GENE EXPRESSION AND SUBGENOMIC RNAs OF CUCUMBER MOSAIC VIRUS.

Thesis submitted to the University of Adelaide in fulfillment of the requirements for the degree of Doctor of Philosophy

by

Karl H. J. Gordon B.Sc. (Hons.) (Melb.)

Department of Biochemistry,
University of Adelaide,
South Australia.

December, 1983.

CONTENTS

		Page
SUMMARY		
STATEMENT		
ACKNOWLEDGEMENTS		
ABBREVIATIONS		
CHAPTER 1	GENERAL INTRODUCTION AND SCOPE OF THIS THESIS	
A.	RNA virus genomes	1
B.	Cucumber mosaic virus	2
C.	Virus-induced RNA-dependent RNA polymerases	4
D.	Satellite RNAs of CMV	6
E.	Scope of this thesis	7
CHAPTER 2	IN VITRO TRANSLATION PRODUCTS FROM FOUR STRAINS OF CMV	
INTRODUCTION		9
MATERIALS AND METHODS		
A.	Virus and viral RNA	10
B.	Translation of CMV RNAs	10
C.	Discontinuous SDS-polyacrylamide gel electrophoresis of proteins	10
D.	Fluorography and autoradiography	11
E.	Purification of proteins for peptide mapping	11
F.	Peptide mapping using proteases	12
G.	Non-equilibrium pH gradient electrophoresis	12
RESULTS		
A.	Translation products from CMV RNAs 1 and 2	13

B. Translation products from CMV RNA 3 14

C. Partial proteolytic digestion of RNA 3 translation products 14

DISCUSSION 15

CHAPTER 3 HIGHLY PURIFIED CMV-INDUCED RNA DEPENDENT RNA POLYMERASE DOES NOT CONTAIN ANY OF THE FULL LENGTH TRANSLATION PRODUCTS OF THE GENOMIC RNAs

INTRODUCTION 18

MATERIAL AND METHODS

A. Extraction and purification of RNA replicase 18

B. Peptide mapping - CNBr cleavage 19

RESULTS

A. Comparison of polypeptides of CMV-induced particulate RNA replicase with *in vitro* translation products of CMV RNAs 1, 2 and 3 19

B. Different strains of CMV induce RNA replicase with identical polypeptide components. 20

C. Peptide mapping confirms the absence of full length translation products from the highly purified RNA replicase 21

DISCUSSION 22

CHAPTER 4 SUBGENOMIC RNAs FROM CMV RNA 2

INTRODUCTION 25

MATERIALS AND METHODS

A. Materials 25

B. Virus and RNA 25

C. Cloning and sequence analysis of CMV RNA2 26

D. Direct sequence analysis of RNA 2 26

E. RNA blotting hybridizations 27

F. Hybrid-selection of RNA 28

G. Other methods for RNA fractionation 28

RESULTS

A. Nucleotide sequence of CMV RNA 2 29

B. RNA blotting of subgenomic RNAs 29

C. Primer extension mapping 32

D. *In vitro* translation products of the subgenomic RNAs 33

DISCUSSION 35

CHAPTER 5 SUBGENOMIC RNAs FROM RNA 1

INTRODUCTION 40

RESULTS

A. Clones from CMV RNA 1 40

B. Subgenomic RNAs from RNA 1 41

C. *In vitro* translation of subgenomic RNAs 42

DISCUSSION 43

CHAPTER 6 CMV SUBGENOMIC RNAs : SYNTHESIS AND GENERAL DISCUSSION

INTRODUCTION 45

RESULTS 45

DISCUSSION

A. Synthesis of subgenomic RNAs 47

B. CMV subgenomic RNAs and gene expression 49

C. Other plant viral subgenomic mRNAs 51

D. Subgenomic RNAs and the evolution of CMV 52

CHAPTER 7 CMV SATELLITE RNA : SECONDARY STRUCTURE AND REPLICATIVE FORMS

INTRODUCTION 54

MATERIALS AND METHODS

A. Virus and RNA 55

B. Materials 55

C. cDNA clones in M13mp9 55

D. Secondary structure analysis of Sat-RNA 55

E. Detection of (+) and (-) Sat-RNA 58

RESULTS

A. Secondary structure analysis of Sat-RNA 59

B. Secondary structure model of Sat-RNA 60

C. Double stranded and oligomeric Sat-RNA 61

DISCUSSION

A. Nucleotide sequence and secondary structure of Sat-RNA and CARNA 5 62

B. Sequence and structural homology between Sat-RNA and CMV genomal RNAs 63

C. Possible function of 3'-terminal structure and replication of Sat-RNA 65

SUMMARY

The aim of this work was to investigate gene products of the RNAs of cucumber mosaic virus (CMV). The full length translation products of the genomic RNAs were characterised and their possible *in vivo* functions studied. Further work involved the characterization of subgenomic RNAs from CMV RNAs 1 and 2, and of the translation products of these subgenomic RNAs.

The genomic RNAs (1, 2 and 3) from four strains of CMV were analysed by *in vitro* translation. Each RNA yielded one major polypeptide (as already known for Q-CMV), whose size varied slightly among the four strains. Peptide mapping of some translation products showed that the strains fell into 2 distinct, but related groups: one with Q- and P-CMV, the other consisting of T- and M-CMV.

