EXPRESSION AND FUNCTION OF CUCUMOVIRAL GENOMES

Thesis submitted for the degree of
Doctor of Philosophy
at the University of Adelaide

by

Bu-Jun Shi (Master of Science)
January, 1997

TO

my supervisor, Professor Bob Symons (right), and

my co-supervisor, Dr Shou-Wei Ding, who made this thesis possible.

TABLE OF CONTENTS

Summary i
Publications iv
Statement v
Acknowledgments vi

Chapter One Review of Cucumoviruses

1.1 Taxonomy of the cucumoviruses 1
1.2 Biological properties 1
1.3 Virion properties 2
1.4 RNA encapsidation 4
1.5 Genome structure, organisation and expression 5
1.5.1 RNA 1 6
1.5.2 RNA 2 7
1.5.3 RNA 3 9
1.5.4 RNA 4 11
1.5.5 RNA 4A 12
1.5.6 RNA 5 and RNA 6 13
1.5.7 Role of untranslated regions 14
1.5.8 Subgenomic promoters 15
1.6 RNA functions 16
1.7 The encoded protein functions 17
1.7.1 Proteins 1a and 2a 18
1.7.2 Protein 2b 19
1.7.3 Protein 3a 19
1.7.4 Coat protein 20
1.8 Satellite RNAs 22
1.9 Defective RNAs 23
1.10 Aims 24

Chapter Two General Materials and Methods

2.1 Materials 25

2.1.1 Synthetic oligodeoxyribonucleotides....25
2.1.2 Nucleotides and radionucleotides....25
2.1.3 Plasmid vectors....25
2.1.4 Bacterial strains (*Escherichia coli*)....26
2.1.5 Molecular weight markers....26
2.1.6 Common growth media....26
2.1.7 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)....27
2.2 Methods....27
2.2.1 Virus sources....27
2.2.2 Storage of viruses....28
2.2.3 Propagation of viruses....28
2.2.4 Purification of viral particles....28
2.2.5 Isolation of RNAs....29
2.2.6 Synthesis of double-stranded cDNA to partial viral RNA genome....30
2.2.7 Agarose gel electrophoresis....31
2.2.8 Isolation of DNA from agarose gels....32
2.2.9 Restriction digestion of DNA....32
2.2.10 Ligation of cDNA into plasmid vectors....32
2.2.11 Transformation of *E. coli*....33
2.2.12 Isolation of plasmid DNA....34
2.2.13 DNA sequencing....35
2.2.14 Direct sequencing of RNA....36
2.2.15 *In vitro* transcription of plasmid clones by RNA Polymerase Preparation of radioactive probes....36
2.2.16 Northern blot hybridisation....36
2.2.17 Removal of protruding 3' termini of restriction enzyme digested DNA....38
2.2.18 Protein expression in *E. coli*....38

Chapter Three *In Vivo* Expression of an Overlapping Gene Encoded by the Cucumoviruses

3.1 Introduction....39
3.2 Materials and methods....40
3.2.1 Isolation of viruses and the RNAs....40
3.2.2 cDNA synthesis, amplification and cloning....40
3.2.3 Northern blot analyses....41
3.2.4 Sequence determination....41

3.2.5 Protein expression in *Escherichia coli* and antibody preparation.....42
3.2.6 Protein extraction from plants and Western immunoblotting.....43
3.3 Results 44
3.3.1 Characterisation of V-TAV RNA 4A.....45
3.3.1.1 Northern blot analyses 45
3.3.1.2 5'-terminal sequence 46
3.3.2 Characterisation of WAII-CMV RNA 4A 46
3.3.2.1 Northern blot analyses 46
3.3.2.2 Partial sequence of RNA 2 and the 5'-terminal sequence of RNA 4A.....47
3.3.3 *In vivo* detection of the protein 2b.....48
3.4 Discussion.....50
3.4.1 RNA 4A may have a signal for encapsidation 50
3.4.2 Protein 2b may be involved in the movement of the virus.....51

Chapter Four Functional Analysis of the 2b Gene by Constructing Hybrid Viruses Based on Q-CMV

4.1 Introduction 53
4.2 Materials and methods 53
4.2.1 Acknowledgments.....53
4.2.2 Plasmid constructs 54
4.2.3 Plant inoculation and virus progeny analysis 55
4.2.4 Preparation and transfection of tobacco protoplasts.....56
4.3 Results 57
4.3.1 Characterisation of the hybrid viruses in the infected tobacco plants.....57
4.3.2 Virulence of interspecies hybrid virus in infected plants.....59
4.3.3 Accumulation of the hybrid virus RNAs in protoplasts.....60
4.3.4 Protein 2b determines the severity of the symptoms 61
4.3.5 Infectivity of the intraspecies hybrid virus 62
4.4 Discussion.....63
4.4.1 Interaction between the heterologous protein 2b and Q-CMV 64
4.4.2 Protein 2b may affect virus movement.....64
4.4.3 Protein 2b may correlate with virulence 65

Chapter Five Construction of the Infectious cDNA Clones of Tomato Aspermy Cucumovirus

5.1 Introduction 67
5.2 Materials and methods 69
5.2.1 Construction of pCass2 vector 69
5.2.2 cDNA syntheses 69
5.2.3 Construction of TAV genomic cDNA clones in pCass2 70
5.2.4 Biological assay 72
5.2.5 Analyses of virus progeny 73
5.3 Results 74
5.3.1 Molecular cloning of TAV genomic cDNAs 74
5.3.2 Biological activity of the cDNA clones 75
5.3.3 5'-terminus of genomic cDNA-derived *in vivo* transcript 76
5.3.4 Comparative infectivity assay 76
5.3.5 Facility studies with the infectious cDNA clones 77
5.4 Discussion 77
5.4.1 Construction of infectious TAV cDNA clones 78
5.4.2 High infectivity by the use of a partially duplicated CaMV 35S promoter 79

Chapter Six Two Novel Subgenomic RNAs Derived from RNA 3 of Tomato Aspermy Cucumovirus That do not Function as mRNAs

6.1 Introduction 80
6.2 Materials and methods 81
6.2.1 Virus sources 81
6.2.2 Purification of RNAs 3B and 5 81
6.2.3 Poly-A tailing of RNAs 3B and 5 82
6.2.4 Enzymatic sequencing of RNAs 3B and 5 83
6.2.5 cDNA synthesis, cloning and sequencing 84
6.2.6 RNA dideoxynucleotide sequencing and primer extension 85
6.2.7 Computer analysis of the sequence 85
6.2.8 Pseudorecombinants between V-TAV and Q-CMV 85
6.2.9 Deletion cDNA clone of V-TAV RNA 3 85
6.3 Results 86
6.3.1 RNA 3B and RNA 5 are both associated with V-TAV 86
6.3.2 RNAs 3B and 5 are subgenomic RNAs derived from RNA 3 87
6.3.3 RNAs 3B and 5 are derived from RNA 3 as shown by genetic studies ... 89

6.4 Discussion....90
6.4.1 Origin of RNAs 3B and 5....90
6.4.2 Possible roles of the tandem repeats....91
6.4.3 RNA 3B and RNA 5 may function as regulation RNAs....92

Chapter Seven General Discussion

7.1 Generation of cucumovirus subgenomic RNAs....93
7.2 Encapsidation of cucumovirus subgenomic RNAs....93
7.3 Functions of cucumovirus subgenomic RNAs and their encoded proteins....96
7.4 Infectious cDNA clones of V-TAV....99
7.5 Future work....100

References....104

SUMMARY

The general aim of this thesis was to characterise subgenomic RNAs of cucumoviruses and the functions of their encoding genes. The genus *Cucumovirus* contains three species, cucumber mosaic virus (CMV), tomato aspermy virus (TAV) and peanut stunt virus (PSV). Strains of CMV are further classified into two major subgroups (I and II) on the basis of nucleotide sequence homology. The viruses from the same subgroup share 92-98% sequence similarities and those from different subgroups share 71-79% sequence similarities. Cucumoviruses contain a single-stranded RNA genome of messenger sense divided into three species, RNAs 1, 2 and 3 that encode four proteins, 1a, 2a, 3a and coat protein (CP). Proteins 1a, 2a and 3a are translated from their respective genomic RNA whereas CP is expressed via RNA 4, a subgenomic RNA of RNA 3. RNA 2 of a subgroup II CMV strain was shown recently to encode an additional 2b protein which is translated from RNA 4A, a subgenomic RNA of RNA 2.

Firstly, the V strain of TAV (V-TAV) and a subgroup I CMV strain (WAII) were chosen to determine whether the 2b genes encoded by these viruses are expressed *in vivo*. Using northern blot and western blot analyses, both RNA 4A and its encoded 2b protein were detected in extracts of plants infected with TAV. RNA 4A was also detected in WAII-CMV infected plants. Considering the fact that ORF 2b is conserved in all cucumoviruses sequenced to date, it is therefore concluded that the 2b gene expression is a common feature of all cucumoviruses. Interestingly, RNA 4A of V-TAV is encapsidated in virions as found previously for the Q strain of subgroup II (Q-CMV) whereas WAII-CMV contains very little RNA 4A in virions, suggesting a varying specificity of *in vivo* encapsidation of WAII-CMV RNAs.

The function of the 2b gene of TAV and WAII-CMV was investigated by replacing the 2b coding sequence of Q-CMV with that of V-TAV and WAII-CMV. Both interspecies and

intraspecies hybrid viruses were significantly more virulent, induced earlier onset of systemic symptoms and accumulated to a higher level in most host species tested than their parental viruses. Mutational analysis of the Q-CMV/V-TAV hybrid virus revealed that it was the 2b protein of V-TAV which interacted synergistically with Q-CMV, thus leading to the increased virulence. Moreover, the synergistic interaction most likely resulted from an increased efficacy of the hybrid virus in systemic spread in infected plants because CMV and the hybrid virus accumulated to a similar level in infected tobacco protoplasts. As a Q-CMV mutant expressing no 2b protein either failed to establish systemic infection or induced essentially no systemic symptoms, the above results also indicate that the 2b gene of both V-TAV and WAII-CMV was functionally similar to the 2b gene of Q-CMV.

A prerequisite for further investigation of the 2b gene function was to construct infectious cDNA clones of three genomic RNAs of V-TAV. A new plasmid vector pCass containing the 35S promoter and terminator from cauliflower mosaic virus had been demonstrated previously in our laboratory to be an effective cloning vector for constructing infectious cDNA clones of Q-CMV. The full-length cDNAs of the three genomic RNAs of V-TAV were cloned into pCass as well as pCass2 which contained a partially duplicated 35S promoter. Both sets of the V-TAV cDNA clones were highly infectious; the host range, symptoms, morphology of viral particles and viral progeny RNAs as obtained using plasmid inocula were identical to those induced by the wildtype virus. Furthermore, the pCass2-based cDNA clones gave a threefold higher infectivity than the pCass-based clones in the lower ranges of inocula concentration, suggesting that the higher level of transcription as predicted from the partially duplicated 35S promoter became critical when only a limited amount of plasmid inocula were present.

V-TAV was found to encapsidate two extra RNAs, designated RNA 3B and RNA 5, in addition to the five known RNAs (RNAs 1, 2, 3, 4 and 4A). The complete nucleotide

sequence of both RNA 3B and RNA 5 were determined; RNAs 3B and 5 were 485 and 323 nucleotides long, respectively. The 3'-terminal 323 nt of RNA 3B was identical in sequence to that of RNA 5, whereas its 5'-terminal 163 nt was a direct repeat (one nt difference) of the 5'-half of RNA 5, and that both RNAs were completely homologous to the 3'-terminal untranslated region of V-TAV RNA 3. In addition, both RNAs 3B and 5 of V-TAV were detected in plants infected with a pseudorecombinant virus consisting of V-TAV RNA 3 and RNAs 1 and 2 from Q-CMV. Using the infectious cDNA clones of V-TAV, a mutant virus containing only one of the two repeats was constructed. While this deletion mutant was infectious, only RNA 5 and not RNA 3B was detected in infected plants. These results provided further strong evidence that both RNA species were derived from V-TAV RNA 3. In contrast to RNAs 4 and 4A, both RNAs 3B and 5 of V-TAV encode no ORF longer than 21 codons and thus are unlikely to function as subgenomic mRNAs.

PUBLICATIONS

(arising from this thesis)

1. Ding, S.W., Shi, B.J., Li, W.X., & Symons, R.H. (1996). An interspecies hybrid RNA virus is significantly more virulent than either parental virus. *Proc. Natl. Acad. Sci. USA* **93**, 7470-7474.

2. Shi, B.J., Ding, S.W., & Symons, R.H. (1996). *In vivo* expression of an overlapping gene encoded by the cucumoviruses. *Journal of General Virology*, In Press.

3. Shi, B.J., Ding, S.W., & Symons, R.H. (1996). Two novel subgenomic RNAs derived from RNA 3 of tomato aspermy cucumovirus. *Journal of General Virology*, In Press.

4. Shi, B.J., Ding, S.W., & Symons, R.H. (1996). Plasmid Vector for Cloning Infectious cDNAs from Plant RNA Viruses: High Infectivity of cDNA Clones of Tomato Aspermy Cucumovirus. *Journal of General Virology*, In Press.

STATEMENT

This thesis contains no material which has been accepted for the award of any other degree or diploma in any University. To the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference has been made in the text.

I consent to this thesis being made available for photocopying and loan.

Bu-Jun Shi

January, 1997

ACKNOWLEDGMENTS

Many people within the Waite Campus have been of great help to me throughout the course of my PhD. It is difficult to acknowledge the full extent of each contribution from all involved, however, I would like to pay special attention and indebted gratitude to my supervisor, Professor Bob Symons, for providing this opportunity and for his guidance, support and encouragement.

Also, I must give special thanks to my co-supervisor, Dr Shou-Wei Ding, for his guidance, encouragement and thesis correction.

Dr Ying-Hong Liu and Wan-Xiang Li must be also thanked for their friendship, encouragement, discussion and assistance.

I am grateful to all members of the Symons' laboratory and the Waite community past and present (in the alphabetical order): Ali Afsharifar, Rob Bennet, Helina Bogacz, Rod Bonfiglioli, Graham Collins, Nick Collins, Sonya Dayan, Andrew Ewart, Geoff Fincher, Chris Ford, Helen Healy, Paul Ingram, Alex Jeffries, Liang-Hui Ji, Brendon King, Chris Kohler, Peter Langridge, Richard Leach, Jing Li, Xin-Min Li, Anthony Mowles, Ernie Nagy, Mandy Pace, Nick Paltridge, John Randles, John Rathjen, Masoud Shams-bakhsh, Neil Shirley, Carole Smith, David Tattersall, Yan Fong Wan Chow Wah, David Warrillow, Daryl Webb, Wendy Winnall for their friendship, support and advice.

I owe a special thanks for an Australian Overseas Postgraduate Research Scholarship (OPRS) and to the Australian Research Council (ARC) Special Research Centre for Basic and Applied Plant Molecular Biology for generous financial support.

Finally, I would like to dedicate this thesis to my immediate and extended families, whose constant support and unlimited encouragement were essential in the completion of this PhD work.

Chapter 1

REVIEW OF CUCUMOVIRUSES

1.1 TAXONOMY OF THE CUCUMOVIRUSES

The genus *Cucumovirus*, first organised by the International Committee on Taxonomy of Viruses (ICTV) in 1971 (Harrison *et al.*, 1971), contains three definitive species, cucumber mosaic virus (CMV), tomato aspermy virus (TAV) and peanut stunt virus (PSV) (Murphy *et al.*, 1995). The properties of this genus have been reviewed by Tolin (1977), Kaper and Waterworth (1981), Francki (1985), Francki and his colleagues (1985), Edwardson and Christie (1991), and Palukaitis and his colleagues (1992).

Many strains of CMV have been described and they differ from each other in symptoms, host range, vector specificity and/or serological relationship (Kaper and Waterworth, 1981). On the basis of serological relationships, peptide mapping of the viral coat protein and nucleic acid hybridisation, most of the known CMV strains fall into two major subgroups, subgroup I and subgroup II (Palukaitis *et al.*, 1992). Incidentally, most subgroup I strains such as Fny-CMV incite severe symptoms in *Nicotiana tabacum* while subgroup II strains such as Q-CMV produce a mild mosaic (Zhang *et al.*, 1994a).

1.2 BIOLOGICAL PROPERTIES

CMV has the largest host range of any viruses. It infects at least 1241 species in 498 genera of 101 families (Edwardson and Christie, 1991). Although the majority of CMV hosts are dicotyledonous species, it also infects a number of monocotyledonous plants such as corn and bananas. The host range of TAV or PSV so far recorded is not as wide as that of CMV. For example, TAV infects about 100 species in 27 families (Hollings and Stone, 1971). This virus is best known in connection with chrysanthemum diseases throughout the world (Kaper and Waterworth, 1981; Shi *et al.*, 1989; 1993; Shi and Li, 1989).

The most common symptom induced by CMV or TAV is mosaic; the other symptoms include symptomless, necrosis, fernleaf, ringspot, flower breaking, seedless fruits and death (Kaper and Waterworth, 1981; Palukaitis *et al.*, 1992; Shi *et al.*, 1989; 1993; Shi and Li, 1992).

The cucumoviruses are readily mechanically transmissible among herbaceous species (Kaper and Waterworth, 1981). However, the natural transmission of the cucumoviruses is mainly by aphids in a nonpersistent manner; CMV is transmissible by 87 species of aphids (Edwardson and Christie, 1991) and TAV by 10 species. CMV and TAV are also seed-borne in some species and transmissible by some dodder species (Francki *et al.*, 1979; Holling and Stone, 1971; Edwardson and Christie, 1991). A more significant means of spread of TAV in chrysanthemum is by cuttings made from infected mother plants (Kaper and Waterworth, 1981; Shi *et al.*, 1989).

1.3 VIRION PROPERTIES

Two methods (Lot *et al.*, 1972; Peden and Symons, 1973) have been commonly used for purification of CMV, which are modifications of that devised by Scott (1963). CMV is usually propagated in tobacco or cucumber (Kaper and Waterworth, 1981) and purified from fresh leaves 1 to 2 weeks after inoculation. TAV has been propagated also in tobacco and purified by methods basically similar to those adopted for CMV (Kaper and Waterworth, 1981). It appears that most strains of TAV are more readily purified than CMV and yield is also higher than CMV (Kaper and Waterworth, 1981).

CMV and TAV are icosahedral. Particle diameters are affected by the type of negative stain (Edwardson and Christie, 1991) and range from 24 (Waterworth *et al.*, 1975) to 42 nm (Francki *et al.*, 1966) but most are about 30 nm. The viruses have hexamer-pentamer clustering of 180 identical protein subunits with T=3 surface lattice symmetry (Finch *et al.*, 1967; Hollings and Stone, 1971). The stability of particles of CMV and TAV largely

depends on RNA-protein interactions (Kaper, 1975). The particles of both viruses can be readily disrupted by low concentrations of sodium dodecyl sulphate (SDS) (Kaper, 1973; Habili and Francki, 1974b) or high concentrations of neutral chloride salts (Kaper *et al.*, 1965; Francki *et al.*, 1966), and are sensitive to pancreatic RNase as well (Habili and Francki, 1974b). Basic lysine residues are critical in the stabilisation of CMV particles (Kaper, 1976). The presence of Mg^{2+} can result in precipitation of CMV particles, whereas EDTA stabilises CMV particles. TAV is opposite to CMV, stabilised by Mg^{2+} and degraded by EDTA (Habili and Francki, 1974b).

Sedimentation coefficients ($S_{20,w}$) of the CMV particles as a homogeneous component range from 92 to 104S (Edwardson and Christie, 1991). The CMV particle weight estimates vary from 5.0 to $6.7x10^6$ (Francki *et al.*, 1979). The particles are negatively charged and have an isoelectric point of 4.75 (Ehara and Mink, 1984). Within experimental error, the physico-chemical parameters of TAV and PSV are basically similar to those of CMV (Kaper and Waterworth, 1981).

CMV strains can be divided by serology into two subgroups as described as above. Habili and Francki (1975), Rao *et al.* (1982), Shi and Li (1992) and Shi *et al.* (1989; 1993) reported that some of CMV strains are serologically related to TAV while others are not. However, such serological relationships between CMV and TAV were not detected using monoclonal antibodies (Wahyuni *et al.*, 1992).

CMV particles contain about 18% RNA (Kaper and Re, 1974). At least seven RNA species are encapsidated in virions of Q-CMV, RNA s 1, 2, 3, 4, 4A, 5 and 6 (Peden and Symons, 1973; Ding *et al.*, 1994) (Fig. 1.1). RNA 1 and RNA 2 are encapsidated separately, whereas one molecule each of RNAs 3 and 4 is encapsidated in the same particles (Lot and Kaper, 1976). In addition, satellite RNAs and defective interfering RNAs were also found in virus particles (see Sections 1.8 and 1.9 in this chapter). Only

RNAs 1, 2 and 3, are necessary for infection (Peden and Symons, 1973; Lot *et al.*, 1974).

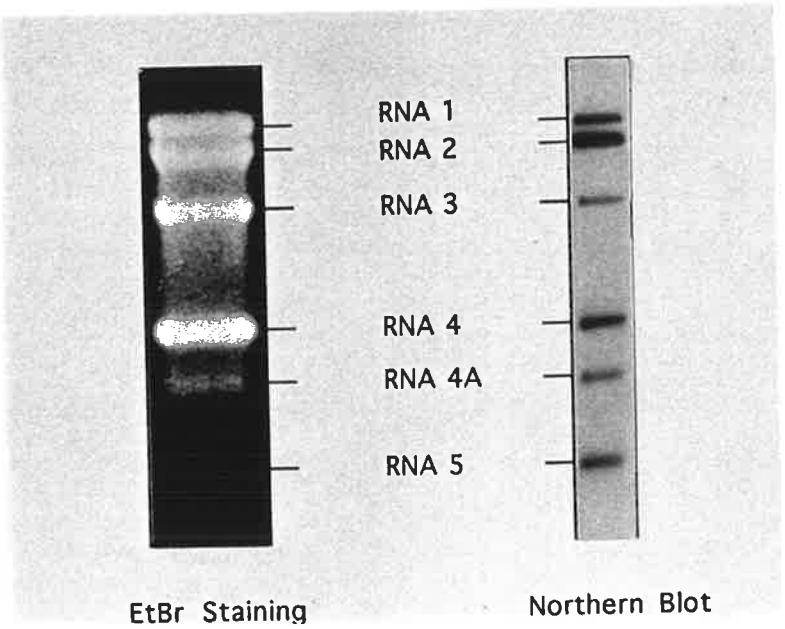


Fig. 1.1 Virion-encapsidated RNAs of Q-CMV

1.4 RNA ENCAPSIDATION

In addition to encapsidating homologous RNAs as mentioned in the above section, cucumoviral coat proteins are able to encapsidate heterologous RNAs such as those of tobacco mosaic tobamovirus (TMV; Chen and Francki, 1990) and turnip yellow mosaic tymovirus (TYMV; Kaper and Geelen, 1971). In return, heterologous virus coat proteins such as alfalfa mosaic virus (AlMV) coat protein (Candelier-Harvey and Hull, 1993) can also encapsidate cucumoviral RNAs. The mechanisms for RNA encapsidation are not well understood. They may be involved in coat proteins-RNA genome interaction (Kaper and Geelen, 1971) through a specific RNA signal. Such an RNA signal has been demonstrated in TMV as a 75-base RNA sequence encompassing a loop 1 formed by the sequence AAGAAGUCG, which is responsible for directing the encapsidation by TMV

coat protein disks of heterologous RNA fragments (Turner *et al.*, 1988). Altering the RNA folding close to the apex of the loop 1 stem reduces the rate of disk binding (Turner *et al.*, 1988), indicating that RNA structure plays an important part in permitting selective protein-RNA recognition.

1.5 GENOME STRUCTURE, ORGANISATION AND EXPRESSION

The cucumoviruses contain tripartite RNA genomes of positive sense (Fig. 1.2).

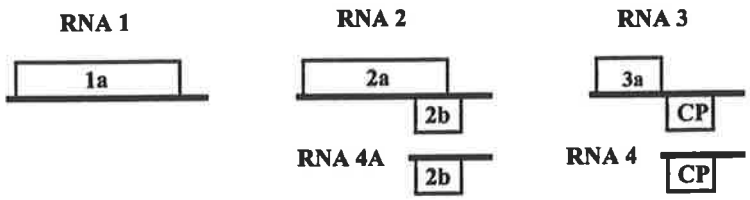


Fig. 1.2 Genome organisation of Q-CMV

The complete genome nucleotide (nt) sequences of five CMV strains from two subgroups have been published; these are strains Q (subgroup II, Rezaian *et al.*, 1984; 1985; Davies and Symons, 1988), Fny (subgroup I, Rizzo and Palukaitis, 1988; 1989; Owen *et al.*, 1990), Y (subgroup I, Kataoka *et al.*, 1990a; 1990b; Nitta *et al.*, 1988), I (subgroup I, Mcgarvey *et al.*, 1995) and NT9 (subgroup I, Hsu *et al.*, 1995). Complete sequence information is also available for two other members of the cucumoviruses, TAV (Bernal *et al.*, 1991; Moriones *et al.*, 1991; O'Reilly *et al.*, 1991) and PSV (Karasawa *et al.*, 1991; 1992).

Each of the genomic RNAs of all cucumoviruses has a 5' cap containing a 7-methylguanosine (Symons, 1975) followed by a guanylate. The 3'-terminal sequence of

about 300 nt in length is highly conserved among the three RNAs. The 3'-terminal 180 nt can be folded into a three-dimensional tRNA-like structure (Wilson and Symons, 1981; Rizzo and Palukaitis, 1989) which accepts tyrosine. The cucumoviral RNAs contain no polyadenylation tail at the 3'-termini.

1.5.1 RNA 1

RNA 1 of all cucumoviruses possesses one long ORF after the first start codon, encoding a polypeptide of about 990 amino acids, which is designated protein 1a. The context for the initiation codon of this putative translation product is the same for all cucumoviruses sequenced, C at position -3 and G at position +4. This differs from the optimal sequence context (A at position -3 and G at position +4) according to Kozak (1984). Codon usage in ORF 1a deviated from random.

The nt sequence comparisons of RNA 1 between cucumoviruses are summarised in Table 1.1.

Table 1.1 Percent sequence identity of RNA 1 and protein 1a between cucumoviruses.

		V-TAV		Fny-CMV		Q-CMV	
		RNA 1	1a	RNA 1	1a	RNA 1	1a
Fny-CMV	RNA 1	61.1					
	1a		72.7				
Q-CMV	RNA 1	58.3		71.4			
	1a		72.5		84.5		
J-PSV	RNA 1	58.3		59.8		60.1	
	1a		69.0		72.6		72.7

(From McGarvey *et al.*, 1995)

RNA 1s of different cucumoviruses share significant homology both at the nt and amino acid levels. The nt identity is closest between Fny-CMV and Q-CMV (71.4%), and the lowest between Q-CMV and V-TAV or between V-TAV and J-PSV (both 58.3%), further confirming that CMV, TAV and PSV are three individual species of the genus *Cucumovirus*. The nt sequence homology of RNA 1 between Fny-CMV and Q-CMV is 71.4%, whereas it is 92-96% between Fny-CMV and two other subgroup I strains, O-CMV and Y-CMV (Palukaitis *et al.*, 1992), further confirming the classification by nt hybridisation.

The homology values for the proteins are slightly higher than those for the RNAs between the cucumoviruses. The regions showing homologies are distributed over the entire molecule in every case (Karasawa *et al.*, 1992). However, three regions in the cucumoviruses, N-terminal (1-220), central (263-494) and C-terminal (570-993), which are separated by two shorter regions (221-262 and 495-569) of low identity (average identity of 32%), are highly homologous (average identity of 79.3%) (Bernal *et al.*, 1991). Hydrophobicity patterns, and the distribution of basic and acidic amino acids, are very similar for V-TAV, Q-CMV and Fny-CMV 1a proteins: basic amino acids predominate in the 150 C-terminal amino acids, acidic ones predominate between residues 500 and 600, and neither type predominates in the rest of the molecule (Bernal *et al.*, 1991).

1.5.2 RNA 2

RNA 2s of Fny-CMV and V-TAV possess one long ORF after the first start codon while RNA 2 of Q-CMV encodes only a 26-nt ORF after the first start codon. A large ORF of Q-CMV RNA 2 appears after the second start codon. V-TAV, Q-CMV and Fny-CMV encode a polypeptide ranging from 829 to 858 amino acids designated the 2a protein within RNA 2s. The context for the initiation codon of this putative translation product in all cucumoviruses compared is also not in the optimal sequence context

(Kozak, 1984). One of two putatively important residues varied between positions -3 and +4. Notably, RNA 2s of all cucumoviruses sequenced to date present an additionally conserved ORF called 2b. The sequence surrounding the initiation AUG codon of protein 2b agrees with the optimal sequence context: an A at position -3 and a G position +4 (Kozak, 1984). The translation product of this ORF 2b has been detected in Q-CMV-infected plants (Ding *et al.*, 1994). Details about the 2b protein and its mRNA are described in Sections 1.8.2 and 1.6.5, respectively. Codon usage in the ORF 2a deviated from random as in the ORF 1a, and codons other than those for Leu, Ile and Val mostly have a U or an A at the third position (Moriones *et al.*, 1991).

The nt sequence comparisons of RNA 2 among cucumoviruses are summarised in Table 1.2.

Table 1.2 Percent sequence identity of RNA 2, proteins 2a and 2b between cucumoviruses.

		V-TAV			Fny-CMV			Q-CMV		
		RNA 2	2a	2b	RNA 2	2a	2b	RNA 2	2a	2b
Fny-CMV	RNA 2	48.3								
	2a		55.9							
	2b			35.1						
Q-CMV	RNA 2	46.6			60.7					
	2a		55.1			73.2				
	2b			36.1			47.5			
J-PSV	RNA 2	44.8			46.2			47.0		
	2a		53.4			53.2			55.3	
	2b			27.1			29.2			19.8

(From McGarvey *et al.*, 1995)

The relationships at the levels of the RNA 2 between different species reveal approximate 46% identity, except for the relationship between Q-CMV and Fny-CMV which shows 60.7% identity, both of which are on the average lower than that between the different RNA 1s. It is unclear whether or not this lower conservation correlates with specific properties of each of the species. The degree of conservation of the RNA 2 sequence within subgroups is higher than those between viruses of different subgroups, indicating that the subgrouping in CMV occurred after the establishment of CMV.

The homology values for the 2a proteins are slightly higher than those for the RNA 2s. The amino acid sequence of TAV 2a protein is shorter than those of Q-CMV and Fny-CMV 2a proteins due to differences at the C ends of the molecule. The regions showing homologies are unevenly distributed in the molecule, being maximal for a central part between amino acids 216 and 757, which shows 71% identity, higher than the other two regions, N- (1-223) and C-terminal (750-857) showing 36 and 23.5% identities, respectively (Moriones *et al.*, 1991). The GDD motif of RNA polymerase is present in the C-terminal half of all cucumovirus 2a proteins. Hydrophobicity patterns of the 2a proteins for TAV, Q-CMV and Fny-CMV were very similar except for the 75 N-terminal residues and the 45 (V-TAV) to 80 (Fny-CMV) C-terminal residues (Moriones *et al.*, 1991), indicating conservative changes for variable positions. Also, the distribution of basis and acidic amino acids in the 2a proteins is the same for TAV, Q-CMV and Fny-CMV (Moriones *et al.*, 1991).

1.5.3 RNA 3

RNA 3 has two large ORFs encoding a putative movement protein, designated 3a, and CP, respectively. RNA 3s of Fny-CMV and Q-CMV possess one long ORF after the first start codon which is also not a strong initiation codon (Kozak, 1984). RNA 3 of V-TAV possesses one long ORF after the third start codon which has an optimal sequence

context: an A at position -3 and an G position +4. The first two AUGs only initiate two shorter ORFs and thereby are unlikely encode any proteins.

Comparisons of cucumovirus RNA 3 sequences are shown in Table 1.3.

Table 1.3 Percent sequence identity of RNA 3, proteins 3a and coat protein (CP) between cucumoviruses.

		C-TAV			Fny-CMV			Q-CMV		
		RNA 3	3a	CP	RNA 3	3a	CP	RNA 3	3a	CP
Fny-CMV	RNA 3	45.4								
	3a		65.1							
	CP			42.0						
Q-CMV	RNA 3	45.3			66.5					
	3a		64.4			82.9				
	CP			44.3			79.9			
J-PSV	RNA 3	52.7			45.8			45.8		
	3a		70.9			65.0			67.5	
	CP			62.4			48.6			49.1

(From McGarvey *et al.*, 1995)

Considerable nt sequence divergence is observed between any two cucumovirus RNA 3s and is higher than that observed between RNAs 1 of cucumoviruses, but lower than that between RNAs 2 of cucumoviruses. The divergence observed between Q-CMV and Fny-CMV is less than that between other cucumoviruses.

There is little divergence within a CMV subgroup (Palukaitis *et al.*, 1992). The sequence divergence at the amino acids level between any two cucumoviruses was low for both ORFs 3a and CP on RNA 3, but was somewhat lower in the 3a protein than the CP

between any two cucumoviruses. The different divergence rates shown for the 3a and CPs has been explained with recombinational events (O'Reilly *et al.*, 1991).

1.5.4 RNA 4

RNA 4 is a subgenomic RNA generated from the 3'-terminus of RNA 3 (see Fig. 1.2). It is for expression of the 3'-terminal gene of RNA 3s, which is not available for translation by these genomic RNAs according to Kozak (1984). RNA 4, unlike genomic RNAs, is not required for infection (Peden and Symons, 1973). RNA 4 in all CMV strains presented in the database (subgroup I: Japanese Y, Y, L, Fny, Ny, C, M, O, Korean, NT9, As; subgroup II: Kin, Q, Trk, Sn, WL) encodes a CP (Schwinghamer and Symons, 1975; Davies and Symons, 1988) of 218 amino acids while that in four TAV strains (B, P, V, and C) sequenced to date encodes a CP of 217 amino acids, 219 amino acids, 219 amino acids and 230 amino acids, respectively.

The level of amino acid sequence identity is lower between CMV and TAV but higher between subgroup I and subgroup II of CMV than the level of nt sequence identity. The level of either amino acid sequence identity or nt sequence identity between CMV and TAV is much lower than the levels of both amino acid and nt sequence identities between subgroup I and subgroup II of CMV. The nt sequence of the CP ORF is the most conserved within subgroups, but not between subgroups or between CMV and TAV. The N-terminal region contains a cluster of basic amino acids (Palukaitis *et al.*, 1992), which is implicated in protein/RNA interactions in the virion (Harrison, 1984), whereas no functional reason is known for the homology in the C-terminus (Palukaitis *et al.*, 1992).

The 5' untranslated region (UTR) is the most variable region of RNA 4 both within and between the CMV subgroups, whereas the 5' UTR of TAV RNA 4s is highly conserved. The variation between TAV and CMV is similar to that between CMV strains. The 3'

UTR of RNA 4 is also a variable region between cucumoviruses. The level of the 3' UTR identity between cucumoviruses is lower than the level of the 5' UTR identity between corresponding cucumoviruses. The overall identity of RNA 4 nt sequence is similar to that of genomic RNAs between cucumoviruses. Interestingly, the identity of the overall RNA 4 nt sequence is 93.6% within subgroup II strains but only 69% within subgroup I. The low homology of RNA 4 nt sequences within subgroup I is due to the divergent nature of the 5' and 3' UTR of strains Korea, NT9 and As (Anderson *et al.*, 1995).

1.5.5 RNA 4A

RNA 4A was found to be encapsidated in Q-CMV in 1973 (Peden and Symons, 1973). However, it was not recognised as a sgRNA till 1994 (Ding *et al.*, 1994). This RNA is generated from 3'-terminus of RNA 2 for expressing a 2b protein encoded by RNA 2 (Fig. 1.2). RNA 4A is capped at the 5' end (Ding *et al.*, 1994), and present in subgroup II strains (Palukaitis *et al.*, 1992). However, it is not clear whether RNA 4A is also present in subgroup I strains of CMV. Q-CMV RNA 4A is 682 nt long and identical to the 3'-terminal 682 nt of Q-CMV RNA 2 (Ding *et al.*, 1994), and encodes a small ORF, called ORF 2b, of 100 codons which overlaps, but out of frame with, the C-terminal portion of the major 2a gene in RNA 2. The translational product of the 2b gene has been detected in Q-CMV infected cucumber plants (Ding *et al.*, 1994). Hence, RNA 4A functions as a mRNA.

On the basis of nt sequence analysis, the ORF 2b is conserved in all cucumoviruses. It encodes proteins ranging from 95 amino acids (V-TAV) to 110 amino acids (Fny-CMV) residues and always occupies the +1 frame of ORF 2a (Rizzo and Palukaitis, 1988; Moriones *et al.*, 1991; Kataoka *et al.*, 1990a; Karasawa *et al.*, 1992). Predicted amino acid sequences of the 2b proteins are also conserved among different cucumoviruses. The percent identities between the protein 2b sequences of the four cucumoviruses are

higher than 27 % except between the 2b proteins of Q-CMV and J-PSV (Table 1.2). Thus, RNA 4A may be present in all cucumoviruses for the *in vivo* expression of the ORF 2b.

1.5.6 RNA 5 and RNA 6

When CMV RNA is fractionated by electrophoresis on 2.4% polyacrylamide gels, four major (RNAs 1, 2, 3 and 4) and at least three minor RNA species, RNA 4A, RNA 5 and RNA 6 in order of decreasing molecular weight, are found (Symons, 1978; references therein). RNA 5 in Q-CMV has now been demonstrated to be a mixed population derived from the conserved 3'-terminal 307 nt and 304 nt of genomic RNAs 2 and 3, respectively (Blanchard *et al.*, 1996) and probably has no coding ability because only small ORFs (19-37 amino acids) are present on both the RNA 2- and RNA 3-like RNA 5 molecules (Blanchard *et al.*, 1996). Similar sgRNAs that have no translational products have been described in sgRNA 3 of barley yellow dwarf luteovirus (BYDV; Kelly *et al.*, 1994) and CMV (Gordon and Symons, 1985). RNA 5 is not found in subgroup I strains of CMV (Palukaitis *et al.*, 1992). RNA 6 has molecular weight of 0.05×10^6 (Peden and Symons, 1973) and is present in both CMV subgroups (Palukaitis *et al.*, 1992). Interestingly, RNA 6 of Q-CMV was shown to be a mixture of plant RNAs contaminated with fragments of the genomic CMV RNAs (Palukaitis *et al.*, 1992). This raises the question whether plant RNAs contain an encapsidational signal, and if so, why these RNAs can not be encapsidated in other plant viruses.

Data on the sgRNAs of TAV are scarce. Only three reports have mentioned a fifth RNA present in TAV to date (Lot *et al.*, 1974; Hull, 1976; Moriones *et al.*, 1992). However, in terms of TAV and CMV genomic RNA similarities, any sgRNA components of TAV would be expected to generally resemble those of CMV.

1.5.7 Role of untranslated regions

The 5' UTRs of cucumovirus RNA 1s are similar in length to the 5' UTRs of cucumovirus RNA 2s. The 5' UTRs of cucumovirus RNA 3s together with the 3' UTRs of all cucumovirus RNAs vary in length. The 5' untranslated regions of V-TAV RNA 3 (F. Garcia-Arenal, personal communication) and the 3' UTRs of all cucumovirus RNA 2 are about 100 nt longer than those of the other cucumovirus RNAs. It is now known that the extra 100 or so nt in the 3' UTR of cucumovirus RNA 2s encodes part of the 2b gene (Ding *et al.*, 1994).

The 3'-termini among cucumoviruses contain a predicted tRNA-like structure (Wilson and Symons, 1981; Rizzo and Palukaitis, 1989). The tRNA-like structure in TAV has been proposed to play a role in recognition by and interaction with aminoacyl tRNA synthetase (Joshi and Haenni, 1986). In CMV, deletions which included the tRNA-like structure of RNA 3 prevented detectable RNA 3 accumulation in tobacco protoplasts (Boccard and Baulcome, 1993). Similar results were seen for brome mosaic bromovirus (BMV) RNAs 1 and 2 (Duggal *et al.*, 1994), TYMV genomic RNA (Tsai and Dreher, 1991) and cowpea chlorotic mottle virus (CCMV) RNA 3 (Pacha *et al.*, 1990), in which aminoacylation-defective mutants severely debilitated viral replication, implying that the 3'-terminal sequence and the tRNA-like structure formed by this sequence are essential for the replication of RNAs.

RNA pseudoknot structures are another striking feature of cucumovirses found in the 3' UTR. In CMV, this RNA pseudoknot structure in the 3'-terminal 180 nt is almost identical between the two subgroups (Palukaitis *et al.*, 1992), indicating that the pseudoknot structures are of functional importance (Palukaitis *et al.*, 1992). In the case of TMV, the RNA pseudoknot structure in the 3' UTR is essential for the regulation of the protein translation (Pleij, 1995) and might be involved in recombination events (Pleij

et al., 1987). In BMV, the RNA pseudoknot structures have been demonstrated to affect viral replication (Duggal *et al.*, 1994).

The 3' UTRs of all cucumovirus RNAs contain the conserved 40 nt: 5'-GAACGGGUUGUCCAUCCAGCUNACGGCUAAAAUGGUCAGU-3', where the underlined N varies with different genomic RNAs and/or viruses (McGarvey *et al.*, 1995). This highly conserved 40 nt is not characterised by an identifiable secondary structure (Boccard and Baulcombe, 1993). Such a high degree of sequence conservation indicates an important function as a structural and/or binding domain (McGarvey *et al.*, 1995). Boccard and Baulcombe (1993) showed that a deletion which included this region in CMV RNA 3 resulted in the absence of a detectable accumulation of the mutant RNA 3 in protoplasts.

Comparative studies show that sequences that closely resemble the internal control regions (ICRs) 1 and 2 of tRNA genes (Marsh and Hall, 1987) are present at the 5' terminus of RNA 1 and RNA 2 of CMV and PSV as well as in the RNA 3 intercistronic region (IR) of all cucumoviruses. A derivative of CMV RNA 3 with a deletion of 250 bases in the IR containing the ICR-like motif did not accumulate in tobacco protoplasts (Boccard and Baulcombe, 1993), indicating that the ICR-like motif in the IR is necessary for efficient synthesis of RNA. At the 5' end of RNA 3 of cucumoviruses, there is a UG tract. Deletion of this tract caused reduction of RNA 3 accumulation (Boccard and Baulcombe, 1993), suggesting that the tract is probably important for either the efficiency of RNA replication or the stability of the viral RNA, rather than for the recognition of the viral RNA by the replicase (Boccard and Baulcombe, 1993).

1.5.8 Subgenomic promoters

Subgenomic RNA 4 production of BMV (Miller *et al.*, 1985) and AlMV (Langereis, 1987) occurred through internal initiation by the demonstration of its *in vitro* synthesis on

a minus-strand RNA 3 template by the respective virus RNA-dependent RNA polymerase (RdRp). It is expected that cucumoviruses share this mechanism for producing subgenomic RNA 4 or RNA 4A, by which RNA 4A is transcribed through internal initiation on the minus-strand of RNA 2.

Marsh *et al.* (1988) demonstrated that a 62-nt sequence immediately upstream of and including the initiation site of the RNA 4 plays a role in initiation of RNA 4 transcript. It is worth noting here that the ICR 2 motif is outside this promoter region for RNA 4 (French and Ahlquist, 1988; Pogue *et al.*, 1992). However, in the case of CMV RNA 3, the ICR 2 motif (Davis and Symons, 1988; Karasawa *et al.*, 1991) is included in the RNA 3 subgenomic promoter which contains 70 nt upstream and 20 nt downstream from the initiation site of RNA 4 (Boccard and Baulcombe, 1993). The similar localisation of the ICR 2 motif relative to the RNA 4 start point has been found in TAV and PSV RNA 3 (Boccard and Baulcombe, 1993). The ICR 2 motif is also in all cucumovirus RNAs 2 immediately upstream of the putative initiation sites of RNA 4As (Ding *et al.*, 1994), indicating that the ICR 2 motif might be required in cucumoviruses for both RNA 4 and RNA 4A syntheses (Ding *et al.*, 1994).

The ICR 2 motif is not found in the upstream region of RNA 5. There was also no sequence homology to the alphavirus promoter motif in this region (Blanchard *et al.*, 1996). In addition, no common RNA secondary structure was found in the upstream region of RNAs 4, 4A and 5 (Blanchard *et al.*, 1996). Hence, whether RNA 5 is generated by transcription from a cryptic promoter or is the result of an endonucleolytic cleavage of RNA 2 and RNA 3 at a specific site is unknown.

1.6 RNA FUNCTIONS

In pseudorecombinant studies, RNA 1 was found to be associated with CMV seed transmission in *Phaseolus vulgaris* (Hampton and Francki, 1992), local infection of

squash and tobacco by NL-CMV (Lakshman and Gonsalves, 1985), CMV transmission and symptom severity (Zitter and Gonsalves, 1991) and differential rate of systemic symptom induction in zucchini squash by Fny-CMV and Sny-CMV (Roossinck and Palukaitis, 1990; Roossinck 1991). RNA 2 was found to be associated with CMV systemic infection of cowpea (Edwards *et al.*, 1983; Hanada and Tochihara, 1980; Marchoux *et al.*, 1974), systemic infection of maize (Rao and Francki, 1982) and the necrosis induction phenotype (Sleat *et al.*, 1994). CMV RNA 3 was found to control a hypersensitive response of Arabidopsis ecotype C24 (Takahashi *et al.*, 1994), serological specificity (Mossop and Francki, 1977; Hanada and Tochihara, 1980; Marchoux *et al.*, 1974), aphid transmissibility (Mossop and Francki, 1977; Zitter and Gonsalves, 1991) and local infection of *Gomphrena globosa* and *Vicia faba* (Rao and Francki, 1982). In pseudorecombinants between Fny- and Sny-CMV or between TAV and CMV, RNA 3 has been implicated in the control symptom types in some host plants (Zitter and Gonsalves, 1991; Shi *et al.*, unpublished results).

Systemic infection of *Lactuca saligna* (Edwards *et al.*, 1983), local infection of cowpea (Marchoux *et al.*, 1974) and of *Datura stramonium* and *Solaum melongena* (Rao and Francki, 1982) and yellow mosaic symptoms in several CMV host species (Rao and Francki, 1982) were shown to be determined by CMV RNAs 2 and 3. The ability to support the systemic movement and accumulation of satellite RNA of CMV is determined primarily by RNAs 1 + 2 and not by RNA 3 (Moriones *et al.*, 1994). Also, in pseudorecombinants between TAV and CMV, RNA 1 or RNA 2 could not be exchanged (Rao and Francki, 1981), indicating that the proteins encoded by RNAs 1 and 2 may interact in concert (Palukaitis *et al.*, 1992).

1.7 THE ENCODED PROTEIN FUNCTIONS

On *in vitro* translation, Q-CMV RNA 1 directs the synthesis of a single polypeptide of Mr 95K, RNA 2 directs a single polypeptide of Mr 110K and RNA 3 directs the

synthesis of a Mr 35K protein (Gordon *et al.*, 1982). The sizes of these proteins are consistent with those deduced from the respective nt sequences (Rezaian *et al.*, 1985; Rezaian *et al.*, 1984; Davies and Symons, 1988). The 24.5K coat protein cistron in RNA 3 is not translated but is translated through RNA 4 (Schwinghamer and Symons, 1975; Davies and Symons, 1988). The new 2b protein translated from RNA 4A was observed in cucumber plants after infection with Q-CMV (Ding *et al.*, 1994). It had MW of 15K (Ding *et al.*, 1994).

1.7.1 Proteins 1a and 2a

RNAs 1 and 2 inoculated alone to tobacco protoplasts replicate (Nitta *et al.*, 1988), whereas RNA 1, RNAs 1 and 3 or RNAs 2 and 3 inoculated alone to tobacco protoplasts fail to replicate, showing that RNAs 1 and 2 contain a replicase function (Nitta *et al.*, 1988, Wood, 1991). Further evidence that the proteins from RNAs 1 and 2 are evolved in RNA replication comes from the amino acid sequence similarities with proteins of other viruses (TMV, BMV and AlMV) known to function as viral RNA polymerases, or genetically implicated in replication functions (Ishikawa *et al.*, 1986; French *et al.*, 1986; Nassuth and Bol, 1983; ; Rizzo and Palukaitis, 1988; Wood, 1991). An active replicase isolated from CMV-infected tissue demonstrated the presence of both 1a and 2a proteins (Hayes and Buck, 1990). Antibodies against the polypeptides of RNAs 1 and 2 inhibited the RdRp activity (Hayes *et al.*, 1994). The greatest inhibition was shown by antibodies to a peptide containing the GDD motif, indicating the functional importance of the identified sequence motifs in CMV RNA replication (Hayes *et al.*, 1994).

