het 170 kD of het 84 kD eiwit alleen.

Nadat diverse aspecten van de genetische expressie van CPMV in vitro waren bestudeerd, werd gestreefd naar de constructie van "infectieuze" cDNA-klonen om ook in vivo mutanten van het virus te kunnen onderzoeken. In hoofdstuk 5 wordt beschreven hoe uiteindelijk cDNA klonen zijn gemaakt, waarvan infectieuze RNA-transcripten konden worden verkregen. Deze RNA-moleculen verschilden van de eerder genoemde niet-infectieuze RNA-transcripten door veel minder extra nucleotiden aan het 5'- en 3'-uiteinde: twee extra G-residuen aan het 5'-uiteinde en 4 à 5 extra nucleotiden achter

de poly(A) staart. Ze waren echter toch nog een factor 100 minder infectieus dan RNA geïsoleerd uit virusdeeltjes. Deze lage infectiositeit wordt waarschijnlijk veroorzaakt door het verschil, hoe gering ook, in de structuur van viraal RNA in vergelijking met *in vitro* RNA. Dit verschil bestaat naast de extra nucleotiden ook uit de afwezigheid van VPg in de *in vitro* gesynthetiseerde RNAs.

De constructie van de hier beschreven infectieuze cDNA-klonen is veelbelovend en laat zien dat van een plante-RNA-virus met een VPg infectieuze klonen kunnen worden gemaakt. Voorlopig is de infectiositeit van de transcripten van deze klonen te laag om de expressie van cDNA-klonen met specifieke mutaties *in vivo* te bestuderen. Het is daarom van het grootste belang om cDNA-klonen te konstrueren, die transcripten opleveren met een hogere infectiositeit. Als dit lukt, betekent dat een aanzienlijke uitbreiding van de mogelijkheden om onze kennis over de levenscyclus van CPMV te vergroten.

Curriculum vitae

Pieter Vos werd geboren op 2 mei 1958 te Geldrop. Tot zijn achttiende jaar woonde hij in Eindhoven waar hij in 1976 het diploma VWO B haalde aan het Van der Putt Lyceum. In hetzelfde jaar werd begonnen met de studie Plantenziektenkunde aan de Landbouwuniversiteit te Wageningen.
In september 1981 werd het kandidaatsdiploma in deze studierichting behaald en in februari 1983 slaagde hij cum laude voor het doctoraalexamen. Uit de drie gekozen hoofdvakken, Erfelijkheidsleer (Prof. Van der Veen), Virologie (Prof. Van der Want) en Moleculaire Biologie (Prof. Van Kammen) blijkt onmiddellijk de bijzondere interesse voor alles wat met DNA, RNA en genexpressie te maken heeft.
Van 1 februari 1983 tot 31 januari 1987 was hij werkzaam op de vakgroep Moleculaire Biologie van de Landbouwuniversiteit, middels een vierjarig contract van de Stichting Scheikundig Onderzoek in Nederland gefinancierd door de Nederlandse Organisatie voor Zuiver Wetenschappelijk Onderzoek. Dit proefschrift rapporteert over de resultaten behaald tijdens deze periode.
Vanaf 1 februari 1987 is hij werkzaam op het Nederlands Instituut voor Zuivelonderzoek te Ede, waar hij onderzoek verricht aan de genetische expressie van melkzuurstreptococcen.