The highly purified RNA-dependent RNA polymerase induced by CMV in cucumber seedlings was investigated, to determine whether the enzyme contains full length translation products of the genomic RNAs. When the enzyme was obtained from plants infected with Q-, P- or T-CMV, no variation was observed in the size of three polypeptides unique to extracts from CMV-infected plants and whose sizes were similar to the full length translation products of Q-CMV. The translation products of RNAs from P- and T-CMV were, however, known to vary in size. It was concluded that the full length translation products of CMV RNAs 1, 2 and 3 were not present in the highly purified enzyme: this was confirmed by peptide mapping.

Possible subgenomic RNAs from RNAs 1 and 2 were investigated using specific probes synthesised from M13 clones of complementary DNA (cDNA) to these RNAs. Blotting studies of RNAs from

either total encapsidated RNA or nucleic acid extracted from CMV-infected plants showed that RNAs 1 and 2 each generated over 20 subgenomic RNAs, most of which were 3'-coterminal with the genomic RNA. Some of the smallest RNAs were specific internal fragments. The subgenomic RNAs were also detected in RNA selected by hybridization to M13 clones containing (-) inserts.

The larger subgenomic RNAs (over 1500 nucleotides) were shown by <u>in vitro</u> translation to be efficient mRNAs. Those from RNA 2 (including one of 880 nucleotides) yielded polypeptides from the carboxyterminal region of its long open reading frame, whereas those from RNA 1 gave a more complex pattern of products. Further analysis must await the completion of nucleotide sequence studies on RNAs 1 and 2.

The secondary structure of a satellite RNA (Sat-RNA) totally dependent on CMV for replication, was investigated using nuclease digestion under non-denaturing conditions. Partial digests of the 5'- or 3'-^{32}P-labelled, 335 residue Sat-RNA with nucleases S_1 or T_1 showed that only about 10% of the residues were cleaved. Further data on base-paired segments were obtained by digestion with RNaseT_1, followed by electrophoretic fractionation of the resulting fragments under first non-denaturing then denaturing conditions. A complete secondary structure model is proposed for Sat-RNA, with 52% of its residues involved in base-pairing. A prominent 3' hairpin shows considerable sequence and structural homology with parts of the 3'tRNA-like structure of CMV genomal RNAs.

Possible Sat-RNA replication intermediates were studied using (+) or (-) probes synthesised from M13 clones of a restriction fragment from cDNA to Sat-RNA. These confirmed the

existence of large amounts of double stranded monomeric Sat-RNA in RNA extracted from plants, and disclosed that small amounts of (+) Sat-RNA were also present in longer, single stranded forms: a dimer and another of about 3,500 nucleotides.

STATEMENT

This thesis contains no material which has previously been submitted for the award of any other degree or diploma in any university. To the best of my knowledge, it contains no material previously published by any other person, except where due reference is made in the text.

I consent to the thesis being made available for photocopying and loan if accepted for the award of the degree.

Karl GORDON

ACKNOWLEDGEMENTS

I wish to thank Prof. W. H. Elliott for permission to work in the Department of Biochemistry. I would also like to express my gratitude to my Supervisor, Bob Symons, for the advice and encouragement he provided during this work. To my colleagues in the CMV genome analysis group (Allan Gould, Ali Rezaian and Rhys Williams) go thanks for the supply of useful and interesting clones, sequences and discussions, essential to the work described in Chapters 4, 5 and 6. The work described in Chapter 3 was done with the willing co-operation of Dalip S. Gill, who taught me a lot about replicases. Paul Keese and George Bruening provided the clones used in Chapter 7, in addition to many interesting discussions about viruses and satellites. Jim Haseloff was a continual source of good advice, reliable solutions (in return for red wine) and many stimulating debates on small circular subjects. Peter Molloy originally taught me translation. The other members of our faction provided cheerful support, occasional lethal threats and even, sometimes, useful help, as did numerous other people in the department. Mr. A. V. Sivaprasad generously provided some very useful computer programs, and Dr. D. Riesner and his colleagues a computer-drawn model of Sat-RNA. Dr. R. I. B. Francki allowed us to use his glasshouse facilities and provided not only plants and virus strains, but also helpful discussion. I am especially grateful to Jenny Rosey, Lisa Waters and Sharon Freund for their expert assistance and help with the thesis figures, to Ali for critical comments on the thesis and to Ms. Hunter for her skilful and speedy interpretation of not only a new field, but also my handwriting. My parents encouraged me most generously, as did Gisela.

ABBREVIATIONS

AlMV	alfalfa mosaic virus
AMV	avian myeloblastosis virus
BMV	brome mosaic virus
bp	base pairs
BSA	bovine serum albumin
cDNA	complementary DNA
CMV	cucumber mosaic virus
cpm	counts per minute
CPMV	cowpea mosaic virus
ddNTP	dideoxynucleoside - 5'- triphosphate
DMSO	dimethyl sulphoxide
dNTP	deoxynucleoside -5'-triphosphate
DTT	1,4-dithiothreitol
EDTA	ethylenediamine tetra-acetate
HEPES	N-2-hydroxyethylpiperazine-N-2-ethanesulphonic acid
M_r	molecular weight
mRNA	messenger RNA
NEPHGE	non-equilibrium pH gradient electrophoresis.
PPO	2,5 - diphenyloxazole
PSV	peanut stunt virus
sat-RNA	satellite RNA of CMV (Adelaide isolate)
SDS	sodium dodecyl sulphate
SSC	1x is 0.15M NaCl, 15mM Na citrate, pH7
TMV	tobacco mosaic virus
Tris	tris (hydroxymethyl) aminomethane
tRNA	transfer RNA
KB	Kilobases