In comparison with Fny-CMV (the fast infection phenotype) and Sny-CMV (the slow infection phenotype), the 1a protein was shown to play some role in both cell-to-cell and systemic movement of virus (Gal-On *et al.*, 1994), which may involve interactions between 1a protein and specific host proteins because the fast vs slow phenomenon is both host- and cultivar-specific (Gal-On *et al.*, 1994). A 3'-terminal truncated CMV RNA

2 expressed in transgenic tobacco also had some effect on both cell-to-cell and systemic movement (Carr *et al.*, 1994; Hellwald and Palukaitis, 1994). These observations indicate that both 1a and 2a may be involved in virus movement. This may be via independent functions of the 1a and 2a proteins or via a complex of these proteins (Gal-On *et al.*, 1994).

1.7.2 Protein 2b

The function of the 2b gene encoded by RNA 2 was investigated by mutational analysis using infectious cDNA clones (Ding *et al.*,1995b). A mutant lacking ORF 2b was unable to spread systematically in cucumber plants, demonstrating involvement of 2b in long-distant movement. The same mutant caused delayed appearance of very mild symptoms in tobacco plants, also demonstrating that 2b contributes to long-distance movement. Further stepwise mutational analyses with the ORF 2b showed that the C-terminal region of the 2b protein is essential for the expression of systemic symptoms whereas the 5 amino acids from the N-terminal half are not required for symptom expression in cucumber, indicating that the 2b gene encodes a pathogenic function in addition to a host-specific long-distance movement function. Deletion of the overlapping C-terminal part of ORF 2a did not change infectivity of the mutant in either host species, ruling out the 2a mutation as the reason for change of the phenotype.

1.7.3 Protein 3a

Studies with deletions in the coding region of CMV 3a protein suggested that the 3a protein is related to transport of CMV because the deletion mutants could replicate in protoplasts but not induce systemic infection on host plants (Suzuki *et al.*, 1991). That the 3a protein was detected in a CMV infected cell wall fraction further supported this suggestion (Burman *et al.*, 1994). Vaquero *et al.* (1994) showed that the 3a protein is involved in the modification of plasmodesmata. This result was confirmed by Ding *et al.*

(1995), and the 3a protein was demonstrated to interact with and dilate plasmodesmata. In contrast, a mutant form of the 3a protein was unable to traffic from cell to cell, to increase the size exclusion limit of plasmodesmata, or to potentiate cell-to-cell trafficking of CMV RNA molecules (Ding *et al.*, 1995; Kaplan *et al.*, 1995), indicating that the CMV 3a protein functions as the cell-to-cell movement protein of CMV.

The Fny-CMV 3a transgenic plants were able to complement both cell-to-cell and long-distance movements of a pseudorecombinant cucumovirus that is defective for such movements in tobacco, in which RNAs 1 and 2 are from TAV and RNA 3 is from M-CMV, and not able to complement the long-distance movement of other related viruses, PSV and BMV (Kaplan *et al.*, 1995). Similar results were obtained recently by Cooper *et al.* (1996) in which the transgenic S-CMV 3a complemented the cell-to-cell and systemic spread of a movement-defective mutant of Fny-CMV but not the local or systemic spread of a movement-defective mutant of TMV, whereas the transgenic TMV 30K movement protein shares some of the functions with the CMV 3a required to transport CMV. By purifying 3a protein from *Escherichia coli* as well as from CMV-infected plants, Li and Palukaitis (1996) showed that the CMV 3a protein contains RNA binding activity similar to that of the TMV 30K movement protein, again directly indicating that the 3a is involved in the movement process.

1.7.4 Coat protein

Besides being the structural protein of virus particles as mentioned above, two lines of evidence suggest that the coat proteins of cucumoviruses facilitate virus movement: (1) Mutants in which parts of the coding region of CMV coat protein were deleted could replicate in protoplasts but did not induce systemic infection on host plants (Suzuki *et al.*, 1991). (2) An intact CMV CP together with 3a mutants complemented the defective long-distance systemic movement of one TAV strain (1-TAV) in cucumbers, whereas an intact 3a protein together with CP mutants failed to do so (Taliansky and Garcia-Arenal, 1995).

Further analyses showed that the plasmodesmata in the vascular system and/or at the boundary between the mesophyll and the vascular system are involved in long-distance movement through the phloem, and those in the mesophyll are involved in cell-to-cell movement (Taliansky and Garcia-Arenal, 1995).

An amino acid at position 129 in CMV CP had been demonstrated to affect the symptom expression, the secondary structure of the CP, virus assembly and virus transport (Shintaku *et al.*, 1992; Suzuki *et al.*, 1995). The symptom expression has been suggested to be affected by a particular local secondary structure surrounding amino acid 129 in the CP rather than the nt sequence itself or the identity of a particular amino acid per se (Shintaku *et al.*, 1992). However, in some other case, symptoms are induced by a combination of the virus coat protein gene and two host recessive genes (Takahashi and Ehara, 1993)

CP is also involved in recognition by the aphid vector for the transmission of CMV (Palukaitis *et al.*, 1992). Particles reassembled from V-TAV coat protein and RNAs from the non aphid-transmissible M-CMV were transmitted by an aphid, *Myzus persicae*, whereas particles reassembled from M-CMV coat protein and RNAs from V-TAV were not transmitted by the aphid (Chen and Francki, 1990). Using chimeric RNA 3s which contain sequences from the efficiently aphid-transmitted strain Fny-CMV and the very poorly aphid-transmissible M-CMV, Perry *et al.* (1994) revealed that two amino acid regions, positions 129 and 162, in the coat protein were critical for transmission by aphids. A single amino acid change at position 129 did not restore aphid transmissibility to M-CMV. Another amino acid at position 168 may influence the efficiency of aphid transmission, but is not essential.

1.8 SATELLITE RNAS

Naturally occurring virions of CMV and PSV, but not TAV, may also package a single-stranded satellite RNA (Kaper and Waterworth, 1977) which is completely dependent on its helper viruses for replication and capped with m^7Gppp at its 5'-termini (Roossinck *et al.*, 1992). CMV and TAV, but not PSV, support the replication of CMV satellite RNAs (Palukaitis *et al.*, 1992). CMV satellite RNAs have been extensively studied and up to now, more than 28 isolates originating from widely separated geographical areas have been characterised and sequenced (Roossinck *et al.*, 1992). They contain between 332 and 342 nt. The sequence homology between CMV satellite RNAs ranges from 70% to 99%. However, no homology between CMV satellite RNAs and CMV genomic RNAs or between CMV satellite RNAs and plant RNAs or genomic DNA has been detected (Palukaitis *et al.*, 1992).

CMV satellites are of particular biological significance and are classified in two groups, benign and necrogenic according to their ability to produce symptoms on tomato (Devic *et al.*, 1990). The necrogenic satellite RNAs induce a systemic necrosis on tomato while benign satellite RNAs induce chlorotic or mosaic symptoms on this plant. The necrogenic capability is affected by nt in the 3' part of the satellite RNA (Devic *et al.*, 1989; Kurath and Palukaitis, 1989). A second symptom-inducing domain is located at the 5' part of the satellite RNA (Devic *et al.*, 1990), which affects the production of a yellow mosaic symptom on tobacco (Devic *et al.*, 1989; Masuta and Takanami, 1989) or a chlorotic symptom on tomato (Kurath and Palukaitis, 1989). A nt change at position 149 within this domain alters the host specificity of chlorosis from tomato to tobacco (Sleat and Palukaitis, 1992). Polypeptides encoded in satellite RNA have no effect on necrosis symptoms (Devic *et al.*, 1990).

The amounts of satellite RNA produced are highly variable depending on both the host species used and the strain of CMV. Some strains increase the accumulation of satellite

RNAs whereas others decrease the accumulation of satellite RNAs (Kaper *et al.*, 1995). In the other way, satellite RNA reduced the accumulation of one strain of CMV while it did not reduce the accumulation of another strain (Gal-On *et al.*, 1995a). It has been shown that some satellite RNAs anneal to RNA 3 and RNA 4, and some others anneal with RNA 1 and RNA 2 as well as with RNA 3 and RNA 4 (Fraile *et al.*, 1993), probably forming an unusual knot-like structure or base-pairing (Rezaian and Symons, 1986) by which the satellite RNA could regulate CMV RNA replication. The combinations of satellite RNA and viral RNA did not affect the *in vitro* mRNA activity of the viral RNA (Fraile *et al.*, 1993), suggesting that the various effects of satellite RNAs are not mediated through the inhibition of the expression of their helper virus genome. A recent finding that satellite RNAs competed with CMV RNAs for RdRp (Wu and Kaper, 1995) suggests that the satellite RNA may regulate CMV RNA replication.

It is not clear why CMV satellite RNAs appear so often under experimental conditions but rarely in nature. A hypothesis proposed by Palukaitis *et al.* (1992) is that many strains of CMV contain some subliminal level of satellite RNA, which, under appropriate conditions, will begin to replicate and become evident because no satellite RNAs were generated during many passages through host by using infectious transcripts from three genomic cDNA clones of CMV (Palukaitis *et al.*, 1992). How the satellite RNAs originally arose is unknown.

1.9 DEFECTIVE RNAS

Defective RNAs, like satellite RNAs, are also completely dependent on their helper viruses for replication and encapsidation although they contain the cis-acting components necessary for efficient replication (Holland, 1990). However, they can be distinguished from satellite RNAs by the presence of significant sequence homology with their helper viruses (Holland, 1990). Defective RNAs have been well characterised for most animal RNA virus groups (Holland, 1990) and several plant RNA virus groups, including

tombusviruses (Burgyan *et al.*, 1993; Hillman *et al.*, 1987), carmoviruses (Li *et al.*, 1989), furoviruses (Bouzoubaa *et al.*, 1991; Chen *et al.*, 1994), potexviruses (White *et al.*, 1991) and bromoviruses (Romero *et al.*, 1993). However, only two defective RNAs were recently characterised for Fny-CMV (Graves and Roossinck, 1995).

These two defective RNAs, designated 3α and 3β were obtained after the fourth passage in a host tobacco. Sequence analysis showed that the 3α and 3β are derived from RNA 3 by single, in-frame deletions with the ORF 3a. The 3α deletion is smaller and resides totally within the region deleted from 3β and is maintained at a very low level while the 3β reaches levels equal to those of the genomic RNAs. Hence, the 3α could be an intermediate in the production of the 3β. Both 3α and 3β can be supported by other CMV strains which did not produce any defective RNAs. These two defective RNAs had no apparent effect upon helper virus yield or symptom production although 3β protein in vitro translation is detected. The mechanism of copy choice has been proposed for formation of the CMV defective RNAs, whereby Fny-CMV RNA 3 could be taken as a primary template.

1.10 AIMS

The aims of this thesis are:

(1) to determine the nucleotide sequences of the unidentified subgenomic RNAs of CMV and TAV.

(2) to construct the infectious cDNA clones of TAV for use as a basic tool in the investigation of gene functions of this virus.

(3) to apply these clones, together with the infectious cDNA clones of CMV that have been constructed in this laboratory (Ding *et al.*, 1995a), to investigate the genetic origin of the above subgenomic RNAs and the functions of their encoding genes.

Chapter 2

GENERAL MATERIALS AND METHODS

The materials and methods described in this chapter are those used routinely during the course of this study. Others more specific are described in the relevant chapter.

2.1 MATERIALS

All materials used in this thesis were at least analytical grade in standard. Solutions were prepared with ultra-pure water, treated as sterile, filtered or autoclaved where appropriate.

2.1.1 Synthetic oligodeoxyribonucleotides

Synthetic oligodeoxyribonucleotides (primers) were prepared on an Applied Biosystems Model 380B DNA synthesiser either by Niel Shirely or by Jing Li. All primers were unphosphorylated and purified from premature termination products by Mono Q column, high pressure liquid chromatography (HPLC).

2.1.2 Nucleotides and radionucleotides

Deoxynucleotide triphosphates (dNTPs; Pharmacia)

Dideoxynucleotide triphosphates (ddNTPs; Pharmacia)

Ribonucleotide triphosphates (NTPs; Pharmacia)

α-^{32}P-dATP (10 mCi/ml; Bresatec)

γ-^{32}P-ATP (10 mCi/ml; Bresatec)

α-^{32}P-UTP (10 mCi/ml; Bresatec)

α-^{35}S-dATP (12.5 mCi/ml; Bresatec)

2.1.3 Plasmid vectors

pBluescript SK$^+$ (Stratagene), pUC19 (New England Biolabs), pSP72 (Promega)

pGEX-2T (AMRAD), pMAL-c2 (New England Biolabs), pQE-30 (QIAexpress), pET-3d (Novagen)

2.1.4 Bacterial strains (*Escherichia coli*)

E. coli JM110: F ', traD36, $lacI^q\Delta$(lacZ)M15, proAB/rpsL, (Str^r), thr, leu, thi, lacY, galK galT, ara, fhuA, tsx, dam, dcm, supE44, Δ(lac-proAB)

E. coli DH5α: F ', ϕ89dlacZΔM15, recA1, endA1, gyrA96, thi-1, hsdR17(r_k^-, m_k^+), supE44, relA1, deoR, Δ(lacZYA-argF)U196

E. coli TB1: ara, Δ(lac-proAB), rpsL, (ϕ80, lacZΔM15) hsdR

E. coli M15(pREP4): Nal^s, Str^s, rif^s, lac^-, ara^-, gal^-, mtl^-, F^-, $recA^+uvr^+$

E. coli HMS174(DE3)pLysE: F^-, recA $r_{K12}^-m_{K12}^+Rif^r$

E. coli BL21(DE3)pLysE: F^-, ompT, $r_B^-m_B^-$

E. coli HMS174(DE3)pLysS: F^-, recA $r_{K12}^-m_{K12}^+Rif^r$

E. coli BL21(DE3)pLysS: F^-, ompT, $r_B^-m_B^-$

2.1.5 Molecular weight markers

DNA (Bresatec): Bacteriophage SPP1 DNA cut with EcoRI, pUC19 DNA cut with HpaII

Protein (Promega): mid-range (14.4K-97.4K), low-range (2.5K-31K)

2.1.6 Common growth media

LB: 1% (w/v) bacto-tryptone (DIFCO), 0.5% (w/v) yeast extract (DIFCO), 1% (w/v) NaCl, pH 7.0.

M9: 0.1% (w/v) NH_4Cl, 0.3% (w/v) KH_2PO_4, 0.6% (w/v) Na_2HPO_4, 0.4% (w/v)

glucose and 0.1% (v/v) 1 M $MgSO_4$.

2.1.7 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

Acrylamide stock (100 ml): 29.2 g acrylamide and 0.8 g bisacrylamide;

4x resolving gel buffer: 1.5 M Tris-HCl, pH8.8 and 0.4% sodium dodecyl sulfate (SDS);

4x stacking gel buffer: 0.5 M Tris-HCl, pH6.8, and 0.4% SDS;

SDS sample buffer (10 ml): 1 ml glycerol, 2 ml 10%SDS, 2 mg bromophenol blue (BDH), 1.25 ml 4x stacking gel buffer, 5.75 ml water, (1 µl β-mercaptoethanol (Sigma) per 20 µl SDS sample buffer is added in just before use)

Running buffer (pH 8.3): 25 mM Tris-base, 192 mM glycine and 0.1% SDS.

2.2 METHODS

All methods were carried out according to Sambrook *et al.* (1989) unless otherwise stated.

2.2.1 Virus sources

Three viruses were chosen for this study. The V strain of tomato aspermy virus (V-TAV) was originally isolated from Chrysanthemum sp. in Victoria, Australia (Habili and Francki, 1974a). Two strains from each of the two subgroups (Owen and Palukaitis, 1988) of cucumber mosaic virus (CMV): Q strain (Q-CMV), a subgroup II strain (Owen and Palukaitis, 1988), was originally isolated from Capsicum sp. in Queensland, Australia (Francki, 1964); WAII strain (WAII-CMV), a subgroup I strain (Wahyuni *et al.*, 1992), was previously described as T strain and originally isolated from Capsicum sp. in Western Australia (Hatta and Francki, 1979).

2.2.2 Storage of viruses

Dried *Nicotiana glutinosa* leaves infected with the viruses listed above were available in the Symons laboratory. All dried leaf inocula were stored at 4°C in sterile glass vials containing fused $CaCl_2$ as the dehydrating agent. Contact between leaf material and $CaCl_2$ was prevented by a sterile cotton wool plug.

2.2.3 Propagation of viruses

The method used here was described by Shi *et al.* (1989) and outlined below.

A small quantity of the dried leaf material was ground to a paste in a few drops of 50 mM phosphate buffer pH 7.2 [72:28 (v/v) 50 mM Na_2HPO_4:NaH_2PO_4]. The mortar and pestle were autoclaved, kept at 4°C and rinsed in distilled water before use. The extract was then rubbed by finger or pestle onto 2 to 3 weeks old *Chenopodium quinoa* leaves using sterile carborundum powder (500 mesh) as the abrasive. The plants used had been heavily watered the night before and kept in the dark for 24 hours prior to inoculation. Inoculated plants were grown in a glasshouse kept at a 15 hour day length at 230 lumens/m^2 and 20-22°C. Five to seven days after inoculation, individual local lesions were isolated and used as the inoculum for infecting *C. quinoa*, and the process repeated three times. Finally a single lesion was inoculated onto *N. glutinosa*. Virus infection was confirmed using the northern blot hybridisation as described below. Infected leaf material was harvested, and was either stored or used to infect large numbers of *N. glutinosa* or other hosts for virus propagation.

2.2.4 Purification of viral particles

The propagation host for Q-CMV was *Cucumis sativus* cv. Green Gem and *N. glutinosa* for WAII-CMV and V-TAV. The infected leaves were harvested 10 to 15 days after inoculation. Q-CMV and WAII-CMV were purified according to the method of Lot *et al.*

(1972) as modified by Peden and Symons (1973). Briefly, to each gram of leaf material, 1 ml of the extraction buffer containing 0.5 M sodium citrate, 5 mM ethylenediaminetetraacetic acid (EDTA) (pH9.0) and 0.5% thioglycollic acid (BDH) and 1 ml of chloroform were added. The mixture was homogenised by a blender and centrifuged at 12,000 g for 10 minutes at 4°C. To every 1 ml of the aqueous phase, 3.3 ml of the buffer containing 5 mM sodium borate, 0.5 mM EDTA (pH 9) and 40% (w/v) polyethylene glycol (PEG, mol. wt 6000) was added and the mixture was stirred slowly for about 30 minutes at 4°C. Viral particles were precipitated by centrifugation at 12,000 g for 10 minutes at 4°C. The pellets were resuspended in 3-4 ml of the suspension buffer containing 5 mM sodium borate, 0.5 mM EDTA (pH 9) and 2% (w/v) Triton X-100 per gram of the infected leaf materials and centrifuged again at 12,000 g for 10 minutes at 4°C. The viral particles were further purified by three cycles of low (12,000 g, 10 minutes) and high speed (100,500 g, 2 hours) centrifugations at 4°C. The virus pellets were resuspended in the suspension buffer without Triton X-100 and stored at 4°C or at -80°C in the presence of 50% glycerol.

TAV was purified essentially as described by Peden and Symons (1973) except that 1% thioglycollic acid was added to the extraction buffer and that the virus was resuspended in 0.02M phosphate buffer pH 7.6 [87:13 (v/v) 20 mM Na_2HPO_4:NaH_2PO_4] (Habili and Francki, 1974a).

The purified virions were examined for morphology under electron microscope or for the isolation of virion RNAs.

2.2.5 Isolation of RNAs

Virion RNAs were isolated from purified virus as described by Peden and Symons (1973). One ml of viral particle suspension containing 0.5% SDS, 0.3 M sodium acetate, pH 5.2 and 10 mM EDTA (pH7.0) was extracted with an equal volume of 50 mM Tris-HCl (pH8.0)-saturated phenol (stored at 4°C). The aqueous phase was reextracted with

an equal volume of the phenol and precipitated in the presence of 2.5 volumes of ethanol by centrifugation at 12,000 g for 30 minutes at 4°C. The precipitated RNAs were dissolved in 0.3 M sodium acetate (pH 5.2) and 1 mM EDTA (pH 7.0) and ethanol reprecipitated again. The RNA pellets were twice washed with 70% ethanol, dried *in vacuo* and suspended in 0.1 mM EDTA (pH 7.0) and stored at -20°C.

Total RNAs were isolated from plants essentially as described by Verwoerd *et al.* (1989) except that cold extraction buffer and a single chloroform extraction were used. Fresh leaves, as little as 100 mg, were collected in 2 ml Eppendorf tubes, frozen by liquid nitrogen and ground with a steel bar. The fine powder was homogenised with 400 µl extraction buffer (stored at 4°C) containing 0.1 M LiCl, 100 mM Tris-HCl (pH8.0), 10 mM EDTA and 1% SDS and 400 µl phenol by vortexing for 1 minute, 400 µl chloroform was then added and the mixture was vortexed for 1 minute. After centrifugation in an Eppendorf microcentrifuge at 14000 rpm for 10 minutes at room temperature, the aqueous phase was removed and mixed with an equal volume of 4M LiCl and left at -20°C overnight. The RNAs were collected by centrifugation at 14,000 rpm for 15 minutes at 4°C and redissolved in 400 µl water containing 0.3 M sodium acetate (pH 5.2), and reprecipitated in the presence of 2.5 volume of ethanol by centrifugation at 14,000 rpm for 15 minutes at 4°C. The RNA pellets were twice washed with cold 70% ethanol, dried *in vacuo* and suspended in 0.1 mM EDTA (pH 7.0) and stored at -20°C.

2.2.6 Synthesis of double-stranded cDNA to partial viral RNA genome

First-strand synthesis: virion RNAs (0.1 µg) or total RNAs (1 µg) from infected plants were annealed to the appropriate primer (200-500 ng) in the presence of 1 mM EDTA by heating at 80°C for 2 minutes followed by snap-cooling on ice. The viral RNA was reverse transcribed in 20 µl reaction mixture containing 50 mM Tris-HCl pH 8.3, 40 mM KCl, 7 mM $MgCl_2$, 1 mM dithiothreitol (DTT), 1 mM dNTPs, 1 unit/µl RNAsin (Promega), 0.1 mg/ml bovine serum albumin (BSA; Promega) and 0.75 units/µl AMV

reverse transcriptase (Promega). The reaction was incubated at 25°C for 5 minutes, 37°C for 5 minutes and finally at 55°C for 30 minutes. The reaction was terminated by extraction with phenol:chloroform (1:1 v:v) and cDNA in the aqueous phase was precipitated with ethanol, as in 2.2.5. The pellet containing the first-strand cDNA was dissolved in 10 µl of water.

Amplification of cDNA by the polymerase chain reaction (PCR) was essentially as described by Ding *et al.* (1995b). One to 5 µl of the first-strand cDNA obtained above and 0.1-0.5 µg of each of the appropriate primers were added to a 0.5 ml tube containing 10 mM KCl, 20 mM Tris-HCl pH 8.8, 10 mM $(NH_4)_2SO_4$, 2 mM Mg_2SO_4. 0.1% Triton X-100, 0.4 mM dNTPs, 0.1 µg/µl BSA, 0.2 units/µl Vent DNA polymerase (New England Biolabs) in a total volume of 50 µl. The reaction was carried out using a thermal cycler machine (Minicycler™, MJ Research). Cycle conditions were : 30 cycles at 94°C for 1 minute, 50°C for 1 minute, 72°C for 2 minutes followed by 1 cycle at 25°C for 5 minutes. No template controls were included in all reactions. The amplified product was analysed by agarose gel electrophoresis, as follows.

2.2.7 Agarose gel electrophoresis

SeaKem GTG agarose (FMC) minigels [0.6-2% (w/v)] were prepared in 1x TAE (40 mM Tris-acetate, 20 M sodium acetate, 1 mM EDTA, pH 8.2). DNA samples were loaded onto a submerged gel in the presence of 1x ULB [0.67 M urea, 3.3 mM Tris-HCl, 6.67% (w/v) sucrose, 3.3 mM EDTA, 0.02% (w/v) xylene cyanol FF, 0.02% (w/v) bromophenol blue]. The gel was electrophoresed at approximately 7.5 V/cm until the bromophenol blue had migrated 3/4 of the length of the gel. DNA was visualised by staining the gel with ethidium bromide [0.5 µg/ml (w/v) in water], and photographed under short wavelength UV light.

2.2.8 Isolation of DNA from agarose gels

DNA bands of interest were excised from the gels under long wavelength UV light and DNA purified using the Geneclean kit of BIO 101 as described below. Each excised gel slice was placed in a 1.5 ml of Eppendorf tube and 3 volumes of 6 M NaI added. After incubation at 55°C for 5 minutes and the gel was completely melted, 5 µl of GLASSMILK was added. The mixture was incubated at room temperature for 5 minutes. The silica matrix with the bound DNA was precipitated by a brief spin of 5 seconds at 14000rpm (Eppendorf) and washed twice with NEW WASH. The DNA was eluted from the GLASSMILK in 10-15 µl water by incubating at 55°C for 5 minutes.

2.2.9 Restriction digestion of DNA

DNA solution was placed in a 1.5 ml Eppendorf tube and mixed with the appropriate restriction endonuclease and buffer recommended by manufacturer and incubated at the appropriate temperature for required period of time. Where appropriate, the restricted DNA was dephosphorylated using calf intestinal alkaline phosphatase (CIAP, Promega) as described below. Digested DNA was added with CIAP (0.01unit/pmol ends) and incubated at 37°C for 1 hour in the presence of 1x CIAP buffer (50 mM Tris-HCl, pH9.0, 1 mM $MgCl_2$, 0.1 mM $ZnCl_2$, 1 mM spermidine). The digested and/or dephosphorylated DNA was purified by agarose gel electrophoresis as described in 2.2.7 and 2.2.8.

2.2.10 Ligation of cDNA into plasmid vectors

For dephosphorylated vectors, approximately 20 ng of the linearised DNA was mixed with the insert DNA at a molar ratio of approximately 1:3 (vector:insert) in the presence of 50 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 1 mM DTT, 1 mM ATP and either 0.25 units (blunt ended insert) or 0.05 units (sticky ended insert) of T4 DNA ligase (Bresatec), in a reaction volume of 10 µl. Reactions were incubated at 25°C for at least 1 hour or at

4°C overnight. Non-dephosphorylated vectors were treated in the same manner as dephosphorylated vectors, except the amount of linearised vector DNA used was decreased to 5 ng.

2.2.11 Transformation of *E. coli*

Preparing competent cells: The desired *E. coli* strain was streaked on LB plates overnight and a single colony picked and grown overnight at 37°C in LB medium. The overnight culture was diluted 100 times with LB and grown at 37°C to log phase (A_{600} 0.4 - 0.6). After chilling on ice, the cells were sedimented by centrifugation at 4,000 g at 4 °C for 10 minutes and resuspended in ice-cold 0.1 M $CaCl_2$ (1 ml $CaCl_2$ per 100 ml culture). The resuspended cells were then incubated on ice for at least 1 hour prior to use.

Transformation: A 5 µl aliquot of the relevant ligation reaction, or 10-20 ng of uncut vector for control, was mixed with 100 µl of competent cells in a pre-cooled Eppendorf tube and incubated on ice for 30 minutes. Cells were heat shocked at 42°C for 2 minutes in a water bath, followed by incubation on ice for 2 minutes. 0.8 ml of LB was added to the tubes and the cells incubated at 37°C in a rotating vertical wheel for 30-45 minutes. 200 µl cell suspension with or without 40 µl 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside(Boehringer Mannheim) [X-gal, 20 mg/ml in dimethylformamide (BDH)] and 20 µl isopropyl-β-D-thiogalactopyranoside (IPTG, 40 mg/ml; Promega) was spread onto LB-agar plates containing ampicillin (100 µg/ml; Boehringer Mannheim). The plates were incubated at 37°C overnight.

Selection of clones: Bacterial colonies or white colonies where the blue/white colour selection is available were selected for minipreparation of plasmid DNA. Restriction enzyme digestion analysis was used to further confirm the presence or absence of the inserted DNA fragment.

2.2.12 Isolation of plasmid DNA

Minipreparation of plasmid DNA: A single colony was picked into 2 ml of LB containing ampicillin (100μg/ml). After growing overnight at 37°C, the bacterial cells were pelleted. After the supernatant was removed, the pellet was resuspended in 200 μl of STET [0.1 M NaCl, 10 mM Tris-HCl (pH8.0), 1 mM EDTA (pH8.0), 5% Triton X-100] containing 0.1 mg/ml lysozyme (Sigma) and the suspension placed in 95°C heating block for 45 seconds, and centrifuged at 14000rpm (Eppendorf) at 4°C for 30 minutes. The "Gloopy" pellet was removed with a toothpick and the supernatant was added with 200 μl ice-cold isopropanol. After 5 minutes on ice, the mixture was re-centrifuged at 14000rpm (Eppendorf) at room temperature for 10 minutes. The supernatant was carefully removed and the DNA pellet was washed with 70% ethanol and resuspended in 30 μl H_2O.

Large scale preparation of plasmid DNA: 500 ml of an overnight culture of the relevant clone was centrifuged at 4000rpm at 4°C for 15 minutes in a Sorvall GS3 rotor. The bacterial pellet was resuspended in 100 ml of ice-cold STE [0.1 M NaCl, 10 mM Tris-HCl (pH8.0), 1 mM EDTA (pH8.0)] and the suspension centrifuged at 4000rpm at 4°C for 15 minutes. The pellet was resuspended in 18 ml GTE [25 mM Tris-HCl (pH 8.0), 10 mM EDTA (pH 8.0), 50 mM glucose] and the suspension then mixed with 2 ml of lysozyme [10 mg/ml in 10 mM Tris-HCl (pH 8.0)], 40 ml of freshly prepared 1% (w/v) SDS, 0.2 M NaOH. The mixture was gently inverted several times and incubated at room temperature for 10 minutes. The mixture was neutralised by the addition of 20 ml of 3 M potassium acetate followed by incubation on ice for 10 minutes. Cellular debris was sedimented by centrifugation at 4000rpm at 4°C for 15 minutes (Sorvall GS3 rotor) and the supernatant was carefully filtered through four layers of cheesecloth into a 250-ml centrifuge bottle and mixed with 0.6 volume of isopropanol and the mixture stored at room temperature for 10 minutes. The mixture was centrifuged as before. The pellet was

resuspended in 4 ml of TE [10 mM Tris-HCl (pH8.0), 1 mM EDTA (pH8.0)] and 4 ml 5 M LiCl/50 mM Tris-HCl (pH8.0) added. The mixture was kept on ice for 15-30 minutes. The supernatant was cleared by centrifugation at 3000rpm (Sorvall SS-34 rotor) at 4°C for 5 minutes and plasmid DNA precipitated by adding 2.5 volume of ethanol and 0.1 volume of 3M sodium acetate. The pellet was resuspended in 5 ml TE and incubated at 37°C for 1 hour in the presence of 50 μg/ml DNase-free RNase A (Promega). The suspension was phenol:chloroform (1:1 v:v) extracted twice and the nucleic acid recovered from the aqueous phase by ethanol precipitation using 0.3 M sodium acetate (pH 5.2) and centrifugation at 10000rpm (Sorvall HB-4 rotor) at 4°C for 40 minutes.

The nucleic acid pellet was resuspended in 1-2 ml of TE and the suspension then filtered through a minipore filter (0.2 μm pore size; Schleicher&Schuell) before being loaded onto a Sepherose 6 HPLC column. The column was washed with the HPLC buffer [0.1 M sodium acetate (pH7.0), 0.05% SDS (v/v), 20% ethanol)]. Fractions containing plasmid DNA were pooled and the plasmid DNA was recovered by ethanol precipitation as in 2.2.5.

2.2.13 DNA sequencing

Dideoxy chain termination sequencing methods (Sanger *et al.*, 1977; 1980) was used to determine DNA sequence. DNA sequencing was performed with Sequenase Version 2.0 T7 DNA polymerase using DNA sequencing kits (USB). Plasmid DNAs purified on a minipreparation or large-scale as described (2.2.12) were denatured by the addition of 5 μl of 1 M NaOH containing 1 mM EDTA to 20 μl of plasmid DNA. The denatured DNAs were purified from solution components by passage through micro-spin columns containing Sepharose CL-6B (Pharmacia), followed by elution in 25 μl TE. Seven μl of the purified denatured plasmid DNA was used in each sequencing reaction.

2.2.14 Direct sequencing of RNA

The method was adapted from Fichot and Girard (1990). 50 ng (about 10 pmol) primer was annealed to 1 µg of purified RNA or 10 µg total plant RNA in 10 µl of 60 mM Tris-HCl (pH 8.3), 75 mM NaCl, 7.5 mM $MgCl_2$, 5 mM DTT. After incubation for 10 minutes at room temperature, 2 µl of labelling mix (2 µM dGTP, dTTP, dCTP; 1 µl [^{35}S]-dATP and 1 µl AMV reverse transcriptase were added. After 5 minutes incubation at room temperature, 2.8 µl of the labelling reaction was dispensed to 5 separate tubes, each containing 2 µl 0.5 mM dATP, dCTP, dTTP, 1 mM dGTP, and 0.25 mM of either ddATP, ddTTP, ddCTP, 0.5 mM ddGTP or no dideoxynucleotide. The mixtures were incubated at 42°C for 30 minutes and then the reactions stopped by the addition of 3 µl of stop solution (95% formamide, 20 mM EDTA, 0.05% bromophenol, 0.05% xylene cyanol FF). Aliquots of the sequencing reactions were electrophoresed as previously described.

2.2.15 *In vitro* transcription of plasmid clones by RNA Polymerase: Preparation of radioactive probes

Linearised DNAs (1-2 µg) were mixed with 40 mM Tris-HCl pH7.5, 6 mM $MgCl_2$, 2 mM spermidine, 10 mM NaCl, 10 mM DTT, 0.5 unit/µl RNasin, 0.5 mM each of dATP, dCTP and dGTP, 12 µM UTP, 50-100 µCi α-^{32}P-UTP, and 20-40 units T3, T7 or SP6 RNA polymerase (Promega) as appropriate, and the mixture incubated at 37°C for 1 hour. The DNA template was destroyed by addition of 1 unit of RNase-free DNase (Promega) and incubation at 37°C for 15 minutes. The transcripts were extracted with phenol:chloroform (1:1 v:v), precipitated in 2.5 M ammonium acetate/2.5 volume of ethanol and the pellet finally resuspended in 100 µl TE.

2.2.16 Northern blot hybridisation

The method was adapted from Ding *et al* (1995b).

Gel preparation: 2.16 g of agarose was boiled in 156.6 ml water and cooled to 60°C. 5.4 ml of 37% formaldehyde and 18 ml of 10x MOPS buffer [0.23 M MOPS (3-(N-Morpholino) propanesulfonic acid; Sigma) (pH7.0), 0.01 M EDTA and 0.05 M NaAc] were then added before the gel was poured into a gel box (14x21cm). The running buffer was 1x MOPS.

Sample preparation: 5.8 µl RNA (0.1-0.5 µg for purified RNA and 2-10 µg for total plant RNA), 2.5 µl 10x MOPS, 4.4 µl formaldehyde and 12.5 µl formamide were mixed each other and the mixture incubated at 65°C for 15 minutes. Five µl formamide loading buffer (FLB: 95% formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol) was then added before loading. The gel was run at 6V/cm until the bromophenol blue reached the bottom 1/4 of the gel.

Northern blotting: The RNA transfer sandwich consisted of the following layers (from bottom to top): 2 sheets of Whatman 3MM prewetted with HETS buffer (CINNA/Biotecx Laboratories), the gel (upside down), the nylon membrane (Hybond$^+$, Amersham) and the dry paper towels. After the transfer was completed in at least 4 hours, the sandwich was disassembled. RNA was fixed onto the membrane by UV light (GS Genelinker, BIORAD) and photographed under short-wave length UV light. Prehybridisation was done at 65°C for at least 4 hours in a bottle containing 10 ml of solution consisting of 5 ml deionied formamide, 2.5 ml 20x SSC [3 M NaCl, 0.3 M sodium citrate (pH7.0)], 1 ml 50x Denhardts [1% (w/v) Ficoll 400 (Pharmacia), 1% (w/v) polyvinylpyrrolidone (Sigma), 1% (w/v) BSA, 0.5% SDS and 20 µg/µl sheared and denatured salmon sperm DNA (BDH)], 1 ml 10% SDS, 0.2 ml phosphate buffer (pH 6.8) [49:51 (v:v) 50 mM Na_2HPO_4:NaH_2PO_4], 1 mg each of denatured salmon sperm DNA and *E. coli* tRNA (BDH). The radioactive probe was then added and the hybridisation continued at the same temperature overnight. The membrane was washed twice in 2x SSC/0.1% SDS at 55°C for 15 minutes and once in 0.2x SSC/0.1% at 75°C

for 45 minutes and then autoradiographed.

2.2.17 Removal of protruding 3' termini of restriction enzyme digested DNA

Removing protruding 3' termini of restriction enzyme digested DNA was performed using T4 DNA polymerase (Promega) in a reaction containing 50 mM NaCl, 10 mM Tris-HCl (pH7.9), 10 mM $MgCl_2$, 1 mM DTT and 2 mM each dNTP. The reaction was incubated at 12°C for 15 minutes, followed by extraction with phenol:chloroform (1:1 v:v) and precipitation by ethanol as described in 2.2.5.

2.2.18 Protein expression in *E. coli*

The ORF 2b coding sequence of either V-TAV or WAII-CMV was ligated into an appropriate vector and the constructs transformed in an appropriate E. *coli* strain, DH5α for pGEXII, M15 for pQE-30, TB1 for pMAL-2c and HMS174 or BL21 for pET-3d. Colonies with the desired insert were grown in 10 ml LB containing appropriate antibiotic(s) overnight at 37°C with shaking, and the overnight culture was used to inoculate 1 litre LB. When A_{600} of the culture reached 0.4-1.0, 1 ml of the culture was removed and centrifuged 14000rpm (Eppendorf) for 1 minute and the supernatant was discarded. The pellet was resuspended in 50 µl of SDS sample buffer and set aside on ice for analysis later. IPTG (0.1 M) was added to the remainder of the culture to a final concentration of 0.1-5 mM and the incubation continued at 37°C for 2-5 hours. The cells were harvested as in 2.2.12, and together with the cells collected prior induction, all were resuspended in SDS sample buffer. SDS-PAGE was performed as described (Laemmi, 1970). After electrophoresis, the protein bands in the gel were visualised by staining with coomassie blue solution containing 50% (v/v) methanol, 10% (v/v) acetic acid and 0.25% (w/v) coomassie blue R-250 (Sigma) and destained with a solution containing 10% (v/v) methanol, 5% (v/v) acetic acid.

Chapter 3

IN VIVO EXPRESSION OF AN OVERLAPPING GENE ENCODED BY THE CUCUMOVIRUSES

3.1 INTRODUCTION

The genus *Cucumovirus* contains three species, cucumber mosaic virus (CMV), tomato aspermy virus (TAV) and peanut stunt virus (PSV). Strains of CMV are further classified into two major subgroups (I and II) based on nucleic acid hybridisation analyses (Palukaitis *et al.*, 1992).

The cucumoviruses contain three genomic RNAs encoding four proteins. Proteins 1a and 2a, encoded by RNAs 1 and 2, respectively (Rezaian *et al.*, 1984; 1985), are involved in replication of the virus (Nitta *et al.*, 1988; Hayes and Buck, 1990). Protein 3a and coat protein (CP) are encoded by RNA 3 (Davies and Symons, 1988). 3a is translated from RNA 3 whereas CP is expressed via RNA 4, a subgenomic RNA of RNA 3 (Schwinghamer and Symons, 1977). Both 3a and CP are required for virus cell-to-cell movement (Suzuki *et al.*, 1991; Boccard and Baulcombe, 1993; Kaplan *et al.*, 1995). It has been shown recently that CP of CMV is able to complement the defective long distance movement of TAV in cucumber plants (Taliansky and Garcia-Arenal, 1995). CP is also involved in virus encapsidation.

Recently, it has been reported that RNA 2 of the Q strain of CMV (Q-CMV), a subgroup II strain, encodes an additional gene (2b) which overlaps the C-terminal region of the 2a gene (Ding *et al.*, 1994). The *in vivo* expression of the 2b gene is most likely via RNA 4A, a subgenomic RNA of RNA 2. RNA 4A is encapsidated in the viral particles, 682 nucleotides (nt) in length, and identical in sequence to the 3' terminal 682 nt of RNA 2 (Ding *et al.*, 1994). Furthermore, the translational product of Q-CMV ORF 2b has also been detected in infected plants by western blot analysis (Ding *et al.*, 1994). It has been further shown that the 2b gene of Q-CMV encodes a host-specific long-distance virus movement function (Ding *et al.*, 1995b).

Importantly, ORF 2b is conserved in RNA 2 of all five cucumoviruses examined (Ding *et al.*, 1994). These include Q-CMV (Rezaian *et al.*, 1984), Fny-CMV (Rizzo and Palukaitis, 1988), Y-CMV (Kataoka *et al.*, 1990a), V-TAV (Moriones *et al.*, 1991) and PSV (Karasawa *et al.*, 1992). ORF 2b is also conserved in other subgroup I CMV strains submitted in Genbank (O strain, Accession D10209; legume strain, Accession D16406) or reported recently (NT9 strain, Hsu *et al.*, 1995). All ORF 2bs are located at the 3'-terminal portion of RNA 2 and always occupy the +1 reading frame of ORF 2a. All ORF 2bs encode a similar size of proteins ranging from 95 amino acid residues in V-TAV to 110 amino acid residues in Fny-CMV. Predicted amino acid sequences of the 2b proteins are also conserved in different cucumoviruses (Ding *et al.*, 1994). However, no equivalent ORF was found in bromoviruses or alfalfa mosaic virus, indicating that the presence of the ORF 2b is a unique feature of the cucumoviruses.

The objective of the work described in this chapter was to determine whether the 2b gene is also expressed in the other members of cucumoviruses in addition to Q-CMV. The V strain of TAV (V-TAV) and the WAII strain of CMV (WAII-CMV), a subgroup I strain (Wahyuni *et al.*, 1992), were chosen for this reason.

3.2 MATERIALS AND METHODS

3.2.1 Isolation of viruses and the RNAs

V-TAV, WAII-CMV and Q-CMV were inoculated and propagated on *Nicotiana glutinosa, N. clevelandii* and *Cucumis sativus cv.* Green Gem, respectively. Virions, virion RNAs and total plant RNAs were isolated as described in detail in Chapter 2.

3.2.2 cDNA synthesis, amplification and cloning

The ORF 2b coding sequence plus the 3' untranslational region (UTR) of V-TAV RNA 2 was obtained by RT-PCR using primers BJ1 and BJ2 (Table 3.1) derived from RNA 2

sequence of V-TAV (Moriones *et al.*, 1991). The PCR fragment was cloned into the *Sma*I site of pBluscript SK$^+$ to produce pBSK2NVT. The cloned V-TAV sequence was obtained either as a *BamH*I-*Kpn*I fragment or as a *BamH*I-*Kpn*I fragment for subcloning into pSP72 (to produce pSP2VT) or the expression vector pGEXII (to produce pGEX2VT).

A similar strategy to that described for V-TAV was used for the cloning of WAII-CMV RNA 4A. The ORF 2b coding sequence plus the 3' UTR of WAII-CMV RNA 2 was obtained by RT-PCR using primers SD18 and SD16 (Table 3.1) derived from the known RNA 2 sequence of Fny strain of CMV (Fny-CMV), also a subgroup I strain (Rizzo and Palukaitis, 1988). The PCR fragment was cloned into the *Sma*I site of pBluscript SK$^+$ to produce pBSK2NWAC and subcloned subsequently either as a *BamH*I-*Stu*I fragment or as a *BamH*I fragment into pSP72 or to pGEXII to give rise to pSP2WAC and pGEX2WC, respectively.

3.2.3 Northern blot analyses

Virion RNAs as well as total RNAs from the respective virus infected plants, *N. glutinosa, N. clevelandii* and *N. tabacum* cv White Burley, were analysed by northern blot hybridisation using four RNA probes transcribed from linearised pBSK2NVT, pBSK2NWAC, pSP2VT and pSP2WAC, respectively, as described in Chapter 2.

3.2.4 Sequence determination

The viral cDNA inserts in pBSK2NVT and pBSK2NWAC were sequenced from both orientations using T3 and T7 primers (Table 3.1) as described in Chapter 2. The 5' UTR of RNA 4A from both V-TAV and WAII-CMV not represented by cDNA clones was determined by RNA dideoxy sequencing as described in Chapter 2. Primers BJ7 and BJ8 (Table 3.1) used are complementary to the 5' terminus of each ORF 2b of V-TAV and WAII-CMV respectively. As a control, virion RNAs and total RNAs from CMV-qt were

also sequenced using BJ7. The 5' terminal sequence was further confirmed by sequencing cDNA clones pBSKVT and pBSKWC obtained by RT-PCR using primer pairs BJ1/BJ3 and SD18/BJ4, respectively (Table 3.1). The cDNA clones pBSKVT and pBSKWC were sequenced with BJ7 and BJ8, respectively, and the reaction products used as a DNA sequence ladder to map the initiation site of RNA 4A.

3.2.5 Protein expression in *Escherichia coli* and antibody preparation

Methods for bacteria growth, induction and detection of protein expression were described in Chapter 2. Purification and antibody preparation of the 2b-GST fusion protein were mainly as described by Ding *et al.* (1994) except that a sonicator and glutathione-agarose beads were used.

Every 100 ml of *E. coli* DH5α cells containing the pGEX2VT construct was pelleted by centrifugation at 5000rpm (4°C, Sovall SS-34 rotor, same conditions below) for 10 minutes and the cells suspended in 5 ml phosphate-buffer saline [PBS (pH7.3): 137 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4-$7H_2O$, 1.4 mM KH_2PO_4] and frozen at -20°C overnight. The cells were lysed using a sonicator with a 2-mm-diameter probe at 50 W for 3 minutes and centrifuged in the presence of 1% TritonX-100 (v/v) at 10000rpm for 5 minutes. The supernatant was gently mixed with 100 µl of a 50% slurry of Glutathione Sepharose 4B (Pharmacia) for more than 2 minutes at room temperature and then microcentrifuged (Eppendorf) for 10 seconds at room temperature. The bound material was eluted with 100 µl 50 mM Tris-HCl (pH 8.0) containing 5 mM glutathione after washing twice with PBS.

As several other non-expected proteins were also present in the eluted proteins, the eluted proteins in the presence of SDS sample buffer were loaded onto a 12% SDS-polyacrylamide resolving gel as described in Chapter 2 for further purification. After staining with Coomassie brilliant blue R250, the band containing the 2b-glutathione-S-

transferase (2b-GST) fusion protein of 40K was excised and electrophoretically electroeluted into a final volume of 0.6 ml (Jacobs and Clad, 1986) using a BIOTRAP apparatus (Schleicher & Schuell). 100 μg of the eluted fusion protein in 400μl of volume was emulsified with an equal volume of complete Freund's adjuvant and injected subcutaneously into a rabbit, and followed by two booster injections with an equal volume of incomplete Freund's adjuvant at two weekly periods. Antiserum was collected two weeks after the last booster injection. The antibody against GST-ORF 4 of BYDV-PAV from the same pGEX II was a gift of Masoud Shams-bakhsh (in this laboratory) and was used as a control.