CHAPTER I

GENERAL INTRODUCTION AND SCOPE OF THIS THESIS

A. RNA Virus Genomes

Viral genomes replicate in host cells through the subversion of cellular processes. The study of such viral infections can tell us a great deal about not only the organization and replication of the viral genomes, but also the molecular biology of the host cell. Most plant viruses have single stranded RNA genomes, which are of (+) sense, i.e. capable of acting directly as mRNAs for protein synthesis. These RNA genomes may comprise a single RNA species of limited size e.g. tobacco mosaic virus (TMV), with 6395 nucleotides (Goelet et al., 1982; Hirth, 1982). Many plant RNA viruses achieve slightly greater overall genome sizes by spreading their genes over 2 or 3 RNA segments (Bruening,1981; Davies and Hull,1982). The only animal viruses with segmented (+) RNA genomes are the Nodaviruses (Friesen and Rueckert, 1981), although segmented (-) RNA genomes also occur; influenza virus (Lamb and Choppin, 1983) and the Bunyaviruses (Bishop et al., 1980).

Replication of these segmented viruses therefore requires co-ordinated translation and replication of the RNA segments. Such regulation may be achieved by novel, virus-specific, protein-protein, protein-RNA or RNA-RNA interactions. Moreover, direct RNA replication has not been observed in virus-free eukaryotic cells, so that the possible involvement of host proteins in viral RNA replication could lead to the discovery of new cellular pathways for the transfer of genetic information (Zimmern, 1982).

Plant RNA viruses show enormous variety in the organization and expression of their genomes (reviewed in Atabekov and Morozov, 1979; Bruening, 1981; Davies and Hull, 1982). The comparative study of plant RNA viral gene expression, and more particularly, nucleotide and amino acid sequence variation, is thus important for the understanding of viral evolution.

The many eucaryotic RNA viruses with (+) single stranded genomes are subject to the constraints affecting translation of eucaryotic mRNAs (Kozak, 1981a,b).

In general, only the protein reading frame nearest the 5'-end of the RNA is available for translation. Some viruses overcome this by linking all their cistrons into one polyprotein (e.g. poliovirus: Kitamura et al., 1981; CPMV: Pelham, 1979). This giant translation product is promptly processed (usually by a portion of itself) into many mature proteins with differing functions. A more common strategy, however, is to express cistrons near the 3'-end of the viral RNA by subgenomic mRNAs whose 5'-termini lie immediately upstream of the relevant reading frame. Such subgenomic mRNAs are generated by many plant RNA viruses (Bruening, 1981; Hirth, 1982) and are the usual mechanism for synthesis of coat protein. The identification of other genes expressed in this way remains an open question.

B. Cucumber Mosaic Virus

With its simple genome, cucumber mosaic virus (CMV) is a very successful RNA virus. Its small (28nm) icosahedral particles (molecular weight: 5.5×10^6) (Francki et al., 1979) infect a broad range of dicotyledonous hosts. The symptoms vary widely, depending on the virus strain and host. Of the many strains known, four will be used in the work described in this thesis. These strains are the Q, (Francki et al., 1966), P, T and M strains (Habili et al., 1974; Mossop et al., 1976; Rao et al., 1982). In contrast to the other three strains, M-CMV is not aphid-transmissible (Mossop and Francki, 1977) a characteristic thought to be due to an altered coat protein. Tomato aspermy virus (TAV) has been shown to have a distant serological relationship to CMV (Rao et al., 1982).

CMV represents one of the best characterised multipartite plant RNA viruses. It has a tripartite genome, and CMV virions contain four major single stranded RNAs, designated 1 to 4 in order of decreasing molecular weight. These RNAs all have 5'-m^7G caps (Symons, 1975) and act as monocistronic messengers in cell-free translation systems (Schwinghamer and Symons, 1977), yielding full-length translation products with molecular weights of 95,000, 110,000, 35,000 and 24,500 (CMV coat protein), respectively (for Q-CMV). RNAs 1, 2 and 3 are sufficient to initiate infection

in plants (Peden and Symons, 1973; Lot et al., 1974), since the complete sequence of RNA 4 is contained within the 3'-terminal half of CMV RNA 3 (Gould and Symons, 1982). The total genetic information of CMV is thus divided among the three largest RNAs.

Recently, the complete nucleotide sequence (2193 residues) of Q-CMV RNA 3 has been determined (Gould and Symons, 1982). The sequence corresponding to RNA 4 is 1027 nucleotides long. The RNA 3 sequence contains two major open reading frames. The cistron in the 5'half encodes a protein of molecular weight 36,700 (p36.7) which is presumably the previously detected M_r 35,000 protein detected upon _in vitro_ translation of this RNA. Another cistron, within the 3' half (which corresponds to the RNA 4 sequence), encodes a protein of molecular weight 26,200 (p26.2) whose amino acid composition corresponds closely, as expected, to that determined for CMV coat protein (c.p.) (Mossop, 1978).

The estimated sizes of Q-CMV RNA 1 and 2 are about 4,000 and 3,400 nucleotides respectively (Peden and Symons, 1973). Hybridization analysis has shown that there is no significant sequence homology among RNAs 1,2 or 3, except for about 300 nucleotides at their 3' ends (Gould and Symons, 1977 ; Symons, 1979).