3.2.6 Protein extraction from plants and Western immunoblotting

Total protein extracts from either healthy or V-TAV infected *N. glutinosa* leaves were performed as described by Von Arnim *et al.* (1993). Two grams of *N. glutinosa* leaves were ground in liquid nitrogen to powder to which 4 ml of GB buffer [10 mM KCl, 5 mM $MgCl_2$, 400 mM sucrose, 10 mM mercaptoethanol, 100 mM Tris-HCl (pH 8.1), 10% glycerol] in the presence of protease inhibitors (Sigma: aprotinin, leupeptin, pepstatin A, and soybean trypsin inhibitor, all at 0.2μg/ml) as well as 2 mM phenylmethylsulfonyl fluoride (PMSF) were added. The suspension was centrifuged for 10 minutes at 1000g. The supernatant (S1) was further centrifuged for 20 minutes at 20000g and the pellet (P1) was washed with GB buffer, then resuspended in GBT (GB buffer containing 1% Triton X-100) and centrifuged for 10 minutes at 1000g into supernatant (S2) and pellet (P2). The P2 was washed with GBT once and resuspended in 0.5 ml of ESB (75 mM Tris-HCl, pH 6.8, 1.5 M urea, 7.5% mercaptoethanol, 4.5% SDS) and boiled for 5 minutes and then centrifuged for 10 minutes at 1000g into pellet (P3) and supernatant (S3). All supernatants obtained were precipitated by 50% acetone and together with all pellet fractions (Figure 3.1) were finally resuspended in 0.5 ml of the SDS sample buffer.

The method for Western immunoblotting assay was adapted from Ding *et al.* (1994) except that a horseradish peroxidase-labelled antibody (HRP; Amersham) was used. Protein extracts (30 µl) were electrophoresed in a 12% SDS-polyacrylamide resolving gel as described in Chapter 2 in a mini vertical electrophoresis unit (BIO-RAD) and electrophoretically transferred to a nitrocellulose membrane (0.2 µm, Schleicher and Schuell) with Towbin transfer buffer (25 mM Tris base, 192 mM glycine, 20% methanol, pH 8.3; Towbin *et al.,* 1979) using a semi-dry Trans-Blot assembly (BIO-RAD). The nitrocellulose membrane was blocked with 5% skim milk in Tris-buffered saline-Tween (TBS-T: 20 mM Tris base, 137 mM NaCl, 0.1% Tween-20, pH7.6) overnight at room temperature (same working temperature below) on an orbital shaker and then incubated with the antibody against the 2b-GST fusion protein or the antibody against GST-ORF 4 of BYDV-PAV overnight. The membrane was briefly rinsed using TBS-T only and then washed twice with the same TBS-T in 20 minutes and incubated overnight with a 1:1000 dilution of the sheep HRP labelled second antibody (Amersham), followed by washing as above. The detection of the 2b protein was performed with the luminol reagent in alkaline conditions as described by the manufacturer (Amersham).

3.3 RESULTS

As reported previously (Ding *et al.*, 1994), ORF 2b of all cucumoviruses is located in the 3' portion of RNA 2, approximately 2.4 Kb downstream from the 5'-terminus of RNA 2. Thus, for ORF 2b to be expressed *in vivo*, it is essential to establish that a subgenomic RNA in which ORF 2b becomes the 5'-proximal ORF is produced in infected plants. In this chapter, northern blot analyses were first applied for the detection of mRNA expressing the 2b gene, RNA dideoxynucleotide sequencing was used to determine the 5'-terminal sequence of RNA 4A, and then antibody against the 2b-fusion protein was raised to detect the product of the 2b gene in infected plants.

3.3.1 Characterisation of V-TAV RNA 4A

3.3.1.1 Northern blot analyses

Two probes which hybridise to different regions along the genomic RNA 2 of V-TAV were used in northern blot hybridisations to determine whether subgenomic RNA 4A (mRNA of the 2b gene) was produced in infected plants. Probe β (Fig. 3.2a) transcribed from pBSK2NVT is complementary in sequence to the 3' terminal 628 nt of V-TAV RNA 2 and was expected to hybridise to all V-TAV RNAs that contain the 3' conserved region and/or the ORF 2b coding region. This probe detected seven major RNA species (Fig. 3.2b, RNAs 1, 2, 3, 4, 4A, 3B and 5) both in total RNAs prepared from V-TAV infected plants (lane 4) and in the encapsidated viral RNAs (lane 3). The four larger RNAs correspond to the known RNAs 1-4 whereas the two smaller RNAs at the bottom of the gel represent two novel subgenomic RNAs of RNA 3, RNAs 3B and 5 (see Chapter 6). The band above RNA 3B most likely corresponded to RNA 4A as it was similar in size to the known RNA 4A (667 nt) of the Q-CMV/V-TAV chimera, CMV-qt (lane 2). The genome structure of CMV-qt is identical to that of Q-CMV except for the ORF 2b coding sequence which was from V-TAV (see Chapter 4).

A duplicate filter of Fig. 3.2b was hybridised with probe α (Fig. 3.2a) transcribed from pSP2VT, which is complementary in sequence to the 5' 234 nt of the V-TAV ORF 2b coding sequence (nt 2447-2680 of RNA 2). As shown in Fig. 3.2b (lanes 7 and 8), this probe strongly hybridised only to two of the seven major RNA species detected above for both CMV-qt and V-TAV, RNAs 2 and 4A (Fig. 1b, lanes 6, 7 and 8). The nature of the bands between RNAs 2 and 4A (lanes 6, 7 and 8) detected by probe α has not been determined. Taken together, the above results indicate that the V-TAV RNA 4A predicted (Ding *et al.*, 1994) was produced in infected plants and was encapsidated in the virions.

3.3.1.2 5'-terminal sequence

To precisely map the 5'-end of RNA 4A of V-TAV identified in the northern blots, RNA dideoxy sequencing (Fichot and Girard, 1990) was done using total RNAs extracted from the infected plants as the templates. Primer BJ7 (Table 3.1) is complementary to nt 2460-2477 of V-TAV RNA 2 and is 10 nt to the 3' side of the start codon of ORF 2b (shown as an arrow in Fig. 3.2a). As primer BJ7 binds to both RNA 2 and RNA 4A, sequencing RNAs 2 and 4A in a mixture using this primer and AMV reverse transcriptase gave rise to continuous sequence ladders plus a termination signal in all four lanes (middle panel in Fig. 3.3) and in the lane (not shown) where no dideoxynucleotide was added, indicating that this cDNA transcript was not terminated by ddNTP incorporation but was the full-length cDNA of RNA 4A. By comparison with the DNA sequence ladder using the same primer BJ7 (data not shown), the 5'-end nucleotide of subgenomic RNA 4A was mapped to nt 2373 of RNA 2 and the complete sequence of V-TAV RNA 4A was determined as 702 nt and is shown in Fig. 3.4. The size of V-TAV RNA 4A correlates well with its relative rate of migration in denaturing agarose gel electrophoresis as compared to the known size of 667 nt for CMV-qt RNA 4A (compare lanes 3 and 2, Fig. 3.2b). Thus, RNA 4A of V-TAV encodes the complete ORF 2b with a 5' UTR of 73 nt and a 3' UTR of 340 nt (Fig. 3.2a). Significantly, the initiation site of V-TAV RNA 4A determined in this work agrees exactly with that predicted previously based on sequence similarities of RNA 2 among five cucumoviruses around the known initiation site of Q-CMV RNA 4A (Ding *et al.*, 1994).

3.3.2 Characterisation of WAII-CMV RNA 4A

3.3.2.1 Northern blot analyses

Northern blot analysis using probe δ (Fig. 3.5a), which was transcribed from pBSK2NWAC and complementary in sequence to the 3' terminal 630 nt of WAII-CMV,

identified in WAII-CMV infected plants an RNA species similar in size to RNA 4A of Q-CMV (compare lanes 2 and 4 in Fig. 3.5b). As the RNA species also hybridised (lane 6) to probe γ whose sequence is complementary only to the 5' 240 nt of the WAII-CMV ORF 2b coding sequence (Fig. 3.5a), it is concluded that this RNA corresponded to the predicted RNA 4A. In contrast to RNA 4As of Q-CMV (Ding *et al.*, 1994) and V-TAV (see above), only a very small proportion of the WAII-CMV RNA 4A molecules accumulated in infected plants (lanes 2 and 6) were encapsidated in the viral particles (lanes 3 and 7), indicating varying specificity of *in vivo* encapsidation of WAII-CMV RNAs.

3.3.2.2 Partial sequence of RNA 2 and the 5'-terminal sequence of RNA 4A

Firstly, the nucleotide sequence of the 3' terminal 779 nt of WAII-CMV RNA 2 was determined (Accession number U64435) by sequencing cDNA clones obtained by RT-PCR and by direct RNA dideoxynucleotide sequencing. Sequence comparisons showed that the 3' terminal sequence of WAII-CMV RNA 2 is 98% identical to the equivalent region of Fny-CMV and 64% identical to that of Q-CMV. This further confirms a previous conclusion (Wahyuni *et al.*, 1992) that WAII-CMV is a subgroup I strain of CMV. Importantly, RNA 2 of WAII-CMV also encodes ORF 2b (Fig. 3.5a). As for Fny-CMV, ORF 2b of WAII-CMV contains 110 codons and their encoded protein sequences differ only by 5 amino acids.

The same strategy used for mapping the 5'-end of V-TAV was used to determine the 5'-end of WAII-CMV RNA 4A. Primer BJ8 (Table 3.1) used is complementary to nt 550-575 away from the 3' terminus of WA II-CMV RNA 4A (shown as an arrow in Fig. 3.5a). Again, a termination product in all four sequencing reactions occurred at a position corresponding to nt 90 of the partially sequenced RNA 2 (right panel in Fig. 3.4). This gave RNA 4A of WAII-CMV a size of 690 nt; the complete sequence of RNA 4A is

shown in Fig. 3.6. Only a very weak termination signal was seen in the RNA dideoxy sequencing using the virion RNAs as template (data not shown) and this indicated that RNA 4A was encapsidated at a very low level, which was consistent with that found in the northern blot analysis. The doublet termination signal found in the sequencing of WA II-CMV RNAs may result from the putative 5'-cap structure of RNA 4A. RNA 4A of Q-CMV is capped (Ding *et al.*, 1994); however, no such doublet was present in the sequencing reactions of either Q-CMV (data not shown) or V-TAV (Fig. 3.3).

Interestingly, the first four nucleotides (5'-GUUU-3') of each RNA 4A of V-TAV, WAII-CMV as well as Q-CMV are identical to the 5'-end of the respective RNA 2. This could be used to predict the transcription start site of subgenomic RNAs in the sequences of genomes (Kelly *et al.*, 1994). A similar occurrence of subgenomic RNA, whose initiation sequence is identical to the 5'-end of genomic RNA, has been reported in many other plant viruses (Lommel *et al.*, 1991; Cornelissen *et al.*, 1986; Symons, 1985; Miller and Mayo, 1991; Kelly *et al.*, 1994; Brooks and Bruening, 1995).

3.3.3 *In vivo* detection of the protein 2b

The 2b protein encoded by V-TAV was expressed in *E.coli* as a fusion protein with glutathione-S-transferase (GST). The antibody raised in rabbits against the 2b-GST fusion protein was used to detect expression of the 2b gene *in vivo*. Total proteins were extracted from V-TAV infected *N. glutinosa* plants harvested two weeks after inoculation using three different extraction buffers as described above (Von Arnim *et al.*, 1993).

The results (Fig. 3.7) showed that the 2b protein (marked by an arrowhead) accumulated in *N. glutinosa* plants infected with V-TAV (lanes 5 and 14). This protein was not detectable by the GST-2b antiserum in uninfected plants (lanes 1, 4, 7, 10 and 13) or in plants infected with Q-CMV (data not shown), but it was detected in plants infected with CMV-qt (lanes 6 and 15), a Q-CMV chimera expressing the V-TAV 2b protein (see

Chapter 4). Further, this protein was not detected when the preimmune serum or an antiserum against an unrelated GST fusion protein (GST fused with ORF 4 of barley yellow dwarf virus-PAV) was used as the first antibody (data not shown). Importantly, the V-TAV 2b protein was not detected in protein extracts prepared with GB buffer (lanes 2 and 8) or GBT buffer (GB buffer plus Triton X-100) (lane 11). The 2b protein only became soluble (fractions S4 and P4) after boiling in the ESB buffer containing urea and SDS (Von Arnim *et al.* 1993). The S4 represented the cell wall-enriched fraction (Von Arnim *et al.* 1993) and contained less than 5% of the total cellular protein as judged by staining with Coomassie brilliant blue (data not shown), indicating that the 2b protein is localised to the cell walls. The 2b protein was not detected in the viral particles (data not shown), suggesting that it may not be a structural protein.

The size of the 2b protein, estimated as Mr 14K under the conditions used, is larger than the 11K predicted from the sequence. A similar difference was found previously for the Q-CMV 2b protein (Ding *et al.*, 1994); such a difference might be caused by extraction conditions or post-translational modification (Hassan *et al.*, 1994) or be a consequence of abnormal migration in SDS-gels. In addition, it is likely that the other larger protein species present in plants infected with V-TAV or CMV-qt is related to the V-TAV 2b protein. It is not clear whether this larger species represents an aggregate of the 2b protein with other proteins or simply a dimeric form of the 2b protein.

The pGEX2WC inserted with the ORF 2b gene failed to express 2b protein encoded by WAII-CMV in *E. coli* as a fusion protein with glutathione-S-transferase (GST) (data not shown). Repeated attempts also failed to express the WAII-CMV 2b protein in several other protein expression systems such as pQE30, pMAL-c2 and pET-3d, indicating a possible toxic effect of the protein to the bacteria. In addition, the WAII-CMV 2b protein was not detectable by the antiserum against the Q-CMV 2b-GST fusion protein despite the fact that the 2b proteins of the two strains are 53% identical in sequence. Since the 2b

protein was detected for both Q-CMV and V-TAV and that RNA 4A accumulated in WAII-CMV infected plants, it was not pursued further to demonstrate the accumulation of the WAII-CMV 2b protein.

3.4 DISCUSSION

It has been shown in this work that the 2b gene encoded by two cucumoviruses (V-TAV and WAII-CMV) was expressed *in vivo* by detecting the accumulation of the mRNA (V-TAV and WAII-CMV) and protein product (V-TAV) in infected plants. The expression of the Q-CMV 2b gene in infected plants has been reported previously (Ding *et al.*, 1994). Furthermore, the 2b gene is conserved in RNA 2 of the 11 cucumoviruses or strains (strains Y, Fny, WAII, L, O, TW, K, I and Q of CMV, V-TAV and J-PSV) presented in the database. It is therefore concluded that the 2b gene is a common feature of the *Cucumovirus* genus, and that the genome of cucumoviruses encodes five genes, two of which are translated *in vivo* from subgenomic mRNAs (Fig. 3.8).

3.4.1 RNA 4A may have a signal for encapsidation

In contrast to RNAs 1-4, RNA 4A of WAII-CMV was hardly encapsidated in the viral particles (Fig. 3.5b). However, the amount of RNA 4A encapsidated was comparable to that of the four other viral RNAs for Q-CMV (Fig. 3.5b) and the amount of V-TAV RNA 4A in virions was significantly lower than that of RNAs 1-4 (Fig. 3.2b). Several possibilities can account for the observed differences in the encapsidation level of RNA 4A:

- There is a difference in the amount of RNA 4A transcribed in plants infected with different cucumoviruses due to the effectiveness of the respective subgenomic RNA promoters. This is likely true in the case of V-TAV, which produced only small amount of RNA 4A *in vivo*, and a lower level of RNA 4A present in the virions is the result of less RNA 4A molecules available for encapsidation.

- The location of the origin of assembly (OAS) in RNA 2 may vary in different cucumoviruses. The OAS sequence of RNA 2 from subgroup II CMV strains may be located within the RNA 4A-equivalent part of RNA 2, whereas in subgroup I CMV strains the OAS sequence of RNA 2 may be located the 5'-end to the initiation site of RNA 4A. This is the reason for the differential encapsidation of the coat protein subgenomic RNA of two different strains of tobacco mosaic tobamovirus (Turner *et al.*, 1988). For the chimeric virus between Q-CMV and WAII-CMV in which the ORF 2b coding region was replaced with that from WAII-CMV (see Chapter 4), a large amount of the chimeric RNA 4A was encapsidated (data not shown). This suggests that the ORF 2b coding sequence of Q-CMV RNA 4A is not essential for encapsidation and the OAS sequence of Q-CMV RNA 4A may be located in the untranslated regions.

The sequence that determines the specificity of encapsidation might be located at the 5'-end like the genomes of bacteriophage ϕ and hepatitis B virus (Gottlieb *et al.*, 1994; Kawamoto *et al.*, 1994), or at the 3'-end like the L-A viral RNA (Fujimura *et al.*, 1990), or at both the 5'- and the 3'-ends of the RNA like defective interfering particles of reovirus (Zou and Brown, 1992). It is also possible that successful encapsidation of RNA 4A may require the coat protein to interact with the RNA at a subcellular location as has been proposed for the encapsidation of poliovirus RNA (Ansardi *et al.*, 1996). Thus, it is necessary to construct a few more subgroup I/II chimeric viruses as well as those based on subgroup I CMV before any specific conclusions can be made.

3.4.2 Protein 2b may be involved in the movement of the virus

Total proteins extracted from infected plants were fractionated by centrifugation as well as by extraction using three different buffers (Fig. 3.1). The 2b protein was solubilized from the pellet fractions only when treated with denaturants such as urea, SDS as well as boiling, but not with the detergent Triton X-100, indicating that the 2b protein is localised

to the cell wall. Q-CMV 2b protein has also been localised to the same fraction (S.W. Ding, personal communication). The movement proteins of tobacco mosaic virus (TMV), alfalfa mosaic virus (AlMV) and cauliflower mosaic virus (CaMV) (Godefroy-Colburn *et al.*, 1986; Deom *et al.*, 1987; Albrecht *et al.*, 1988; Ding *et al.*, 1995) have been found in cell wall by using essentially identical cell wall enrichment protocols. Compared with putative movement proteins of viruses belonging to 17 groups (Mushegian and Koonin, 1993), the C-terminal of the TAV 2b protein shows a weak homology to the identified conserved motif consisting of approximately 30 amino acid residues of CMV and BMV 3a proteins (data not shown). Taken together, the 2b protein might share a function in common with these movement proteins, which is supported by the recent evidence that the 2b gene of Q-CMV encodes a host-specific long-distance virus movement function (Ding *et al.*, 1995b).

Table 3.1. Primers used for either cDNA cloning or sequencing of RNA and cDNA

Primer	Primer sequence	Primer position	Primer usage
BJ1	5'-AGGATCCTGGGACCCCTAGGGGGAACCTACGGA-3'	*V-TAV RNA 2 nt 3049-3074*	construction of pBSK2NVT
BJ2	5'-TGGATCC**ATG**GCAAGCATCGAGATCCCTCTACA-3	V-TAV RNA 2 nt 2447-2472	construction of pBSK2NVT
BJ3	5'-ACGCCTATAAAGAGTTGTA-3'	V-TAV RNA 2 nt 2318-2336	construction of pBSKVT
BJ4	5'-TTTTCAGTCCTACAAAGAACT-3'	nt 730-750 from the 3'-end of WAII RNA 2	construction of pBSKWC
BJ7	5'-TCTCGTGTAGAGGGATCT-3'	*V-TAV RNA 2 nt 2460-2477*	sequencing
BJ8	5'-AGCCAGTTGGAGTTCGACGT-3'	*nt 594-603 from the 3'-end of WAII RNA 2*	sequencing
SD18	5'-TGGATCCTGGTCTCCTTTTGGAGGCCCCA-3'	*Fny-CMV RNA 2 nt 3029-3050*	construction of pBSK2NWC
SD16	5'-AGGATCC**ATG**GAATTGAACGTAGGTGCAA-3'	Fny-CMV RNA 2 nt 2419-2440	construction of pBSK2NWC
T3	5'-ATTAACCCTCACTAAAG-3'	flanking in MCS of pBluescript SK^+	sequencing
T7	5'-AATACGACTCACTATAG-3'	flanking in MCS of pBluescript SK^+	sequencing

The single underlined nucleotides in the primer sequences indicate BamHI and NcoI (the second underline in BJ2 and SD16) sites, respectively.

The nucleotides in bold indicate the initiation codon of ORF 2b.

Viral RNAs and residue numbers in italics indicate that the primer sequences are complementary to the positive strand of the respective RNA 2s while the others correspond to the positive strand of the respective RNA 2s.

MCS in the table is the abbreviation of multi-cloning site.

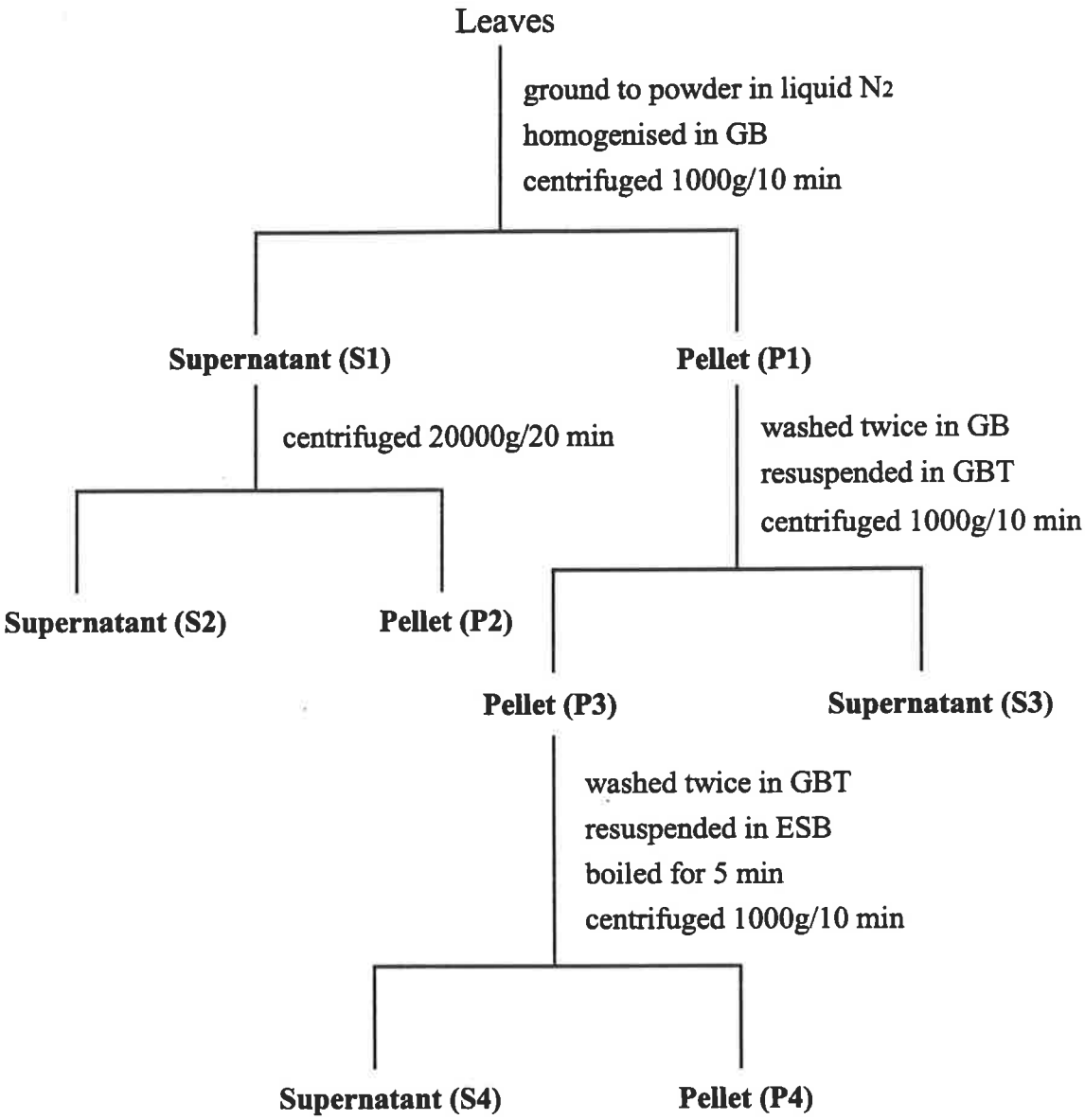


Fig. 3.1 Fractionation scheme for total proteins extracted from plants. Descriptions in detail are given in the text.

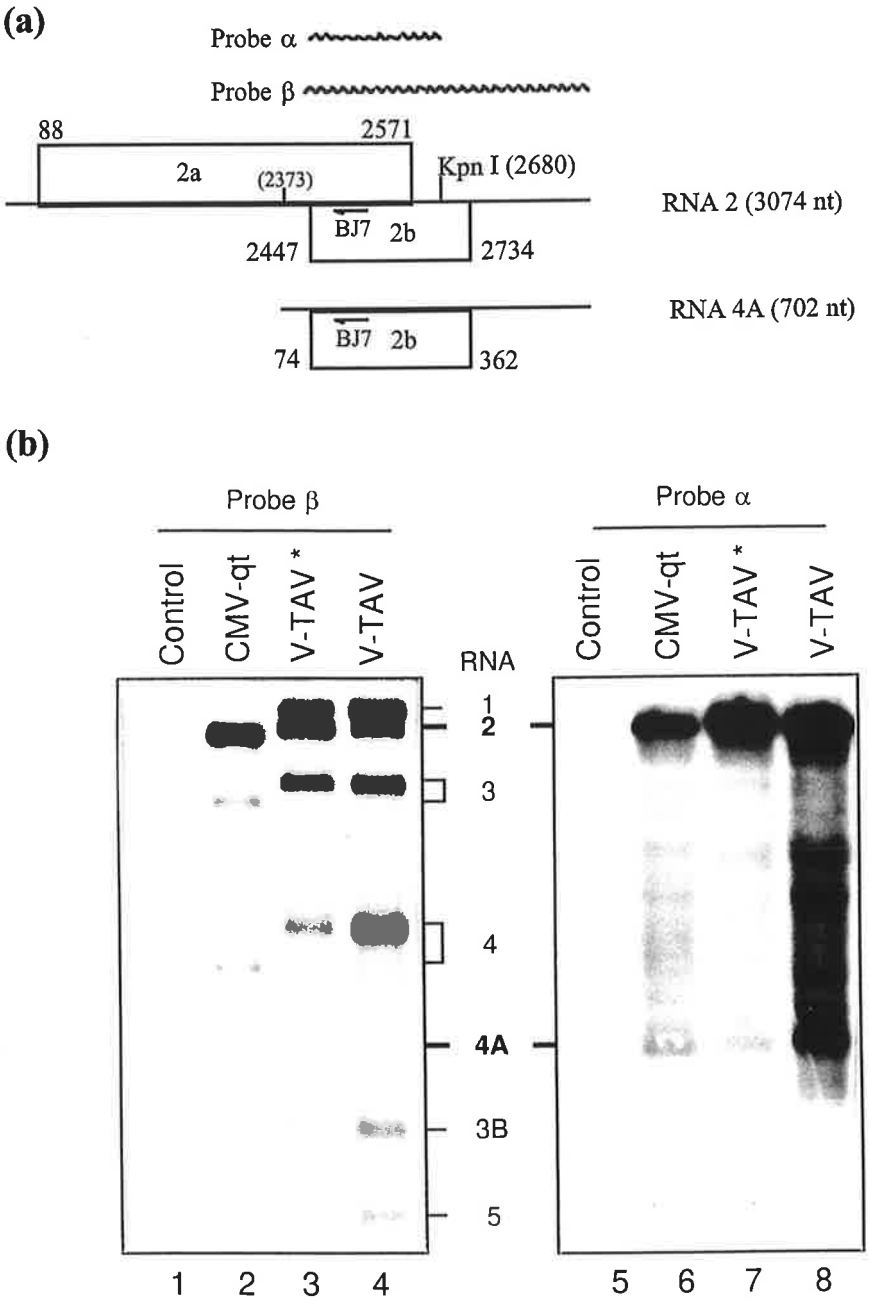


Fig. 3.2 Structure and expression of V-TAV RNA 2. (a) Structure of V-TAV RNA 2 and RNA 4A. Open reading frames (ORFs) for both 2a and 2b genes are shown as open boxes. The wavy lines indicate the positions of strand-specific RNA probes, used in the northern blot hybridisation, which are complementary to 627 nt (probe α, between 2b initiation site and 3'-end) and 204 nt (probe β, between 2b initiation site and *Kpn*I site) of RNA 2. The arrowhead lines within RNA 2 and RNA 4A ORFs indicate the positions of primers used in RNA dideoxynucleotide sequencing. (b) Northern blot analysis of viral RNAs. V-TAV virion RNAs marked by an asterisk (lanes 3 and 7) and total RNAs extracted from *N. glutinosa* inoculated with sterile water (lanes 1 and 5), CMV-qt (lanes 2 and 6) or V-TAV (lanes 4 and 8) were hybridised with probe β (lanes 1-4) or probe α (lanes 5-8). The positions of viral RNAs 1, 2, 3, 4, 4A, 3B and 5 are indicated. RNAs 3B and 5 are the newly discovered subgenomic RNAs derived from RNA 3 (see Chapter 6).

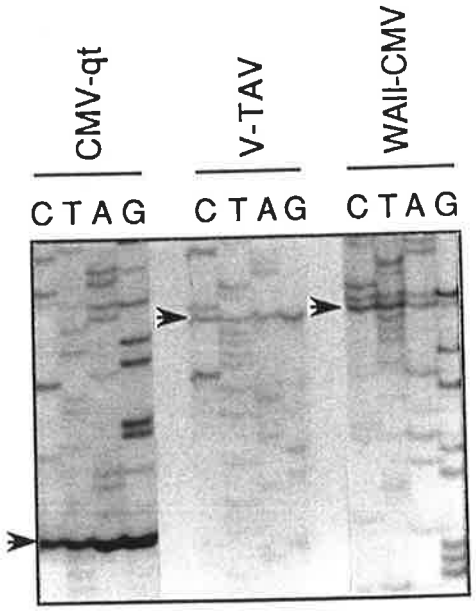


Fig. 3.3 Sequence analysis of the 5'-end of RNA 4A of V-TAV and WAII-CMV using the RNA dideoxy nucleotide sequencing method. The total RNAs extracted from *N. glutinosa* infected with CMV-qt (left panel), V-TAV (middle panel) and WAII-CMV (right panel), respectively, were used as templates and annealed with primers BJ7 and BJ8 shown in Figs 3.2(a) and 3.5(a), respectively. G, A, T, and C represent single sequencing reactions to which ddGTP, ddATP, ddTTP and ddCTP was added respectively. The band indicated by the arrow in each panel represents the full-length cDNA transcript of RNA 4As and thereby the 5' terminal sequence of RNA 4A was obtained from this sequencing gel and by reference to the nucleotide sequence of the corresponding region of the genomic RNA.

```
(2373)
GUUUUCUAUAAGUGAUUUGAUUGUUGGUCGUGGAAAAUCUUCGAAAGCUUCUAGGAGGA
1
AAGCGGUUGAGAGUA AUG GCA AGC AUC GAG AUC CCU CUA CAC GAG AUC
                Met ala ser ile glu ile pro leu his glu ile

AUU CGA AAG UUG GAA CGG AUG AAU CAA AAG AAA CAA GCA CAG AGG
ile arg lys leu glu arg met asn gln lys lys gln ala gln arg

AAA CGA CAC AAA CUG AAC CGC AAG GAG CGG GGU CAC AAA AGU CCA
lys arg his lys leu asn arg lys glu arg gly his lys ser pro

AGU GAA CAA AGG CGA UCG GAG UUA UGG CAC GCG CGU CAA GUU GAA
ser glu gln arg arg ser glu leu trp his ala arg gln val glu

CUU UCU GCC AUU AAU UCC GAU AAU UCU UCA GAU GAG GGU ACC ACU
leu ser ala ile asn ser asp asn ser ser asp glu gly thr thr

CUG UGU CGC UUU GAC ACA UUU GGU UCC AAG UCU GAU GCU AUU UGU
leu cys arg phe asp thr phe gly ser lys ser asp ala ile cys

GAU CGC UCU GAC UGG UGU CUC GAU CAA UGA UAAUAUCCGAUGUUGUCAU
asp arg ser asp trp cys leu asp gln

GUCCGAAGACGUUAAACUACGCUUGAACCGUGUUCGAGUGUCUGAGUUGGUAGUAUUGC

UCUAAACUAUCUGAAGUCACUAAACGCUUGUGCGGUGAACGGGUUGUCCAUCCAGCUAA

CGGCUAAAAUGGUCAGUCAUGUCGUAAGACAUGCCGUCGGUCUUUGAUCGAUGAGGUGC

CUUUGAACCCUUUAUCCCGGGGUUCUUCGGAAGGUGAGACUUGAAUUCCAUGUAGAGUC

UCGCCGUGCACGGUAUCACACUGAUGAUACCUUCAGAGUGCAGGCAUCGCCUACGGUUU

                                (3074)
UCCGUAGGUUCCCCCUAGGGGUCCCA
                                702 nt
```

Fig. 3.4 Complete sequence of subgenomic RNA 4A of V-TAV together with the predicted amino acid sequence of ORF 2b. Residue numbers in parentheses refer to the equivalent position in RNA 2 of V-TAV.

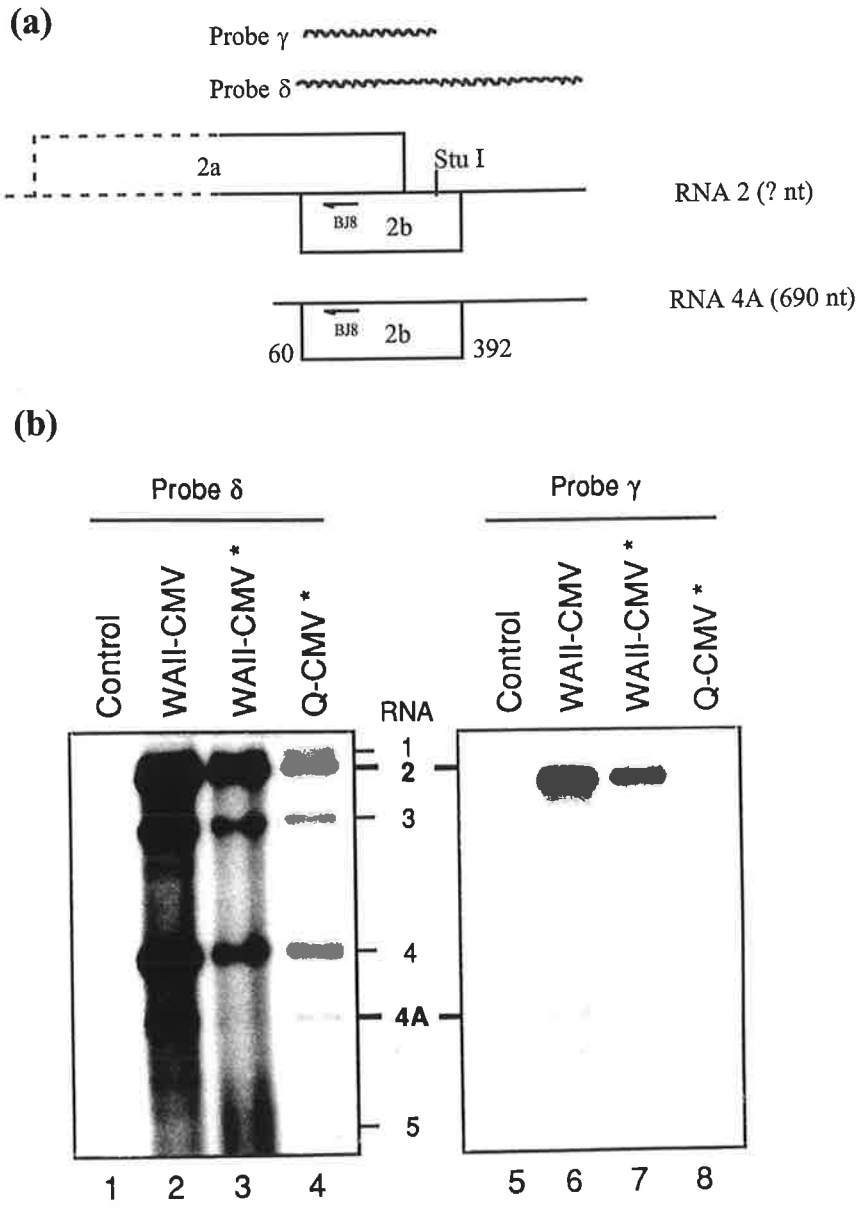


Fig. 3.5 Structure and expression of WAII-CMV RNA 2. (a) Structure of WAII-CMV RNA 2 and RNA 4A. Only 3'-terminal sequence of 779 nt of RNA 2 was determined. The remain undetermined sequence is indicated by the dotted line. Open reading frames (ORFs) for both 2a and 2b genes as shown as open boxes. The wavy lines indicate the positions of strand-specific RNA probes, used in the northern blot hybridisation, which are complementary to 630 nt (probe δ, between 2b initiation site and 3'-end) and 251 nt (probe γ, between 2b initiation site and *Stu*I site) of RNA 2. The arrowhead lines within RNA 2 and RNA 4A ORFs indicate the positions of primers used in RNA dideoxynucleotide sequencing. (b) Northern blot analysis of viral RNAs. Virion RNAs (marked by an asterisk) of WAII-CMV (lanes 3 and 7) or Q-CMV (lanes 4 and 8) and total RNAs extracted from *N. glutinosa* inoculated with sterile water (lanes 1 and 5) or WAII-CMV (lanes 2 and 6) were hybridised with probe δ (lanes 1-4) or probe γ (lanes 5-8). The positions of viral RNAs 1, 2, 3, 4, 4A and 5 are indicated.

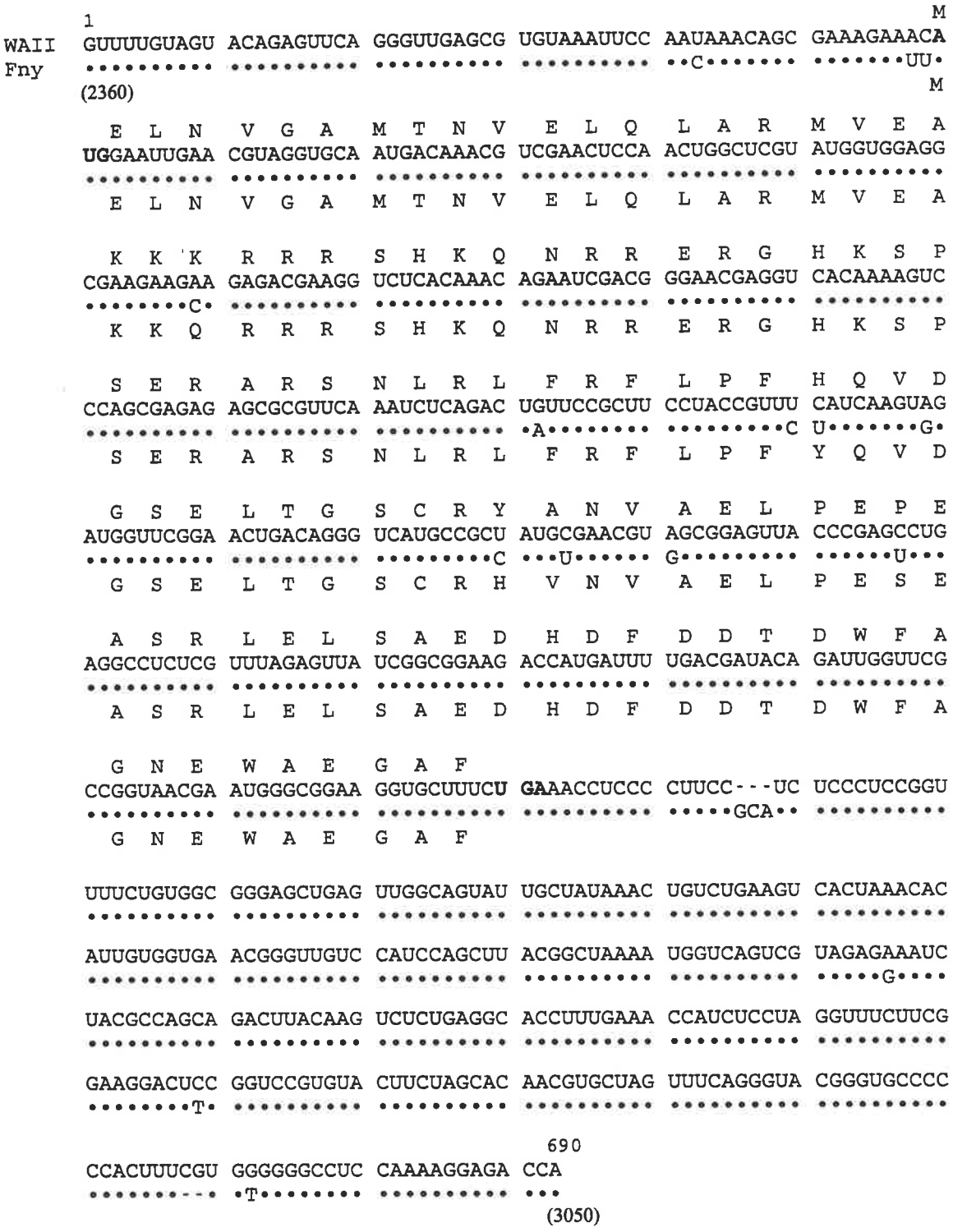


Fig. 3.6 Complete sequence of subgenomic RNA 4A of WAII-CMV together with alignment with 3' 691 nt of Fny-CMV RNA 2. The predicted amino acid sequences of both ORF 2bs are indicated, respectively. The differences between the two nucleotide sequences are indicated. Dots indicate an identical sequence and dashed lines indicate a missing sequence. Residue numbers in parentheses refer to the position in Fny-CMV RNA 2.

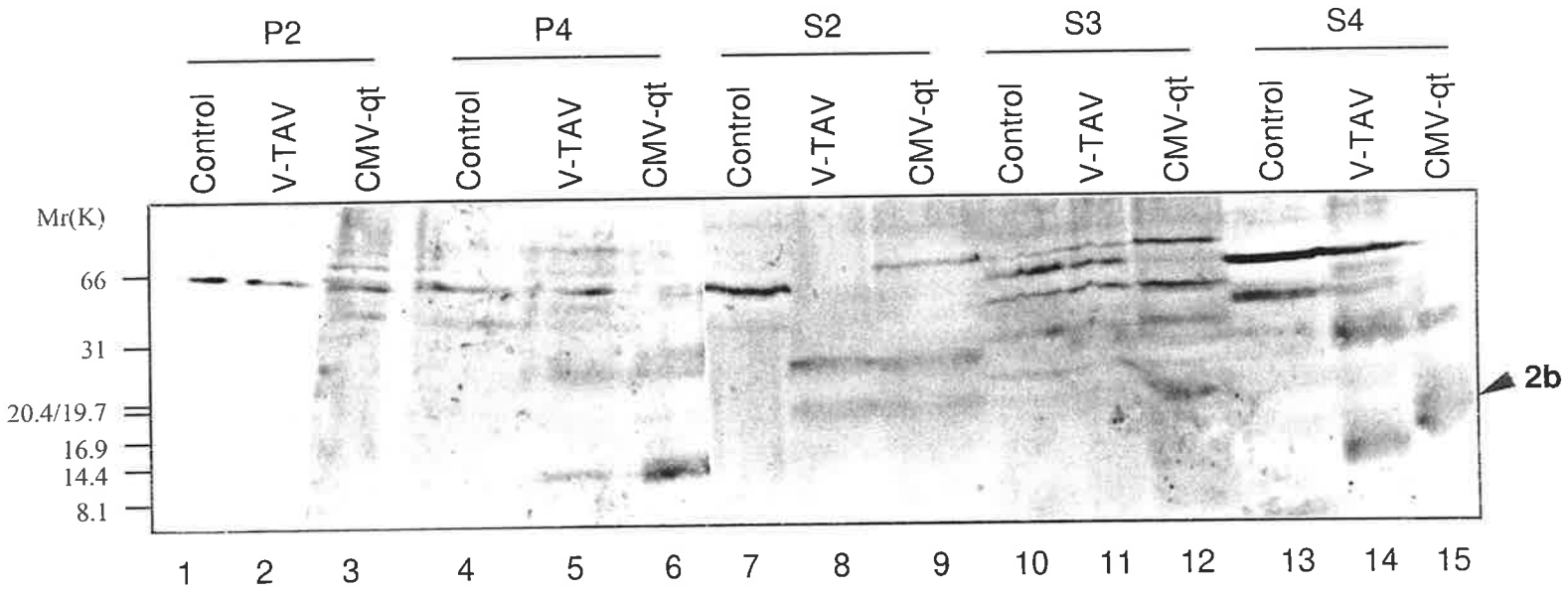


Fig. 3.7 Western blotting analysis of total proteins extracted from *N. glutinosa* using antibody against the V-TAV 2b protein fused with GST. Total proteins were extracted from mock-inoculated plants (lanes 1, 4, 7, 10 and 13), or V-TAV-infected plants (lanes 2, 5, 8, 11 and 14) or CMV-qt-infected plants (lanes 3, 6, 9, 12 and 15) two weeks after inoculation and fractionated according the scheme shown in Fig. 3.1. Letters P and S represent pellet fraction and supernatant fraction, respectively. The position of the 2b protein detected is indicated by an arrowhead. Protein molecular weight markers (Promega) are shown to the left of the gel in kilodalton (K).

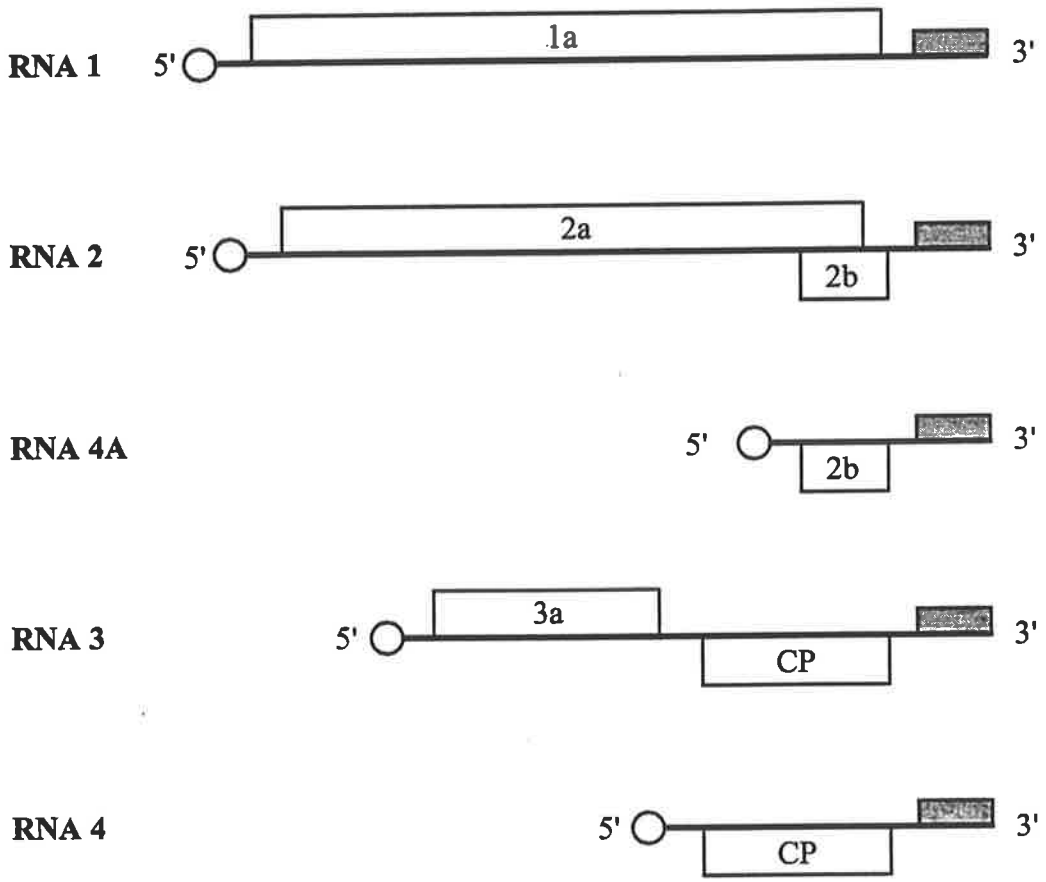


Fig. 3.8 Genome organization and expression of the cucumoviruses. Three genomic RNAs (RNAs 1, 2 and 3), together with represent two subgenomic RNAs (RNAs 4 and 4A) act as messenger RNAs for in vivo expression of five virus-encoded proteins, 1a, 2a, 3a, CP and 2b, respectively. Open reading frames are shown as open boxes. 3' terminal homologogous regions are indicated by filled rectangles. Cap structures are indicated by circles.