CMV RNA 1 is unique among the larger genomic RNAs of the tripartite viruses in one respect - the low proportion (65%) of its length required to encode the full length _in vitro_ translation product of M_r 95,000. The other large genomic RNAs of tripartite viruses generally require up to 90% of their sequence to encode the full length translation products, as does CMV RNA 2 (reviewed in Davies and Hull, 1982). This observation raises the question of whether RNA 1 contains a second, silent cistron at its 3'-end. In fact, some 1000 nucleotides are unaccounted for, based on coding requirements and the length of the RNA.

The only comparative sequence information available for different strains of CMV and TAV is for the 3' conserved 300 nucleotides (Symons, 1979;

Wilson and Symons, 1981; Barker, Wilson and Symons, unpublished). The 3' 125 nucleotides form a distinctive tRNA-like structure and are highly conserved among the 3 RNAs of each strain, and to a lesser extent between strains. These tRNA-like structures can be aminoacylated by aminoacyl: tRNA synthetases (Kohl and Hall,1974). Their possible role in the initiation of viral RNA replication has been discussed (Hall, 1979; Symons et al., 1982), but few data are available. Other tripartite viruses (the bromovirus group) and several other plant and animal RNA viruses possess 3' tRNA-like structures (reviewed in Haenni et al., 1982). An intact tRNA-like structure was found to be required for BMV RNA replication in vivo (Loesch-Fries and Hall, 1982).

The genomic RNAs of another tripartite plant virus, alfalfa mosaic virus (AlMV) lack 3' tRNA-like structures (Pinck and Pinck, 1979; Köper-Zwarthoff et al., 1979; Gunn and Symons 1980), requiring instead viral coat protein, or its effective mRNA, RNA 4 to initiate RNA replication (reviewed in van Vloten-Doting and Jaspars, 1977).

C. Virus-induced RNA-dependent RNA polymerases

Plant RNA viruses are thought to use both host and virus encoded proteins for the replication of their genomal RNAs, as do other RNA viruses. The purification and subunit characterization of the RNA-dependent RNA polymerase (RNA replicases) induced by the plant viruses represents a major unresolved problem. In contrast, the RNA replicases induced by some RNA bacteriophages have been well characterized. These enzyme systems show an interplay of host and phage encoded subunits with the RNA polymerase activity provided by proteins of phage origin (Kamen, 1975; Fedoroff, 1975; Blumenthal and Carmichael, 1979). More recently, studies on animal RNA viruses have shown a similar requirement for host and virus encoded subunits, with the RNA polymerase activity residing on virus encoded proteins (Traub et al., 1976; Newman et al., 1979; Dasgupta et al., 1980; van Dyke and Flanagan, 1980; Pallansch et al., 1980; Polatnick and Wool, 1981).

Studies involving the infection of cucumber seedlings with cucumber mosaic

virus (CMV) have shown that CMV induces a high level of RNA replicase in both soluble and particulate fractions of plant extracts, whereas no such activity is found in non-inoculated seedlings (Kumarasamy and Symons, 1979a; Gill et al., 1981). Furthermore, the specific activity of our most highly purified RNA replicase preparations, at 900 units/mg protein (1 unit of enzyme activity equals 1 nmol of GMP incorporated into RNA per minute at 37°C) using poly (C) as template, is about as high as that of the 95% purified RNA replicase induced by bacteriophage Qβ (800units/mg) with poly (C) as template (Blumenthal, 1979). This high specific activity of the CMV RNA replicase suggests that further extensive purifications will not be possible.

With other RNA virus infected-plant systems, the situation is complicated by the presence in healthy plants of low levels of RNA-dependent RNA polymerase activity which increase on virus infection; e.g. in tobacco (Clerx and Bol, 1978; Ikegami and Fraenkel-Conrat, 1978a; Romaine and Zaitlin, 1978; Duda, 1979; Chifflot et al., 1980), in cowpea (Ikegami and Fraenkel-Conrat, 1978b; White and Dawson, 1978; Dorssers et al., 1982, and in tomato (Boege and Sänger, 1980).

The best characterized of these enzymes is that of cowpea (Dorssers et al., 1982). Upon infection of cowpea with cowpea mosaic virus (CPMV), the level of membrane-bound RNA replicase activity increases at least 10-fold. Extensively purified enzyme preparations from healthy and infected plants appear identical in containing only one major polypeptide of M_r 130,000 (Dorssers et al., 1982). This protein is different to that encoded by RNA 1 of CPMV (CPMV B-component RNA) which is required for CPMV replication (Goldbach et al., 1980). In fact, the viral RNA replication complex has been shown to be functionally different to this host-encoded RNA replicase (Dorssers et al., 1983). Antibodies to both the host and viral coded polypeptides have confirmed that the latter is actually involved in viral RNA replication (Dorssers, 1983).

RNA replicase preparations from brome mosaic virus (BMV)-infected barley

leaves contain a polypeptide which comigrates, and shares tryptic peptides, with the translation product of BMV RNA 1 (Hall et al., 1982). This polypeptide is found associated with RNA replicase activity in pellets obtained after centrifugation of leaf extracts on sucrose gradients. These pellets presumably contain membranes and bound proteins but it is not known what role the BMV RNA 1 translation product plays in enzyme activity.

After extensive purification over phosphocellulose and poly (C)-cellulose, both the soluble and particulate forms of the CMV-induced RNA replicase show the same polypeptide composition on SDS-polyacrylamide gel electrophoresis (Kumarasamy and Symons, 1979a; Gill et al., 1981; Gill, 1983). The major polypeptide present is one of M_r 100,000 and is not found in similarly purified extracts from healthy plants. Two other polypeptides of M_r 110,000 and 35,000 are also unique to extracts from infected plants but are present in varying, smaller amounts. The M_r 100,000 polypeptide is slightly larger than the translation product of CMV RNA 1 (Q strain) whereas the M_r110,000 and M_r 35,000 polypeptides comigrate with the translation products of Q-CMV RNAs 2 and 3, respectively (Gill et al., 1981; and unpublished data).

All evidence indicates that the major polypeptide of M_r 100,000 which copurified with both CMV RNA-copying and poly (C)-copying activities of RNA replicase up to the final step of the sequential chromatographic purification procedure is the catalytic subunit of CMV-induced RNA replicase, (Gill, 1983). All other peptides could be removed by various chromatographic techniques and may therefore either not be components of the enzyme or perhaps play some regulatory role in RNA replication.

D. Satellite RNAs of CMV

CMV is unique among the tripartite RNA viruses, in that several satellite RNAs have been found encapsidated with viral RNA in some, but not all, isolates (Kaper et al., 1976; Gould et al., 1978; Murant and Mayo, 1982). These satellite RNAs are small (about 335 residues) linear molecules and so are much smaller than the three CMV genomal RNAs, with which they have essentially no sequence homology detectable by hybridization analysis with complementary

DNA (cDNA) (Diaz-Ruiz and Kaper, 1977; Gould et al., 1978). They are totally dependent on the helper virus for their replication, encapsidation and transmission.

Sequence and structural studies of such satellite RNAs are important for several reasons. They may provide information about their possible origins and, when compared to known sequence and structural features of the helper virus RNAs, about the minimum requirements for viral RNA replication. Also of interest are possible gene products since the CMV satellite RNAs have m^7G caps, but do not encode a satellite-specific coat protein as do some other plant virus satellite RNAs (Murant and Mayo, 1982). Further, the satellite RNA can modify the disease symptoms of the parent virus (Mossop and Francki, 1979; Waterworth et al., 1979) which indicates specific interactions with the viral and/or host genomes.

The nucleotide sequences of three isolates of satellite RNA have now been determined. These include two isolates of CMV associated RNA 5, one of which is necrotic (the 336 nucleotide (N) CARNA 5; Richards et al., 1978) and the non-necrotic (L) CARNA 5(335 residues; Collmer et al., 1983) form. The third isolate is from Adelaide (Sat-RNA) and has 335 nucleotides (R.H. Symons, P. Palukaitis and A. C. Iasiello, personal communication) which show 88% sequence homology with (N) CARNA 5. Point mutations result in the three RNAs all encoding different potential polypeptides. (N) CARNA 5 has very different biological properties to either Sat-RNA or (L) CARNA 5.

Other satellite RNAs have been reported for CMV (Takanami, 1981; P. Palukaitis, personal communication) and for the related peanut stunt virus (PSV) (Kaper et al., 1978). No other tripartite virus has been found to support a satellite RNA, although several other plant viruses do so (Murant and Mayo,1982).

E. Scope of this Thesis

The first part of the work described in this thesis describes the in vitro translation products of genomic RNAs (1,2 and 3) from Q-, P- , T- and M-CMV. This knowledge, together with partial proteolysis studies, provided a basis for the investigation of whether the highly purified CMV-induced RNA replicase

(Gill et al., 1981) contains the full length translation products of these RNAs (Gordon et al., 1982).

The second part of this thesis describes the characterization of sub-genomic mRNAs generated from both RNAs 1 and 2 of CMV. These RNAs were detected using probes from cDNA restriction fragments cloned in the bacteriophage vector M13, and RNA blotting (Thomas, 1980) of viral RNAs fractionated by agarose gel electrophoresis, after glyoxal denaturation (McMaster and Carmichael, 1977). The implications of these results for our work on the proteins involved in the replication of viral RNA are discussed.

The final part of this thesis describes experiments leading to a complete secondary structure model of CMV satellite RNA and its sequence and structural homology with part of the genomic RNAs' 3' tRNA-like structure (Gordon and Symons, 1983). Further work in this section concerns a possible replication mechanism of Sat-RNA, which has been compared with models for other satellite RNAs, and viroids (Symons et al., 1983).

CHAPTER 2

IN VITRO TRANSLATION PRODUCTS FROM FOUR STRAINS OF CMV

Introduction

The (+) sense, single stranded RNAs of many plant viruses have been extensively characterised as messenger RNAs, using in vitro translation systems. The optimum conditions found for translation of these viral RNAs have recently been reviewed (Davies, 1979; Atabekov and Morozov, 1979; van Vloten-Doting and Neeleman, 1982). The conditions which have been used for translation of CMV RNAs 1, 2 and 3 (see Methods) are similar to those which have been found suitable for many viral RNAs especially for mRNAs other than the subgenomic mRNAs for the coat protein. In vitro translation of a mixture of genomic and subgenomic RNAs led to the conclusion that subgenomic coat protein mRNAs of several plant viruses were much more efficient mRNAs than the genomic RNAs. (Gerlinger et al., 1977; Rutgers, 1977; Benicourt and Haenni, 1978; Zagorski, 1978; van Tol and van Vloten-Doting, 1979). CMV RNA4 was also found to be a very efficient mRNA (Schwinghamer and Symons, 1977), but translation of total CMV viral RNA at concentrations up to 150 μg/ml in the rabbit reticulocyte lysate, still results in efficient translation of RNAs 1, 2 and 3 (see Chapter 4).