Chapter 4

FUNCTIONAL ANALYSIS OF THE 2B GENE BY CONSTRUCTING HYBRID VIRUSES BASED ON Q-CMV

4.1 INTRODUCTION

Chapter 3 described the detection of the *in vivo* expression of the 2b gene encoded by V-TAV and WAII-CMV which is a subgroup I strain of CMV. Considering the fact that ORF 2b is conserved in all cucumoviruses sequenced to date, it is therefore concluded that the 2b gene expression is a common feature of all cucumoviruses.

A mutational analysis of the 2b gene encoded by Q-CMV using the infectious cDNA clones has shown that the 2b gene of Q-CMV encodes a host-specific long distance virus movement function (Ding *et al.*, 1995b). It is thus of great interest to determine whether the 2b gene encoded by V-TAV or WAII-CMV has a similar function. The infectious cDNA clones of both V-TAV and WAII-CMV were not available at the time when this study commenced. To circumvent this problem, hybrid viruses expressing the 2b protein of either of the two viruses were constructed based on the infectious cDNA clones of Q-CMV. It has been shown that the Q-CMV mutant expressing no 2b protein failed to systemically infect cucumber and caused essentially no systemic symptoms in *Nicotiana glutinosa* (Ding *et al.*, 1995b). Thus, any functional complementation provided by the substituting heterologous 2b gene will be readily recognised. Furthermore, heterologous genes encoded by hybrid viruses have been shown to function to some extent although not as efficiently as the authentic genes (De Jong *et al.*, 1992; Pletnev *et al.*, 1992; Hilf and Dawson, 1993; Mise *et al.*, 1993; Burgyan et al., 1993; Peng *et al.*, 1995; Solovyev *et al.*, 1996; Kuhn *et al.*, 1996).

4.2 MATERIALS AND METHODS

4.2.1 Acknowledgments

The work presented in this chapter involving the construction of infectious interspecies hybrid cDNA clones and their derivatives, and the studies conducted using these

constructs, was carried out in collaboration with Dr Shou-Wei Ding and Ms Wan-Xiang Li.

4.2.2 Plasmid constructs

Plasmids pQCD1, pQCD2 and pQCD3 used for efficient Q-CMV infection have been described by Ding *et al.* (1995a). Plasmid pQCD2qt, a derivative of pQCD2, was constructed using three overlapping PCR fragments I, II and III (Fig. 4.1). Fragment II, representing the coding sequence of TAV ORF 2b [nucleotide (nt) 2447-2734 of TAV RNA 2, Moriones *et al.*, 1991], was obtained from the virion RNAs by reverse transcription and polymerase chain reaction (RT-PCR) as described by Ding *et al.* (1995b) using primers SD22 and SD23 (Table 4.1). Fragments I and III flank Q-CMV ORF 2b (nt 2410-2712) and represent nt 2050-2411 and nt 2713-3035 of Q-CMV RNA 2, respectively; these two fragments were obtained by two separate PCRs using pQCD2 as the template and primer pairs SD45/SD24 and SD25/SD6, respectively. SD6 and SD24 are complementary to nt 3011-3035 and nt 2388-2409 whereas SD25 and SD45 correspond to nt 2711-2730 and nt 2050-2075 of Q-CMV RNA 2, respectively (Ding *et al.*, 1995a). It is to be noted that the primers SD6, SD24, SD25 and SD45 have not been put in Table 4.1 because they are not mutagenic primers, whereas Table 4.1 contains only mutagenic primers.

Fragments I, II and III were assembled into the chimeric fragment of approximately 970 bp long in a final PCR using primers SD45/SD6 since fragment II contained sequences at its termini that overlap I and III. This overlap was defined by the chimeric primers SD22 and SD23 (Table 4.1) used to generate fragment II. This PCR product was digested by *Asp*718 and the resultant *Asp*718 fragment was used to substitute the wildtype *Asp*718 fragment (corresponds to nt 2094 and nt 2904 of Q-CMV RNA 2) of pQCD2 to yield pQCD2qt.

A similar strategy was adapted to construct pQCD2qw (Fig. 4.2), in which the 2b sequence of pQCD2 was replaced with the homologous sequence from WAII-CMV using three sets of primers, SD24 and SD45, BJ14 and BJ15 (Table 4.1), and SD6 and SD25.

Specific mutations were introduced into pQCD2qt to create pQCD2qt1 and pQCD2qt2 (Fig. 4.2) by a PCR mutagenesis protocol as described by Ding *et al.* (1995b). Each set of two fragments with the desired mutation was amplified using primer set pairs SD6 and SD43 (Table 4.1), SD45 and SD44 (Table 4.1), and primer set pairs SD6 and SD47 (Table 4.1), SD45 and SD48 (Table 4.1), respectively. The third fragment was amplified using the two previous fragments and primers SD6 and SD45, and digested by *Asp*718. Each of the resulting *Asp*718 fragments was used to substitute for the wildtype *Asp*718 fragment of pQCD2 to yield mutant plasmid constructs of pQCD2qt1 and pQCD2qt2.

The *Asp*718 fragments in all four constructs (pQCD2qt, pQCD2qt1, pQCD2qt2 and pQCD2qw) that had been manipulated by PCR were completely sequenced before use.

4.2.3 Plant inoculation and virus progeny analysis

Methods for inoculating plants with inocula, extraction of RNAs from plants, northern blot hybridisation, virus progeny RNA analysis by direct RNA dideoxynucleotide sequencing and by sequencing the cloned cDNA obtained by RT-PCR were as described in Chapter 2. Virions (0.2mg/ml) of Q-CMV, V-TAV, WAII-CMV, CMV-qt (derived from inoculum pQCD2qt+pQCD1+pQCD3, in which pQCD1 and pQCD3 are the wild type cDNA clones of Q-CMV RNAs 1 and 3) and CMV-qw (derived from inoculum pQCD2qw+pQCD1+pQCD3) were purified from the *N. glutinosa* plants infected by respective virus or plasmids and used to inoculate seedlings of *Datura stramonium, Gomphrena globosa, Lycopersicon esculentum, Physalis floridana, Stellaria media, N. tabacum* cv. White Burley and *N. glutinosa*.

4.2.4 Preparation and transfection of tobacco protoplasts

Protoplasts were prepared as described by Aryan *et al.* (1991). Exponentially growing tobacco (*N. tabacum*) NT-1 suspension cells (three days after subculturing at 26°C with shaking) were precipitated by centrifugation at 100g for 10 minutes. The pellets were resuspended in an enzyme solution [0.1% Pectolyase Y-23, and 1% Cellulase (Onozuka RS) in 0.4 M mannitol, pH 5.8], and incubation was continued for 30 minutes at 37°C without shaking. The digest was centrifuged at 100g for 10 minutes. The protoplast pellets were washed (100g, 5 minutes) twice in 0.4 M mannitol containing 20 mM MES, pH 5.5. The protoplast pellet was resuspended in a desired volume of 0.4 M mannitol to give a concentration of approximately 300,000 - 400,000 protoplasts per ml and divided into 1 ml aliquots and centrifuged (100g, 5 minutes) again.

The transfection of protoplasts followed the protocol described by van der Vossen *et al.* (1993). The protoplasts were carefully resuspended in 30 µl of water containing 4 µg virion RNAs. 100 µl PEG solution (40% PEG and 3 mM $CaCl_2$) was added to the protoplast suspension and the mixture incubated for 20-30 seconds, followed by washing with 0.4 M mannitol. The protoplasts were collected immediately after transfection and washing or after incubation in a protoplast incubation medium (Aryan *et al.*, 1991) for 24 hours at 26°C in dark.

Total RNAs were extracted from the protoplasts using TRIzol Reagent (GibcoBRL). The protoplasts were resuspended in 0.25 ml TRIzol Reagent and lysed by repeated pipetting. After incubation at room temperature for 5 minutes, 0.05 ml chloroform was added and the suspension incubated at room temperature for a further 3 minutes, and then centrifuged for 10 minutes at 14000rpm (Eppendorf) at 4°C. Iso-propanol (0.125 ml) was added to the upper aqueous colourless phase which was incubated for 10 minutes and then centrifuged for 10 minutes at 14000rpm (Eppendorf) at 4°C. The RNA pellet

was dissolved in 10 μl of 0.1 mM EDTA at 55°C for about 10 minutes after washing by 70% ethanol and analysed by northern blot hybridisation.

4.3 RESULTS

The structures of the two hybrid RNA 2 cDNA clone constructed, pQCD2qt and pQCD2qw, are shown in Fig. 4.2. Their structures were the same as that of pQCD2 (the RNA 2 cDNA clone of wildtype Q-CMV as shown in Fig. 4.2) except for the ORF 2b coding sequence which was from either V-TAV or WAII-CMV. The amino acid sequences encoded by the ORF 2b of Q-CMV are 22.6% identical to V-TAV, the most different pair of known cucumoviruses (Ding *et al.*, 1994), and 53.4% to WAII-CMV.

4.3.1 Characterisation of the hybrid viruses in the infected tobacco plants

pQCD2qt, like wildtype pQCD2, was biologically active when co-inoculated with pQCD1 and pQCD3 as the resultant hybrid virus (CMV-qt) accumulated in systematically infected leaves (upper uninoculated leaves) of the inoculated tobacco plants (Fig. 4.3c, lane 3). To examine the nature of the progeny RNA 2, total RNAs extracted from systemic leaves three weeks post-inoculation were hybridised with the probe ω (Fig. 4.3a) specific for the 3' ends of all Q-CMV positive-strand RNAs and the probe α (Fig. 4.3b; see details in Chapter 3) specific for the positive strands of the TAV 2b gene. Three genomic RNAs (RNAs 1, 2 and 3) and two subgenomic RNAs (RNAs 4 and 4A) were detected in the tobacco plants infected either with Q-CMV (Fig. 4.3c, lane 2) or with CMV-qt (Fig. 4.3c, lane 3). Of these viral RNAs, RNAs 2 and 4A also hybridised to the probe α (Fig. 4.3d, lane 8), indicating the presence of the engineered TAV 2b gene in the progeny. The specificity of the probe α was demonstrated by the absence of hybridisation signal with any of Q-CMV RNAs (Fig. 4.3d, lane 7) and by its hybridisation strongly with RNAs 2 and 4A of TAV (Fig. 4.3d, lane 9). The nature of

the bands between RNAs 2 and 4A (lanes 8 and 9) detected by the probe α has not been determined.

The 3' terminal sequence of 667 nt (including the inserted TAV ORF 2b sequence) of the progeny RNA 2 derived from pQCD2qt (Fig. 4.2) was completely determined from three independent cDNA clones obtained by RT-PCR. The data showed that the progeny RNA 2 contained the complete ORF 2b coding sequence from TAV, but not the Q-CMV ORF 2b coding sequence, and that the nucleotide sequence of this region of the progeny RNA 2 was identical to the corresponding region of pQCD2qt. This result not only further confirmed the chimeric nature of the progeny RNA 2 but also indicated that the engineered TAV sequence had been stably and accurately maintained as an integrated part of the hybrid viral RNA genome.

The chimeric RNAs 2 and 4A were also detected by northern blot analysis in virions purified from plants infected with CMV-qt. A primer (5'-TCTCGTGTAGAGGGATCT-3') that binds to a region of the chimeric RNA 2 which is from nt 14 to nt 31 from the start codon of the TAV ORF 2b was used in a primer extension experiment which showed that the 5' initiation site of the chimeric RNA 4A was identical to that of Q-CMV RNA 4A (see Chapter 3, Fig. 3.3). Thus, replacing the Q-CMV ORF 2b coding sequence with the homologous TAV sequence had no obvious effect on either the encapsidation or the transcriptional initiation of chimeric RNA 4A.

An antiserum raised against the TAV 2b protein which was fused with glutathione S-transferase (GST) and produced in *Escherichia coli* detected the 2b protein of M_r 14K in the TAV infected plants by Western blotting (see Chapter 3, Fig. 3.7, lanes 5 and 14). A protein species of the same mobility as the TAV 2b protein was also detected by this antiserum in protein extracts of the tobacco plants infected by CMV-qt (see Chapter 3, Fig. 3.7, lanes 6 and 15); this unequivocally demonstrated the *in vivo* expression of TAV ORF 2b from CMV-qt.

pQCD2qw, which contains the 2b gene sequence of WAII-CMV RNA 2, was also biologically active when co-inoculated with pQCD1 and pQCD3. The resultant hybrid virus (CMV-qw) accumulated in systematically infected leaves (upper uninoculated leaves) of the inoculated tobacco plants (Fig. 4.4c, lane 2) as detected by the probe ω (see Figs. 4.3a and 4.4a). As expected, RNAs 2 and 4A of CMV-qw (Fig. 4.4d, lane 8) were also detected by the WAII-CMV 2b gene-specific probe, probe γ (Fig. 4.4b; see details in Chapter 3). DNA sequencing analysis showed that the progeny RNA 2 contained the complete ORF 2b coding sequence from WAII-CMV without any nucleotide variations (data not shown), indicating that the engineering of the WAII-CMV 2b gene was successful and that it was stable.

4.3.2 Virulence of interspecies hybrid virus in infected plants

In the glasshouse conditions used, systemic infection of tobacco plants (*N. tabacum* cv Samsun NN) with Q-CMV produced only light mottling that was hardly visible and the infected plants appeared healthy (Fig. 4.5). However, tobacco plants infected with CMV-qt showed severe symptoms (Fig. 4.5); leaves were distorted and the plants dwarfed. At later stages of infection, the CMV-qt infected plants produced fern-leaf symptoms. CMV-qt was also significantly more virulent than TAV, although TAV induced clearly visible symptoms as it also did in *P. floridana* (Fig. 4.5, see below). Importantly, CMV-qt induced systemic symptoms to appear rapidly, at least three days earlier than those induced by TAV. Further, the genomic RNAs of CMV-qt clearly accumulated to a much greater concentration in infected tobacco plants than did those of Q-CMV (data not shown). This increase in viral RNA accumulation correlates with the hypervirulence of CMV-qt.

The susceptibility of six additional host species belonging to three families (Table 4.2) was tested to determine whether the synergistic responses associated with CMV-qt infection were specific to particular host species. Q-CMV did not cause obvious

symptoms in any of the hosts listed in Table 4.2 except for *N. glutinosa* although it infected them all. TAV induced clearly visible symptoms in *P. floridana, L. esculentum*, and *N. glutinosa* (Fig. 4.5 for *P. floridana*; and data not shown) in addition to *N. tabacum* as described above. The results obtained, as summarised in Table 4.2, showed that CMV-qt induced rapid appearance of severe systemic symptoms (Fig. 4.5 for *P. floridana*) and the viral RNAs accumulated to high levels (data not shown) in all of these plants. Thus, these six host species responded to CMV-qt infection in the same way as the tobacco plants described above. In the host species which displayed visible symptoms when infected with Q-CMV or V-TAV (Table 4.2), the systemic symptoms induced by CMV-qt were significantly more severe and appeared at least three days earlier than those caused by infection with the parental viruses. The observed features of the CMV-qt infection in *N. glutinosa* also indicate that the TAV 2b gene is functionally similar to the 2b gene of Q-CMV since a Q-CMV mutant expressing no 2b protein induced a delayed appearance of the very mild systemic symptoms in this host as compared to wildtype Q-CMV (Ding *et al.*, 1995b).

4.3.3 Accumulation of the hybrid virus RNAs in protoplasts

The molecular basis of increased viral RNA accumulation in infected plants was followed in tobacco protoplasts inoculated with RNAs prepared from purified viral particles of Q-CMV and CMV-qt. Infectious plasmid DNAs were not used here because viral RNA transcripts must firstly be transcribed from these plasmids, presumably in the nuclei, and transported back to the cytoplasm before a normal cycle of viral gene expression/replication can begin; these additional processes may have an effect on the profile of virus replication if monitored for only a short period post-inoculation. Northern blot analysis of total RNAs extracted from the protoplasts immediately after inoculation detected only trace amounts of viral genomic RNAs. However, after 24 hours incubation at 26°C, a significant amount of viral genomic and subgenomic RNAs accumulated in the

protoplasts that had been inoculated with Q-CMV virion RNAs (Fig. 4.6, lane 2), showing that the introduced viral RNAs replicated in these protoplasts. Clearly, CMV-qt and Q-CMV accumulated the three genomic (RNAs 1, 2 and 3) and two subgenomic RNAs (RNAs 4 and 4A) to a similar level (Fig. 4.6, lanes 3 and 2), thus indicating that the greater amount of CMV-qt RNAs compared with Q-CMV RNAs in infected plants was unlikely to result from differences in the relative rates of virus replication.

4.3.4 Protein 2b determines the severity of the symptoms

As ORF 2b overlaps ORF 2a, the V-TAV RNA sequence of the hybrid RNA 2 derived from pQCD2qt was expected to be translated in two reading frames, and give the V-TAV 2b protein as well as the the unrelated C-terminal 41 amino acids of the 92 K chimeric 2a protein (Fig. 4.2). Therefore, two mutants of pQCD2qt were constructed to determine whether the severe symptoms in CMV-qt infected plants are caused by i) the V-TAV 2b protein, ii) the chimeric 2a protein, or iii) the TAV RNA sequence. A stop codon was introduced in ORF 2a immediately to the 5' side of the TAV sequence of pQCD2qt to yield pQCD2qt1 (Fig. 4.2). The second codon (GCA) of ORF 2b in pQCD2qt1 was further changed to a stop codon (UAA) to give pQCD2qt2 (Fig. 4.2).

CMV-qt1 (derived from inoculum pQCD2qt1+pQCD1+pQCD3) also induced severe systemic symptoms in tobacco (Fig. 4.5), and was thus phenotypically similar to CMV-qt. In contrast, CMV-qt2 (derived from inoculum pQCD2qt2+pQCD1+pQCD3) was similar in virulence to Q-CMV and caused no visible symptoms in tobacco. Further, CMV-qt1, like CMV-qt, accumulated to a much greater concentration than Q-CMV did in infected tobacco (data not shown) in spite of a similar level of accumulation of all viruses in infected protoplasts (Fig. 4.6). Therefore, the severe symptoms in CMV-qt infected plants is determined by the TAV 2b protein, not the inserted TAV 2b nucleotide sequence or the chimeric 2a protein, that led to the hypervirulence of CMV-qt.

The results of CMV-qt1 and CMV-qt2 infection of the six other host species (Table 4.2) agreed with those from infection experiments in tobacco plants described above (Fig. 4.3c); the hypervirulence and enhanced accumulation of CMV-qt RNAs in infected plants correlated with the expression of TAV 2b protein because CMV-qt1, but not CMV-qt2, accumulated to high levels and induced severe systemic symptoms (data not shown) in all of the six hosts. This is also consistent with the previous finding that protein 2b is essential but that the overlapping C-terminal portion of protein 2a is dispensable for symptom expression of Q-CMV in *N. glutinosa* and cucumber (Ding *et al.*, 1995b).

Protoplast infection experiments further indicated that the C-terminal overlapping portion of the 2a protein (Fig. 4.2) was not required for viral replication because of a similar amount of viral RNA accumulated in protoplasts (Fig. 4.6) infected with Q-CMV (lane 2), CMV-qt (lane 3), CMV-qt1 (lane 4) or CMV-2a' (lane 5). CMV-2a', a mutant of Q-CMV, contained a point mutation that resulted in an early termination of ORF 2a immediately before the start of ORF 2b (Ding *et al.*, 1995b).

4.3.5 Infectivity of the intraspecies hybrid virus

CMV-qw is an intraspecies hybrid virus made by exchanging the 2b coding sequence between the two subgroup strains of Q-CMV (Fig. 4.2). The genome of CMV-qw is identical to that of Q-CMV except for the 2b gene which constitutes less than 4% of the complete genome. As summarised in Table 4.2, CMV-qw was also significantly more virulent than Q-CMV in all the seven species as found previously for the interspecies hybrid, CMV-qt. This again shows that the 2b gene of WAII-CMV was functional as encoded by the hybrid genome. In *N. tabacum* cv White Burley (Fig. 4.7), *D. stramonium* (data not shown) and *P. floridana* (Fig. 4.7), CMV-qw also caused more severe symptoms than WAII-CMV from which the 2b gene was obtained. However, CMV-qw showed no difference to WAII-CMV in symptom severity in the other four plant species. It was clear that WAII-CMV was more virulent than Q-CMV in all of the

seven host species (Table 4.2) and it is likely that the 2b gene plays an important role in this difference of virulence between the two CMV strains.

Significantly, CMV-qw systemically infected cucumber (*Cucumis sativus* cv. Green Gem) and the systemic symptoms induced were very similar to those induced by Q-CMV or WAII-CMV. Northern blot analysis showed that CMV-qw accumulated in systemic leaves to a level similar to that of Q-CMV (data not shown). These results therefore clearly show that the 2b gene of WAII-CMV was able to functionally substitute for the 2b gene of Q-CMV. This is because a Q-CMV mutant expressing no 2b protein was restricted to the site of primary infection and failed to invade upper uninoculated parts of the cucumber plants (Ding *et al.*, 1995b).

4.4 DISCUSSION

The genome structure of CMV-qt described in this chapter is identical to that of Q-CMV except for the small overlapping gene 2b. This gene was transferred from V-TAV and constitutes only about 3.3% of the complete genome of CMV-qt. It has been shown that CMV-qt was significantly more virulent and CMV-qt induced earlier onset of systemic symptoms than either of the parental viruses in seven plant species belonging to three families (Table 4.2). CMV-qt RNAs clearly accumulated to a higher level in the infected plants than did those of Q-CMV (data not shown). This is the first example of an interspecies hybrid made from plant (De Jong and Ahlquist, 1992; Hilf and Dawson, 1993; Mise *et al.*, 1993; Burgyan *et al.*, 1993) or animal (Pletnev *et al.*, 1992; Lopez *et al.*, 1994; Peng *et al.*, 1995; Kuhn *et al.*, 1996) RNA viruses which is more efficient in systemic infection of a number of hosts than the naturally selected parents. It would be of interest to determine whether CMV-qt is also more competitive than its parental viruses and becomes the dominant isolate when plants are doubly infected with CMV-qt and Q-CMV or V-TAV.

4.4.1 Interaction between the heterologous protein 2b and Q-CMV

The results described indicate that the hypervirulence of these hybrid viruses most likely resulted from a synergistic interaction between the heterologous 2b protein and Q-CMV (or precisely, Q-CMV in the absence of its own 2b). It is known that certain plant viruses can interact with each other synergistically leading to enhanced virulence (Kassanis, 1963). However, as double infection with Q-CMV and TAV induced no synergistic response (data not shown), the interaction observed represents a new type of virus synergism, namely, one plant virus can interact synergistically with a protein derived from a second virus and encoded in *cis* even though no virus synergy occurs in double infections.

What is the molecular basis of the observed synergism? Infection experiments in tobacco protoplasts showed that substituting the 2b gene in RNA 2 of Q-CMV with that of TAV had no detectable effect on accumulation of the genomic RNAs 1 and 3 of Q-CMV, indicating that the synergism was unrelated to changes in viral replication. The heterologous 2b protein may act as a cytotoxin, thus contributing to the hypervirulence of hybrid viruses. However, this seems unlikely since the toxicity theory alone would not explain the increased accumulation of either CMV-qt or CMV-qw in infected plants.

4.4.2 Protein 2b may affect virus movement

It is most likely that the 2b protein of V-TAV and WAII-CMV is functionally superior to the Q-CMV 2b protein in facilitating virus movement within the plant, at least in the host species examined, and as a result, the hybrid virus spreads much more quickly in infected plants. This is partly supported by differences in the timing when *N. glutinosa* plants first showed systemic symptoms after inoculation with CMV-qt, CMV-qw or Q-CMV since both CMV-qt and CMV-qw induced the appearance of systemic symptoms at least three days earlier than did Q-CMV. Rapid appearance of systemic symptoms and a high level

of RNA accumulation have been found for the Fny strain of CMV, which induced symptoms faster than the Sny strain of CMV (Gal-On, *et al.*, 1994).

The 2b gene of Q-CMV has been shown to encode a long-distance virus movement function (Ding *et al.*, 1995b); thus, it is highly likely that its homologue encoded by either V-TAV or WAII-CMV has a similar function. As all seven species examined are also hosts of V-TAV or WAII-CMV (Table 4.2), the 2b gene from V-TAV or WAII-CMV should have been well adapted to function in these hosts. Moreover, it has been demonstrated that symptom severity is largely determined by the relative rate of virus movement and the rate of plant growth (Dawson and Hilf, 1992; Gal-On, *et al.*, 1994); further, efficient virus movement will lead to an increased number of cells being infected and thus to increased viral accumulation. It has been shown that single amino acid changes in the movement proteins of other plant RNA viruses can dramatically alter both viral accumulation and virulence (Tsai and Dreher, 1993; Rao and Grantham, 1995). Therefore, the synergistic responses associated with infection of these hybrid viruses, including hypervirulence, enhanced virus accumulation and rapid appearance of systemic symptoms, can all result from an increased speed of CMV-qt or CMV-qw systemic spread through infected plants as compared to Q-CMV.

4.4.3 Protein 2b may correlate with virulence

CMV-qt was also much more virulent than V-TAV. This is most likely to result from a greater amount of the V-TAV 2b protein being expressed by the CMV-qt genome than by the V-TAV genome as shown by Western blot analysis (see Chapter 3, Fig. 3.7). A greater amount of the chimeric RNA 4A (mRNA of 2b) accumulated in plants infected with CMV-qt than that of RNA 4A in V-TAV infected plants (data not shown). This may reflect the relative strength of Q-CMV and V-TAV or WAII-CMV subgenomic RNA promoters (Ding *et al.*, 1994) controlling transcription of RNA 4A from RNA 2. This difference in the amount of V-TAV 2b protein produced by V-TAV and by CMV-qt may

also explain why no synergistic interaction occurred in double infection with Q-CMV and V-TAV where expression of V-TAV 2b protein was controlled by V-TAV genome.

Table 4.1 Mutagenic primers used for generating mutant constructs of Q-CMV RNA 2

Primer	Primer sequence	Primer position	Construct
SD22	5'-ACCGTTAAGAAGAAGAAGA**ATGGCAAGCATCGAGATCCCT**-3'	Q-CMV RNA 2 nt 2391-2409 V-TAV RNA 2 nt 2447-2467 (in bold)	pQCD2qt
SD23	5'-CGGACGACGAAGGGTCGGAAA**TCATTGATCGAGACACCAGT**-3'	*Q-CMV RNA 2 nt 2713-2733* *V-TAV RNA 2 nt 2715-2734 (in bold)*	pQCD2qt
BJ14	5'-AAGGGTCGGAAAT**CAGAAAGCACCTT**-3'	*Q-CMV RNA 2 nt 2712-2724* *nt 300-312 from the 3'-end of WAII-CMV RNA 2 (in bold)*	pQCD2qw
BJ15	5'-AAGAAGAAGAAGA**ATGGAATTGAACGT**-3'	Q-CMV RNA 2 nt 2397-2409 nt 618-631 from the 3'-end of WAII-CMV RNA 2 (in bold)	pQCD2qw
SD43	5'-ACCGTTAAGAAGTAGAAGAAT-3'	CMV-qt RNA 2 nt 2391-2411	pQCD2qt1
SD44	5'-ATTCTTCTACTTCTTAACGGT-3'	*CMV-qt RNA 2 nt 2391-2411*	pQCD2qt1
SD47	5'-TAAGAAGAAGTAGAATGTAAAGCATCGAGA-3'	CMV-qt RNA 2 nt 2396-2425	pQCD2qt2
SD48	5'-TCTCGATGCTTTACATTCTACTTCTTCTTA-3'	*CMV-qt RNA 2 nt 2396-2425*	pQCD2qt2

The underlined nucleotides in the primer sequences specify the mutations introduced into each of the mutant constructs. Viral RNAs and residue numbers in italics indicate that the primer sequences are the complementary strand of RNA 2; all others are plus sense.

Table 4.2 Virulence of viruses in plant species

	Solanaceae					*Caryophyllaceae*	*Amaranthaceae*
Virus	*N. tabacum*	*N. glutinosa*	*D. stramonium*	*L. esculentum*	*P. floridana*	*S. media*	*G. globosa*
Q-CMV	+	++	+	+	+	+	+
V-TAV	++	++	+	++	++	+	+
CMV-qt	++++	++++	++++	++++	++++	++++	++++
WAII-CMV	+++	+++	+	+++	+++	+++	+
CMV-qw	++++	+++	+++	+++	++++	+++	+

Four levels of virulence were scored: +, systemic infection but almost symptomless; ++, mosaic; +++, severe mosaic; ++++, severe mosaic with distortion.

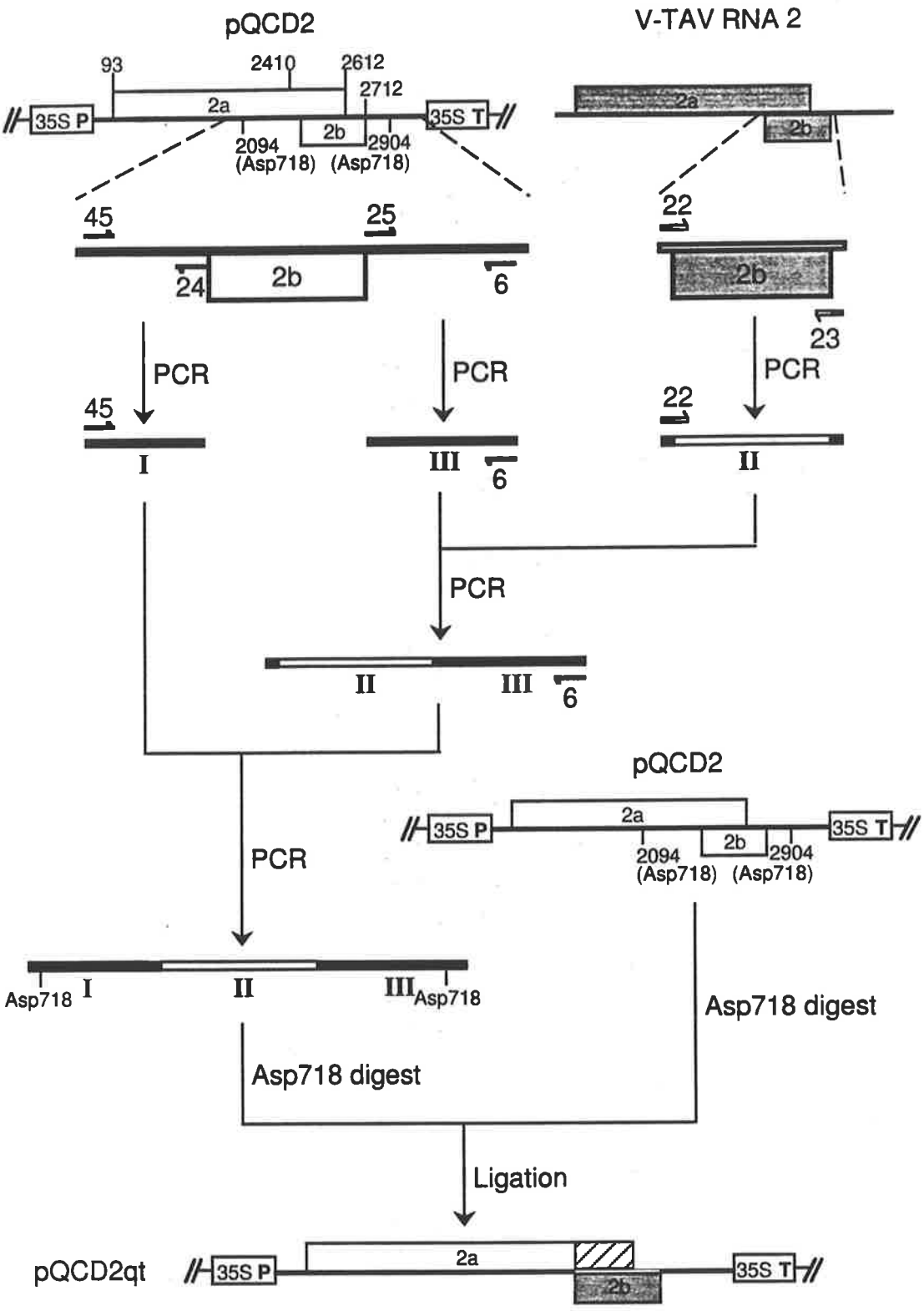


Fig. 4.1 Schematic representation of the construction of mutant cDNA clones used for virus infection. Descriptions in detail are given in the text. At each step, the diagram is simplified to illustrate only how the new fragment obtained during the previous step is utilised. The regions within dotted lines are amplified into the three PCR fragments (enlarged for clarity) used for constructing pQCD2qt. The numbers with an arrow above or underneath represent the names of primers in the SD or BJ series. The filled bars represent the Q-CMV RNA sequence whereas the opened bars depict the V-TAV RNA sequence encoding the 2b protein. The cross-hatched area represents amino acid sequence unrelated to 2a or 2b sequence.

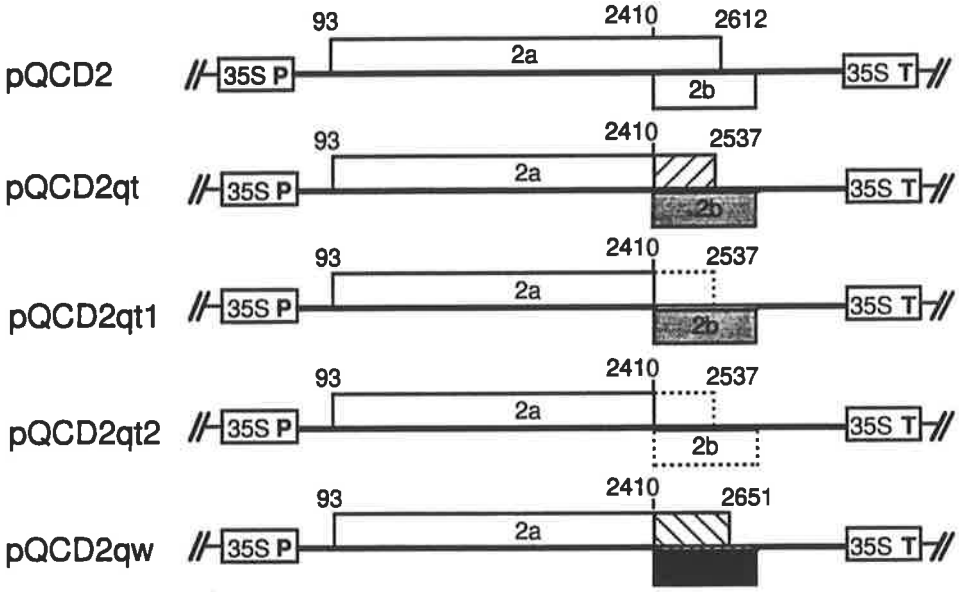


Fig. 4.2 The structure of plasmids pQCD2 and its derivatives. Plasmids pQCD2 (or its derivatives) contains the full-length DNA copy of RNA 2 under the transcriptional control of the 35S promoter (P) and terminator (T) of cauliflower mosaic virus. The blank area surrounded by dashed lines in ORF 2a (pQCD2qt1 and pQCD2qt2) or ORF 2b (pQCD2qt2) represents deletions in the ORF only by single nucleotide substitution(s). The two areas, cross-hatched in different directions, represent different amino acid sequences and both of which are also unrelated to 2a or 2b sequences.

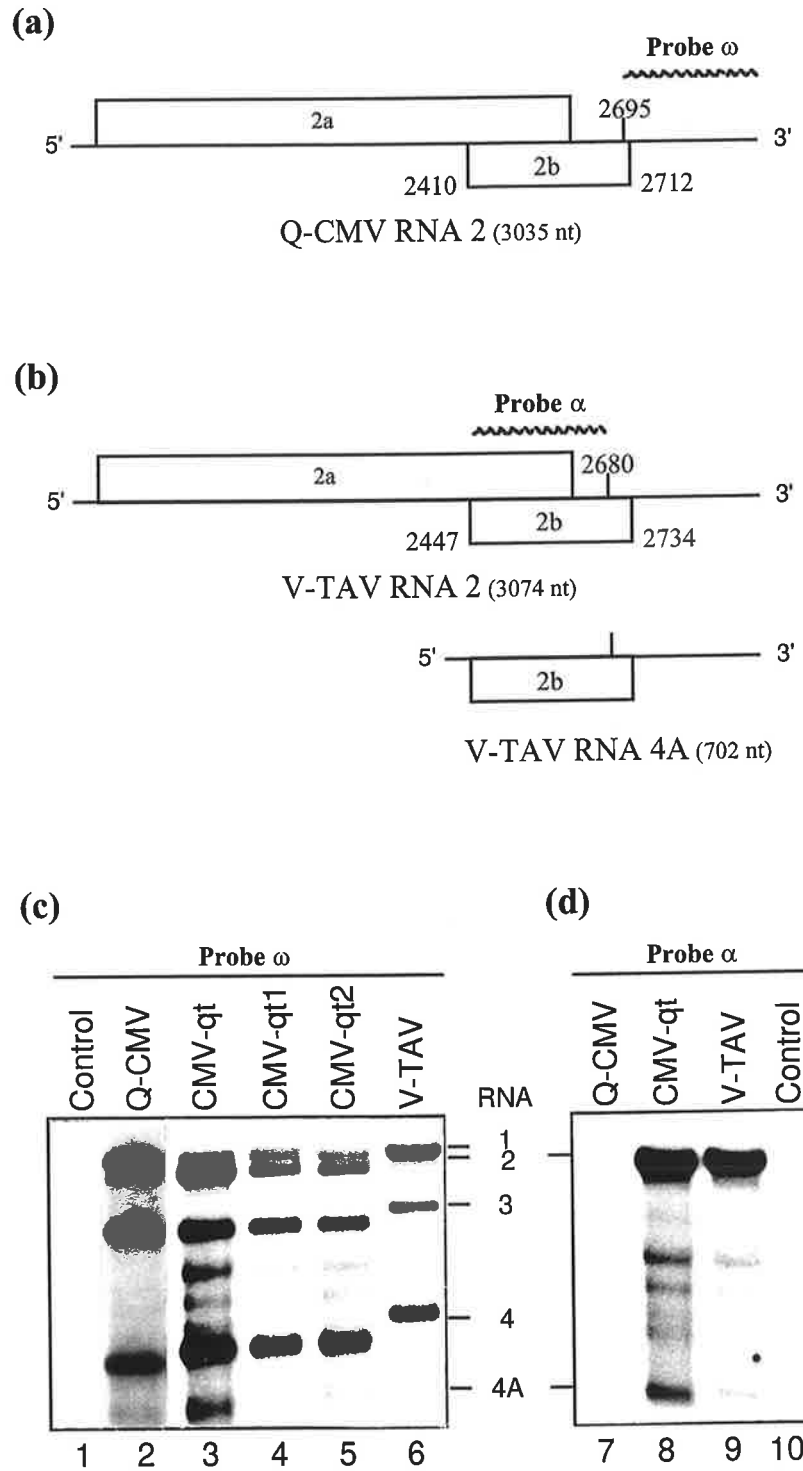
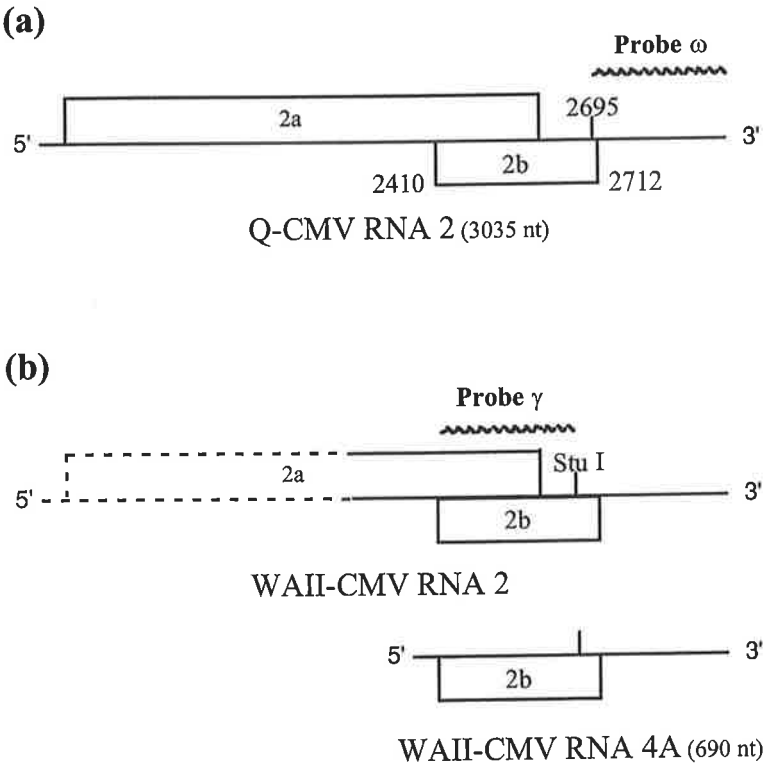


Fig. 4.3 Northern blot hybridisation of the RNAs of Q-CMV, TAV and the hybrid viruses. Total RNAs isolated from tobacco leaves inoculated with sterile water (lanes 1 and 10), Q-CMV (lanes 2 and 7), CMV-qt (lanes 3 and 8), CMV-qt1 (lane 4), CMV-qt2 (lane 5) and V-TAV (lanes 6 and 9) were hybridised with probe ω (c) which is complementary to the 3' terminal 625 nt of Q-CMV RNA 2 as shown by the wavy lines in (a) or hybridised with probe α (d), a V-TAV 2b-specific probe, as used in Chapter 3 and shown in (b). The positions of viral RNAs 1, 2, 3, 4 and 4A are indicated. The open boxes in (a) and (b) represent the designated open reading frames.



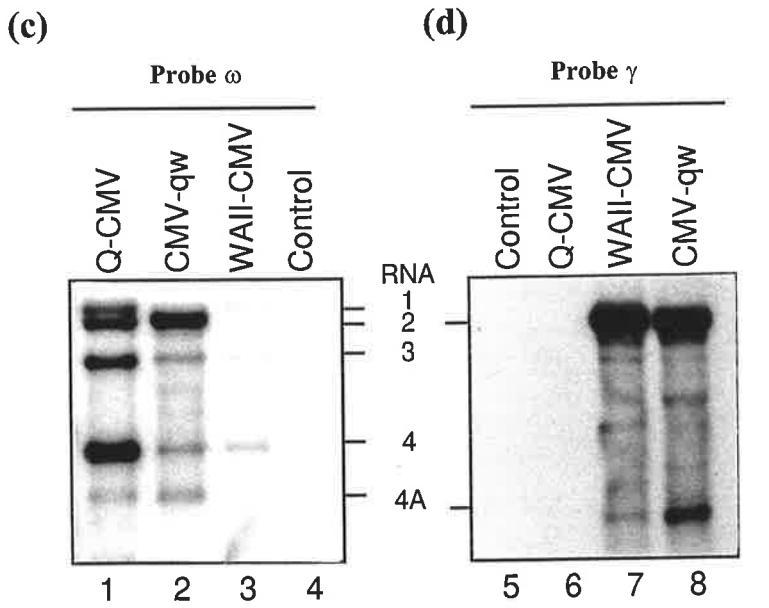


Fig. 4.4 Northern blot hybridisation of the RNAs of Q-CMV, WAII-CMV and the hybrid viruses. Total RNAs isolated from tobacco leaves inoculated with Q-CMV (lanes 1 and 6), CMV-qw (lanes 2 and 8), WAII-CMV (lanes 3 and 7) and sterile water (lanes 4 and 5) were hybridised with probe ω (c) as used in Fig. 4.3 and shown in (a) or hybridised with probe γ (d), a WAII-CMV 2b-specific probe, as used in Chapter 3 and shown in (b). The positions of viral RNAs 1, 2, 3, 4 and 4A are indicated. The open boxes in (a) and (b) represent the designated open reading frames.

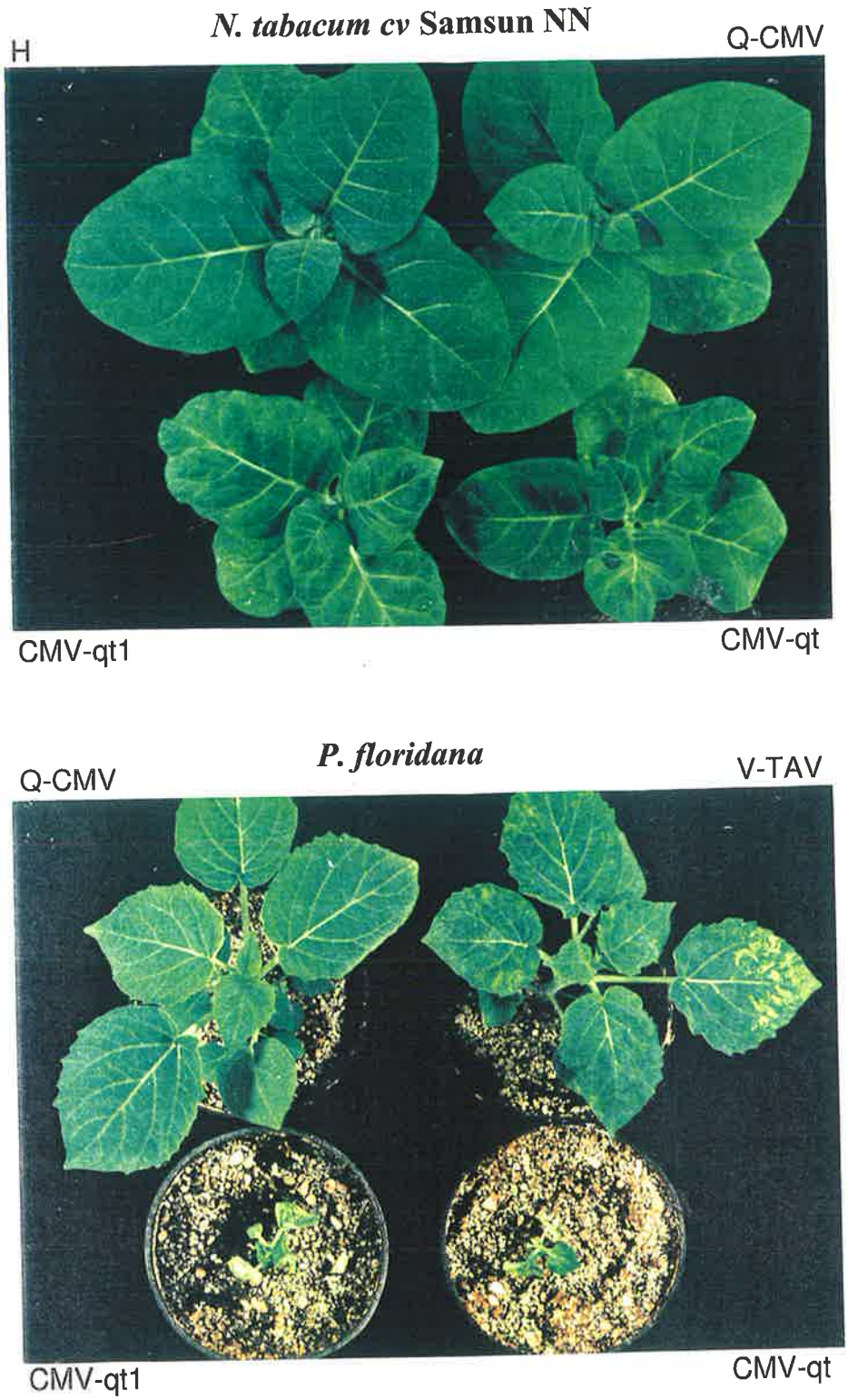
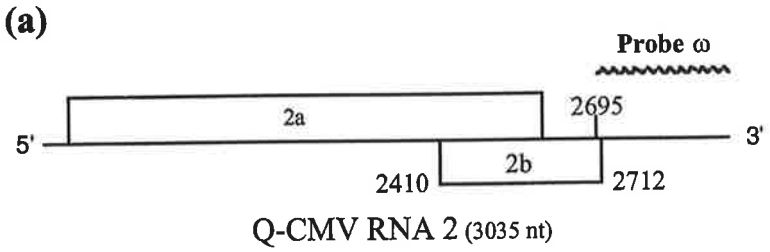


Fig. 4.5 Symptoms of plants infected with Q-CMV, V-TAV, or with the hybrid viruses CMV-qt and CMV-qt1. Top: Healthy (H) or infected tobacco (*Nicotiana tabacum* cv Samsun NN) plants. Bottom: Infected *Physalis floridana* plants. The infected plants were photographed 3 weeks post-inoculation.



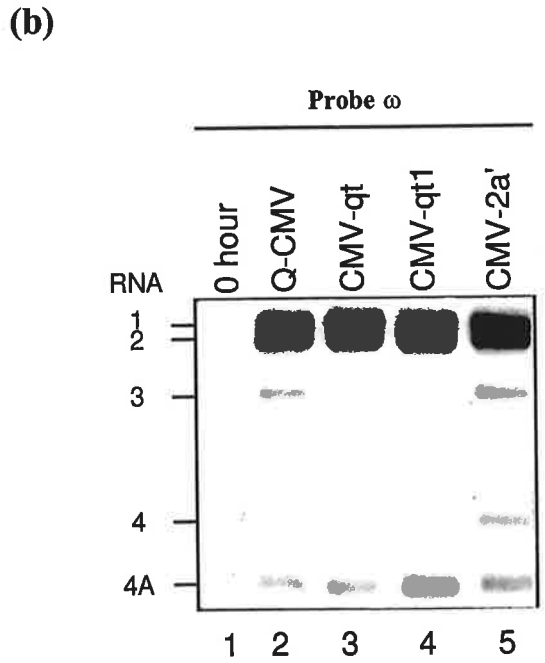


Fig. 4.6 Accumulation of the RNAs of Q-CMV, TAV and the hybrid viruses in tobacco protoplasts. Each lane contained RNA extracted from about 200,000 protoplasts after inoculation and 24 hours of incubation and were hybridised with probe ω (b) as used in Fig. 4.3 and shown in (a). The letter(s) above each lane indicates the source of inoculum. 0 hour represents an uninoculated control. CMV-2a', a mutant of CMV, contained a point mutation that resulted in an early termination of ORF 2a immediately before the start of ORF 2b (see Section 4.3.4 in the text).

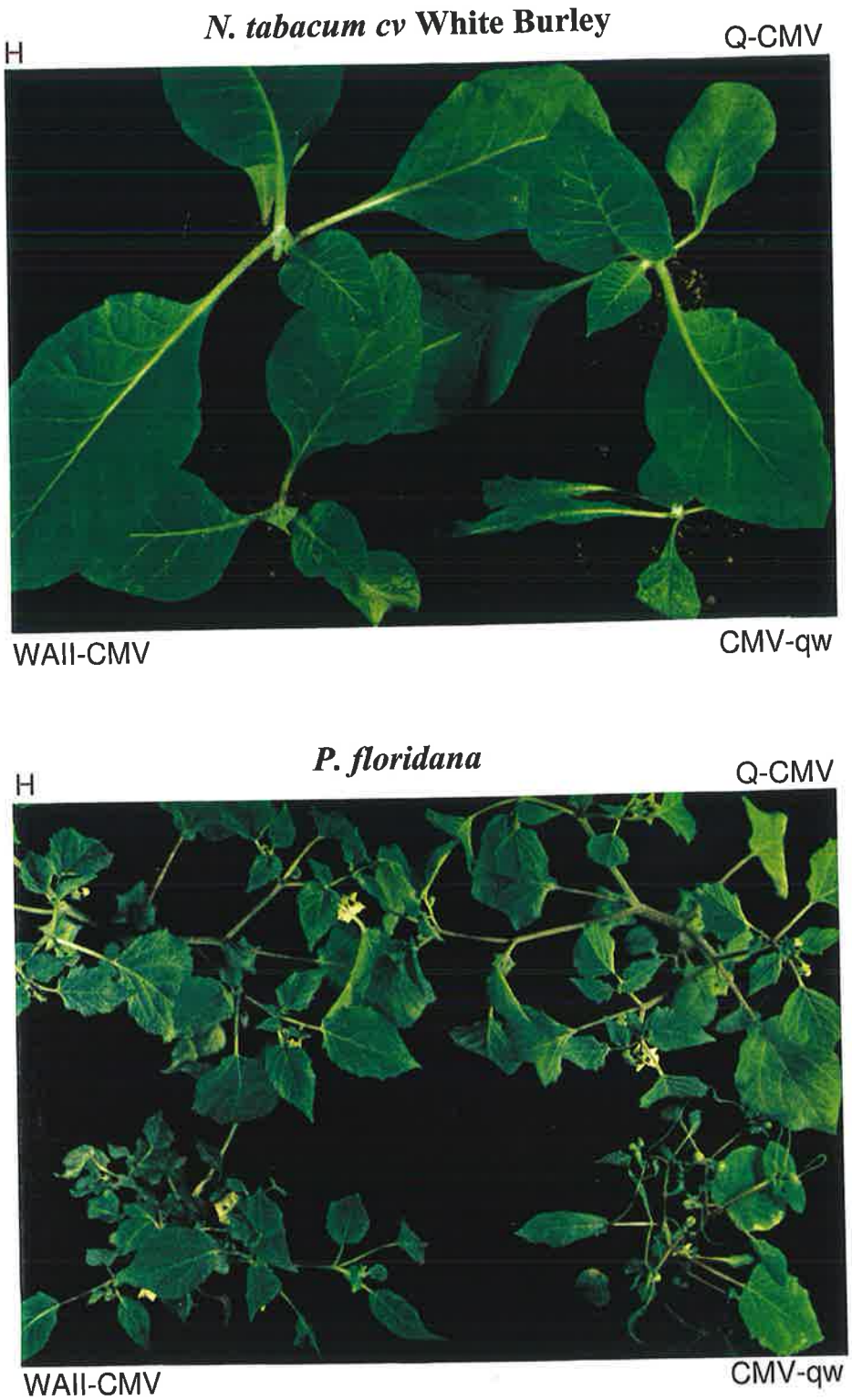


Fig. 4.7 Symptoms of plants infected with Q-CMV, WAII-CMV, or with the hybrid virus CMV-qw. Top: Healthy (H) or infected tobacco (*Nicotiana tabacum* cv White Burley) plants. Bottom: Infected *Physalis floridana* plants. The infected plants were photographed 3 weeks post-inoculation.

Chapter 5

CONSTRUCTION OF THE INFECTIOUS cDNA CLONES OF TOMATO ASPERMY CUCUMOVIRUS

5.1 INTRODUCTION

In Chapter 4, both interspecies hybrid virus made by replacing the 2b gene of CMV with the 2b gene of TAV and intraspecies hybrid virus made by replacing the 2b gene of Q-CMV with the 2b gene of WAII-CMV have been shown to be significantly more virulent than either of the respective parental viruses in at least six host plant species tested. However, the systemic infection of the interspecies hybrid was very inefficient in cucumber, which is a systemic host for Q-CMV but immune to V-TAV (Habili and Francki, 1974a). This host specificity shown by homologous genes from different viruses may constitute the molecular basis of different but overlapping host ranges of different cucumoviruses.

To further study the specificity of control of genes encoded by different cucumoviruses, it is essential to establish an efficient infection system for the cloned viral genomes of several representive species such as TAV. In addition, investigating functions of the two newly discovered V-TAV-associated RNAs (see Chapter 3, 3.3.1.1) will also require the development of novel experiments and approaches, which will inevitably involve the use of infectious full-length cDNA clones of V-TAV.

Infectious full-length cDNA clones have been proved to be useful to characterise plant RNA viruses at the molecular level. The first infectious clone corresponding to the RNA genome of Qβ bacteriophage was obtained in 1978 by Taniguchi *et al.*. Since then, infection from cloned cDNAs has been achieved for a large number of plant RNA viruses using a variety of strategies including a few of truly infectious cDNAs, rather than DNA templates from which infectious RNA molecules must be transcribed *in vitro* before inoculation onto plants (Boyer and Haenni, 1994; Ding *et al.*, 1995a). However, it has to be admitted that some of those clones had no practical use due to various problems (Boyer and Haenni, 1994). It is generally believed that two factors, vectors

and cloning strategies used, determine the successful construction of infectious cDNA clones.

A plasmid vector, pCass, has been recently described (Ding, *et al.*, 1995a). This vector is specifically designed for cloning full-length cDNAs of plant RNA viruses whose transient *in planta* expression is controlled by the 35S promoter and terminator from cauliflower mosaic virus (CaMV). The full-length cDNA clones of CMV constructed in this vector have been shown to be stable (stably maintained and in *Escherichia coli* DH5α cells for nearly four years) and to be an effective inoculum for direct mechanical inoculation of host plants (Ding, *et al.*, 1995a). They have been also used for the successful mutational analysis of the 2b gene of cucumoviruses (Ding *et al.*, 1995b; see Chapter 4). Hence, it was considered that the pCass vector could be used for the construction of infectious cDNA clones of V-TAV or WAII-CMV.

Previous reports have described the higher level of expression of foreign genes in some transgenic plants under the control of a duplicated cauliflower mosaic virus (CaMV) 35S promoter than that under the control of a single CaMV 35S promoter (Kay *et al.*, 1987; Mogen *et al.*, 1990; Deng *et al.*, 1992; Tourneur *et al.*, 1993), but no information was provided on virus infectivity in plants achieved by mechanical inoculation of infectious clone(s) with such a duplicated CaMV 35S promoter. This chapter describes the latest version of the pCass vector with a clearly improved efficiency of infection of the TAV genomic cDNA clones by partial duplication of the CaMV 35S promoter. The infectious cDNA clones of WAII-CMV have not been constructed as the genome sequence of this virus has not been completed.

5.2 MATERIALS AND METHODS

5.2.1 Construction of pCass2 vector

Construction of pCass2 vector under transcriptional control of the partially duplicated CaMV 35S promoter followed the basic strategy described by Kay *et al.* (1987). The upstream region containing base pairs -416 to -90 of the CaMV 35S promoter was excised from the pCass vector as a *Hind*III-*EcoR*V fragment and inserted into the pCass vector between the sites of *Hind*III and *Pst*I blunt-ended by T4 DNA polymerase to make an intermediate vector (Fig. 5.1). The unrearranged and partially duplicated 35S promoters were then excised from the pCass vector and the intermediate vector, respectively, as *Hind*III-*Bgl*II fragments and inserted between the *Hind*III and *BamH*I sites of pUC19, in which the polycloning restriction sites from *EcoR*I to *Sma*I have been removed in advance, to make pUC19-based pCass vectors, pCass1 with a single 35S promoter (Fig. 5.1), and pCass2 with a partially duplicated 35S promoter (Figs. 5.1 and 5.2), respectively.

5.2.2 cDNA syntheses

Virion and viral RNAs of the V strain of TAV (V-TAV, Habili, and Francki, 1974a) were purified as in Chapter 2 (Sections 2.2.4 and 2.2.5). cDNAs were synthesised as shown in Fig. 5.3. First strand was synthesised in a reaction volume of 20 µl containing 1 µg purified virion RNAs, 0.5µg primer BJ1 (5'-AGGATCTGGGACCCCTAGGGGGAACCTACGGA-3') complementary to the 3'-end of all three genomic RNAs, 1x avian myeloblastosis virus (AMV) reverse transcriptase reverse transcription buffer (Promega), 10 mM DTT, 1 mM each of dATP, dCTP, dGTP and dTTP, 40 units RNasin and 16 units AMV-RT (Promega) for 30 minutes at 43°C. The reaction is continued by directly adding the following reagents: 3.5 µl 0.1 M $MgCl_2$, 5 µl 1 M Tris-HCl (pH7.5), 0.75 µl 1 M $(NH_4)_2SO_4$, 0.4 µl RNAase H (15

units/μl, Promega), 3.3 μl DNA polymerase I (7 units/μl, Promega) to a volume of 71 μl. Incubation was for 4 hours at 16°C. The reaction mixture was further incubated for 15 minutes at room temperature after adding 1 μl of 50 mM ßNAD, 2 units *E. coli* DNA ligase (Promega) and 7 units bacteriophage T4 polynucleotide kinase (PNK; New England Biolabs). The reaction was terminated by adding EDTA (pH 8.0) to 20 mM, and extracted twice with phenol:chloroform (1:1 v:v). The resultant double-stranded cDNA was precipitated with 2 volumes of ethanol in the presence of 0.3 M sodium acetate (pH5.2), dissolved in 20 μl of H_2O and cloned into the *Sma*I site of pBluescript SK^+. Screening of cDNA clones was initially carried out by digestion with several appropriate restriction enzymes, and then by dideoxynucleotide chain termination sequencing from both ends using primers T3 or T7 that flanked the multicloning site. The cDNA clones containing the longest viral cDNA inserts were selected for constructing the full-length cDNAs of each of the three genomic RNAs (see below).

5.2.3 Construction of TAV genomic cDNA clones in pCass2

The 5' terminal sequences of the three genomic RNAs, which were not represented by cDNA clones obtained above, were obtained by reverse transcription and polymerase chain reaction (RT-PCR) with one of the following primer pairs: BJ9 (5'-GTTTGTCTA TCAAGAGCGTACGGTTCA-3') homologous to nucleotides 1-27 of RNA 1 and BJ10 (5'-TCCCGGGATGTAGACCTTAAGGCCTCTACGT-3') complementary to nt 217-242 of RNA 1, BJ9 and BJ11 (5'-AGGTACCTCCGGAAGCTTAGTATACTCACA-3') complementary to nt 865-889 of RNA 2, BJ12 (5'-GTTTACCAACCAACCAACCACTACTAT-3') homologous to nucleotides 1-27 of RNA 3 and BJ13 (5'-GGCGCGCATCCAGTGGTGGCGACCGA-3') complementary to nt 211-236 of RNA 3, under the same program (25 cycles, each at 92°C for 1 minutes, 55°C for 1 minutes, and 72°C for 2 minutes) using Vent DNA polymerase (New England Biolabs). The three PCR products were individually cloned into the *Stu*I

site of pCass1 and pCass2, to give pCass1T1$^{5'}$, pCass1T2$^{5'}$, pCass1T3$^{5'}$, pCass2T1$^{5'}$, pCass2T2$^{5'}$, and pCass2T3$^{5'}$, respectively (Figs 5.4, 5.5 and 5.6). The Arabic numbers 1, 2 and 3 after the T correspond to RNA 1, RNA 2 and RNA 3 of TAV. The correct fusions of each of clones were confirmed by dideoxynucleotide sequencing as described in Chapter 2 (Section 2.2.13) before these clone were adapted for further cloning.

Two overlapping cDNA clones pSK2145 and pSK2303 (Fig. 5.4), which cover nt 20-2165 and nt 1107-3410 of RNA 1, respectively, were selected to construct infectious full-length cDNA clone of RNA 1. Firstly, a *Pst*I fragment was excised from pSK2303 and inserted into the *Pst*I site of pSK2145 to give pSK3390. The pSK3390 was further digested with *Stu*I and *Sac*I to release a 3390 nt fragment. This fragment was then inserted between the *Stu*I and *Sac*I sites of pCass1T1$^{5'}$ or pCass2T1$^{5'}$ to produce the full-length cDNA clones of RNA 1, designated as pCass1T1 (with a single CaMV 35S promoter) and pCass2T1 (with a partially duplicated CaMV 35S promoter), respectively (Fig. 5.4).

Similarly, the *Sph*I-*Sac*I fragment released from pSK1541 was inserted between the *Sph*I and *Sac*I sites of pSK1511 to give pSK2531 (Fig. 5.5). The *Acc*III-*Sac*I fragment excised from pSK2531 was inserted between the *Acc*III and *Sac*I sites of pCass1T2$^{5'}$ or pCass2T2$^{5'}$ to yield the full-length cDNA clones of RNA 2, called pCass1T2 and pCass2T2, respectively (Fig. 5.5).

The cDNA clone pSK2346 contained the nearly full-length cDNA clones of RNA 3 and lacked only 40 nt at the 5' terminus (Fig. 5.6). The *BssH*II-*Sac*I fragment excised from this clone was ligated between the *BssH*II and *Sac*I sites of pCass1T3$^{5'}$ or pCass2T3$^{5'}$ to produce the full-length cDNA clones of RNA 3, pCass1T3 and pCass2T3, respectively (Fig. 5.6).

5.2.4 Biological assay

Infectivity of the cDNA clones was carried out initially using two hosts of TAV, *Nicotina glutinosa* and *N. clevelandii*. The seedlings were grown to a stage of a total of four true leaves respectively under natural lighting conditions and placed in the dark for 24 hours prior to inoculation. The full-length cDNA clones of each of the three genomic RNAs either with the single or partially duplicated CaMV 35S promoter were purified by Superose 6 gel filtration chromatography as described by Skingle *et al.* (1990). The purified plasmid DNAs were digested with *Hae*II for cDNAs of RNA 1 and RNA 3, and with *Pvu*II for cDNA of RNA 2. The reaction mixtures were extracted with phenol:chloroform (1:1 v:v) once and DNAs precipitated with 2 volumes of ethanol in the presence of 0.3 M sodium acetate (pH 5.2).

For each plant, each of the three digested genomic plasmid cDNAs (pCass1T1, pCass1T2 and pCass1T3 or pCass2T1, pCass2T2 and pCass2T3) was mixed in equal amounts in a final volume of 10 µl sterile water and mechanically inoculated on the leaves dusted with carborundum powder. Initially, 10 µg plasmid DNA of each of the three digested genomic cDNA clones was applied to each plant. For determining the relative infectivity of the V-TAV cDNA clones constructed in pCass1 and pCass2, each of the 13 dilutions (0.001, 0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 1, 2, 3, 4, 5 µg each of the three digested genomic cDNA clones for each plant) from either set of the TAV cDNA clones was inoculated onto 30 seedlings (two replicates of 15 plants each) of each of *N. glutinosa* and *N. clevelandii* and the percentages of plants infected were recorded 45 days after inoculation. The six additional host plant species of TAV, *Chenopodium amaranticolor, Datura strasmonium, Gomphrena globosa, Lycopersicon esculentum, Physalis floridana* and *N. tabacum cv.* White Burley were inoculated with either set of the TAV cDNA clones (5, 15 and 45 µg each of the three digested genomic cDNA clones for each plant).

5.2.5 Analyses of virus progeny

Virus samples were prepared as described by Shi *et al.* (1989; 1993) for examing the morphology of viral particles. Leaf tissues (0.5 g), either from infected or from mock-inoculated plants, were homogenised in 1 ml 0.02 M phosphate buffer (pH7.6) as described in Chapter 2 (Section 2.2.4). The drops of the supernatant obtained by centrifugation at 14000rpm for 1 minute (Eppendorf) as well as purified virions were placed on copper grids which had been coated with collodion membranes and reinforced with carbon. The excess liquid was removed from the grids with filter paper (Whatman 3MM) and the preparations were stained with 1.5% phosphotungstic acid (PTA; Kodak) for 2 minutes. The specimens were examined in electron microscope (CM-100, Phillips).

For RNA analyses, total plant RNAs were isolated from either infected or mock-inoculated plants and virion RNAs were isolated from purified viral particles. Both RNAs were analysed by northern blot hybridisation (see Chapter 2, 2.2.16) using probe β as described (see Chapter 3, Fig. 3.2).

Direct RNA dideoxynucleotide sequencing was carried out using both purified virion RNAs and total plant RNAs extracted from infected *N. glutinosa* and *N. clevelandii* as templates. These RNAs were annealed with one of the following primers BJ20 (5'-GGAATTCAGGAATTAGAAACAGA-3') complementary to nt 79-94 of RNA 1, BJ21 (5'-GGAATTCAATGAACAAACGA-3') complementary to nt 75-87 of RNA 2, or BJ22 (5'-GGAATTCTGCCATAGGAGGATT-3') complementary to nt 81-95 of RNA 3. Sequencing reactions and dideoxynucleotides were carried out using AMV reverse transcripase as described in Chapter 2 (Section 2.2.14).

5.3 RESULTS

5.3.1 Molecular cloning of TAV genomic cDNAs

The procedures employed for the construction of pCass2 are described in detail under Fig. 5.1 and further details of pCass2, the new version of the pCass vector, are shown in Fig. 5.2. This pCass2 has two important features. Firstly, a partially duplicated 35S promoter replaces the single 35S promoter in the pCass. This duplication does not contain sequence of a TATA element which determines correction initiation of transcription, so it does not affect the insertion of the viral cDNA at the +1 transcription site according to Kay *et al.* (1987). Conversely, it is predicted to increase the rate of transcription and thereby enhance the transcriptional activity (Kay *et al.*, 1987). Secondly, most of the restriction sites in polylinkers outside the expression cassette have been eliminated which will facilitate manipulation of the cloned viral genome.

The strategy of the use of cDNA libraries (Fig. 5.3) was introduced to obtain a large numbers of clones so that cDNAs close in size to the full-length genomic RNAs could be selected. The result showed that 23 clones had an insert of 1.5 Kb or longer as initially judged by restriction digestion and then confirmed by DNA dideoxynucleotide sequencing. Ten of these 23 clones contained cDNA 1 inserts, five clones contained cDNA 2 and eight clones contained cDNA 3.

The missing 5'-terminal sequence of each of three genomic cDNAs obtained by RT-PCR was cloned into the *Stu*I site of pCass1 or pCass2 prior to cDNA obtained from the cDNA library. This cloning strategy provided genomic cDNA precisely at the +1 transcription site because nonviral residues at the transcript 5'-terminus have been found to interfere with transcript infectivity in other systems (Boyer and Haenni, 1994). At the 3'-terminus, all genomic cDNAs had the same 40-residue nonviral extension. The risk that the 3' nonviral extensions would interfere with transcript infectivity of

TAV was judged to be minimal in view of experience with other viruses (Boyer and Haenni, 1994; Brault *et al.*, 1995; Prüfer *et al.*, 1995; Ding *et al.*, 1995a).

The integrity of the promoter fusions and identity to the published sequences were confirmed by sequencing at each of the 5'- and 3'-terminal regions. No variations were observed in the sequence of the cDNA 2 clone as compared to the published RNA 2 sequence of the same strain of this virus (Moriones *et al.*, 1991). However, several variations were observed in the 3' untranslated regions of RNA 1 and RNA 3. The differences between nucleotide sequences of the cDNA 1 clone and the published RNA 1 sequence (Bernal *et al.*, 1991) included one insertion (AUU between nucleotides 3312 and 3313) in all cDNA clones of RNA 1 sequenced. Importantly, the insert creates a recognition site for *EcoR*I confirmed by digestion with this enzyme. The changes to the cDNA clone of RNA 3 included two substitutions (G^{1974} to U and U^{1975} to G) and one U insertion between nucleotides 2369 and 2371 (F. Garcia-Arenal, personal communication).

5.3.2 Biological activity of the cDNA clones

For assaying infectivity, each of three purified genomic plasmid cDNAs was digested with the appropriate restriction endonuclease to release the cloned insert, a step previously shown to enhance infectivity of cDNA clones of CMV containing a single 35S promoter (Ding *et al.*, 1995a). Equal amounts (10 μg per plant of each of the digested three genomic plasmid cDNAs in 10 μl of water) were inoculated on *N. glutinosa* plants which had been in the dark for 24 hours. Both sets of the V-TAV cDNA clones were infectious and produced systemic virus infections indistinguishable from each other or from infections resulting from inoculation with virions as judged by symptomology (Fig. 5.7a), particle morphology (not shown), ethidium bromide staining (Fig. 5.7b) and northern blot analysis (Fig. 5.7c) of virion RNAs. The results thus

further show the usefulness of pCass as a general cloning vector for constructing infectious cDNA clones of plant RNA viruses.

5.3.3 5'-terminus of genomic cDNA-derived *in vivo* transcript

The 5' terminal sequences of the three progeny viral RNAs isolated both from *N. glutinosa* and *N. clevelandii* plants infected with pCass1T1T2T3 or pCass2T1T2T3 as determined by RNA dideoxynucleotide sequencing (Ding *et al.*, 1995b) were all identical to each other as well as to those of the respective plasmid DNA inocula with one following exception. One U insertion was found between nucleotides 2 and 3 of both progeny RNAs 1 and 2 derived from pCass2T1T2T3 (data not shown). This extra U had been persistently maintained in several subsequent passages and had no apparent effect on symptomology. It is not clear how this insertion occurred or whether it is related to the partial duplication of the CaMV 35S promoter.

5.3.4 Comparative infectivity assay

The relative infectivity of the V-TAV cDNA clones constructed in pCass1 and pCass2 was examined in two *Nicotiana* species, *N. glutinosa* and *N. clevelandii*. Each of the 13 dilutions from either set of the TAV cDNA clones was inoculated onto 30 seedlings (two replicates of 15 plants each) of either plant species and the percentages of plants infected were recorded 45 days after inoculation. The following points are evident from the data presented in Fig. 5.8. (1) When using lower amounts of plamids as inocula the relative infectivity of the pCass2-based TAV cDNA clones was clearly higher (3 fold) than that of the pCass1-based clones; however, the difference was not obvious for higher amounts of plasmid DNAs as inocula. It is most likely that the relatively higher infectivity of pCass2-based TAV clones resulted from a higher level of transient transcription as directed by the partially duplicated promoter inside the inoculated cells. It should be pointed out that the data do not exclude the extra U found in RNAs 1 and 2

derived from pCass2-based clones being the cause of higher infectivity. (2) There is an obvious difference in susceptibility to virus infection from plasmids inocula between the two host species. For the same amounts of plasmids inocula, the percentage of *N. clevelandii* plants became infected is always higher than that of *N. glutinosa* plants. (3) Low amounts of 0.2 μg to 2 μg per plasmid per plant were sufficient to achieve an infection rate over 50%. Thus, a large-scale plasmid DNA preparation can generate a substantial amount of virus inocula.

5.3.5 Facility studies with the infectious cDNA clones

Further infectivity assays showed that both sets of cDNA clones, as virus inocula, were as effective as viral particles to six additional host plant species: *Chenopodium amaranticolor, Datura strasmonium, Gomphrena globosa, Lycopersicon esculentum, Physalis floridana* and *N. tabacum cv.* White Burley (data not shown). However, up to 45 μg per plasmid per plant of either cDNA clones failed to infect cucumber (data not shown) which is a non-host of V-TAV, indicating that the host range of the plasmids inocula reflected the true host range of V-TAV at least when the 9 plant species examined are concerned.

5.4 DISCUSSION

The expression cassette consisting of CaMV 35S promoter and terminator from the newly established cloning vector pCass (Ding *et al.*, 1995a) was transferred into the modified pUC19 and the promoter upstream region of 326 base pairs has been duplicated in the pUC19-based pCass (called pCass2). Full-length TAV cDNA clones constructed in pCass2 gave a higher level of infectivity than that did those constructed in pCass1. This is the first report that a partially duplicated CaMV 35S promoter was used to construct cDNA clones of plant RNA viruses for mechanical inoculation.

5.4.1 Construction of infectious TAV cDNA clones

The investigations in this chapter were undertaken initially with the intention of duplicating CaMV 35S promoter sequences in the pCass-based vector to enhance transcriptional activity. A second goal was the construction of infectious cDNA clones of TAV to facilitate molecular and genetic studies of the virus. The partially duplicated 35S promoter constitutes 326 base pairs of upstream sequences of the promoter and the intact promoter of 416 base pairs. The duplicated region can exist compatibly with the intact promoter and does not affect correct initiation of transcription from the promoter (Kay *et al.*, 1987). The expression cassette containing the partially duplicated promoter or the single promoter was transferred into a pUC19 vector. This is because the full-length cDNA clone corresponding Q-CMV RNA 1 failed to be assembled in pCass (Ding *et al.*, 1994). The utility of such pUC19-based vectors has been successfully demonstrated in construction of each of three full-length TAV genomic cDNAs. The infectivity of transcribed RNAs *in vivo* through the partially duplicated promoter achieved by mechanical inoculation appeared, as expected, higher than that through the single promoter, which makes the vector with the partially duplicated promoter useful as a general and special-purpose infectious cloning vector. The viral progeny RNAs transcribed from these cDNAs *in vivo* through the pUC19-based vector with either the partially duplicated or the single 35S promoter are identical to those derived from the wildtype virus.

Compared with other strategies currently used to obtain progeny virus from cDNA clones such as infectious RNAs transcribed *in vitro* from cDNA clones (Boyer and Haenni, 1994) and particle bombardment of cDNA clones (Gal-On *et al.*, 1995; Fakhfakh *et al.*, 1996), pCass2, the vector with the partially duplicated 35S promoter, offers several unique features including easy preparation of plasmid inocula and plant

infection by conventional mechanical inoculation, and thus is simple to use and cost-effective.

5.4.2 High infectivity by the use of a partially duplicated CaMV 35S promoter

The TAV cDNA clones constructed in pCass2 under the control of the partially duplicated 35S promoter in this study gave higher infectivity than did those under the control of the single 35S promoter. However, the increased infectivity was not as high as previous reports that a partially duplicated CaMV 35S promoter can give 10 to 12-times higher level of expression of foreign genes in transgenic plants than that under the control of a single 35S promoter (Kay *et al.*, 1987; Mogen *et al.*, 1990; Deng *et al.*, 1992; Tourneur *et al.*, 1993). Furthermore, these cDNA clones showed higher infectivity than those under the control of single promoter only when less than 0.5 μg each of three plasmid DNAs in a final volume of 10 μl was used for inoculation. The infectivity of two sets clones with an amount over 0.5 μg each of three plasmid DNAs in a final volume of 10 μl of inoculum showed no apparent difference.

The cause of the higher infectivity of TAV cDNA clones driven by the partially duplicated promoter than those driven by the single promoter, which is only revealed when low amounts of each of two sets of infectious cDNA clones are inoculated, are unknown. It could be due to integration of the expression cassette containing TAV genomic cDNA insert into the host genome where activity of the partially duplicated promoter could be hampered by some complicated secondary structures (Boyer and Haenni, 1994) or due to the efficiency of the *in vivo* transcription from TAV genomic cDNA, which determines the infection requiring a minimum concentration of the viral RNAs when low amounts of the cDNAs are inoculated and which is not critical if more RNAs are transcribed when high amounts of the cDNAs are inoculated.

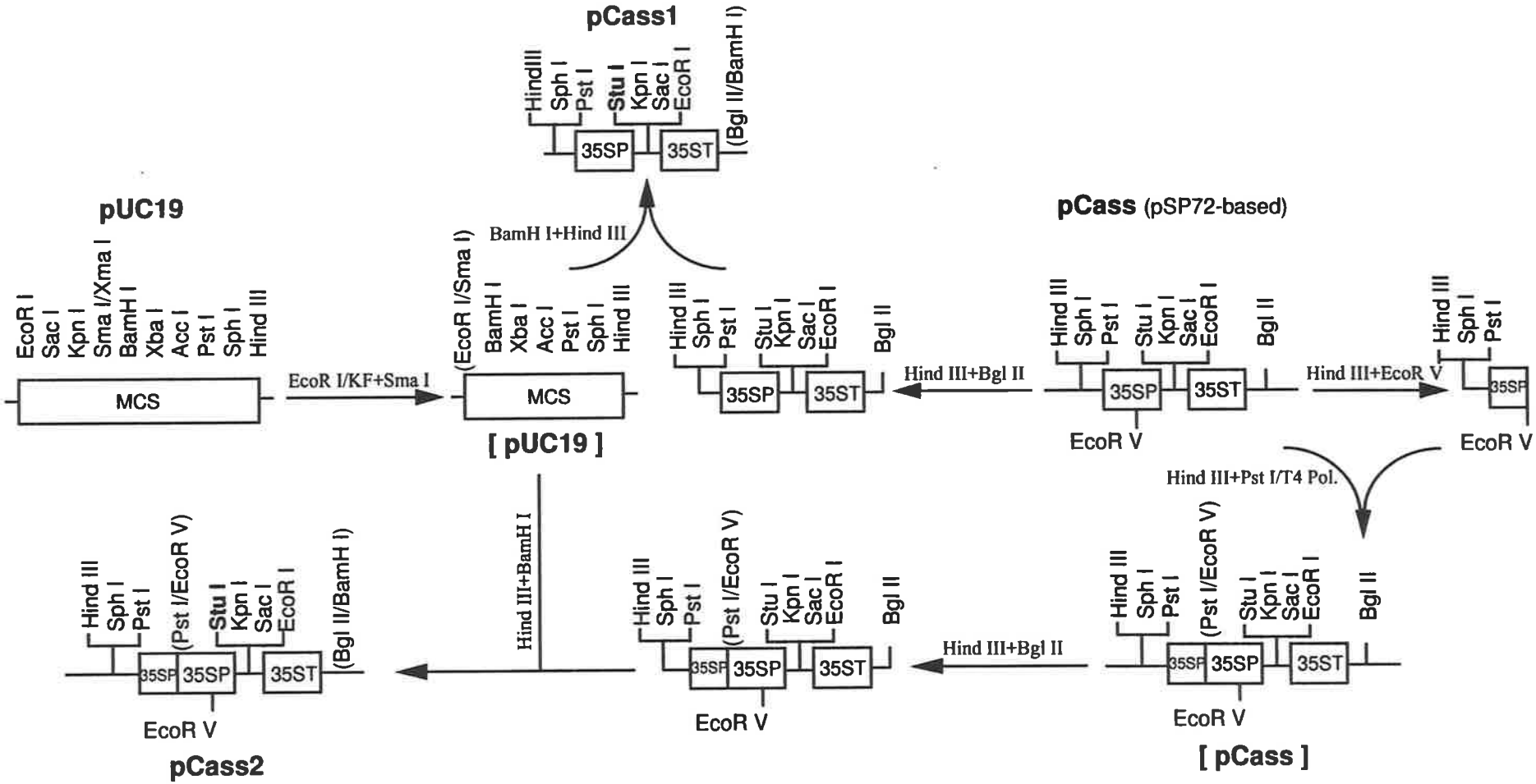


Fig. 5.1 Schematic diagram of a partial CaMV 35S promoter duplication. The pUC19 in brackets is an intermediate vector in which the restriction sites between *EcoR*I and *Sma*I in the multicloning site (MCS) have been removed from pUC19. Similarly, the pCass in brackets is the one with a partial CaMV 35S promoter inserted upstream of CaMV 35S promoter in pCass. The restriction sites in brackets are a fusion site of two restriction enzymes in the brackets. *EcoR*I and *Pst*I-restricted site are made blunt-ended with DNA polymerase [Klenow fragment (KF)] and T4 DNA polymerase (T4 Pol.), respectively. A *Hind*III-*EcoR*V fragment of CaMV 35S promoter excised from pCass is from positions -416 to -90.

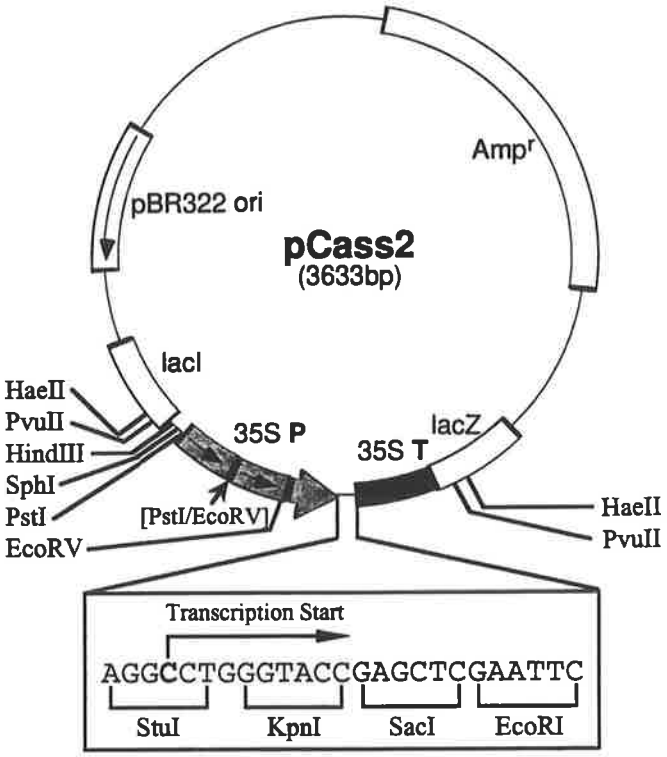


Fig. 5.2 Structure of the cloning vector pCass2. pCass2 is constructed on the basis of the modified pUC19 as shown in Fig. 5.1 with a partially duplicated CaMV 35S promoter. The multicloning site used for insertion of a foreign fragment and the transcription start site controlled by the partially duplicated CaMV 35S promoter are indicated within the box. The fusion site of the partial and the intact promoters is indicated by a short and thick arrow. Restriction by *Hae*II is used for release of the 35S transcriptional cassette containing the viral cDNAs corresponding to TAV RNAs 1 and 3 while *Pvu*II is used for excising the cassette containing TAV cDNA 2. The *Stu*I site from which transcription starts is methylation sensitive and therefore pCass2 should be propagated in dcm- strains of *E. coli,* such as JM110.

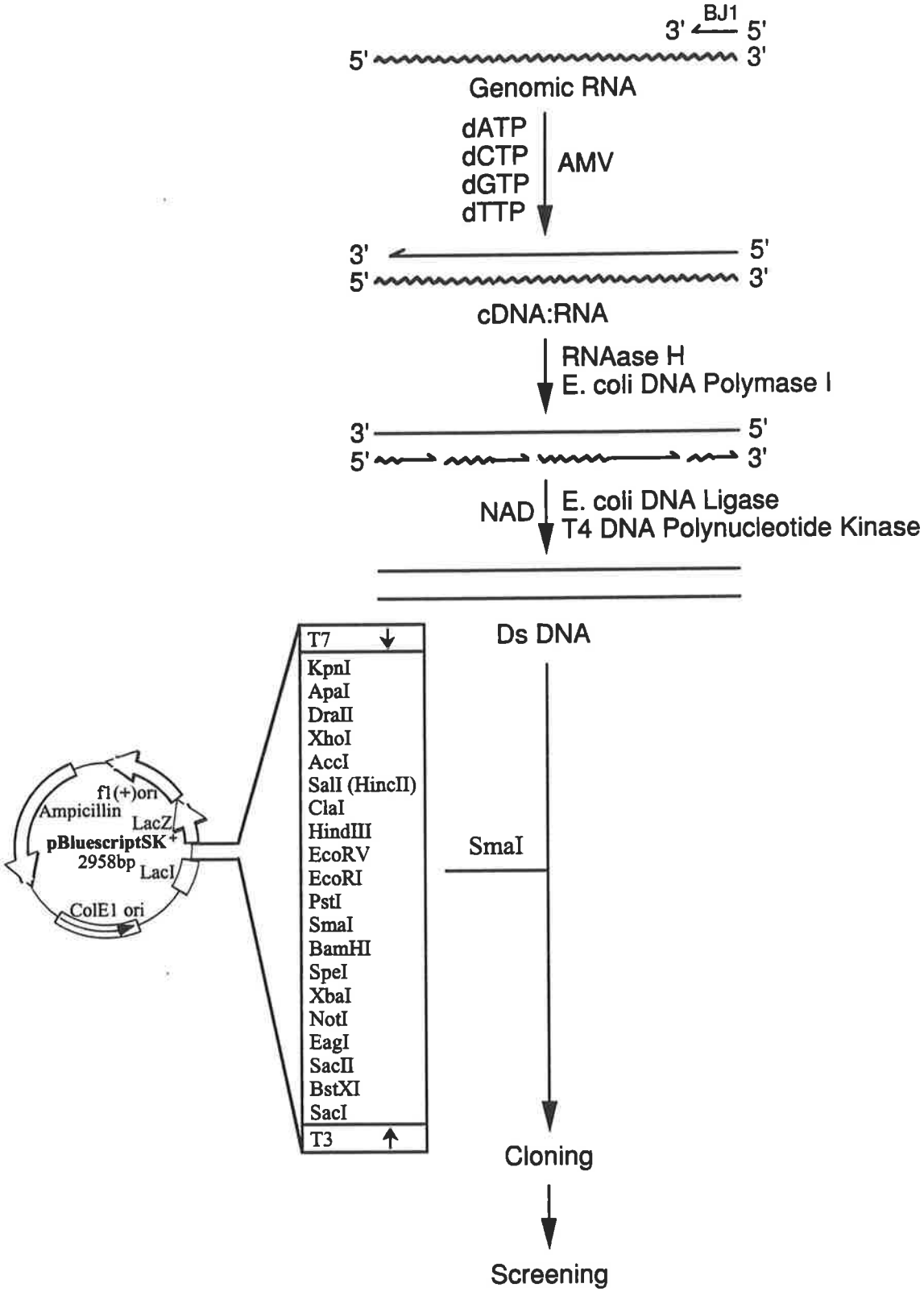


Fig. 5.3 Schematic diagram of TAV cDNA synthesis and cloning procedure. The resultant double-stranded (ds) DNA from cDNA libraries using a TAV 3'-end-specific primer as shown was ligated into a *Sma*I-digested pBluescript SK$^+$ and transferred into *E. coli* DH5α cells. Screening of cDNA clone was carried out by a combination of restriction enzyme digestion and DNA dideoxynucleotide sequencing.

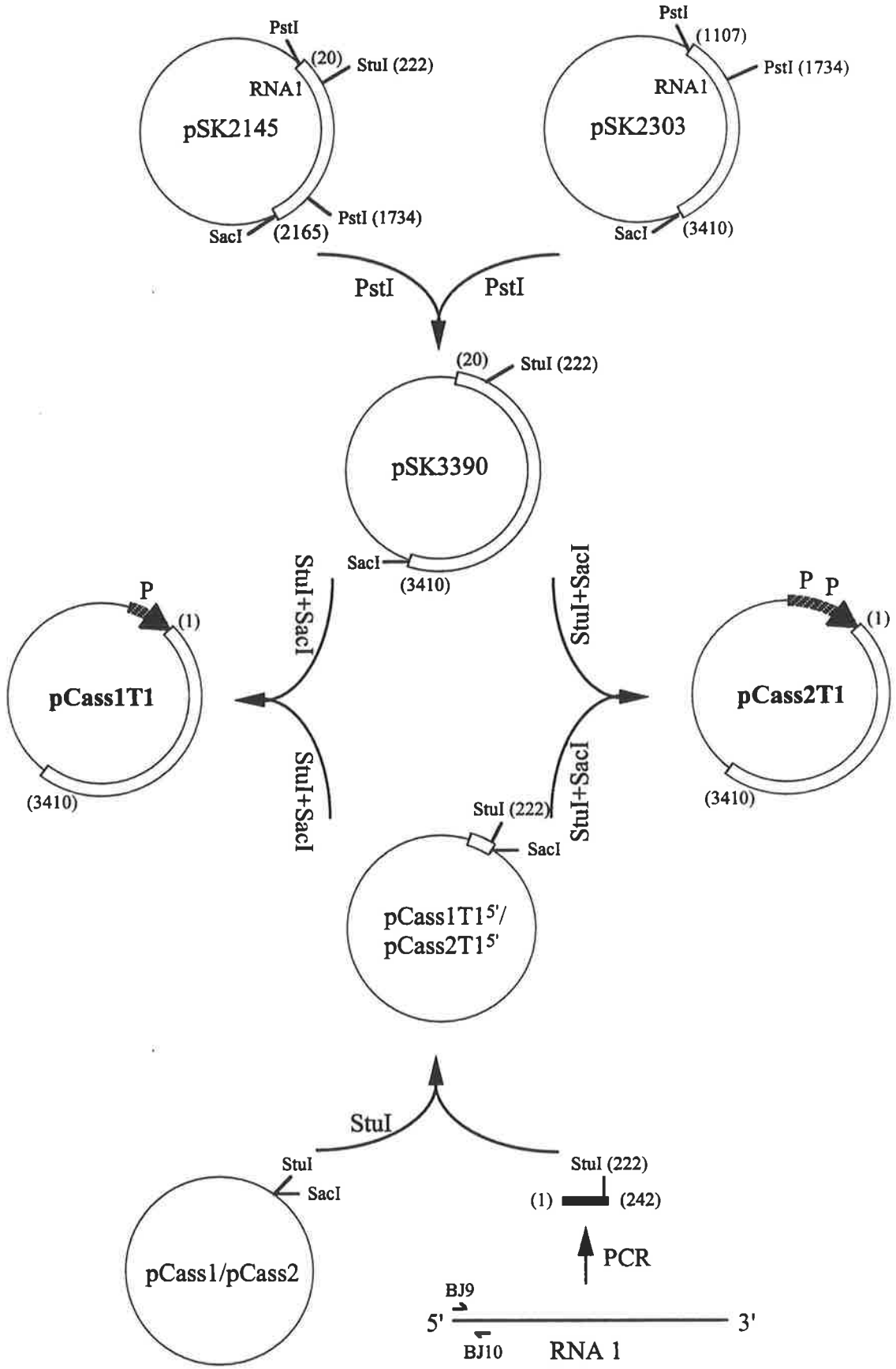


Fig. 5.4 Cloning of a full-length cDNA corresponding to TAV RNA 1 into pCass1 and pCass2. The numbers in brackets represent the nucleotide positions in each of three genomic RNAs 1, 2 and 3. The restriction enzymes used for cloning are indicated. The single and duplicated promoter are indicated by a P and PP, respectively.

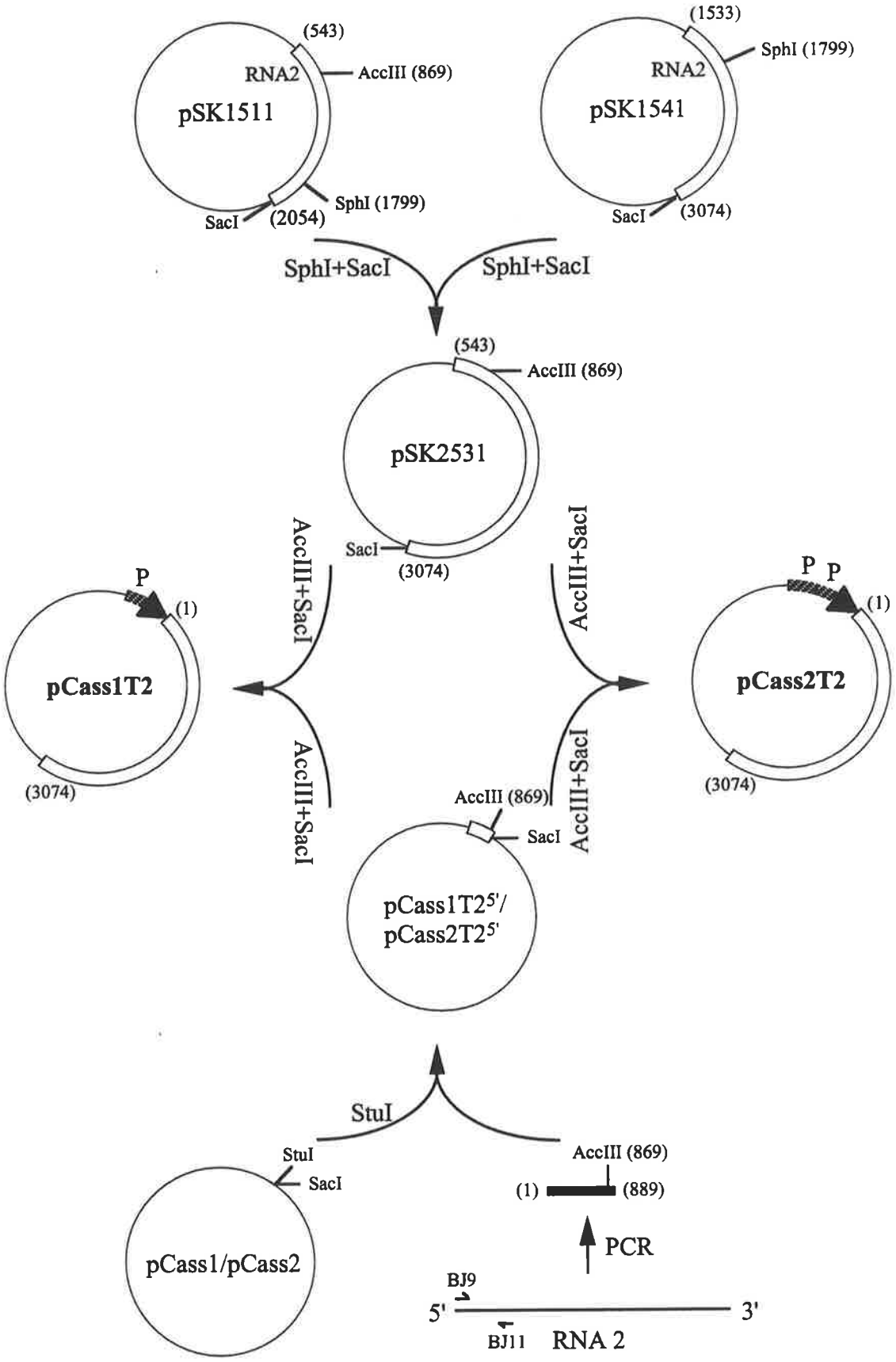


Fig. 5.5 Cloning of a full-length cDNA corresponding to TAV RNA 2 into pCass1 and pCass2. The numbers in brackets represent the nucleotide positions in each of three genomic RNAs 1, 2 and 3. The restriction enzymes used for cloning are indicated. The single and duplicated promoter are indicated by a P and PP, respectively.