Previous studies have shown that, in vitro, the four major CMV RNAs function as monocistronic mRNAs, each yielding a single major translation product (Schwinghamer and Symons, 1975, 1977; P. Molloy, unpublished). This work has related mainly to Q-CMV, and, to a lesser extent, to P-CMV RNAs. In this chapter, I have compared the sizes of all in vitro translation products (other than the Coat Protein) for Q-, P-, T- and M-CMV. The translation products of RNA 3, especially, were compared using partial proteolytic peptide mapping. This information, on the variation in size of translation products, provided a basis for the determination of whether the purified RNA-dependent RNA polymerase from CMV-infected cucumber plants included viral gene products (Chapter 3).

MATERIALS AND METHODS

A. Virus and viral RNA

Cucumber (Cucumis sativus cv. Polaris) seedlings were grown and infected with CMV as described by Clark et al., (1974). The Q, P, T, and M strains of CMV (Habili and Francki, 1974; Mossop et al., 1976; Rao et al.,1982) were kindly provided by Dr. R.I.B. Francki. RNA was isolated from purified virus as in Peden and Symons (1973). Total viral RNA was fractionated using one or two preparative 2.8% polyacrylamide slab gels (Symons, 1978) and the purified RNAs freed of contaminating acrylamide by preparative sucrose gradient centrifugation (Schwinghamer and Symons, 1975).

B. Translation of CMV RNAs

A cell-free lysate of rabbit reticulocytes was prepared as described by Ranu and London (1979), then treated with micrococcal nuclease (Pelham and Jackson, 1976) and used for translation of viral RNAs. Reaction mixtures (10 - 50 µl) contained 65% (v/v) lysate (20 µM in haemin), 10 µg/ml creatine kinase, 10mM creatine phosphate, 1 mM Mg acetate, 70 mM K acetate, 20 mM HEPES-KOH, pH 7.5, 100 µM spermidine-HCl, 1 mM dithiothreitol, 100 µg/ml rat liver tRNAs, 75 µM of each of the unlabelled amino acids, 20 - 40 µg/ml purified viral RNA and 200 - 500 µCi/ml of ^{3}H-leucine (120 Ci/mmol) or ^{35}S-methionine (1300 Ci/mmol) (both from NEN). For the synthesis of more highly labelled translation products, the concentration of labelled amino acid was increased to 1.0 mCi/ml.

C. Discontinuous SDS-Polyacrylamide Slab Gel Electrophoresis of Proteins

Discontinuous polyacrylamide gel electrophoresis was carried out in the Tris-glycine-SDS buffer system of Laemmli (1970) as described by Schwinghamer and Symons (1977). The slab gels (15 x 12 x 0.2 cm) were cast between vertical glass plates. The separating gel was made from a stock solution of 40% acrylamide (w/v) and 0.33% bis-acrylamide (w/v) (deionised with mixed bed resin, Amberlite MB-1) in glass distilled water. The stacking gel, which contained deionised 3% acrylamide (w/v) and 0.08% bisacrylamide, was poured above the separating gel to a height of 3.0 cm. A comb was then

inserted between the two glass plates with the teeth immersed to a depth of 3 mm. Protein samples, which had been boiled for 2 min in 40 µl of sample buffer (Laemmli, 1970), were loaded in the wells between the teeth of the comb. Electrophoresis was performed at 25 mA constant current at room temperature until the bromophenol blue dye reached near the bottom.

D. Fluorography and Autoradiography

After electrophoresis, the gels were fixed in isopropanol : acetic acid: water (25:10:65, by volume), for at least 1 h. After this, gels were impregnated in 20% (w/v) naphthalene and 0.5% (w/v) PPO in DMSO (150 ml per gel) for 3 h on a slow shaker followed by soaking in distilled water for 1h. Gels were then placed on a sheet of Whatman 3 MM paper and dried under vacuum and heat using a BIO-RAD gel slab dryer.

Fluorography was carried out at -80°C using Fuji Rx medical X-ray film (Bonner and Laskey, 1974; Laskey and Mills, 1975).

E. Purification of Proteins for Peptide Mapping

Radiolabelled proteins were fractionated by SDS poly-acrylamide gel electrophoresis on 10% slab gels. Either dansylated proteins of comparable mobility to the translation products or ^{35}S-labelled translation products (detected by autoradiography of the wet gels for 16 - 20 h at 4°C) were used as markers to locate the position of the proteins labelled with ^{3}H. The bands were excised and protein recovered by electroelution using a method adapted from the discontinuous buffer system of Laemmli (1970).

Gel slices containing each sample were equilibrated for 30-40 min at room temperature in 1 ml of 62.5 mM Tris-HCl, pH 6.8, 1 mM EDTA, 0.1%SDS, and placed in a 1 ml disposable plastic syringe (internal diameter 5mm), immediately above a 10 mm plug of 1.5% agarose in 0.125 M Tris-HCl, pH 6.8, 1 mM EDTA, 0.1% SDS. The gel slices were then sealed in place with about 0.5 ml of 1% agarose in 62.5 mM Tris-HCl, pH 6.8, 1 mM EDTA, 0.1% SDS, 0.05% bromophenol blue (a sufficient volume to just cover them). The bottom of a 1.5 ml Eppendorf tube was cut off (care being taken to ensure a perfectly smooth cut), a small piece of dialysis membrane was glued on to cover the

opening at the bottom of the tube using cyanoacrylate ester glue (Symons, 1978), and about 200µl of 0.125 M Tris-HCl, pH 6.8, 1 mM EDTA, 0.1% SDS was added. A hole with diameter equal to the outer diameter of the 1 ml syringe (about 8 mm) was cut in the lid of the Eppendorf tube, allowing the syringe to be just pushed through the lid until the 1.5% agarose plug just made contact with the buffer over the dialysis membrame. The syringe was then glued in place to provide an airtight seal.

The upper tank (cathode) contained 25 mM Tris-glycine, pH 8.3, 0.1% SDS, and the lower (anode) 0.125 M Tris-HCl, pH 6.8, 1 mM EDTA, 0.1% SDS. Electroelution was carried out at 2-3 mA/tube (150 - 200 V) for about 1 h after all the blue dye had entered the Eppendorf tube (about 3 h total).

Twentyfive µg of bovine serum albumin (BSA) was added as carrier to each Eppendorf tube and the sample dialysed for 2 -3 h against 0.125 M Tris-HCl, pH 6.8, 1 mM EDTA, 0.1% SDS, to remove glycine which would otherwise be precipitated at the next step. After addition of four volumes of acetone, the protein samples were precipitated overnight at -15°C. In control experiments, overall recovery was found to be essentially complete by comparison of the amount of radio-activity eluted from gel slices with that present in other gel slices containing identical protein samples.

F. Peptide Mapping using Proteases

Labelled protein samples (20,000 - 40,000 cpm of ^{3}H- or 5,000 - 10,000 cpm of ^{35}S-labelled protein) were digested with either the Staphylococcus aureus V8 protease (Miles), trypsin or chymotrypsin (both from Sigma). Reactions were carried out with the V8 protease (500ug/ml) trypsin (1mg/ml) or chymotrypsin (200 mg/ml) in a reaction volume of 40 ul, containing 0.125 M Tris-HCl, pH 6.8, 1 mM EDTA, 10% (v/v) glycerol, 2% SDS and 25µg of unlabelled BSA, at 37°C for 3 h. After addition of 0.05% bromophenol blue, the digests were fractionated on 18% polyacrylamide SDS slab gels (Laemmli, 1970), before fluorography.

G. Non-equilibrium pH Gradient Electrophoresis

The labelled (^{35}S) translation products were purified from SDS gels and

recovered by acetone precipitation (which also removed SDS). Samples (10,000 cpm) of each were electrophoresed on a 12x10x0.1 cm, 4% acrylamide 10.2% bisacrylamide gel, containing 5% (v/v) 3-10 ampholines (LKB, Sweden) and 9M urea, as described for NEPHGE by O'Farrell et al., (1977). The gel was then rinsed with water and dried under vacuum, before being soaked briefly in a solution of 25% (w/v) napthalene, 1% (w/v) PPO in ether, and allowed to dry. Fluorography was carried out at -80°C as described above.

RESULTS

A. Translation Products from CMV RNAs 1 and 2

RNAs 1 and 2 from all four strains of CMV functioned as efficient monocistronic mRNAs using the *in vitro* translation conditions given above (Fig.2.1). The sizes of the full length translation products of each RNA are summarized in Table 2.1. Translation products of very similar sizes were observed for each RNA, among the four strains. Note that T-CMV RNA 1 was heavily cross-contaminated with RNA 2 (Fig. 2.1, tracks 5,6) and that M-CMV RNA 1 was slightly contaminated with RNA 2 (Fig. 2.1, tracks 7,12).

For each virus, RNA 2 encoded the largest translation product. These had identical mobilities for Q- and T- CMV (Fig. 2.1, tracks 2,6). The P- and M- CMV RNAs 2 translation products were slightly smaller (Fig. 2.1, tracks 4,11,12). The deletion in the full-length translation product of P-CMV RNA 2 appears to be reflected in the faint, shorter polypeptides visible in Fig.2.1, (tracks 2 and 4). Since these polypeptides are presumed to arise from early termination of translation, this observation is in agreement with the suggestion (P.Molloy, personal communication) that this deletion is from the amino-terminal part of the M_r 110,000 gene on Q-CMV RNA 2. Among the RNA 1 translation products, only that from T-CMV, (M_r 105,000, see Fig. 2.1, track 5) differed in size from that encoded by Q-CMV RNA 1 (Table 2.1).

The V8 protease from *S. aureus* was used to digest the RNA1 and 2 translation products from Q- and P-CMV. Comparison of the resulting peptide fragments showed that the RNA 1 translation products (which have the same

FIGURE 2.1.

In vitro translation products of RNAs 1 and 2 from Q (tracks 1, 2, 7, 10), P (tracks 3,4, 8, 11), T (tracks 5, 6) and M (tracks 9, 12) strains of CMV. RNA 1 translation products are shown in tracks 1, 3, 5, 7, 8, and 9; RNA 2 translation products are shown in tracks 2, 4, 6, 10, 11 and 12. All RNAs were purified by one or two cycles of gel electrophoresis and translated as described in Methods, using ^{3}H-leucine (tracks 1 to 6) or ^{35}S-methionine (tracks 7 to 12). The products were analysed on two different 12% polyacrylamide SDS slab gels (one for tracks 1 to 6, the other for tracks 7 to 12). 'E' refers to an endogenous lysate polypeptide detected only with ^{35}S-methionine. The molecular weights of the Q-CMV translation products are indicated.