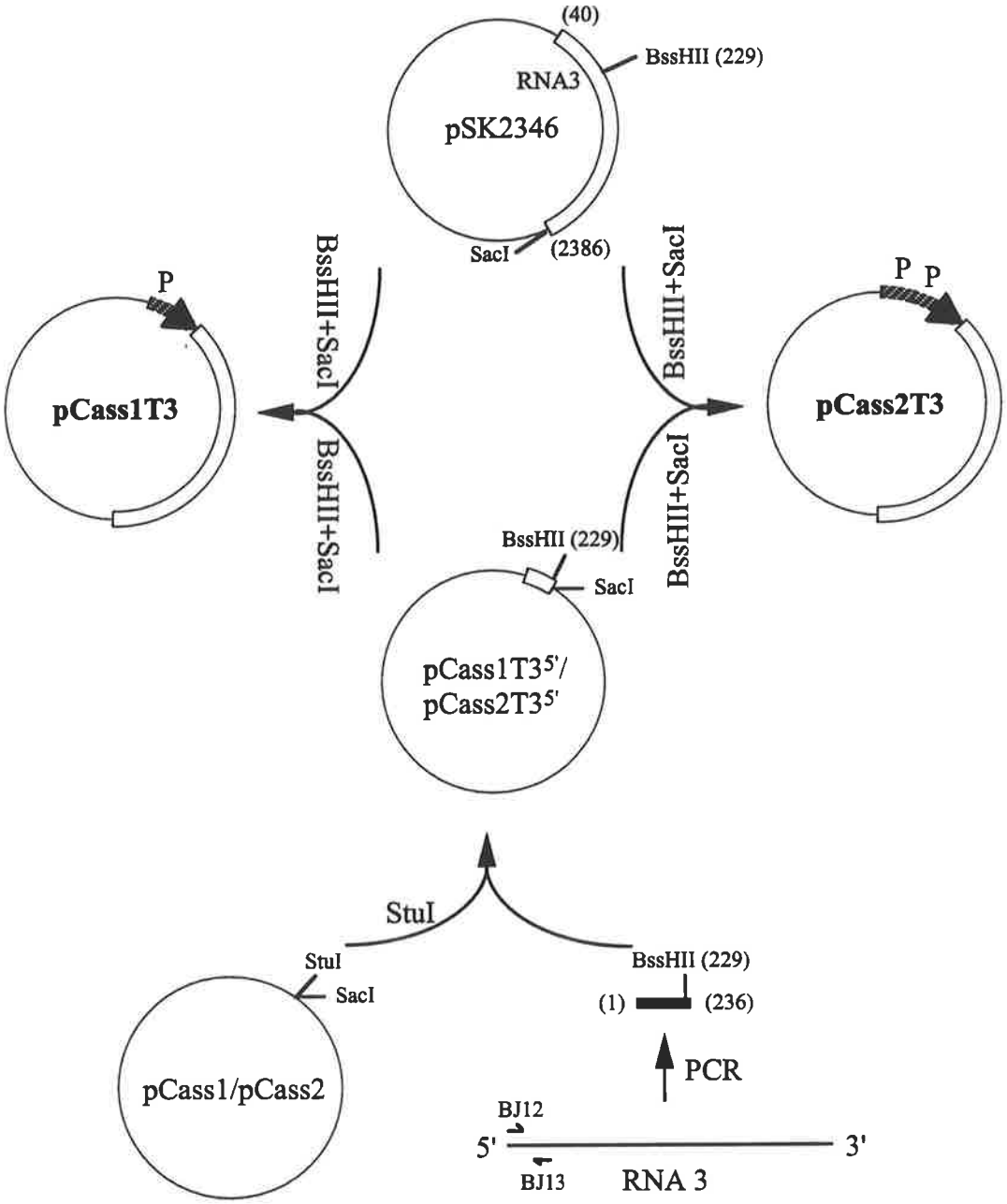


Fig. 5.6 Cloning of a full-length cDNA corresponding to TAV RNA 3 into pCass1 and pCass2. The numbers in brackets represent the nucleotide positions in each of three genomic RNAs 1, 2 and 3. The restriction enzymes used for cloning are indicated. The single and duplicated promoter are indicated by a P and PP, respectively.

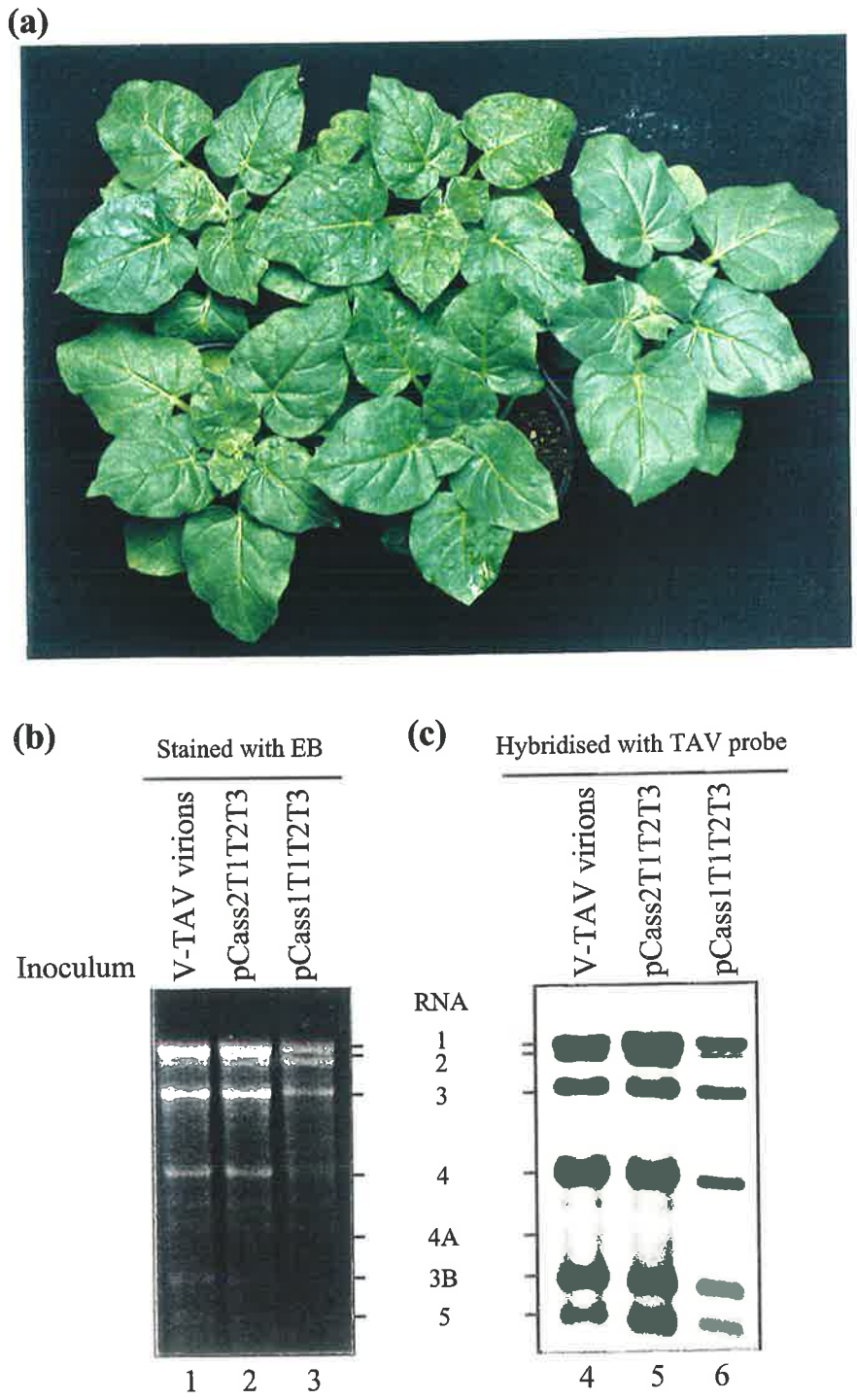


Fig. 5.7 Resulting RNAs of *N. glutinosa* plants inoculated with cDNA clones and virions of V-TAV, respectively. (a) Systemic mosaic symptoms were induced with pCass2T1T2T3 (left top) or pCass1T1T2T3 (right top) or the wildtype virions (left bottom) or pCass2T1T2T3(Δ163) (right bottom) by manual inoculation. A control healthy *N. glutinosa* with mock-inoculation is shown in the middle on the right. (b) Electrophoresis analysis of encapsidated RNAs extracted from *N. glutinosa* inoculated with the wildtype V-TAV (lane 1), pCass2T1T2T3 (lane 2) or pCass1T1T2T3 (lane 3) on a 1.2% agarose gel stained with ethidium bromide (EB). (c) Northern blot analysis of encapsidated RNAs. Virion RNAs extracted from *N. glutinosa* inoculated with the wildtype V-TAV (lane 4), pCass2T1T2T3 (lane 5) or pCass1T1T2T3 (lane 6) were electrophoresed and hybridised with a strand-specific RNA probe complementary to the 3' terminal 128 nt of all the three genomic RNAs. The positions of viral RNAs 1, 2, 3, 4, 4A, 3B and 5 are indicated.

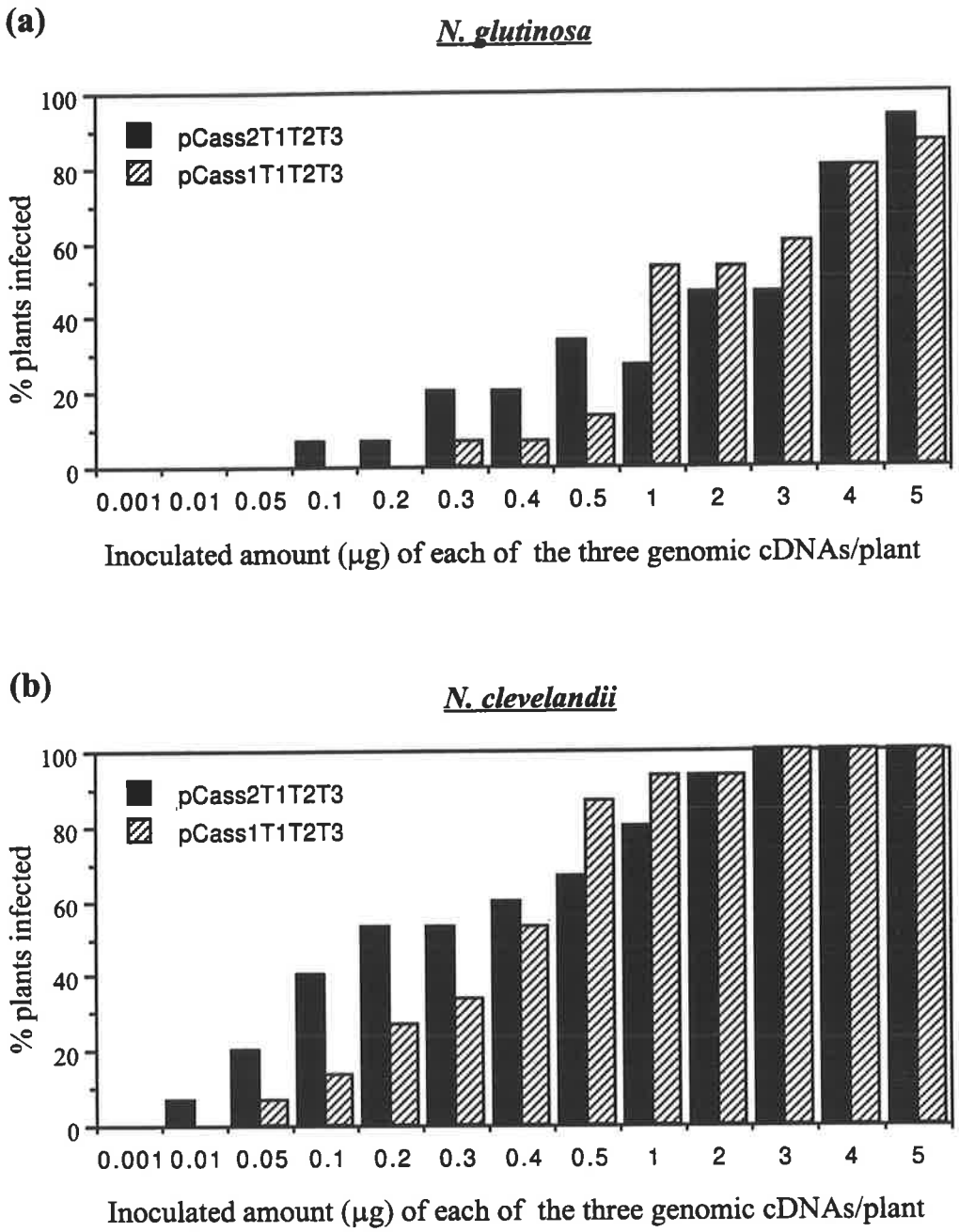


Fig. 5.8 Dose response of plants to pCass1T1T2T3 and pCass2T1T2T3 45 days after inoculation. (a) *N. glutinosa* seedlings (two true leaves) were inoculated with pCassTs and pCassITs. All points are the sum of two separate experiments with 15 plants each. (b) *N. clevelandii* seedlings (two true leaves) were inoculated with pCassTs and pCassITs. All points are the average of the results of two separate experiments with 15 plants each.

Chapter 6

TWO NOVEL SUBGENOMIC RNAS DERIVED FROM RNA 3 OF TOMATO ASPERMY CUCUMOVIRUS THAT DO NOT FUNCTION AS MESSENGER RNAS

6.1 INTRODUCTION

The previous chapters showed that tomato aspermy cucumovirus (V-TAV), a member of alpha-like superfamily (Haseloff *et al.*, 1984; Ahlquist *et al.*, 1985; Goldbach, 1987; Habili and Symons, 1989; Mattews, 1991), encapsidated two extra small RNAs as well as RNAs 1, 2, 3, 4 and 4A. These two extra small RNAs strongly hybridised to the probe which was transcribed from the 3'-coterminus of TAV genomic RNAs. Interestingly, one of these two extra RNAs has not been reported in any other cucumoviruses, whereas the other corresponded in size to the fifth RNA of another TAV strain (Lot *et al.*, 1974; Hull, 1976) or to RNA 5 of CMV (Peden and Symons, 1973). It is worth noting here that RNA 5 of CMV is not CARNA 5 (a CMV associated satellite RNA named by Kaper and Tousignant, 1973) although both named RNAs were found in one strain of CMV (Q-CMV; Gould *et al.*, 1978).

Viruses within the alpha-like superfamily often have subgenomic (sg) RNAs for expressing the internal ORFs. Three models have been proposed for the mechanism by which sgRNAs are produced (Ding, 1989): (1) Internal initiation by the RNA replicase on the minus-strand genomic RNA (Kennedy, 1980; Nassuth *et al.*, 1981), which has been shown to occur *in vitro* for brome mosaic virus (BMV; Miller *et al.*, 1985) and alfalfa mosaic virus (AlMV; van der Kuyl *et al.*, 1990) and *in vivo* for turnip yellow mosaic virus (Gargouri, *et al.*, 1989), (2) premature termination during minus strand synthesis followed by independent replication of subgenomic RNAs (Goelet and Karn, 1982), (3) processing by nuclease cleavage of genome-length RNAs (Gonda and Symons, 1979). A comparison of the sequences of genomic RNAs of a number of alphaviruses in the region spanning the start of the sgRNA, the junction region, identified a conserved 21-nucleotides (nt) domain encompassing 19 nt upstream and 2 nt downstream of the initiation start site (Ou *et al.*, 1982; van der Vossen *et al.*, 1995).

The aim of the work presented in this chapter was to complete the sequences of these two extra small TAV-associated RNAs so the origin of these RNAs could be determined and to compare these nongenomic RNAs with those of other cucumoviruses. This would help to establish the relationship, particularly in nongenomic RNAs, between TAV and other cucumoviruses, and possibly to explain the biological differences between TAV and CMV.

6.2 MATERIALS AND METHODS

6.2.1 Virus sources

V-TAV (Habili and Francki, 1974a) was propagated in *Nicotiana glutinosa* and purified as described by Peden and Symons (1973). Infectious full-length cDNA clones of each of the three genomic RNAs of the Q strain of CMV (Q-CMV) (Ding *et al.*, 1995a) and infectious cDNA clones of TAV three genomic RNAs, which are under the control of the partially duplicated CaMV 35S promoter as shown in Chapter 5, were chosen and inoculated on *N. glutinosa* and *N. clevelandii,* as described in Chapter 5. Total RNAs were extracted from plants by the method of Verwoerd *et al.* (1989) as described in Chapter 2 (Section 2.2.5).

6.2.2 Purification of RNAs 3B and 5

RNAs 3B and 5 were initially purified with a commercial kit (RNaid; Bio 101) as outlined below. Purified virion RNAs (20 μg) were heated at 65°C for 10 minutes in the presence of 50% deionized formamide, 29.4% formaldehyde, 1x gel buffer [200 mM MOPS (pH 7.0), 10 mM EDTA (pH 7.0), 10 mM NaOAc (pH 5.2)] and 0.02μg/μl ethidium bromide and then loaded on a 1.2% agarose gel containing 6.6% formaldehyde and 1x gel buffer. RNAs were visualized with long wave UV after electrophoresis (3-5V/cm) and gel slices containing individual RNAs 3B and 5 were excised and placed in 1.5 ml Eppendorf tubes. To 0.1 g gel slice, 0.3 ml RNA BINDING SALT containing

0.02% ACETIC ACID was added. When the gel slices were completely melted by incubation at 37°C for about 10 minutes, 1-2 µl of RNAMATRIX™ per µg of RNA was added. The mixtures were incubated at room temperature for further 10 minutes with occasional vortexing, centrifuged for 1 minute at 14000rpm (Eppendorf) and the pellets resuspended in 500 µl RNA WASH solution. After repeated centrifugations and resuspensions twice, the pellets were resuspended in 10-20 µl of 1 mM EDTA (pH 7.0) per 5 µl RNAMATRIX™. The suspensions were incubated at 80°C for 10 minutes and then centrifuged for 1 minute at 14000rpm (Eppendorf). The supernatants containing individual RNAs 3B and 5 were removed to new 1.5 ml Eppendorf tubes.

Further purification of RNAs 3B and 5 was as described by Symons (1978). Each of purified RNAs 3B and 5 from the above agarose gel was heated at 65°C for 15 minutes in the presence of 50% formamide to release aggregated and nicked RNA, and mixed with 0.5 volume FLB loading buffer. Each of the mixtures was then loaded on a 3% (0.15% bisacrylamide) polyacrylamide slab gel containing 7 M urea and 1x TBE buffer. After electrophoresis (2V/cm) the gel was stained with 0.1% toludine blue for 10 minutes and destained with repeated changes of water. Gel slices containing individual RNAs 3B and 5 were excised and the RNAs eluted by a mixture of 500 mM NH_4Ac, 1 mM EDTA and 0.1% SDS. The RNAs were then precipitated with ethanol as described in Chapter 2 (Section 2.2.5). The pelleted RNAs 3B and 5 were finally resuspended in sterile water to a concentration of 1µg/µl.

6.2.3 Poly-A tailing of RNAs 3B and 5

Purified RNAs 3B and 5 were 3'-poly(A) tailed using poly(A) polymerase (Pharmacia Biotech) as described by the manufacturer. One µg purified RNA 3B or RNA 5 was denatured by heating at 80°C for 1 minute followed by snap-cooling on ice. Poly-A addition was carried out on the denatured RNA by adding: 40 mM Tris-HCl pH 8.0, 250 mM NaCl, 2.5 mM $MnCl_2$, 10 mM $MgCl_2$, 250 µM ATP, 50µg/ml BSA, 1 µl

(0.85unit/µl) poly-A polymerase to a volume of 20 µl. The reaction was incubated at 37°C for 60 minutes and stopped by the addition of EDTA to 10 mM. The RNAs 3B and 5 were recovered, after phenol/chloroform (1:1 v:v) extraction, by ethanol precipitation, as in Chapter 2 (Section 2.2.5).

6.2.4 Enzymatic sequencing of RNAs 3B and 5

The method was adapted from Haseloff and Symons (1981) and Forster *et al.* (1990). 0.5 µg of the purified RNA 3B or RNA 5 was 5'-dephosphorylated using calf intestinal alkaline phosphatase (CIAP) as in Chapter 2. For 5'-radiolabelling, the dephosphorylated RNA 3B or RNA 5 was added to 200 µCi [γ-^{32}P]ATP, reduced to 6 µl *in vacuo* (Jouan), and then 1 µl of 12 mM spermidine added. The mixtures were heated at 80°C for 1 minute followed by snap-cooling on ice and then 2 µl of 5x T4 polynucleotide kinase (PNK) buffer and 1 µl (5 units) T4 PNK were added and the solution incubated at 37°C for 30 minutes. For 3'-radiolabelling, the dephosphorylated RNA 3B or RNA 5 was heated at 80°C for 1 minute followed by snap-cooling on ice and 50 µCi [5'-^{32}P]dpCp, which was synthesised from [γ-^{32}P]ATP and 3' dCMP using T4 PNK as described above, 5 units T4 RNA ligase (New England Biolabs), 1 mM ATP and 1x T4 RNA ligase buffer containing 50 mM Tris-HCl (pH7.8), 10 mM $MgCl_2$, 10 mM 2-mercaptoethanol added. After incubation at 4°C for 16 hours, the 3'-radiolabelling reaction and the 5'-radiolabelling reaction were separately heated at 80°C for 30 seconds and loaded on a 3% (0.15% bisacrylamide) polyacrylamide slab gel and each of radiolabelled RNAs recovered by elution from the gel as in 6.2.2. After each of these steps, each of the 5'- and 3'-radiolabelled RNAs of each of RNAs 3B and 5 was mixed with 10 µg of *E. coli* tRNA (phe) and the mixtures dried down *in vacuo*. The RNAs were resuspended in 6 µl of water and the suspensions divided into 6 separate 1 µl aliquots.

To each aliquot was added one of the following reaction mixtures: (no enzyme) 9 µl 20 mM sodium citrate, pH 5.0, 7 M urea, 1 mM EDTA; (ladder) 5 µl 50 mM

$Na_2CO_3/NaHCO_3$, pH9.0; (RNase T_1) 9 µl 20 mM sodium citrate, pH 5.0, 7 M urea, 1 mM EDTA, 1 µl RNase T_1 (5units/µl); (RNase Phy M) 9 µl 20 mM sodium citrate, pH 5.0, 7 M urea, 1 mM EDTA, 1 µl RNase Phy M extract; (RNase U_2) 9 µl 20 mM sodium citrate, 7 M urea, pH 3.5, 1 mM EDTA, 1 µl RNase U_2 (5units/µl); (*Bacillus cereus* RNase) 5 µl 20 mM sodium citrate, 1 mM EDTA, pH 5.0, 1 µl *B. cereus* RNase extract. The ladder tube was incubated at 100°C for 90 seconds while the remaining tubes were incubated at 50°C for 20 minutes. Reactions were stopped by the addition of an equal volume of formamide loading solution. Samples were heated at 80°C for 1 minute followed by snap cooling on ice just prior to loading onto 0.5 mm thick, 1x TBE, 7 M urea, 10 % polyacrylamide gels. Sequencing patterns were visualised by autoradiography using an intensifying screen at -80°C.

6.2.5 cDNA synthesis, cloning and sequencing

The rapid amplification of cDNA ends (RACE) technique (Frohman *et al.*, 1988) used for cDNA syntheses of RNAs 3B and 5 is outlined in Fig. 6.1. One µg each of polyadenylated RNAs 3B and 5 was primed with 0.5 µg of an oligo-dT primer [5'-GACTCGAGTCGACATCGA$(T)_{17}$-3'; a gift of David Warrilow (in this laboratory)] and first strand cDNA was synthesised by reverse transcription (RT) with avian myeloblastosis virus (AMV) reverse transcriptase as in Chapter 2. After phenol/chloroform (1:1 v:v) extraction, the solution was passed through a spin column containing a chromatography matrix (Sephacryl S-400; Promega) as described by the manufacturer to remove unincorporated dNTPs, and then 3' dA-tailed using 5 units terminal deoxynucleotidyl transferase (Promega) in the presence of 100 mM cacodylate buffer (pH 6.8), 5 mM $CoCl_2$, 0.5 mM DTT, 0.25 mM dATP. The polymerase chain reaction (PCR) was carried out using 0.1 µg of the single oligo-dT primer and 2 units of Taq DNA polymerase (Promega) in the presence of 1.5 mM each of dNTPs, 50 mM KCl, 10 mM Tris-HCl (pH 9.0), 0.1% Triton X-100 and 1.5 mM $MgCl_2$ at 94°C for 1

minute, and then 30 cycles (94°C, 10 seconds; 50°C, 10 seconds; 72°C, 1 minute) followed by a 5-minute final extension at 72°C. The resultant product was cloned into the *Sma*I site of plasmid pBluescript SK^+ and sequenced from both orientations by dideoxynucleotide chain termination as described in Chapter 2.

6.2.6 RNA dideoxynucleotide sequencing and primer extension

Direct RNA dideoxynucleotide sequencing and primer extension in which no ddNTPs were added (Fichot and Girard, 1990) were carried out using virion RNAs, total V-TAV-infected plant RNAs, RNA 3B and RNA 5, respectively, and an oligonucleotide primer complementary to a known sequence of the 5' terminal region of RNA 3B and RNA 5 obtained as above.

6.2.7 Computer analysis of the sequence

Program DNA Strider was used for analysing the coding capability while GCG was used to describe the secondary structures and similarities between sequences of RNAs 3B and RNA 5, and all three genomic RNAs of TAV published to date.

6.2.8 Pseudorecombinants between V-TAV and Q-CMV

Infectious cDNA clones of RNA 3 of V-TAV and Q-CMV were exchanged between infectious cDNA clones of V-TAV and Q-CMV to make two different pseudorecombinant plasmid mixtures, C1C2T3 and T1T2C3, in which numbers 1, 2 and 3 correspond to genomic RNAs 1, 2 and 3 of each virus and C and T correspond to Q-CMV and V-TAV, respectively. The mixtures were then inoculated directly onto two common host plants, *N. glutinosa* and *N. clevelandii*.

6.2.9 Deletion cDNA clone of V-TAV RNA 3

The deletion of first copy of the two 163 nt tandem repeats in V-TAV RNA 3 (F. Garcia-

Arenal, personal communication) was constructed by sequentially digesting pCass2T3 as follows. pCass2T3 was digested with *Tth111*I and *Spe*I to produce a *Tth111*I-*Spe*I fragment which was followed by digestion with *Dde*I to give a *Tth111*I-*Dde*I fragment and a *Dde*I-*Spe*I fragment. After purification on a 6% native polyacrylamide gel, the *Tth111*I-*Dde*I fragment and the *Dde*I-*Spe*I fragment were ligated with the *Tth111*I-*Spe*I restricted pCass2T3 using T4 DNA ligase (Bresatec). After confirmation by digestion and sequencing, the resulting deletion (designated as pCass2T3Δ163) and wildtype pCass2T3 were separately coinoculated unless otherwise described with pCass2T1 and pCass2T2 onto *N. glutinosa* plants with an equal amount (10 μg each plasmid in a final volume of 10 μl) as in Chapter 5.

6.3 RESULTS

6.3.1 RNA 3B and RNA 5 are both associated with V-TAV

RNAs 1, 2 and 3 (genomic RNAs) and RNA 4 (a subgenomic RNA of RNA 3 and the mRNA for coat protein) of V-TAV have been reported (Habili and Francki, 1974a; Moriones *et al.*, 1991; Bernal *et al.*, 1991; Garcia-Arenal, personal communication). It has shown in Chapter 3 that TAV RNA 2 also has a subgenomic RNA, RNA 4A (702 nt), which most likely functions as a mRNA for a newly discovered overlapping gene 2b for which the translational protein product has been detected in infected plants.

As shown by the northern blot hybridisation analysis in Fig. 6.2b, two additional smaller RNA species, designated as RNA 3B and RNA 5 (see below), were found encapsidated in virions of V-TAV (lane 4) and these hybridised to an RNA probe (probe T) complementary to the 3' terminal 100 residues of V-TAV RNA 3 (Fig 6.2a), indicating that these two RNA species are of a TAV origin and that they most likely contain the conserved 3' terminal sequences. These two RNAs accumulated in infected plants to levels comparable to the other known viral RNAs (lane 5). An RNA of similar size to that

of RNA 3B has not been detected in viral RNAs of either Q-CMV (a subgroup II strain) or WAII strain of CMV (WAII-CMV, a subgroup I strain) (Fig. 6.2b, lanes 2 and 3) using an RNA probe (probe C in Fig. 6.2a) complementary to the 3' terminal 495 residues of Q-CMV RNA 3. In addition, these two RNAs were detected at early stages of virus infection (6 days after inoculation) in leaves inoculated with plasmid DNAs of the infectious cDNA clones of V-TAV (data not shown), indicating that these two RNAs are probably not defective interfering (DI) RNAs as the appearance of such RNAs usually requires several passages at high multiplicity of infection (Hillman *et al.*, 1987). Therefore, it is most likely that RNAs 3B and 5 are novel subgenomic RNAs of TAV.

6.3.2 RNAs 3B and 5 are subgenomic RNAs derived from RNA 3

To determine the nature and the molecular structure of RNAs 3B and 5, complete nucleotide sequences of both RNAs were determined after purification from virions by a combination of three sequencing strategies described in Materials and Methods. Firstly, gel purified RNA 3B or RNA 5 were either 5'- or 3'-end labelled and sequenced enzymatically. Secondly, full-length cDNA clones of RNA 3B and RNA 5 were prepared by the RACE technique and seven cDNA clones of RNA 3B and five clones of RNA 5 were sequenced in both orientations. Lastly, the 5'-terminal sequences of purified RNAs 3B and 5 were further confirmed by RNA dideoxynucleotide sequencing and primer extension and sequencing experiments (Fig. 6.3) using internal primers designed from sequences obtained in the two other approaches.

The complete nucleotide sequences of both RNAs are presented in Fig. 6.4; RNA 3B is 486 nt long and RNA 5 323 nt long. As the 5' end-labelling of both RNAs required dephosphorylation but not decapping, it is likely that both RNAs 3B and 5 are 5'-phosphorylated and not capped. Sequence comparisons revealed that the 3'-terminal 323 nt of RNA 3B is identical in sequence to RNA 5 whereas its 5'-terminal 163 nt is a direct repeat of the 5'-half (163 nt) of RNA 5 (Fig. 6.4). Thus, the 5'-terminal sequence (326

nt) of RNA 3B consists of two tandem repeats of 163 nt that differ only by a single nt (marked by a asterisk in Fig. 6.4) while RNA 5 contains only one copy of the repeat. In other words, the 5'-terminal sequences (163 nt) of both RNAs 3B and 5 differ only by one nt. Importantly, both RNAs encode no open reading frame of 21 codons or longer, and the tandem repeats of RNA 3B have not been found in the other strains of this virus sequenced to date (O'Reilly *et al.*, 1991; 1994; Salanki *et al.*, 1994).

The nucleotide sequences of RNAs 3B and 5 were further compared to those of V-TAV RNAs 1 and 2 (Moriones *et al.*, 1991; Bernal *et al.*, 1991) as well as to the unpublished sequence of RNA 3 of V-TAV (Garcia-Arenal, personal communication). It has been found that RNA 3B and RNA 5 show 99.4% and 99.7% identity to the 3'-terminal 486 and 323 of RNA 3 of V-TAV, respectively. The differences between either RNA 3B or RNA 5 and the RNA 3 sequences include two substitutions (G^{1974} of RNA 3 to U^{73} of RNA 3B; U^{1975} of RNA 3 to G^{74} of RNA 3B) and one insertion (U^{469} of RNA 3B or U^{306} of RNA 5 between nucleotides 2369/2371 of RNA 3). However, RNA 3B is completely identical in sequence to the 3'-terminal 486 nt, and RNA 5 to the 3'-terminal 323 nt, of V-TAV RNA 3 determined from four RNA 3 cDNA clones (data not shown). As RNA 3 contains a 3' untranslated region of 502 nt which encodes no open reading frame of 21 codons or longer (Garcia-Arenal, personal communication), it can be concluded that both RNAs 3B and 5 are completely derived from the 3' untranslated region of RNA 3 and are unlikely to function as mRNAs. The lack of 5'-capping of RNAs 3B and 5 further supports a non-mRNA role for these RNAs.

Structure computer-aided analyses using a GCG program showed that both RNAs 3B and 5 had a tRNA-like structure at the 3' terminal region (data not shown), which was similar to that of the same strain RNA 4 proposed by Wilson and Symons (1981). Such tRNA-like structure at the 3'-terminal region has been shown to be involved in minus strand synthesis in brome mosaic virus (BMV), a virus closely related with TAV (Miller

et al., 1986), and in several other plant viruses (Zaccomer *et al.*, 1995).

6.3.3 RNAs 3B and 5 are derived from RNA 3 as shown by genetic studies

The following two sets of genetic studies further support the conclusion that both RNAs 3B and 5 are derived from RNA 3. Using infectious cDNA clones of the three genomic RNAs of V-TAV (see Chapter 5) and Q-CMV (Ding *et al.*, 1995a), two pseudorecombinant viruses, T1T2C3 and C1C2T3, were reconstituted by exchanging RNA 3 cDNA clones between the two viruses. Both pseudorecombinants were viable as systemic virus infections were established in *N. glutinosa* plants, in agreement with earlier work using purified RNAs to prepare pseudorecombinants (Habili and Francki, 1974b). Total RNAs were extracted from systematically infected leaves and analysed by northern blot hybridisation using probes T and C specific for the 3' conserved untranslated region of V-TAV and that of Q-CMV, respectively (Fig. 6.2a). RNA 3B was only detected in C1C2T3 infections (Fig. 6.5b, lane 2) and was not detected in T1T2C3 infections (lane 3), demonstrating that TAV RNA 3 is the genetic source of RNA 3B.

In contrast to RNA 3B, RNA 5 of V-TAV was detected in infections of both pseudorecombinants (Fig. 6.5b, lanes 2 and 3); the specificity of the probe only for the 3' conserved untranslated region of V-TAV was demonstrated by the absence of hybridisation with any of the six known RNAs of Q-CMV RNAs (lane 6). Hence, this indicates that V-TAV RNA 5 is also derived from RNAs 1 and/or 2 in addition to RNA 3. RNA 5 of CMV has recently been demonstrated to be a mixed population derived from the conserved 3'-terminal regions of genomic RNAs 2 and 3 (Blanchard *et al.*, 1996).

In the second approach, a derivative (pCass2T3Δ163) of the V-TAV RNA 3 cDNA clone was constructed by deleting the first copy of the tandem repeats. Co-inoculation of

pCass2T3Δ163 with the cDNA clones of V-TAV RNAs 1 and 2 gave rise to a viable mutant virus designated T1T2T3(Δ163) that systemically infected *N. glutinosa* (Fig. 6.5b, lane 1). As expected, the progeny RNAs 3 and 4 of T1T2T3(Δ163) (lane 1) migrated faster than those of the wildtype V-TAV (lane 5) in the denaturing agarose gel electrophoresis. However, only RNA 5, not RNA 3B, was detected in the plants infected with T1T2T3(Δ163), thus unequivocally demonstrating that RNA 3B originated from RNA 3.

6.4 DISCUSSION

Sequence analysis showed that RNA 3B is 486 nt in length containing two tandem repeats of 163 nt whereas RNA 5 is 323 nt in length containing only one copy of the repeats, and that both RNAs are identical in sequence to the 3'-terminal 486 and 323 nt of RNA 3, respectively, and have no ability to encode any proteins. Genetic studies demonstrated that these two novel subgenomic RNAs, RNAs 3B and 5, are derived from TAV RNA 3. RNA 5 of V-TAV was also detected in infections of the pseudorecombinant (T1T2C3) suggesting that RNA 5 may be a mixed population derived from RNA 3 and RNAs 1 and/or 2 as found in CMV (Blanchard *et al.*, 1996).

6.4.1 Origin of RNAs 3B and 5

RNA 3B and RNA 5 have the same initiation site, suggesting that these two RNAs might have the same mechanism for generation. Using a radioactive probe complementary to the whole of RNA 3 except for the region common to RNA 3B, only the full-length RNAs 3 and 4 were detected (Fig. 6.6). No other fragments homologous to RNA 3 could be detected even after long exposure (not shown), indicating that RNA 3B and RNA 5 are not the result of an endonucleolytic cleavage of RNA 3 at a specific site. Hence, these subgenomic RNAs are presumably transcribed from the (-) RNA 3 sequence (Miller *et al.*, 1985). The promoter region required for subgenomic RNA synthesis has been

investigated with some other members of the *Bromoviridae*, brome mosaic bromovirus (BMV) (French and Ahlquist, 1987; 1988; Marsh *et al.*, 1988; Smirnyagina *et al.*, 1994), cowpea chlorotic mottle bromovirus (Allison *et al.*, 1989; Pacha and Ahlquist, 1992), alfalfa mosaic alfamovirus (AlMV) (van der Kuyl *et al.*, 1990; 1991, van der Vossen *et al.*, 1995) and CMV (Boccard and Baulcome, 1993). Most of these promoter regions have been shown to lie predominantly upstream of the transcription initiation site, eg. of BMV and CMV subgenomic RNA 4. As RNAs 3B and 5 have the same initiation sequence, they may have the same promoter-like sequences. However, sequence analysis showed that the sequence upstream of RNA 3B on RNA 3 is different from that upstream of RNA 5 on RNA 3, suggesting that the organisation of the RNA 3 promoter for subgenomic RNA 3B and RNA 5 might be, like beet necrotic yellow vein virus RNA 3 subgenomic RNA promoter (Balmori *et al.*, 1993), situated downstream, and within, each repeat of RNAs 3B and 5 on RNA 3.

The mechanism of generation of repeats remains unresolved. TAV RNA 3 as well as RNA 3B contain two tandem repeats of 163 nt, suggesting an association with replication slippage (Hancock *et al.*, 1995). A single nucleotide difference in two repeats, A^{1966} and G^{2129} could arise by a copy error of an RNA polymerase replicating a single RNA species (Keese *et al.*, 1988). However, due to the sequence from residue G^{2754} to A^{2821} of RNA 2 (Moriones *et al.*, 1991) or from U^{3131} to A^{3161} of V-TAV RNA 1 (Bernal *et al.*, 1991) corresponding to the region around A^{1966} in the first repeat of V-TAV RNA 3, the two tandem repeats of RNA 3 could also have arisen by RNA polymerase jumping (Keese and Symons, 1985; 1987) from the primary template RNA 3 to the nascent complementary RNA 1 or RNA 2. Such a replicase-mediated copy choice-type model has been proposed to explain the generation of DI RNAs (Pogany *et al.*, 1995).

6.4.2 Possible roles of the tandem repeats

Direct repeats have been investigated in other plant RNA viruses (Ding *et al.*, 1989). In

AlMV, a tandem repeat at the 5' untranslated region of RNA 3 affected accumulation of P3 protein (van der Vossen *et al.*, 1993) and may be involved with ribosome binding (Pinck *et al.*, 1981). In *ononis* yellow mosaic virus, the repeat of the 5' untranslated region has been proposed to form a stem-loop structure to modulate initiation of translation of the tymoviral overlapping genes (Ding *et al.*, 1989). In this study, a TAV mutant in which one of the repeats has been precisely removed from RNA 3 did not apparently affect the symptoms and host range as compared to those of the wildtype TAV (data not shown), indicating that the repeat may not be of major biological significance. In addition, the TAV mutant was stable in the presence of the wildtype TAV (data not shown), suggesting that the repeat did not interact with the other viral RNAs. However, it is feasible that the repeat sequence may provide some selection advantage under specific host and/or environmental conditions.

6.4.3 RNA 3B and RNA 5 may function as regulation RNAs

Sequence analysis showed that both RNA 3B and RNA 5 contain no open reading frames longer than 63 nt. It is unlikely that these two RNAs encode any protein although they have been clearly shown to be two genuine subgenomic RNAs. A similar subgenomic RNA with no coding ability has been found in barley yellow dwarf luteovirus (Kelly *et al.*, 1994). Pseudorecombinant experiments showed that the level of accumulation of RNA 3B is higher than that of RNA 3 in C1C2T3 (Fig. 6.5, lane 2) while it is lower than that of RNA 3 in wild type V-TAV (Fig. 6.5, lane 5). Combined with the fact that RNAs 3B and 5 are abundantly transcribed during infection, this leads us to speculate that both RNAs might have a regulatory role in viral RNA replication.

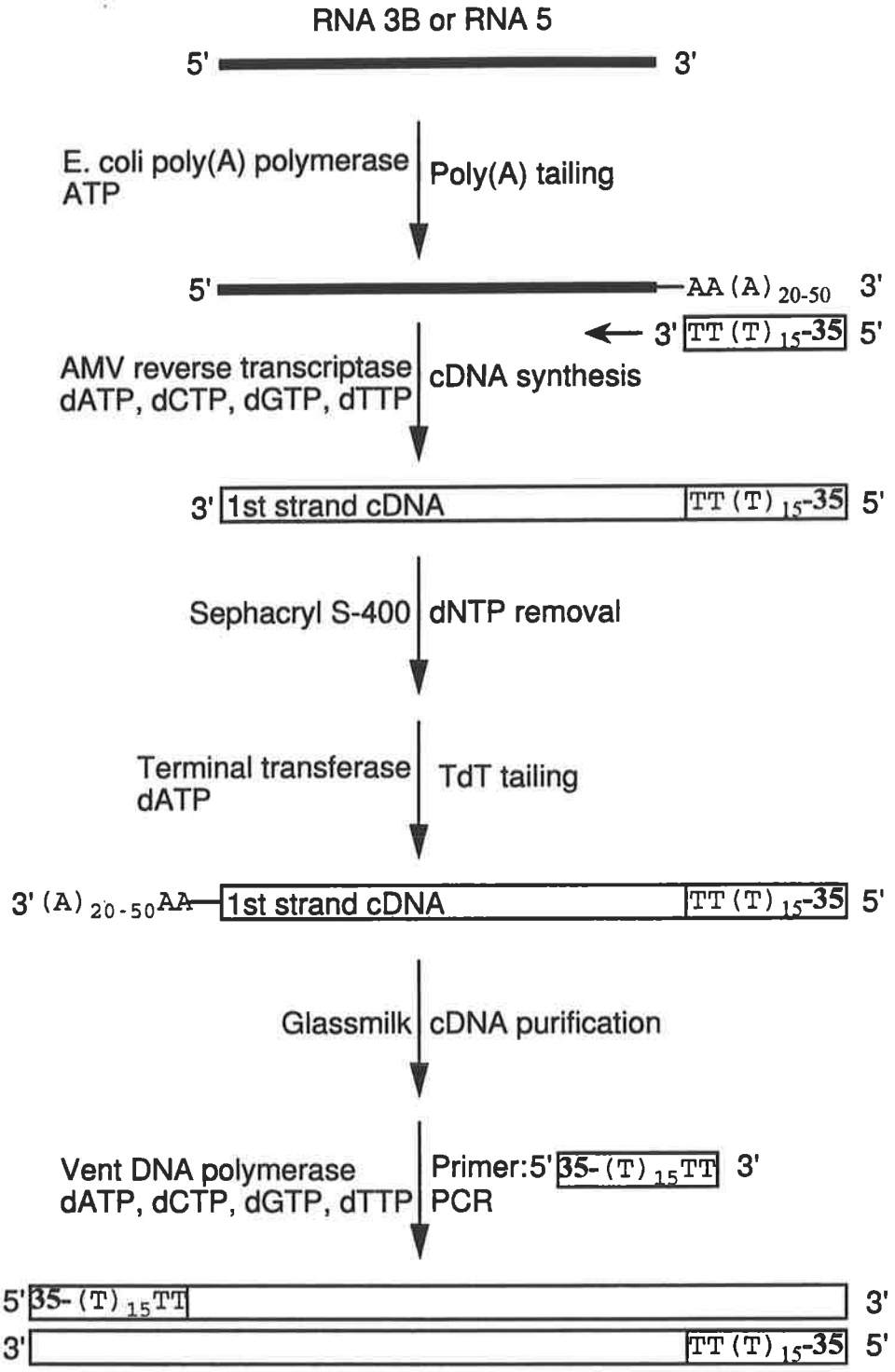


Fig. 6.1 Schematic representation of cDNA syntheses of RNAs 3B and 5. Explanations are given in the text. At each step, the diagram is simplified to illustrate only how the new product formed during the previous step is utilized. The 35-nucleotide primer [35-$(T)_{15}$TT] used for either first strand synthesis or PCR contains a 17-nucleotide oligo(dT) sequence at the 3'-end following a 18-nucleotide sequence encoding *PaeR*71 and *Hinc*II recognition sites as described in the text.

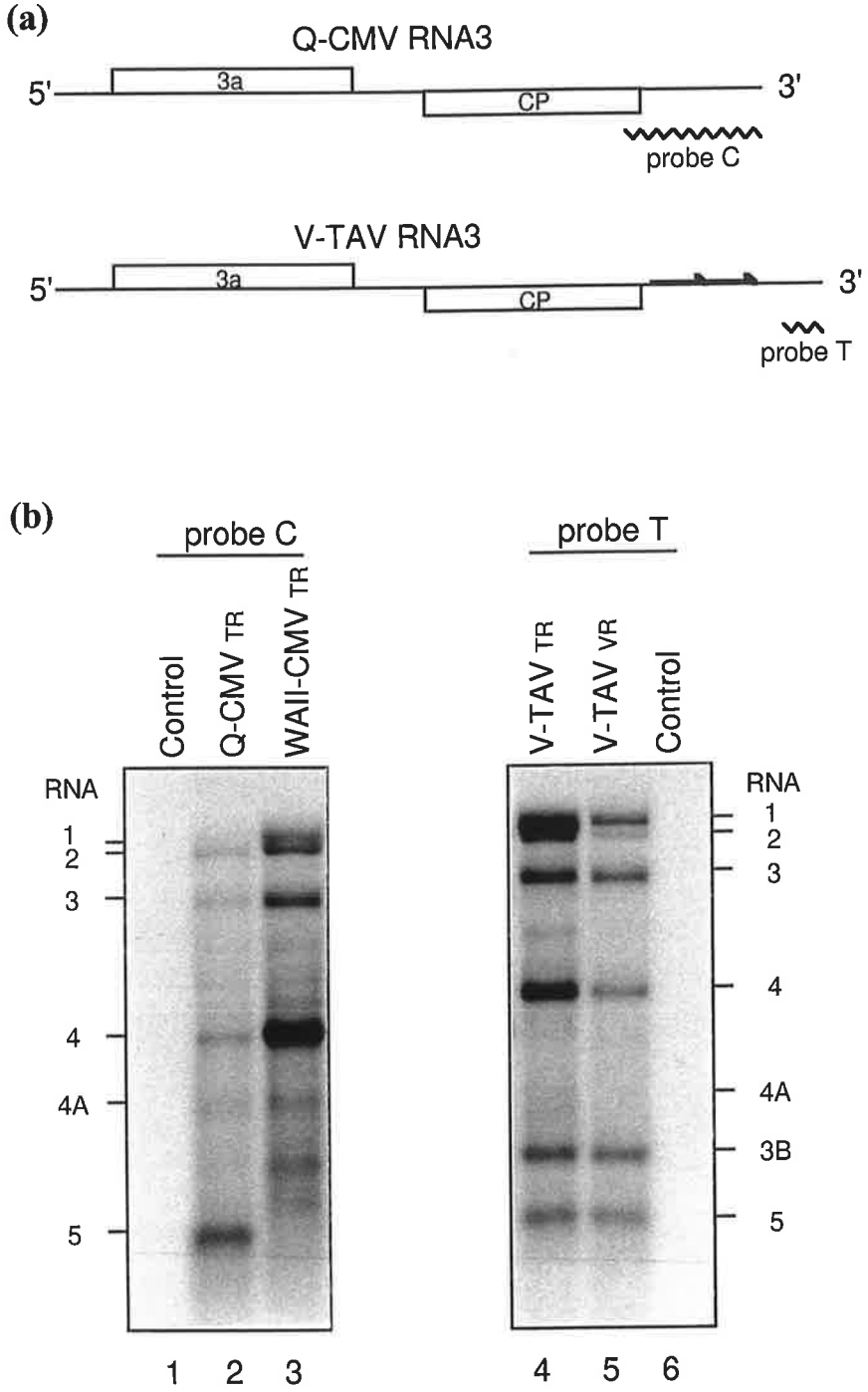


Fig. 6.2 Northern blot analysis of viral RNAs of CMV and TAV. (a) Positions of the positive-strand-specific RNA probes used for northern blot hybridisation. Probe C is complementary to nt 1702 to 2197 of RNA 3 of Q-CMV, while probe T is complementary to nt 2287 to 2386 of RNA 3 of V-TAV, respectively. (b) Northern blot analysis of viral RNAs of CMV and TAV. Total RNAs (5μg) extracted from N. glutinosa inoculated with sterile water (lanes 1 and 6), Q-CMV (lane 2), WAII-CMV (lane 3) or V-TAV (lane 4), and V-TAV virion RNAs (0.2μg, lane 5) were hybridised with probe C (lanes 1-3) or probe T (lanes 4-6). The positions of viral RNAs 1, 2, 3, 4, 4A, 3B and 5 are indicated. RNA 3B was detected only in virion RNAs of V-TAV or total RNAs extracted from V-TAV-infected plants.

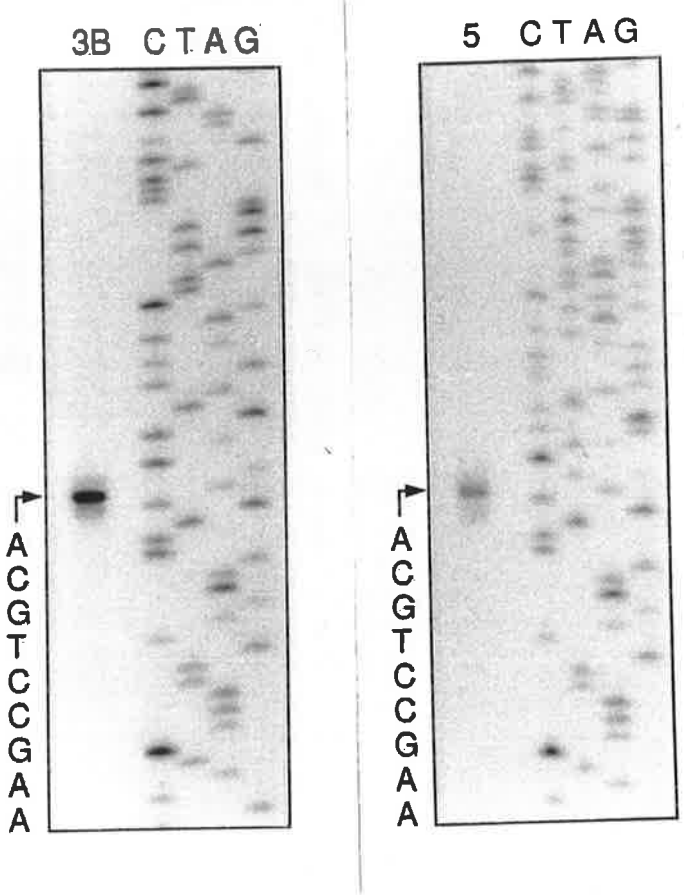


Fig. 6.3 Primer extension analysis of purified RNA 3B and RNA 5. Reverse transcriptase-catalysed extension with the oligonucleotide primer complementary to the 5'-terminal region of either RNA 3B or RNA 5 was in the absence of ddNTPs as described in the text. The arrows indicate the unique primer extension product with the starting position of RNA 3B (left panel, lane 3B) or RNA 5 (right panel, lane 5). A dideoxynucleotide DNA sequencing ladder generated by the same primer using the full-length cDNA of RNA 3 as template is shown in lanes C, T, A and G, respectively. The 5'-end sequences of RNA 3B and RNA 5 are respectively given to the left of each panel. The first A residue in each sequence corresponds to the initiation site of each RNA.

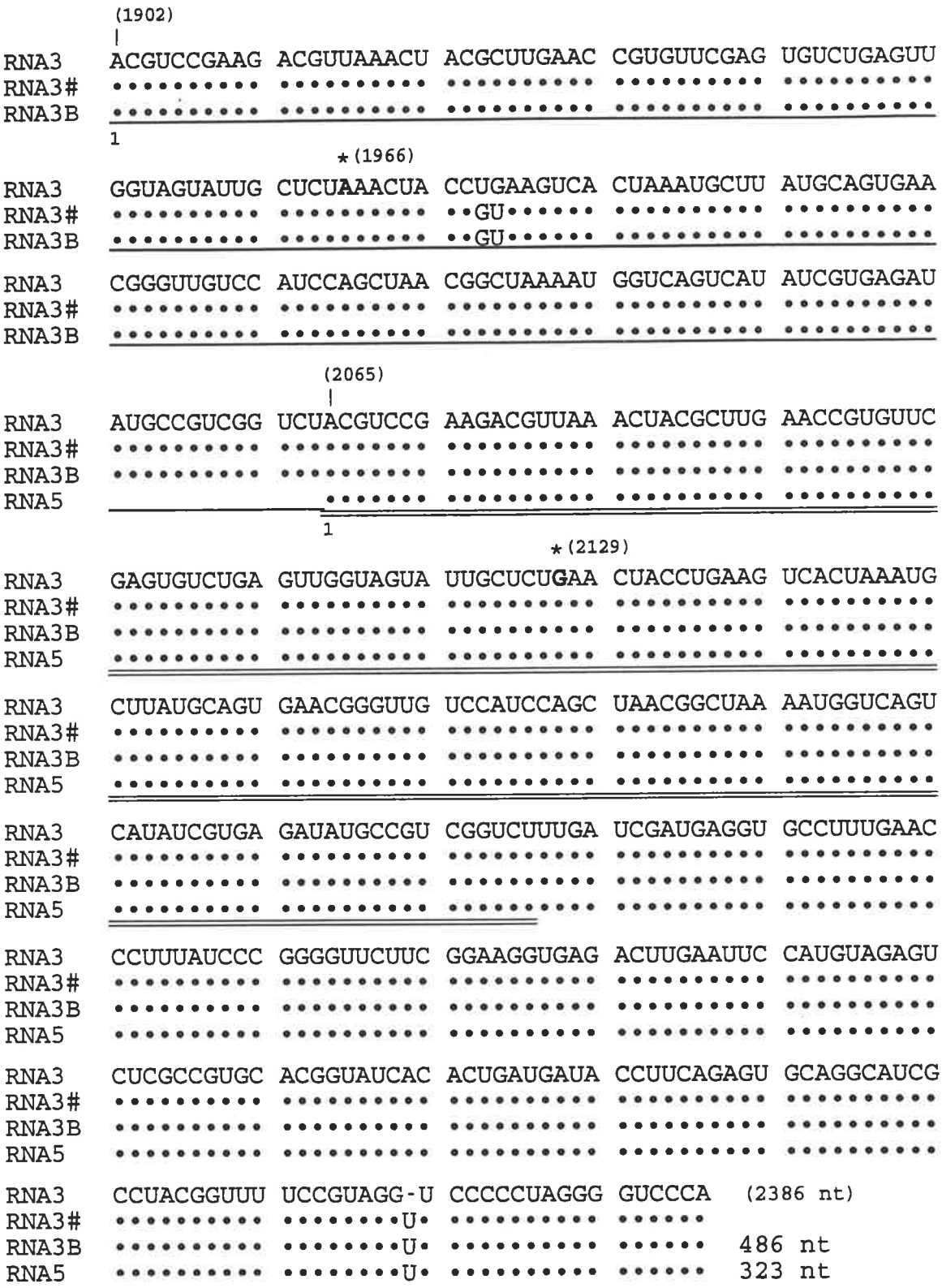


Fig. 6.4 Complete nucleotide sequences of RNA 3B and RNA 5 and their alignment with the 3'-terminal sequence of V-TAV RNA 3 (kindly provided by F. Garcia-Arenal) or the sequence deduced from our own cDNA clones (RNA 3#). The differences between the two RNA 3 sequences are indicated. Residue numbers in parentheses refer to the equivalent position in V-TAV RNA 3. Positions of identity are indicated by dots. A dash represents one missing nucleotide. The repeat sequences at the 5' end of either RNA 3B or RNA 5 are underlined (first repeat with single underline and second repeat with double underlines). The two single nucleotide differences between the two repeats are indicated by an asterisk above. The lengths of RNA 3, RNA 3B and RNA 5 are given at the 3'-end, The full length sequence of RNA 3# has yet to be determined.

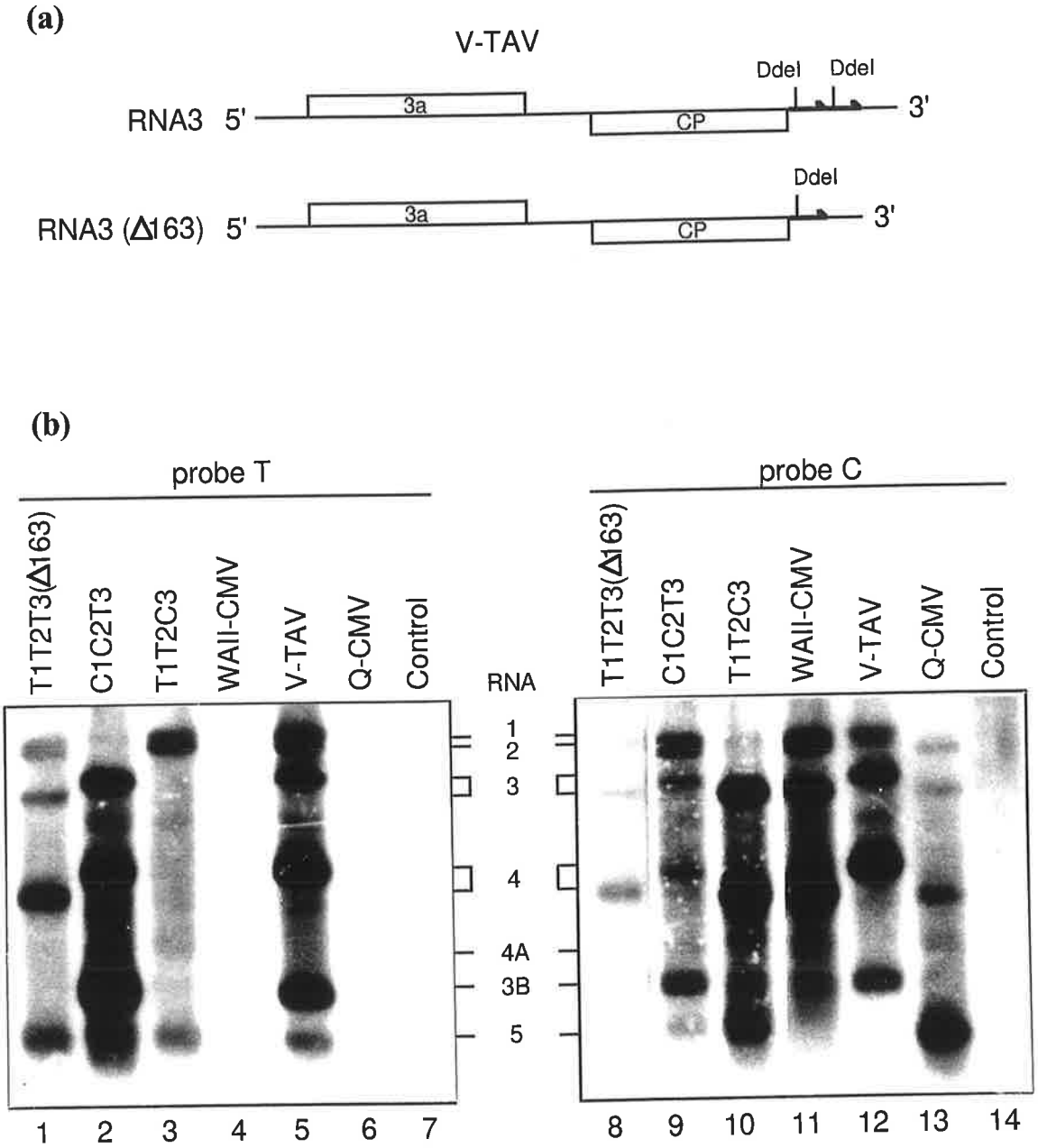


Fig. 6.5 Northern blot analysis of total RNAs extracted from *N. glutinosa* inoculated with different inocula. Total RNAs extracted from *N. glutinosa* inoculated with a TAV mutant [T1T2T3(Δ163)] (lanes 1 and 8), C1C2T3 (lanes 2 and 9), T1T2C3 (lanes 3 and 10), WAII-CMV (lanes 4 and 11), V-TAV (lanes 5 and 12), Q-CMV (lanes 6 and 13) or sterile water (lanes 7 and 14) were electrophoresed and hybridised with probe T (lanes 1-7) or probe C (lanes 8-14). Probes T and C are shown in Fig. 6.2a. The positions of viral RNAs 1, 2, 3, 4, 4A, 3B and 5 are indicated.

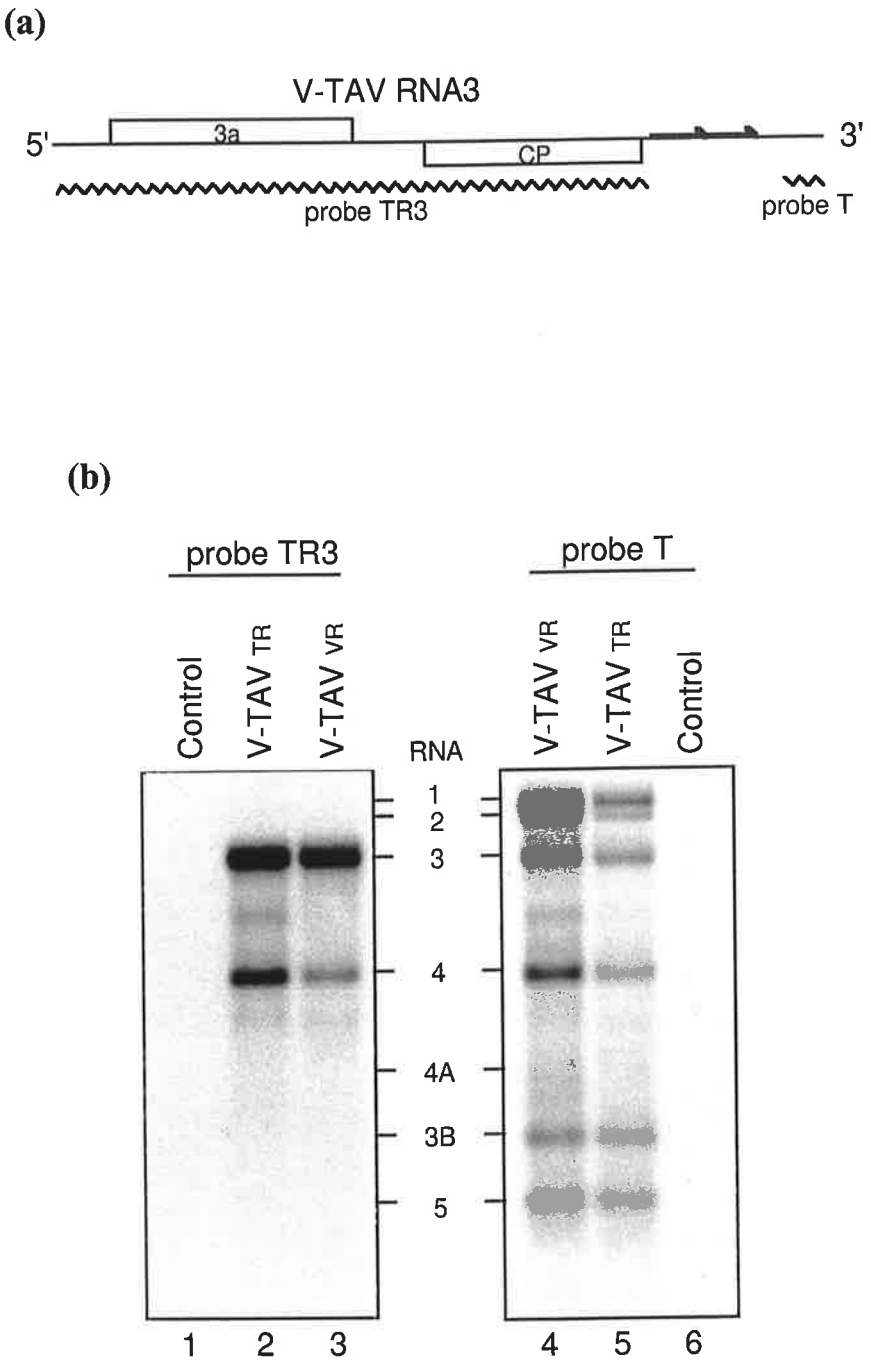


Fig. 6.6 Northern blot analysis of viral RNAs of TAV. Total RNAs (5μg) extracted from healthy *N. glutinosa* (lanes 1 and 6) or V-TAV infected *N. glutinosa* (lanes 2 and 5) and V-TAV virion RNAs (0.2μg, lanes 3 and 4) were electrophoresed and hybridised with probe TR3 (lanes 1-3) or probe T (lanes 4-6). Probe TR3 is complementary to nt 1 to 1903 of V-TAV RNA 3 and probe T is as used in Fig. 6.2a and both probes are shown in (a). The positions of viral RNAs 1, 2, 3, 4, 4A, 3B and 5 are indicated.

Chapter 7

GENERAL DISCUSSION

7.1 GENERATION OF CUCUMOVIRUS SUBGENOMIC RNAS

Three new subgenomic RNAs (RNAs 4A, 3B and 5) have been characterised in this study. V-TAV RNA 4A is 702 nucleotides long while WAII-CMV RNA 4A is 690 nucleotides in length. Both RNA 4As are derived from the 3'-terminal of RNA 2 and encode the ORF 2b, which, in RNA 2, overlaps the C-terminal portion of the major 2a gene, and have been suggested to be transcribed from the (-) RNA 2 sequence (Ding *et al.*, 1994). RNAs 3B and 5 are derived from the 3'-terminal 486 and 323 nucleotides of V-TAV RNA 3, respectively, and are not the result of an endonucleolytic cleavage of RNA 3 at a specific site. Hence, RNAs 3B and 5 may be also transcribed from the (-) RNA 3 sequence (Miller *et al.*, 1985). .

Sequence analysis showed that a sequence UUG(A/G)UCAA(C/U)(U/A)C immediately before the initiation site of RNA 4A in RNA 2 is highly conserved in all cucumoviruses to date. This sequence resembles the 'promoter-like' element on RNA 3 upstream to that of RNA 4 (Boccard and Baulcombe, 1993), suggesting that this conserved sequence may play a role in subgenomic RNA 4A transcription. However, this sequence does not exist on RNA 3 upstream of either RNA 3B or RNA 5. Also, no common RNA secondary structure was found upstream of RNAs 4, 4A, 3B and 5 (Blanchard *et al.*, 1996). These considerations indicate that the promoters for transcribing subgenomic RNAs 3B and 5 may be different from those for subgenomic RNAs 4 and 4A. As RNAs 3B and 5 have the same initiation sequence, they may have the same promoter-like sequences, which could be situated downstream, and within, each repeat of RNAs 3B and 5 on RNA 3 because the sequence upstream of RNA 3B on RNA 3 is different from that upstream of RNA 5 on RNA 3.

7.2 ENCAPSIDATION OF CUCUMOVIRUS SUBGENOMIC RNAS

In this study, the subgenomic RNAs, RNAs 4A, 3B and 5, were found to be

encapsidated in varying amounts. WAII-CMV RNA 4A was hardly encapsidated in the viral particles, whereas Q-CMV RNA 4A was encapsidated abundantly as compared to RNAs 1, 2, 3, and 4. The level of V-TAV RNA 4A encapsidated was between that of WAII-CMV RNA 4A and Q-CMV RNA 4A. Similar amounts of RNAs 3B and 5 were encapsidated to those of RNAs 1, 2, 3, and 4. These subgenomic RNAs encapsidated in varying amounts may reflect a different interaction between the region of the coat protein involved in subgenomic RNA binding and the region of the subgenomic RNA specifically recognised by the coat protein in individual cucumoviruses.

In the experiments with pseudorecombinants in which RNA 3 was exchanged between Q-CMV and V-TAV, both Q-CMV and V-TAV coat proteins could encapsidate heterologous genomic RNAs and subgenomic RNAs that were produced during the process of genomic RNA replication *in vivo* (Shi *et al.*, unpublished results). Particularly, Q-CMV coat protein has been shown to encapsidate RNA 6 which is a mixture of plant RNAs contaminated with fragments of the CMV genomic RNAs (Palukaitis *et al.*, 1992). These results indicate that cucumovirus coat proteins interact with RNAs in a less specific manner not dependent on nucleotide sequence, and thereby cucumoviral RNA encapsidations may mainly depend on cis-acting elements.

The postulated cis-acting elements could comprise RNA conformation or the nucleic acid encapsidation signals. They can be located at the 5'-end (Gottlieb *et al.*, 1994; Kawamoto *et al.*, 1994), at the 3'-end (Fujimura *et al.*, 1990), at both 5'- and 3'-ends of the RNA molecules (Zou and Brown, 1992) or in the interior (Turner *et al.*, 1988). In WAII-CMV, RNA 4A was hardly encapsidated in the viral particles, indicating that WAII-CMV RNA 4A may lack the cis-acting elements, or in the other words, the cis-acting elements may be located in RNA 2 beyond the RNA 4A sequence. In Q-CMV, RNA 4A was encapsidated abundantly, indicating that RNA 4A contains the postulated cis-acting elements. A chimeric Q-CMV RNA 4A in which the ORF 2b coding region was replaced

with the homologous ORF 2b of WAII-CMV (see Chapter 4) was also encapsidated in the coat protein of Q-CMV (data not shown), and the level of the chimeric RNA 4A encapsidated was similar to that of the wildtype RNA 4A of Q-CMV encapsidated, indicating that the cis elements required for Q-CMV RNA 4A encapsidation are not located within the ORF 2b coding sequence. Hence, they may be located in either the 3'- or 5'-untranslated regions or both regions.

Sequence analyses showed that the 5'- and 3'-terminal noncoding regions of Q- and WAII-CMV RNA 4As have 56.1% and 65.9% identity, respectively. Furthermore, the 3'-terminal noncoding region of both RNAs can form very similar secondary structures, whereas the 5'-terminal noncoding region of both RNAs can not form very similar stem-loop structures which were found to be necessary for RNA encapsidation in many other viruses (Gottlieb *et al.*, 1994; Kawamoto *et al.*, 1994; Fujimura *et al.*, 1990; Zou and Brown, 1992; Turner *et al.*, 1988; Mansky *et al.*, 1995). Hence, the specific RNA sequence at the 5'-terminal noncoding region of Q-CMV RNA 4A may have a primary encapsidation signal.

The level of V-TAV RNA 4A encapsidated was between that of WAII-CMV RNA 4A and Q-CMV RNA 4A, suggesting that V-TAV RNA 4A may have a weak signal for encapsidation, or the cis elements for V-TAV RNA 4A encapsidation may consist of more than two separate regions (Mansky *et al.*, 1995), one of which may be located within the V-TAV RNA 4A and the other may be located beyond RNA 4A sequence on RNA 2. It is also possible that the cis elements for V-TAV RNA 4A encapsidation, like the duck virus (Hirsch *et al.*, 1991), may cover long sequences outside these of RNA 4A.

The differences in the relative levels of RNA 4As to other viral RNAs may also reflect an inherent difference in the levels of RNA transcription. This is possibly true in the case of V-TAV, which produced only small amounts of RNA 4A *in vivo*, and the lower level of

RNA 4A present in the virions could be due the result of less RNA 4A molecules available for encapsidation. This may eventually reflect the effectiveness of the respective subgenomic RNA promoters. Analysis of the putative promoter-like sequence on RNA 2 revealed that two nucleotide variations, positions 2352 and 2353 in Q-CMV RNA 2, are specifically associated with Q-CMV which had the highest level of RNA 4A, whereas other cucumoviruses which had lower levels of RNA 4A had no nucleotide variations within the putative promoter-like sequence on RNA 2. However, computer-aided analysis revealed that these two nucleotide variations, positions 2352 and 2353 in Q-CMV RNA 2, did not change any original putative secondany structure which is probably required for RNA encapsidation. Hence, whether these two nucleotide variations affect the transcriptional levels of the RNA 4As is uncertain. In addition, northern blot data in this study does not rule out differential stability of the subgenomic RNA 4As.

7.3 FUNCTIONS OF CUCUMOVIRUS SUBGENOMIC RNAS AND THEIR ENCODED PROTEINS

In RNA viruses, since translation generally cannot be initiated from an internal site in an mRNA strand, polyproteins within the viral genome are translated through subgenomic RNA(s). RNA 4As of V-TAV and WAII-CMV encoding ORF2b of 95 codons and 110 codons, respectively, which are subgenomic RNAs derived from the respective RNA 2, are of a size appropriate for subgenomic RNAs synthesised to place internal ORFs in a 5' proximal position for translation (Singer and Berg, 1991). The 5'-end of each RNA 4A is most likely capped, which is believed to enhance translation by promoting the formation of the translation initiation complex. Computer-aided analysis predicts a tRNA-like structure present at the 3' non-coding region of RNA 4As of V-TAV and WAII-CMV as well as Q-CMV (data not shown), which is essential for biological activity. In addition, RNA 2 of Q-CMV translated into only one protein *in vitro* (Schwinghamer and

Symons, 1977; Gordon *et al.*, 1982). Furthermore, one of the subgenomic RNAs from Q-CMV, which was encapsidated in viral particles and homologous in sequence and corresponded in size to RNA 4A, was shown to be a strong messenger RNA *in vitro* (Gordon and Symons, 1985). All these indicate that RNA 4A may function as a mRNA for expressing the 2b gene encoded by RNA 2.

The 2b gene product has been detected in V-TAV-infected plants. This, together with the result that the 2b gene product has been detected in Q-CMV-infected plants as well (Ding *et al.*, 1994), clearly indicate that the 2b gene is a common feature of the genus *Cucumovirus*. By using infectious full-length cDNA clones of Q-CMV by which specific mutations were introduced into the ORF 2b, the 2b gene was demonstrated to be essential for the virus virulence and for long distance virus movement. V-TAV systemically infected tobacco plants and induced obvious symptoms on these plants. A hybrid virus (CMV-qt2; see Chapter 4) lacking the V-TAV 2b gene, however, infected these plants systemically with a much reduced virulence and delayed appearance of symptoms. In contrast, hybrid viruses (CMV-qt or CMV-qt1; see Chapter 4) without deletion of the V-TAV 2b gene induced the rapid appearance of severe systemic symptoms in these plants. This, combined with the fact that the V-TAV 2b protein was found in the infected plant cell walls, indicates that the 2b gene of V-TAV also contributes to long distance virus movement in these plants.

Previous studies have shown that a Q-CMV mutant lacking ORF 2b was not capable of systemic infection in cucumber (Ding *et al.*, 1995b), indicating a role of the 2b protein in long-distance virus movement. In this study, CMV-qt failed to systemically infect cucumber which is not a host for V-TAV, whereas CMV-qw (see Chapter 4) systemically infected cucumber as efficiently as did either Q-CMV or WAII-CMV, indicating that the 2b protein functions in a host-specific manner. Taken together, it is possible that the 2b protein forms a stable and specific transport-competent complex with

host factors necessary for long distance virus movement. This may be the reason why the interspecies hybrid virus in which the Q-CMV 2b gene has been replaced by the V-TAV 2b gene and one of the Q-CMV mutants in which the Q-CMV 2b gene has been deleted (Ding *et al.*, 1995b) failed to systemically infect cucumber. This is the case for tomato bushy stunt tombusvirus whose p19 protein was required for long distance virus movement in spinach and pepper but not in *Nicotiana benthamiana* plants (Scholthof *et al.*, 1995).

RNAs 3B and 5 were consistently detected in virion RNAs purified from all the V-TAV-infected host species used in this study. RNA 3B appears to be specifically associated with V-TAV, whereas RNA 5 appears to be associated not only with V-TAV but also with the other strains of TAV (Moriones *et al.*, 1992). Sequence analysis showed that the 3'-terminal 323 nt of RNA 3B was identical to RNA 5, whereas its 5'-terminal 163 nt was a direct repeat (one nt difference) of the 5'-half of RNA 5 and that both RNAs are derived from V-TAV RNA 3. Interestingly, the two tandem repeats were not found in the other strains of TAV sequenced to date. RNAs 3B and 5 only contain small open reading frames less than 21 codons. Hence, it is unlikely that RNAs 3B and 5 could express any viral gene product. The lack of 5'-capping of RNAs 3B and 5 further supports a non-mRNA role for these RNAs.

Precise deletion of one of the repeats changed, as expected, RNA 3B into RNA 5 *in vivo* and resulted in increased amounts of RNA 5 relative to those of wild-type V-TAV. However, these changes did not apparently affect the symptoms and host range compared to wild-type V-TAV (data not shown), indicating that RNA 3B may not be essential for virus replication and movement and may not be of major biological significance either and that RNAs 3B and 5 may not interact with each other. This, combined with the fact that RNAs 3B and 5 were transcribed in large amounts in the infected cell, leads us to speculate that these two RNAs may have a regulatory role in

viral RNA replication. Of course, it is also possible that the presence of RNAs 3B and 5 may provide some selective advantage under specific host and/or environmental conditions. To reveal the properties of RNAs 3B and 5, direct comparative differences in biological aspects such as symptoms, transmission and host range between V-TAV and the other strains which lack RNA 3B by using their full-length infectious clones would be required.

7.4 INFECTIOUS cDNA CLONES OF V-TAV

Infectious cDNA clones of RNA viruses can be divided into two groups. On one hand, infection results from inoculation with plasmids carrying viral cDNA. On the other hand, infection requires inoculation with *in vitro*-transcribed RNA copies. The latter make up the majority of infectious cDNA clones of RNA viruses prepared to date. This is despite the fact that the first infectious cDNA clones of RNA viruses were the former. Although a number of infectious cDNA clones of RNA viruses have been obtained (Boyer and Haenni, 1994), problems are also often associated with some of these infectious cDNA clones. An example is the high instability of full-length cDNA clones in bacteria. In the cases of yellow fever flavivirus (Rice *et al.*, 1989) and Japanese encephalitis flavivirus (Sumiyoshi *et al.*, 1992), full-length cDNA clones were successfully obtained. However, multiplication of these clones in bacteria always resulted in the introduction of mutations in the viral sequence. This may be the reason why no infectious cDNA clones have been obtained for some important viruses. Similar problems have been encountered for many other viruses (Boyer and Haenni, 1994; references therein). Another problem found is the tendency not to show the normal host range (Weber *et al.*, 1992) or symptoms (MacFarlanes *et al.*, 1992) of the virus concerned. This is most strikingly shown by the L strain of tobacco mosaic tobamovirus (TMV) whose infectious cDNA clones failed to infect some normal hosts including *Lycopersicon esculentum* (Weber *et al.*, 1992).

However, the cDNA clones of TAV constructed in the modified pCass (pCass2; see Chapter 5) in this study have been stably maintained in *Escherichia coli* DH5α for more than two years and are highly infectious in plants. No differences in symptoms, morphology, host range and progeny RNA sizes and patterns as judged by northern blot hybridisation were seen between the infectious cDNA clones of V-TAV and the wildtype V-TAV. Further advantages of the directly infectious cDNA clones of V-TAV include: 1) Infectivity is less dependent on RNA degradation since *in vitro* transcription is not required. 2) No nonviral nucleotides are present at the 5'-end of viral transcripts which may reduce infectivity. 3) The infectious cDNA clones of V-TAV are simple to use, time-saving and cost-effective as compared to other strategies for infection such as particle bombardment (Gal-On *et al.*, 1995b, Fakhfakh *et al.*, 1996). 4) The cDNA transcription in pCass2 is driven by a partially duplicated CaMV 35S promoter which would be expected to increase infectivity of the cDNA clones on plants according to Kay *et al.* (1987).

7.5 FUTURE WORK

This study characterised a small overlapping gene (2b) encoded by RNA 2 of cucumoviruses. This 2b protein is nonstructural and has been demonstrated to facilitate long-distance virus movement (Ding *et al.*, 1995b). This is the first example that a non-structural protein encoded by an RNA virus is essential for long-distance virus movement. However, it is not clear how the 2b protein facilitates long-distance virus movement in the infected plants. A previous report that a Q-CMV mutant lacking ORF 2b was restricted to the localised chlorotic lesions in the inoculated cucumber cotyledons (Ding *et al.*, 1995b) suggests that the 2b protein may function as a suppressor of the hypersensitive reaction (HR) induced by virus infection (Linthorst, 1991).

To examine such a possible role of the 2b protein, it is necessary to transform the 2b gene, if possible, together with some marker gene (eg., GUS gene) that would not affect

the expression of the 2b protein and be easily detected, into *N. tabacum* cv Samsun NN and then challenge the transformed plants with TMV to see whether the 2b protein can overcome the HR induced by TMV. This is because TMV is a necrotising pathogen to this tobacco plant. In addition, inoculation of cucumovirus 2b-deficient mutants (eg., CMV-qt2) or movement-defective unrelated viruses/mutants onto the 2b transgenic plant would determine whether the 2b protein can complement the rapid spread and severe symptoms of the cucumovirus mutants and complement the unrelated viruses/mutants other than the cucumovirus mutants in movement function.

A better understanding of the role of the 2b protein in long-distance virus movement is to explore the interaction between the 2b protein and plasmodesmata through which plant viruses move from cell to cell (Gibbs, 1976; Hull, 1989; Robards and Lucas, 1990; Maule, 1991). This can be achieved by ultrastructural studies and microinjection studies performed on the 2b transgenic plant, or on virus-infected plants, using electron microscopy, immunodetection and *in situ* localisation techniques.

Identification of putative host factors (proteins) that may form the complex required for long-distance movement would provide a foundation upon which to dissect the molecular interaction between the viral-encoding protein and host factors. This would involve localisation and isolation of the complex.

Further investigation for the role of the cucumovirus 2b protein is to construct V-TAV mutants, similar to the constructs of Q-CMV mutants used in this study, on the basis of infectious cDNA clones of V-TAV constructed in this study. Construction of infectious cDNA clones of WAII-CMV and of mutants based on these cDNA clones would also allow a better understanding of the role of the cucumovirus 2b protein.

Other work relevant to this thesis is to identify cis-acting elements required for RNA encapsidation. As coat proteins of both CMV and TAV were shown to encapsidate

heterologous RNAs in a non-specific manner *in vitro* (Chen and Francki, 1990), the identification of the cis-acting elements for cucumoviral RNA encapsidations must be carried out *in vivo*. This may be difficult given the requirements for identification without infection of two CMV mutants in plants. These two mutants should contain a WAII-CMV mutant in which WAII-CMV RNA 4A sequence should be replaced by its homologous sequence from Q-CMV and a Q-CMV mutant in which Q-CMV RNA 4A sequence should be replaced by its homologous sequence from WAII-CMV. If plant infection with these two CMV mutants is possible, then stepwise deletions and or replacements between two RNA 4A sequence regions could be made. These directed mutagenesis experiments would allow the discovery of the sequence responsible for the encapsidation of RNA 4A in viral particles.

Finally, the promoter for the transcription of subgenomic RNAs (RNAs 4A, 3B and 5) characterised in this study could also be investigated. The 5'-terminal sequence of RNA 4A on RNA 2 is located at the 3'-portion of the 2a gene. Deletion of the sequence upstream, or of the 5'-terminal nucleotides, of RNA 4A on RNA 2 could affect viral RNA replication. Hence, it would be better to map the promoter of RNA 4A in *vitro* rather than *in vivo*. This will require a full-length cDNA clone of cucumovirus RNA 2 by which minus-stranded RNA 2 can be transcribed *in vitro* and thereby deletions and or substitutions would be made on the minus-stranded RNA 2 using the RNA-dependent RNA polymerase complex extracted from cucumovirus infected plants.

Mapping the promoter for RNAs 3B and 5 on RNA 3 could be carried out *in vivo* as RNA 3 has no trans-acting role in viral RNA replication. Deletions and insertions on a RNA 3 cDNA clone would be necessary. In addition, successfully abolishing production of RNAs 3B and 5 will help to determine the role of these non-messenger subgenomic RNAs and their repeat sequences.

Additional experiments including substitutions of two different nucleotide in each of the

two repeats at the 5'-terminal of RNA 3B could be done as well. Such substitutions would help to illustrate the mechanism of generation of the two repeats and eventually find the relationships among the three genomic RNA of cucumoviruses.

REFERENCES

Ahlquist, P., French, R. Janda, M., Loesch-Fries, L.S. (1984). Multicomponent RNA plant virus infection derived from cloned viral cDNA. *Proc. Natl. Acad. Sci. USA.* **81**, 7066-7070.

Ahlquist, P., Strauss, E., Rice, C.M., Strauss, J. and Zimmern, D. (1985). Sindbis virus proteins nsP1 and nsP2 contain homology to nonstructural proteins from several RNA plant viruses. *Journal of Virology* **53**, 536-542.

Albrecht, H., Geldreich, A., De Murcia, J.M., Kirchherr, D., Mesnard, J. and Lebeurier, G. (1988). Cauliflower mosaic virus gene 1 product detected in a cell wall-enriched fraction. *Virology* **163**, 503-508.

Allison, R.F., Janda, M. and Ahlquist, P. (1989). Sequence of cowpea chlorotic mottle virus RNAs 2 and 3 and evidence of a recombination event during bromovirus evolution. *Virology* **172**, 321-330.

Anderson, B.J., Boyce, P.M. and Blanchard, C.L. (1995). RNA 4 sequences from cucumber mosaic virus subgroups I and II. *Gene* **161**, 293-294.

Ansardi, D.C., Porter, D.C., Anderson, M.J. and Morrow C.D. (1996). Poliovirus assembly and encapsidation of genomic RNA. *Advances in Virus Research* **46**, 1-68.

Aryan, A.P., An, G. and Okita, T.W. (1991). Structural and functional analysis of promoter from gliadin, an endosperm-specific storage protein gene of *Triticum aestivum* L.. *Mol. Gen. Genet.* **225**, 65-71.

Balmori, E., Gilmer, D., Richards, K., Guilley, H. and Jonard, G. (1993). Mapping the promoter for subgenomic RNA synthesis on beet necrotic yellow vein virus RNA 3. *Biochimie* **75**, 517-521.

Banerjee, N., Wang, J.Y. and Zaitlin, M. (1995). A single nucleotide change in the coat protein gene of tobacco mosaic virus is involved in the induction of severe chlorosis. *Virology* **207**, 234-239.

Bernal, J.J., Moriones, E. and Garcia-Arenal, F. (1991). Evolutionary relationships in the cucumoviruses: nucleotide sequence of tomato aspermy virus RNA 1. *Journal of General Virology* **72**, 2191-2195.

Blanchard, C.L., Boyce, P.M. and Anderson, B.J. (1996). Cucumber mosaic virus

RNA 5 is a mixed population derived from the conserved 3'-terminal regions of genomic RNAs 2 and 3. *Virology* **217**, 598-601.

Blencowe, J.W. and Caldwell, J. (1949). Aspermy-a new virus disease of the tomato. *Ann. Appl. Biol.* **36**, 320-326.

Boccard, F., and Baulcombe, D. (1993). Mutational analysis of cis-acting sequences and gene function in RNA 3 of cucumber mosaic virus. *Virology* **193**, 563-578.

Bouzoubaa, S., Niesbach-Klosgen, U., Jupin, I., Guilley, H., RichardS, K. and Jonard, G. (1991). Shortened forms of beet necrotic yellow vein virus RNA-3 and RNA-4: Internal deletions and a subgenomic RNA. *Journal of General Virology* **72**, 259-266.

Boyer, J.C. and Haenni, A.L. (1994). Infectious transcripts and cDNA clones of RNA viruses. *Virology* **198**, 415-426.

Brault, V., van den Heuvel, J.F.J.M., Verbeek, M., Ziegler-Graff, V., Reutenauer, A., Herrbach, E., Garaud, J.C., Guilley, H., Richards, K. and Jonard, G. (1995). Aphid transmission of beet western yellows luteovirus requires the minor capsid read-through protein P74. *The EMBO Journal* **14**, 650-659.

Brooks, M. and Bruening, G. (1995). A subgenomic RNA associated with cherry leafroll virus infections. *Virology* **211**, 33-41.

Burgyan, J., Tavazza, M., Dalmay, T., Lucioli, A. and Balazs, E. (1993). Consequences of gene transfer between distantly related tombusviruses. *Gene* **129**, 191-196.

Burman, A.J.,Osman, T.A.M. and Buck, K.W. (1994). Detection of the 3a protein of cucumber mosaic virus in a cell wall fraction from infected *Nicotiana clevelandii* plants. *Journal of Phytopathology-Phytopathologische Zeitschrift* **142**, 317-323.

Candelier-Harvey, P. and Hull, R. (1993). Cucumber mosaic virus genome is encapsidated in alfalfa mosaic virus coat protein expressed in transgenic tobacco plants. *Transgenic Research* **2**, 277-285.

Carr, J.P., Gal-On, A., Palukaitis, P. and Zaltin, M. (1994). Replicase-mediated resistance to cucumber mosaic virus in transgenic plants involved suppression of both virus replication in the inoculated leaves and long-distance movement. *Virology* **199**,

439-447.

Chen, B. and Francki, R.I.B. (1990). Cucumovirus transmission by the aphid *Myzus persicae* is determined solely by the viral coat protein. *Journal of General Virology* **71**, 939-944.

Chen, J., MacFarlane, S.A. and Wilson, T.M.A. (1994). Detection and sequence analysis of a spontaneous deletion mutant of soil-borne wheat mosaic virus RNA 2 associated with increased severity. *Virology* **202**, 921-929.

Cooper, B., Lapidot, M., Heick, J.A., Dodds, J.A. and Beachy, R.N. (1995). A defective movement protein of TMV in transgenic plants confers resistance to multiple viruses whereas the functional analog increases susceptibility. *Virology* **206,** 307-313.

Cooper, B., Schmitz, I., Rao, A.L.N., Beachy, R.N. and Dodds, J.A. (1996). Cell-to-cell transport of movement-defective cucumber mosaic and tobacco mosaic viruses in transgenic plants expressing heterologous movement protein genes. *Virology* **216**, 208-213.

Cornelissen, B.J.C., Linthorst, H.J.M., Brederode, F.T.H. and Bol, J.F. (1986). Analysis of the genome structure of tobacco rattle virus strain PSG. *Nucleic Acids Research* **14**, 2157-2169.

Davies, C. and Symons, R. H. (1988). Further implications for the evolutionary relationships between tripartite plant viruses based on cucumber mosaic virus RNA 3. *Virology* **165**, 216-224.

Dawson, W.O. and Hilf, M.E. (1992). Host-range determinants of plant viruses. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **43,** 527-555.

Dawson, W.O., Bubrick, P. and Grantham, G.L. (1988). Modifications of the tobacco mosaic virus coat protein gene affecting replication, movement, and symptomatology. *Phytopathology* **78**, 783-789

De Jong, W, Chu, A. and Ahlquist, P. (1995). Coding changes in the 3a cell-to-cell movement gene can extend the host range of brome mosaic virus systemic infection. *Virology* **214**, 464-474.

De Jong, W. and Ahlquist, P. (1992). A hybrid plant RNA virus made by transferring

the noncapsid movement protein from a rod-shaped to an icosahedral virus is competent for systemic infection. *Proc. Natl. Acad. Sci. USA* **89**, 6808-6812.

Deng, W.L., Grayburn, W.S., Hamilton-Kemp, T.R., Collins, G.B. and Hildebrand, D.F. (1992). Expression of soybean-embryo lipoxygenase 2 in transgenic tobacco tissue. *Planta* **187**, 203-208.

Deom, C.M., Oliver, M.J. and Beachy, R.N. (1987). The 30 KDa gene product of tobacco mosaic virus potentiates virus movement. *Science* **237**, 389-394.

Devergne, J.C. and Cardin, L. (1973). Contribution a l'étude du virus de la mosaique du concombre (CMV). IV. Essai de classification de plusieurs isolats sur la base de leur structure antigénique. *Ann. Phyptopathologie* **5**, 409-430.

Devergne, J.C. and Cardin, L. (1975). Relations sérologiques entre cucumovirus (CMV, TAV, PSV). *Ann. Phyptopathologie* **7**, 255-276.

Devergne, J.C., Cardin, L., Burckard, J. and van Regenmortel, M.H.V. (1981). Comparison of direct and indirect ELISA for detecting antigenically related cucumoviruses. *Journal of Virological Methods* **3**, 193-200.

Devic, M., Jaegle, M. and Baulcombe, D. (1989). Symptom production on tobacco and tomato is determined by two distinct domains of the satellite RNA of cucumber mosaic virus (strain Y). *Journal of General Virology* **70**, 2765-2774.

Devic, M., Jaegle, M. and Baulcombe, D. (1990). Cucumber mosaic virus satellite RNA (strain Y): Analysis of sequences which affect systemic necrosis in tomato. *Journal of General Virology* **71**, 1443-1449.

Ding, B.A., Li, Q.B., Nguyen, L., Palukaitis, P. and Lucas, W.J. (1995). Cucumber mosaic virus 3a protein potentiate cell-to-cell trafficking of CMV RNA in tobacco plants. *Virology* **207**, 345-353.

Ding, S.W. (1989). Molecular biology and evolution of tymoviruses. *PhD Thesis*. Australian National University, Australia.

Ding, S.W., Anderson, B.J., Haase,H.R. and Symons R.H. (1994). New overlapping gene encoded by the cucumber mosaic virus genome. *Virology* **198**, 593-601.

Ding, S.W., Keese, P. and Gibbs, A. (1989). Nucleotide sequence of the ononis yellow

mosaic tymovirus genome. *Virology* **172**, 555-563.

Ding, S.W., Li, W.X. and Symons, R.H. (1995b). A novel naturally occuring hybrid gene encoded by a plant RNA virus facilitates long distance virus movement. *The EMBO Journal* **14**, 5762-5772.

Ding, S.W., Rathjen, J.P., Li, W.X., Swanson, R., Healy, H. and Symons, R.H. (1995a). Efficient infection from cDNA clones of cucumber mosaic cucumovirus RNAs in a new plasmid vector. *Journal of General Virology* **76**, 459-464.

Doolittle, S.P. (1916). A new infectious mosaic disease of cucumber. *Phytopathology* **9**, 326-327.

Duggal, R., Lahser, F.C. and Hall, T.C. (1994). Cis-acting sequences in the replication of plant viruses with plus-sense RNA genomes. *Annu. Rev. Phytopathol.* **32**, 287-309.

Edwards, M.C. and Gonsalves, D. (1983). Grouping seven biological defined isolates of cucumber mosaic virus by peptide mapping. *Phytopathology* **73**, 1117-1120.

Edwards, M.C., Gonsalves, D. and Provvidenti, R. (1983). Genetic analysis of cucumber mosaic virus in relation to host resistance: Location of determinants for pathogenicity to certain legumes and *Lactuca saligna*. *Phytopathology* **73**, 269-273.

Edwardson, J.R. and Christie, R.G. (1991). CRC Handbook of Viruses Infecting Legumes. pp. 293-319. CRC Press, Boca Rato.

Ehara, Y. and Mink, G.I (1984). Very early events in the infection of cowpea leaf epidermal cells by cucumber mosaic virus. *Tohoku J. Agric. Res.* **34**, 31-46.

Fakhfakh, H., Vilaine, F., Makni, M. and Robaglia, C. (1996). Cell-free cloning and biolistic inoculation of an infectious cDNA of potato virus Y. *Journal of General Virology* **77**, 519-523.

Fichot, O. and Girard, M. (1990). An improved method for sequencing of RNA templates. *Nucleic Acids Research* **18**, 6162.

Finch, J.T., Klug, A. and Van Regenmortel, M.H.V. (1967). The structure of cucumber mosaic virus. *Journal of Molecular Biology* **24**, 303-305.

Foster, A.C., Davies, C., Hutchins, C.J. and Symons, R.H. (1990). Characterisation of self-cleavage of viroid and virusoid RNAs. *Methods in Enzymology* **181**, 583-607.

Fraile, A., Garcia-Arenal, F. and Palukaitis, P. (1993). In vitro interactions between cucumoviral satellite RNAs and RNAs of helper and non-helper viruses. *Physiological and Molecular Plant Pathology* **42**, 195-204.)

Francki, R.I.B. (1964). Inhibition of cucumber mosaic virus infectivity by leaf extracts. *Virology* **21**, 193-199.

Francki, R.I.B. (1985). The Plant Viruses. pp. 1-18. Plenum Press, New York.

Francki, R.I.B., Milne, R.G. and Hatta, T. (1985). An Atlas of Plant Viruses. pp. 53-68. CRC Press, Florida.

Francki, R.I.B., Mossop, D.W. and Hatta, T. (1979). CMI/AAB Description of Plant Viruses No 213. Commonwealth Agricultural Bureau, Kew, Surrey, England.

Francki, R.I.B., Randles, J.W., Chambers, T.C. and Wilson, S.B. (1966). Some properties of purified cucumber mosaic virus (Q strain). *Virology* **28**, 729-741.

French, R. and Ahlquist, P. (1987). Intercistronic as well as terminal sequences are required for efficient amplification of brome mosaic virus. *Journal of Virology* **61**, 1457-1465.

French, R. and Ahlquist, P. (1988). Characterisation and engineering of sequences controlling *in vivo* synthesis of brome mosaic virus subgenomic RNA. *Journal of Virology* **62**, 2411-2420.

French, R., Janda, M. and Ahlquist, P. (1986). Bacteria gene inserted in an engineered RNA virus: efficient expression in monocotyledonous plant cells. *Science* **231**, 1294-1297.

Frohman, M.A., Dush, M.K., and Martin, G.R. (1988). Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. USA* **85**, 8998-9002.

Fuentes, A.L. and Hamilton, R.I. (1991). Sunn-hemp mosaic virus facilitates cell-to-cell spread of southern bean mosaic virus in a nonpermissive host. *Phytopathology* **81**, 1302-1305.

Fujimura, T., Esteban, R., Esteban, L.M. and Wickner, R.B. (1990). Portable encapsidation signal of the L-A double-stranded RNA virus of *S. cerevisiae*. *Cell* **62**, 819-828.

Gal-On, A., Kaplan, I. and Palukaitis, P. (1995a). Differential effects of satellite RNA on the accumulation of cucumber mosaic virus RNAs and their encoded proteins in tobacco vs Zucchini squash with two strains of CMV helper virus. *Virology* **208**, 58-66.

Gal-On, A., Kaplan, I., Roossinck, M.J. and Palukaitis, P. (1994). The kinetics of infection of zucchini squash by cucumber mosaic virus indicate a function for RNA 1 in virus movement. *Virology* **205**, 280-289.

Gal-On, A., Meiri, E., Huet, H., Hua, W.J., Raccah, B. and Gaba, V. (1995b). Particle bombardment drastically increases the infectivity of cloned DNA of zucchini yellow mosaic potyvirus. *Journal of General Virology* **76**, 3223-3227.

Gargouri, R., Joshi, R.L., Bol, J.F., Astier-Manifacier, S. and Haenni, A.L. (1989). Mechanism of synthesis of turnip yellow mosaic virus coat protein subgenomic RNA *in vitro*. *Virology* **171**, 386-393.

Gibbs, A. (1976). Viruses and plasmodesmata. *In* "Intercellular Communication in Plants: Studies on Plasmodesmata." (B.E.S. Gunning and R.W. Robards, Eds.), pp. 149-164. Springer-Verlag Press, New York.

Godefroy-Colburn, T., Gagey, M.J., Berna, A.and Stussi-Garaud, C. (1986). A nonstructural protein of alfalfa mosaic virus in the walls of infected tobacco cells. *Journal of General Virology* **67**, 2233-2239.

Goelet, P. and Karn, J. (1982). Tobacco mosaic virus induces the synthesis of a family of coterminal messenger RNAs and their complements. *Journal of Molecular Biology* **154**, 541-550.

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

Gonda, T.J. and Symons, R.H. (1979). Cucumber mosaic virus replication in cowpea protoplasts: time course of virus, coat protein and RNA synthesis. *Journal of General Virology* **45**, 723-736.

Gordon, K.H.J. and Symons, R.H. (1985). Subgenomic RNA with nucleotide sequences derived from RNA 1 and RNA 2 of cucumber mosaic virus can act as messenger RNA *in vitro*. *Virology* **142**, 144-157.

Gordon, K.H.J., Gill, D.S. and Symons, R.H. (1982). Highly purified cucumber mosaic virus-induced RNA-dependent RNA polymerase does not contain any of the full length translation products of the genomic RNAs. *Virology* **123**, 284-295.

Gottlieb, P., Qiao, X., Strassman, J., Frilander, M. and Mindich, L. (1994). Identification of the packaging regions within the genomic RNA segments of bacteriophage f. *Virology* **200**, 42-47.

Gould, A. R., Palukaitis, P., Symons, R. H. and Mossop, D. M. (1978). Characterisation of satellite RNA associated with cucumber mosaic virus. *Virology* **84**,443-455.

Graves, M.V. and Roossinck, M.J. (1995). Characterisation of defective RNAs derived from RNA 3 of the Fny strain of cucumber mosaic cucumovirus. *Journal of Virology* **69**, 4746-4751.

Gubler, U. and Hoffman, B.J. (1983). A simple and very efficient method for generating cDNA libraries. *Gene* **25**, 263-269.

Haase, A., Richter, J. and Rabenstein, F. (1989). Monoclonal antibodies for detection and serotyping of cucumber mosaic virus. *Journal of Phytopathology* **127**, 129-136.

Habili, N. and Francki, R.I.B. (1974a). Comparative studies on tomato aspermy and cucumber mosaic viruses. I. Physical and chemical properties. *Virology* **57**, 392-401.

Habili, N. and Francki, R.I.B. (1974b). Comparative studies on tomato aspermy and cucumber mosaic viruses. III. Further studies on relationship and construction of a virus from parts of the two viral genomes. *Virology* **61**, 443-449.

Habili, N. and Francki, R.I.B. (1975). Comparative studies on tomato aspermy and cucumber mosaic viruses. IV. Immunogenic and serological properties. *Virology* **64**, 421-429.

Habili, N. and Symons, R.H. (1989). Evolutionary relationship between luteoviruses and other plant viruses based on the sequence motifs in their putative RNA polymerases and nucleic acid helicases. *Nucleic Acids Research* **17**, 9543-9555.

Hampton, R.O. and Francki, R.I.B. (1992). RNA-1 dependent seed transmissibility of cucumber mosaic virus in *Phaseolus vulgaris*. *Phytopathology* **82**, 127-130.

Hanada, K. and Tochihara, H. (1980). Genetic analysis of cucumber mosaic, peanut stunt and chrysanthemum mild mottle viruses. *Ann. Phytopath. Soc. Japan* **46**, 159-168.

Hancock, J.M., Chaleeprom, W, Dale, J. and Gibbs, A. (1995). Replication slippage in the evolution of potyviruses. *Journal of General Virology* **76**, 3229-3232.

Harrison, B.D, Finch, J.T., Gibbs, A.J., Hollings, M., Shephered, R.J., Valenta, V. and Wetter, C. (1971). Sixteen groups of plant viruses. *Virology* **45**, 356-363.

Harrison, S.C. (1984). Multiple modes of subunit association in the structures of simple spherical viruses. *Trends Biochem. Sci.* **9**, 345-351.

Haseloff, J. and Symons, R.H. (1981). Chrysanthemum stunt viroid: Primary sequence and secondary structure. *Nucleic Acids Research* **9**, 2741-2752.

Haseloff, J., Goelet, P., Zimmern, D., Ahlquist, P., Dasgupta, R. and Kaseberg, P. (1984). Striking similarities in amino acid sequence among nonstructural proteins encoded by RNA viruses that have dissimilar genomic organisation. *Proc. Natl. Acad. Sci. USA* **81**, 4358-4362.

Hassan, H., Borkhardt, B. and Albrechtsen, M. (1994). A surprising effect of extraction conditions on the mobility of some plant virus coat proteins in SDS-PAGE. *Journal of Virological Methods* **46**, 255-261.

Hatta, T. and Francki, R.I.B. (1979). Enzyme cytochemical method for identification of cucumber mosaic virus particles in infected cells. *Virology* **93**, 265-268.

Hayes, R.J. and Buck, K.W. (1990). Complete replication of a eukaryotic virus RNA *in vitro* by a purified RNA-dependent RNA polymerase. *Cell* **63**, 363-368.

Hayes, R.J., Pereira, V.C.A., Mcquillin, A. and Buck, K.W. (1994). Localisation of functional regions of the cucumber mosaic virus RNA replicase using monoclonal and polyclonal antibodies. *Journal of General Virology* **75**, 3177-3184.

Hellwald, K.H. and Palukaitis, P. (1994). Nucleotide sequence and infectivity of cucumber mosaic cucumovirus (strain K) RNA 2 involved in breakage of replicase-

mediated resistance in tobacco. *Journal of General Virology* **75**, 2121-2125.

Hilf, M.E. and Dawson, W.O. (1993). The tobamovirus capsid protein functions as a host-specific determinant of long-distance movement. *Virology* **193**, 106-114.

Hillman, B.I, Carrington, J.C. and Morris, T.J. (1987). A defective interfering RNA that contains a mosaic of a plant virus genome. *Cell* **51**, 427-433.

Hirsch, R.C., Loeb, D.D., Pollack, J.R. and Ganem, D. (1991). Cis-Acting sequences required for encapsidation of duck hepatitis B virus pregenomic RNA. *Journal of Virology* **65**, 3309-3316.

Holland, J.J. (1990). Defective viral genomes. *In* "Virology." (B.N. Fields and D.M. Knipe, Eds), pp. 151-165. Raven Press, New York.

Hollings, M. and Stone, O.M. (1971). C.M.I./A.A.B. Descriptions of Plant Viruses, No. 79. 4pp. Commonwealth Agricultural Bureau, Kew, Surrey, England.

Holt, C.A., Hodgson, R.A.J., Coker, F.A., Beachy, R.N. and Nelson, R.S. (1990) Characterisation of the masked strain of tobacco mosaic virus: Identification of the region responsible for symptom attenuation by analysis of an infectious complementary DNA clone. *Molecular Plant-Microbe Interactions* **3**, 417-423.

Hsu, Y.H., Wu, C.W., Lin, B.Y., Chen, H.Y., Lee, M.F. and Tsai, C.H. (1995). Complete genomic RNA sequences of cucumber mosaic virus strain NT9 from Taiwan. *Archives of Virology* **140**, 1841-1847.

Hull, R. (1976). The behaviour of salt-labile plant viruses in gradients of cesium sulphate. *Virology* **75**, 18-25.

Hull, R. (1989). The movement of viruses in plants. *Annu. Rev. Phytopathol.* **27**, 213-240.

Ishikawa, M., Meshi, T., Motoyoshi, F., Takamatsu, N. and Okada, Y. (1986). *In vitro* mutagenesis of the putative replicase genes of tobacco mosaic virus. *Nucleic Acids Research* **21**, 8291-8305.

Jacobs, E. and Clad, A. (1986). Electroelution of fixed and stained membrane proteins from preparative sodium dodecyl sulfate-polyacrylamide gels into a membrane trap. *Anal. Biochem.* **154**, 583-589.

Jagger, I.C. (1916). Experiments with the cucumber mosaic disease. *Phytopathology* **6**, 148-151.

Joshi, R.L. and Haenni, A.L. (1986). Search for tRNA-like properties in tomato aspermy virus RNA. *FEBS Letters* **194**, 157-160.

Kaper, J. M. and Tousignant, M. E. (1977). Cucumber mosaic virus-associated RNA 5: I. Role of host plant and helper strain in determining amount of associated RNA 5 with virions. *Virology* **80**, 186-195.

Kaper, J.M. (1973). Arrangement and identification of simple isometric viruses according to their dominating stabilising interactions. *Virology* **55**, 299-304.

Kaper, J.M. (1975). The Chemical Basis of Virus Structure, Dissociation and Reassembly. pp. 485. Elsevier/North-Holland Press, New York.

Kaper, J.M. (1976). Molecular organisation and stabilising forces of simple RNA viruses. V. The role lysyl residues in the stabilisation of cucumber mosaic virus strain S. *Virology* **71**, 185-198.

Kaper, J.M. and Geelen, J.L.M.C. (1971). Studies on the stabilising forces of simple RNA viruses. II Stability, dissociation and reassembly of cucumber mosaic virus. *Journal of Molecular Biology* **56**, 277-294.

Kaper, J.M. and Re, G.G. (1974). Redetermination of the RNA content and the limiting RNA size of three strains of cucumber mosaic virus. *Virology* **60**, 308-311.

Kaper, J.M. and Waterworth, H.E. (1977). Cucumber mosaic virus associated RNA 5: causal agent for tomato necrosis. *Science* **196**, 429-431.

Kaper, J.M. and Waterworth, H.E. (1981). Cucumoviruses. Handbook of Plant Virus Infections and Comparative Diagnosis (E. Kurstak, Ed.), Elsevier/North-Holland Press, New York.

Kaper, J.M., Dienner, T.O. and Scott, H.A. (1965). Some physical and chemical properties of cucumber mosaic virus (strain Y) and its isolated ribonucleic acid. *Virology* **27**, 54-72.

Kaper, J.M., Geletka, L.M., Wu, G.S. and Tousignant, M.E. (1995). Effect of temperature on cucumber mosaic virus satellite-induced lethal tomato necrosis is helper

virus dependent. *Archives of Virology* **140**, 65-74.

Kaplan, I.B., Shintaku, M.H., Li, Q.B., Zhang, L., Marsh, L.E. and Palukaitis, P. (1995). Complementation of virus movement in transgenic tobacco expressing the cucumber mosaic virus 3a gene. *Virology* **209**, 188-199.

Kararawa, A., Nakaho, K., Kakutani, T., Minobe, Y. and Ehara, Y.(1991). Nucleotide sequence of RNA 3 of peanut stunt cucumovirus. *Virology* **185**, 464-467.

Karasawa, A., Nakaho, K., Kakutani, T., Minobe, Y. and Ehara, Y. (1992). Nucleotide sequence analyses of peanut stunt cucumovirus RNAs 1 and 2. *Journal of General Virology* **73**, 701-707.

Kassanis, B. (1963). Interactions of viruses in plants. *Advances in Virus Research* **10**, 219-255.

Kataoka, J., Masuta C., and Takanami, Y. (1990a). Complete nucleotide sequence of RNA 2 of cucumber mosaic virus Y strain. *Ann. Phytopath. Soc. Japan* **56**, 495-500.

Kataoka, J., Masuta, C. and Takanami, Y. (1990b). Complete nucleotide sequence of RNA 1 of cucumber mosaic virus Y strain and evolutionary relationships among genome RNAs of the virus strains. *Ann. Phytopath. Soc. Japan* **56**, 501-507.

Kawamoto, S., Ueda, K., Mita, E. and Matsubara, K. (1994). The packaging signal in hepatitis B virus pregenome functions only at the 5' end. *Journal of Virological Methods* **49**, 113-128.

Kay, R., Chan, A., Daly, M. and McPherson, J. (1987). Duplication of CaMV 35S promoter sequences creates a strong enhancer for plant genes. *Science* **236**, 1299-1302.

Keese, P. and Symons, R.H. (1985). Domains in viroids: evidence of intermolecular RNA rearrangements and their contribution to viroid evolution. *Proc. Natl. Acad. Sci. USA.* 82, 4582-4586.

Keese, P. and Symons, R.H. (1987). The structure of viroids and virusoids. *In* "Viroids and Viroid-like Pathogens." (J.S. Semancik, Ed.), pp. 1-47. CRC Press, Florida.

Keese, P., Visvader, J.E. and Symons, R.H. (1988). Sequence variability in plant viroid RNAs. *In* "RNA Genetics." (E. Domingo, J.J. Holland and P. Ahlquist, Eds.), pp.

71-98. CRC Press, Florida.

Kelly, L., Gerlach, W.L. and Waterhouse, P.M. (1994). Characterisation of the subgenomic RNAs of an Australian isolate of barley yellow dwarf luteovirus. *Virology* **202**, 565-573.

Kennedy, S.I.T. (1980). Synthesis of alphavirus RNA. *In* "The Togaviruses." (R.W. Schlesinger, Ed.), pp. 359-361. Academic Press, New York.

Kozak, M. (1984). Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. *Nucleic Acids Research* **12**, 857-872.

Kuhn, R.J., Griffin, D.E., Owen, K.E., Niesters, H.G.M. and Strauss, J.H. (1996). Chimeric sindbis-ross river viruses to study interactions between alphavirus nonstructural and structural regions. *Journal of Virology* **70**, 7900-7909.

Kurath, G. and Palukaitis, P. (1989). Satellite RNAs of cucumber mosaic virus: recombinants constructed in vitro reveal independent functional domain for chlorosis and necrosis in tomato. *Molecular Plant-Microbe Interactions* **2**, 91-96.

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

Lakshman, D.K. and Gonsalves, D. (1985). Genetic analyses of two large-lesion isolates of cucumber mosaic virus. *Phytopathology* **75**, 758-762.

Langereis, K. (1987). Expression of alfalfa mosaic virus cDNA. *PhD Thesis*. State University of Leiden, The Netherlands.

Lawson, R.H. (1967). Relationships among tomato aspermy, aspermy-related viruses from chrysanthemum, and two strains of cucumber mosaic virus. *Virology* **32**, 357-362.

Lewandowski, D.J. and Dawson, W.O. (1993). A single amino acid change in tobacco mosaic virus replicase prevents symptom production. *Molecular Plant-Microbe Interactions* **6**, 157-160.

Li, Q. and Palukaitis, P. (1996). Comparison of the nucleic acid- and NTP-binding properties of the movement protein of cucumber mosaic cucumovirus and tobacco mosaic tobamovirus. *Virology* **216**, 71-79.

Li, X.H., Heaton, L.A., Morris, T.J. and Simon, A.E. (1989). Turnip crinkle virus defective interfering RNAs intensify viral symptoms and are generated *de novo*. *Proc. Natl. Acad. Sci. USA* **86**, 9173-9177.

Linthorst, H.J.M., Danhash, N., Brederode, F.T., Van Kan, J.A.L., De Wit, P.J.G.M. and Bol, J.F. (1991). Tobacco and tomato PR proteins homologous to win and pro-hevein lack the "hevein" domain. *Molecular Plant-Microbe Interactions* **4**, 586-592.

Lommel, S.A., Kendall, T.L., Xiang, Z. and Nutter, R.C. (1991). Identification of the maize chlorotic mottle virus capsid protein citron and characterisation of its subgenomic messenger RNA. *Virology* **181**, 382-385.

Lopez, S., Yao, J-S., Kuhn, R.J., Strauss, E.G. and Strauss, J.H. (1994). Nucleocapsid-glycoprotein interactions required for assembly of alphaviruses. *Journal of Virology* **68**, 1316-1323.

Lot, H. and Kaper, J.M. (1976). Further studies on the RNA component distribution among the nucleoproteins of cucumber mosaic virus. *Virology* **74**, 223-226.

Lot, H. and Kaper, J.M. (1976a). Physical and chemical differentiation of three strains of cucumber mosaic virus and peanut stunt virus. *Virology* **74**, 209-222.

Lot, H., Marchoux, G., Marrou, J., Kaper, J.M., West, C.K., van Vloten-Doting, L. and Hull, R. (1974). Evidence for three functional RNA species in several strains of cucumber mosaic virus. *Journal of General Virology* **22**, 81-93.

Lot, H., Marrou, J., Quiot, J.B. and Esvan, C. (1972). Contribution a l'étude du virus de la mosaique du concombre (CMV). II. Méthode de purification rapide du virus. Ann. *Phytopathol.* **4**, 25-38.

MacFarlane, S.A., Gilmer, D. and Davies, J.W. (1992). Efficient inoculation with CaMV 35S promoter-driven DNA clones of the tobravirus PEBV. *Virology* **187**, 829-831.

Maeda, T. and Inouye, N. (1987). Differentiation of two serotypes of cucumber mosaic virus in Japan by F(ab')2 ELISA with cross absorbed antibodies. *Ber. Ohara Inst. Landw. Biol. Okayama University* **19**, 149-157.

Mansky, L.M., Krueger, A.E. and Temin, H.M. (1995). The bovine leukemia virus encapsidation signal is discontinuous and extends into the 5' end of the gag gene.

Journal of Virology **69**, 3282-3289

Marchoux, G., Marrou, J. and Quiot, J.B. (1974). Complementation entre ARN de differentes souches du virus de la mosaique du concomber. Mise en evidence d'une interaction entre deux ARN pour determiner un type de symptom. *Compt rendus hebdomadaires des seances de l'Academie des scinces, Serie D* **278**, 1943-1946.

Marrou, J., Quiot, J.B., Marchoux, G. and Duteil, M. (1975). Caractérisation par la symptomatologie de quatorze souches du virus de la mosaique du concombre et de deux autres cucumovirus. *Tentative de classification. Meded. Fac. Landbouww. Rijksuniv. Gent* **40**, 107-121.

Marsh, L.E. and Hall, T.C. (1987). Evidence implicating a tRNA heritage for the promoters of (+) strand RNA synthesis in brome mosaic virus and related viruses. *Cold Spring Harbour Symp. Quant. Biol.* **52**, 331-341.

Marsh, L.E., Dreher, T.W. and Hall, T.C. (1988). Mutational analysis of the core promoter and modulator sequences of the BMV RNA 3 subgenomic promoter. *Nucleic Acids Research* **16**, 981-995.

Masuta, C. and Takanami, Y. (1989). Determination of sequence and structural requirements for pathogenicity of a cucumber mosaic virus satellite RNA (Y-satellite RNA). *The Plant Cell* **1**, 1165-1174.

Matthews, R.E.F. (1991). Plant Virology. Academic Press, New York.

Maule, A.J. (1991). Virus movement in infected plants. *Crit. Rev. Plant Sci.* **9**, 457-473.

Mcgarvey, P., Tousignant, M., Geletka, L., Cellini, F. and Kaper, J.M. (1995). The complete sequence of a cucumber mosaic virus from Ixora that is deficient in the replication of satellite RNAs. *Journal of General Virology* **76**, 2257-2270.

Meshi, T., Motoyoshi, F., Adachi, A., Watanabe, Y., Takamatsu, N. and Okada, Y. (1988). Two concomitant base substitutions in the putative replicase genes of tobacco mosaic virus confer the ability to overcome the effects of a tomato resistance gene, Tm-1. *The EMBO Journal* **7**, 1575-1582.

Meshi, T., Watanabe, Y., Saito, T., Sugimoto, A., Maeda, T. and Okada, Y. (1987). Function of the 30-kd protein of tobacco mosaic virus: Involvement in cell-to-cell

movement and dispensability for replication. *The EMBO Journal* **6**, 2557-2564.

Miller, J.S. and Mayo, M.A. (1991). The location of the 5' end of the potato leafroll luteovirus subgenomic coat protein mRNA. *Journal of General Virology* **72**, 2633-2638.

Miller., W.A., Bujarski, J.J., Dreher, T.W. and Hall, T.C. (1986). Minus-strand initiation by brome mosaic virus replicase within the 3' tRNA-like structure of native and modified RNA templates. *Journal of Molecular Biology* **187**, 537-546.

Miller., W.A., Dreher, T.W. and Hall, T.C. (1985). Synthesis of brome mosaic virus subgenomic RNA *in vitro* by internal initiation on (-) sense genomic RNA. *Nature* **313**, 68-70.

Mise, K., Allison, R.F., Janda, M. and Ahlquist, P. (1993). Bromovirus movement protein genes play a crucial role in host specificity. *Journal of Virology* **67**, 2815-2823.

Mogen, B.D, MacDonald, M.H, Grabosch, R. and Hunt, A.G. (1990). Upstream sequences other than AAUAAA are required for efficient messenger RNA 3'-end formation in plants. *The Plant Cell* **2**, 1261-1272.

Moriones, E., Diaz, I., Fernandezcuartero, B., Fraile, A., Burgyan, J. and Garcia-Arenal, F. (1994). Mapping helper virus functions for cucumber mosaic virus satellite RNA with pseudorecombinants derived from cucumber mosaic and tomato aspermy viruses. *Virology* **205**, 574-577.

Moriones, E., Diaz, I., Rodriguez-Cerezo, E., Fraile, A. and Garcia-Arenal, F. (1992). Differential interactions among strains of tomato aspermy virus and satellite RNAs of cucumber mosaic virus. *Virology* **186**, 475-480.

Moriones, E., Roossinck, M.J. and Garcia-Arenal, F. (1991). Nucleotide sequence of tomato aspermy virus RNA 2. *Journal of General Virology* **72**, 779-783.

Mossop, D.W. and Francki, R.I.B. (1977). Association of RNA 3 with aphid transmission of cucumber mosaic virus. *Virology* **81**, 177-181.

Murphy, F.A., Fauquet, C.M., Bishop, D.H.L., Ghabrial, S.A., Jarvis, A.W., Martelli, G.P., Mayo, M.A. and Summers, M.D. (1995). Virus Taxonomy-sixth report of the international committee on taxonomy of viruses. *Archives of Virology*

Supplement 10.

Mushegian, A.R. and Koonin, E.V. (1993). Cell-to-cell movement of plant viruses: insights from amino acid sequence comparisons of movement proteins and from analogies with cellular transport systems. *Archives of Virology* **133**, 239-257.

Nassuth, A. and Bol, J.F. (1983). Altered balance of the synthesis of plus- and minus-strand RNAs induced by RNAs 1 and 2 of alfalfa mosaic virus in the absence of RNA 3. *Virology* **124**, 75-85.

Nassuth, A., Alblas, F. and Bol, J.F. (1981). Localisation of genetic information involved in the replication of alfalfa mosaic virus. *Journal of General Virology* **53**, 207-204.

Neeleman, L., van der Kuyl, A.C. and Bol, J.F. (1991). Role of alfalfa mosaic virus coat protein gene in symptom formation. *Virology* **181**, 687-693.

Nitta, N., Takanami, Y., Kuwata, S. and Kubo, S. (1988). Comparative studies on the nucleotide sequence of cucumber mosaic virus RNA 3 between Y strain and Q strain. *Ann. Phytopathol. Soc. Japan* **54**, 516-522.

O'Reilly, D., Thomas, C.J.R. and Coutts, R.H.A. (1991). Tomato aspermy virus has an evolutionary relationship with other tripartite RNA plant viruses. *Journal of General Virology* **72**, 1-7.

O'Reilly, D., Thomas, C.J.R. and Coutts, R.H.A. (1994). Nucleotide sequence of RNA 3 of the British type isolate (Blencowe strain) of tomato aspermy virus. *Virus Genes* **8**, 79-81.

Ou, J.H., Rice, C. M., Dalgarno, L., Strauss, E. G. and Strauss, J. H. (1982). Sequence studies of several alphavirus genomic RNAs in the region containing the start of the subgenomic RNA. *Proc. Natl. Acad. Sci. USA* **79**, 5235-5239.

Owen, J. and Palukaitis, P. (1988). Characterisation of cucumber mosaic virus: I. molecular heterogeneity mapping of RNA 3 in eight CMV strains. *Virology* **166**, 495-502.

Owen, J., Shintaku, M., Aeschleman, P., Tahar, S. B. and Palukaitis, P. (1990) Nucleotide sequence and evolutionary relationships of cucumber mosaic virus (CMV) strains; CMV RNA 3. *Journal of General Virology* **71**, 2243-2249.

Pacha, R.F. and Ahlquist, P. (1992). Substantial portions of the 5' and intercistronic noncoding regions of cowpea chlorotic mottle virus RNA 3 are dispensable for systemic infection but influence viral competitiveness and infection pathology. *Virology* **187**, 298-307.

Pacha, R.F., Allison, R.F. and Ahlquist, P. (1990). Cis-Acting sequences required for *in vivo* amplification of genomic RNA 3 are organised differently in related bromoviruses. *Virology* **174**, 436-443.

Palukaitis, P., Roossinck, M.J., Dietzgen, R.G. and Francki, R.I.B. (1992). Cucumber mosaic virus. *Advances in Virus Research* **41**, 281-348.

Peden, K.W.C. and Symons, R.H. (1973). Cucumber mosaic virus contains a functionally divided genome. *Virology* **53**, 487-492.

Peng, D., Koetzner, C.A., McMahon, T., Zhu, Y. and Masters, P.S. (1995) Construction of murine coronavirus mutants containing interspecies chimeric nucleocapsid proteins. *Journal of Virology* **69**, 5475-5484.

Perry, K.L., Zhang, L., Shintaku, M.H. and Palukaitis, P. (1994). Mapping determinants in cucumber mosaic virus for transmission by *Aphis gossypii*. *Virology* **205**, 591-595

Petty, I.T.D., Edwards, M.C. and Jackson, A.O. (1990). Systemic movement of an RNA plant virus determined by a point substitution in a 5' leader sequence. *Proc. Natl. Acad. Sci. USA* **87**, 8894-8897

Piazzolla, P., Diaz-Ruiz, J.R. and Kaper, J.M. (1979). Nucleic acid homologies of eighteen cucumber mosaic virus isolates determined by competition hybridisation. *Journal of General Virology* **45**, 361-369.

Pinck, M., Fritsch, C., Ravelonandro, M., Thivent, C. and Pinck, L. (1981). Banding of ribosomes to the 5' leader sequence (N=258) of RNA 3 from alfalfa mosaic virus. *Nucleic Acids Research* **9**, 1087-1100.

Pleij, C.W.A. (1995). Structure and function of RNA pseudoknots. *In* "Genetic Engineering." (J.K. Setlow, Ed.), pp. 67-80. Plenum Press, New York.

Pleij, C.W.A., Abrahams, J.P., Van Belkum, A., Rietveld, K. and Bosch, L. (1987). The spatial folding of the 3' noncoding region of aminoacylatable plant viral RNAs. *In*

"Positive Strand RNA Viruses." (M. Brinton and R. Rueckert, Eds), pp. 299-316. Liss Press, New York.

Pletnev, A.G., Bray, M., Huggins, J. and Lai, C.J. (1992) Construction and characterisation of chimeric tick-borne encephalitis/dengue type 4 viruses. *Proc. Natl. Acad. Sci. USA* **89**, 10532-10536

Pogany, J.P., Romero, J., Huang, Q., Sgro, J.Y., Shang, H.Z. and Bujarski, J.J. (1995). *De novo* generation of defective interfering-like RNAs in broad bean mottle bromovirus. *Virology* **212**, 574-586.

Pogue, G.P., Marsh, L.E., Connell, J.P. and Hall, T.C. (1992). Requirement for ICR-like sequences in the replication of brome mosaic virus genomic RNA. *Virology* **188**, 742-753.

Porta, C., Devergne, J.C., Cardin, L., Briand, J.P. and van Regenmortel, M.H.V. (1989). Serotype specificity of monoclonal antibodies to cucumber mosaic virus. *Archives of Virology* **104**, 271-285.

Prüfer, D., Wipf-Scheibel, C., Richards, K., Guilley, H., Lecoq, H. and Jonard, G. (1995). Synthesis of a full-length infectious cDNA clone of cucurbit aphid-borne yellows virus and its use in gene exchange experiments with structural proteins from other luteoviruses. *Virology* **214**, 150-158.

Quiot, J.B. (1980). Ecology of cucumber mosaic virus in the Rhone valley of France. *Acta Hort.* **88**, 9-21.

Quiot, J.B., Labone, G. and Quiot-Douine, L. (1983). The comparative ecology of cucumber mosaic virus in Mediterranean and tropical regions. *In* "Plant Virus Epidemiology: The spread and control of insect-borne viruses." (R.T. Plumb and J.M. Thresh, Eds.), pp. 177-183. Blackwell Scientific Publ., Oxford.

Rao, A.L.N. and Francki, R.I.B. (1981). Comparative studies on tomato aspermy and cucumber mosaic viruses. VI. partial compatibility of genomes from the two viruses. *Virology* **114**, 573-575.

Rao, A.L.N. and Francki, R.I.B. (1982). Distribution of determinants for symptom production and host range on the three RNA components of cucumber mosaic virus. *Journal of General Virology* **61**, 197-205.

Rao, A.L.N. and Grantham, G.L. (1995). A spontaneous mutation in the movement protein gene of brome mosaic virus modulates symptom phenotype in *Nicotiana benthamiana*. *Journal of Virology* **69,** 2689-2691.

Rao, A.L.N., Hatta, T. and Francki, R.I.B. (1982). Comparative studies on tomato aspermy and cucumber mosaic viruses. VII. serological relationships reinvestigated. *Virology* **116**, 318-326.

Rezaian, M.A. and Symons, R.H. (1986). Anti-sense regions in satellite RNA of cucumber mosaic virus from stable complex with the viral coat protein gene. *Nucleic Acids Research* **14**, 3229-3239.

Rezaian, M.A., Williams, R.H.V. and Symons, R.H. (1985). Nucleotide sequence of cucumber mosaic virus RNA 1: Presence of a sequence complimentary to part of the viral satellite RNA and homologies with other viral RNAs. *European Journal of Biochemistry* **150**, 331-339.

Rezaian, M.A., Williams, R.H.V., Gordon, K.H.J. and Symons, R.H. (1984). Nucleotide sequence of cucumber mosaic virus RNA 2 reveals a translation product significantly homologous to corresponding proteins of other viruses. *European Journal of Biochemistry* **143**, 277-284.

Rice, C.M., Grakoui, A., Galler, R. and Chambers, T.J. (1989). Transcription of infectious yellow fever RNA from full-length complementary DNA templates produced by *in vitro* ligation. *New Biologist* **1**, 285-296.

Richins, R.D., Donson, J., Lewandowsky, D.J., Ducasse, D.A, Dawson, W.O. and Shepherd, R.J. (1993). Temperature-dependent complementation of movement-deficient mutants of tobacco mosaic virus by gene I of peanut chlorotic streak caulimovirus. *Phytopathology* **83**, 1349.

Richter, J. (1983). Differenzierung von stammen des Gurkenmosaik-Virus (cucumber mosaic virus). *Archiv. Phytopathol. und Pflanzenschutz* **19**, 77-79.

Rizzo, T. M. and Palukaitis, P. (1988) Nucleotide sequence and evolutionary relationships of cucumber mosaic virus (CMV) strains: CMV RNA 2. *Journal of General Virology* **69,** 1777-1787

Rizzo, T. M. and Palukaitis, P. (1989) Nucleotide sequence and evolutionary

relationships of cucumber mosaic virus (CMV) strains: CMV RNA 1. *Journal of General Virology* **70**, 1-11.

Robards, A.W. and Lucas, W.J. (1990). Plasmodesmata. *Annu. Rev. Plant Phsiol. Plant Biol.* **41**, 369-419.

Rodriguez-Cerezo, E., Klein, P.G. and Shaw, J.W. (1991). A determinant of disease symptom severity is located in the 3'-terminal noncoding region of the RNA of a plant virus. *Proc. Natl. Acad. Sci. USA* **88,** 9863-9867.

Romero, J., Huang, Q., Pogany, J. and Bujarski, J.J. (1993). Characterisation of defective interfering RNA components that increase symptom severity of broad bean mottle virus infections. *Virology* **194**, 576-583.

Roossinck, M.J. (1991). Temperature-sensitive replication of cucumber mosaic virus in muskmelon (*Cucumis melo*, cultivar Iroquois), maps to RNA 1 of a slow strain. *Journal of General Virology* **72**, 1747-1750.

Roossinck, M.J. and Palukaitis, P. (1990). Rapid induction and severity of symptoms in zucchini squash (*Cucurbita pepo*) map to RNA 1 to cucumber mosaic virus. *Molecular Plant-Microbe Interactions* **3**, 188-192.

Roossinck, M.J., Sleat, D. and Palukaitis, P. (1992). Satellite RNAs of plant viruses: Structures and biological effects. *Microbiological Reviews* **56**, 265-279.

Salanki, K., Balazs, E. and Burgyan, J. (1994). Nucleotide sequence and infectious *in vivo* transcripts of RNA 3 of tomato aspermy virus pepper isolate. *Virus Research* **33**, 281-289.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual. Cold Spring Harbour Laboratory, Cold Spring Harbour, New York.

Sanger, F., Coulson, A.R., Barrell, B.G., Smith, A.J.H. and Roe, B.H. (1980). Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. *Journal of Molecular Biology* **143**, 161-178.

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

Scholthof, H.B., Scholthof, K.B.G., Kikkert, M. and Jackson, A.O. (1995). Tomato

bushy stunt virus spread is regulated by two nested genes that function in cell-to-cell movement and host-dependent systemic invasion. *Virology* **213**, 425-438.

Schwinghamer, M. W. and Symons, R. H. (1977). Translation of the four major RNA species of cucumber mosaic virus in plant and animal cell-free systems and in toad oocytes. *Virology* **79**, 88-108.

Schwinghamer, M.W. and Symons, R.H. (1975). Fractionation of cucumber mosaic virus RNA and its translation in a wheat embryo cell-free system. *Virology* **63**, 252-262.

Scott, H. (1963). Purification of cucumber mosaic virus. *Virology* **20**, 103-106.

Shi, B.J. and Li, Y.Y. (1989). Studies on viral disease of Chrysanthemum (for medicinal usage) in Anhui. II. A viral isolate infecting Chrysanthemum, its physical and chemical properties. *Journal of Anhui Agricultural University* **2**, 157-163.

Shi, B.J. and Li, Y.Y. (1992). Comparison of CMV from tobacco with CMV from other seven plants on some biological, physical and chemical properties. *Chinese Journal of Tobacco* **1**, 36-43.

Shi, B.J., Li, Y.Y. and Zhou, G.H. (1989). Studies on viral disease of chrysanthemum (for medicinal usage) in Anhui. I. A viral isolate infecting chrysanthemum, its biological properties and control experiments. *Journal of Anhui Agricultural University* **2**, 152-156.

Shi, B.J., Li, Y.Y., Zhang, C.L. and Hua, R.M. (1993). Identification of a viral isolate infecting chrysanthemum (for medicinal usage). *Acta Phytopathologica Sinica* **23**, 231-236.

Shintaku, M.H., Zhang, L. and Palukaitis, P. (1992). A single amino acid substitution in the coat protein of cucumber mosaic virus induces chlorosis in tobacco. *The Plant Cell* **4**, 751-757.

Singer, M. and Berg, P. (1991). Genes and Genomes: A Change Perspective. University Science Books, Mill Valley, California.

Skingle, D.C.,McInnes, J.L. and Symons, R.H. (1990). An improved method for eliminating RNA contaminating of plasmid DNA preparations. *BioTechniques* **9**, 314-317.

Sleat, D.E. and Palukaitis, P. (1992). A single nucleotide change within a plant virus satellite RNA alters the host specificity of disease induction. *Plant Journal* **2**, 43-49.

Sleat, D.E., Zhang, Li and Palukaitis, P. (1994). Mapping determinants within cucumber mosaic virus and its satellite RNA for the induction of necrosis in tomato plants. *Molecular Plant-Microbe Interactions* **7**, 189-195.

Smirnyagina, E., Hsu, Y.H., Chua, N. and Ahlquist, P. (1994). Second-site mutations in the brome mosaic virus RNA 3 intercistronic region partially suppress a defect in coat protein mRNA transcription. *Virology* **198**, 427-436.

Smith, D.B. and Johnson, K.S. (1988). Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene* **67**, 31-40.

Solovyev, A.G., Zelenina, D.A., Savenkov, E.I., Grdzelishvili, V.Z., Morozov, S.Yu., Lesemann, D.E., Maiss, E., Casper, R. and Atabekov, J.G. (1996). Movement of a barley stripe mosaic virus chimera with a tobacco mosaic virus movement protein. *Virology* **217**, 435-441.

Studier, F.W., Rosenberg, A.H., Dunn, J.J. and Dubendorff, J.W. (1990). Use of T7 RNA polymerase to direct expression of cloned genes. *Methods in Enzymology* **185**, 60-89.

Sumiyoshi, H., Hoke, C.H. and Trent, D.W. (1992). Infectious Japanese encephalitis virus RNA can be synthesised from in vitro-ligated cDNA templates. *Journal of Virology* **66**, 5425-5431.

Suzuki, M., Kuwata, S., Kataoka, J., Masuta, C., Nitta, N. and Takanami, T. (1991). Functional analysis of deletion mutants of cucumber mosaic virus RNA 3 using an *in vitro* transcription system. *Virology* **183**, 106-113.

Suzuki, M., Kuwata, S., Masuta, C. and Takanami, Y. (1995). Point mutations in the coat protein of cucumber mosaic virus affect symptom expression and virion accumulation in tobacco. *Journal of General Virology* **76**, 1791-1799.

Symons, R.H. (1975). Cucumber mosaic virus RNA contains 7-methylguanosine at the 5'-terminus of all four RNA species. *Molecular Biology Reports* **2**, 277-285.

Symons, R.H. (1978). The two-step purification of ribosomal RNA and plant viral RNA

by polyacrylamide slab gel electrophoresis. *Aust. J. Biol. Sci.* **31**, 25-37.

Symons, R.H. (1985). Viral genome structure. *In* "The Plant Viruses, Vol. 1: Polyhedra Virions with Tripartite Genomes " (R.I.B. Francki, Ed.), pp. 57-81. Plenum Press, New York.

Takahashi, H. and Ehara, Y. (1993). Severe chlorotic spot symptoms in cucumber mosaic virus strain Y-infected tobaccos are induced by a combination of the virus coat protein gene and two host recessive genes. *Molecular Plant-Microbe Interactions* **6**, 182-189.)

Takahashi, H., Goto, N and Ehara, Y. (1994). Hypersensitive response in cucumber mosaic virus-inoculated *Arabidopsis thaliana. Plant Journal* **6**, 369-377.

Takanami, Y. and Tomaru, K. (1969). Effect of EDTA on cucumber mosaic virus and its application in purification. *Virology* **37**, 293-295.

Taliansky, M.E. and Garcia-Arenal, F. (1995). Role of cucumovirus capsid protein in long-distance movement within the infected plant. *Journal of Virology* **69**, 916-922.

Taliansky, M.E., de Jager, C.P., Weillink, J., van Lent, J.W.M. and Goldbach, R.W. (1993). Defective cell-to-cell movement of cowpea mosaic virus mutant N123 is efficiently complemented by sunn-hemp mosaic virus. *Journal of General Virology* **74**, 1895-1901.

Taniguchi, T, Palmieri, M. and Weissmann, C. (1978). Qb DNA-containing hybrid plasmids giving rise to Qb phage formation in the bacterial host. *Nature* **274**, 2293-2298.

Tolin, S.A. (1977). Cucumovirus (cucumber mosaic virus) group. In "The Atlas of Insect and Plant Viruses: Including Mycoplasma Viruses and Viroids." (K. Maramorosch, Ed.), pp. 303-307. Academic Press, New York.

Tourneur, C., Jouanin, L. and Vaucheret, H. (1993). Over-expression of acetolactate synthase confers resistance to valine in transgenic tobacco. *Plant Science* **88**, 159-168.

Towbin, H., Staehelin, T. and Gordon, T. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some application. *Proc. Natl. Acad. Sci. USA* **76**, 4350-4354.

Traynor, P., Young, B. M. and Ahlquist, P. (1991). Deletion analysis of brome mosaic virus 2a protein: Effects on RNA replication and systemic spread. *Journal of Virology* **65,** 2807-2815.

Tsai, C-H. and Dreher, T.W. (1993). Increased viral yield and symptom severity result from a single amino acid substitution in the turnip yellow mosaic virus movement protein. *Molecular Plant-Microbe Interactions* **6,** 268-273.

Turner, D.R., Lorie, E.J. and Butler, P.J.G. (1988). The tobacco mosaic virus assembly origin RNA: Function characteristics defined by directed mutagenesis. *Journal of Molecular Biology* **203**, 531-547.

Van der Kuyl, A.C., Langereis, K., Houwing, C.J., Jaspars, E.M.J. and Bol, J.F. (1990). cis-Acting elements involved in replication of alfalfa mosaic virus RNAs *in vitro*. *Virology* **176**, 346-354.

Van der Kuyl, A.C., Neeleman, L. and Bol, J.F. (1991). Deletion analysis of cis- and trans-acting elements involved in replication of alfalfa mosaic virus RNA 3 *in vivo*. *Virology* **183**, 687-694

Van der Vossen, E.A.G., Neeleman, L. and Bol, J.F. (1993). Role of the 5' leader sequence of alfalfa mosaic virus RNA 3 in replication and translation of the viral RNA. *Nucleic Acids Research* **21**, 1361-1367.

Van der Vossen, E.A.G., Notenboom, T. and Bol, J.F. (1995). Characterisation of sequences controlling the synthesis of alfalfa mosaic virus subgenomic RNA *in vivo*. *Virology* **212**, 663-672.

Vance, V.B., Berger, P.H., Carrington, J.C., Hunt, A.G. and Shi, X.M. (1995). 5' proximal potyviral sequences mediate potato virus X/potyviral synergistic disease in transgenic tobacco. *Virology* **206,** 583-590.

Vaquero, C., Turner, A.P., Demangeat, G., Sanz, A., Serra, M.T., Roberts, K. and Garcialuque, I. (1994). The 3a protein from cucumber mosaic virus increases the gating capacity of plasmodesmata in transgenic tobacco plants. *Journal of General Virology* **75**, 3193-3197.

Verwoerd, T.C., Dekker, B.M.M. and Hoekema, A. (1989). A small-scale procedure for the rapid isolation of plant RNAs. *Nucleic Acids Research* **17**, 2362.

Von Arnim, A., Frischmuth, T. and Stanley, J. (1993). Detection and possible functions of African cassava mosaic virus DNA B gene products. *Virology* **192**, 264-272.

Wahyuni, W.S., Dietzgen, R.G., Hanada, K. and Francki, R.I.B. (1992). Serological and biological variation between and within subgroup I and II strains of cucumber mosaic virus. *Plant Pathology* **41**, 282-297.

Waterworth, H.E. and Povish, W.R. (1975). A virus related to cucumber mosaic virus isolated from imported Ixora plants. *Phytopathology* **65**, 728-729.

Weber, H., Haeckel, P. and Pfitzner, A.J.P. (1992). A cDNA clone of tomato mosaic virus is infectious in plants. *Journal of Virology* **66**, 3909-3912.

Weiland, J.J. and Edwards, M.C. (1994). Evidence that the alpha a gene of barley stripe mosaic virus encodes determinants of pathogenicity to oat (Avena sativa). *Virology* **201**, 116-126.

White, K.A., Bancroft, J.B. and Mackie, G.A. (1991). Defective RNAs of clover yellow mosaic virus encode nonstructural/coat protein fusion products. *Virology* **183**, 479-486.

Wilson, P.A. and Symons, R.H. (1981). The RNAs of Cucumoviruses: 3'-terminal sequence analysis of two strains of tomato aspermy virus. *Virology* **112**, 342-345.

Wilson, T.M.A. (1993). Strategies to protect crop plants against viruses: Pathogen-derived from resistance blossoms. *Proc. Natl. Acad. Sci. USA* **90**, 3134-3143.

Wood, K.R. (1991). Organisation and expression of the genome of cucumber mosaic virus. *In* "Biochemistry and Molecular Biology of Plant-Pathogen Interactions." (C.J. Smith, Ed.).pp. 26-49. Clarendon Press, Oxford.

Wu, G. and Kaper, J.M. (1995). Competition of viral and satellite RNAs of cucumber mosaic virus for replication *in vitro* by viral RNA-dependent RNA polymerase. *Research in Virology* **146**, 61-67.

Zaccomer, B., Haenni, A.L. and Macaya, G. (1995). The remarkable variety of plant RNA virus genomes. *Journal of General Virology* **76**, 231-247.

Zhang, L., Hanada, K. and Palukaitis, P. (1994a). Mapping local and symptom determinants of cucumber mosaic cucumovirus in tobacco. *Journal of General*

Virology **75**, 3185-3191.

Zhang, L., Kim, C.H. and Palukaitis, P. (1994b). The chlorosis-induction domain of the satellite RNA of cucumber mosaic virus-identifying sequences that affect accumulation and the degree of chlorosis. *Molecular Plant-Microbe Interactions* **7**, 208-213.

Ziegler-Graff, V., Guilford, P.J. and Baulcombe, D.C. (1991). Tobacco rattle virus RNA-1 29K gene product potentiates viral movement and also affects symptom induction in tobacco. *Virology* **182**, 145-155.

Zitter, T.A. and Gonsalves, D. (1991). Differentiation of pseudorecombinants of two cucumber mosaic virus strains by biological properties and aphid transmission. *Phytopathology* **81**, 139-143.

Zou, S. and Brown, E.G. (1992). Identification of sequence elements containing signals for replication and encapsidation of the reovirus M1 genome segment. *Virology* **186**, 377-388.