STRAIN: Q P T
RNA: 1 2 1 2 1 2
O –
$M_r \times 10^{-3}$
110 –
95 –
1 2 3 4 5 6

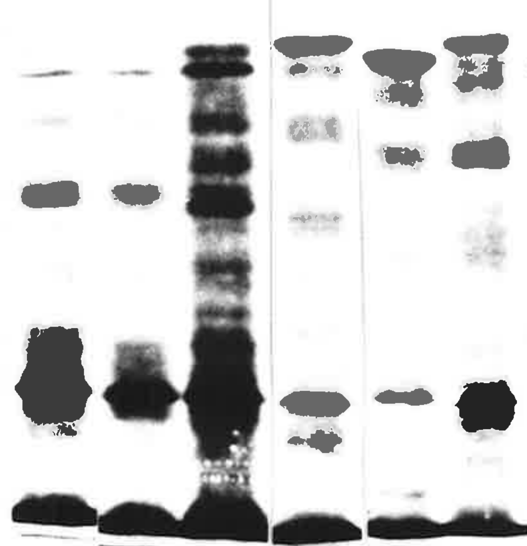

Q P M Q P M
1 2
–O
$M_r \times 10^{-3}$
– 110
– 95
– E
7 8 9 10 11 12

TABLE 2.1

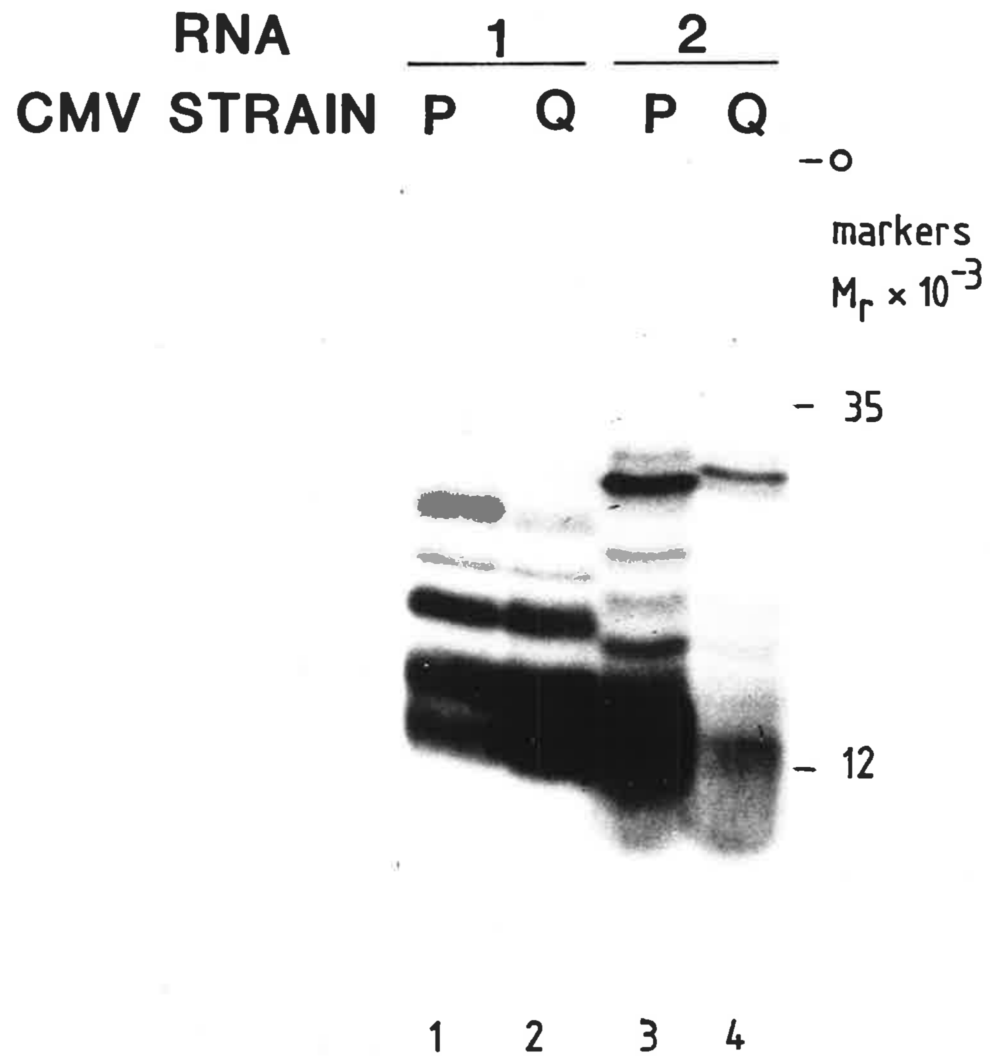
FULL-LENGTH TRANSLATION PRODUCTS FROM FOUR CMV STRAINS

Strain	Molecular weight ($x10^{-3}$) of full length translation product from		
	RNA 1	RNA 2	RNA 3
Q	95	110	36.7[a]
P	95	100	37
T	105	110	40
M	95	105	40

a – Calculated from nucleotide sequence (Gould and Symons, 1982).

FIGURE 2.2.

V8 proteolytic digests of translation products from RNAs 1 (tracks 1, 2) and 2 (tracks 3,4) of CMV strains P (tracks 1, 3) and Q (tracks 2, 4). Translation products (^{35}S-labelled) were purified as described in Methods and aliquots (10,000 cpm) digested with V8 protease. Electrophoresis of the peptides was on a 18% polyacrylamide SDS gel.


RNA
1
2
CMV STRAIN
P
Q
P
Q
-o
markers
$M_r \times 10^{-3}$
- 35
- 12
1
2
3
4

mobility) produce almost indistinguishable patterns, (Fig. 2.2 tracks 1,2). The peptide patterns from the RNA 2 translation products were also very similar, with some differences among the larger peptides (Fig. 2.2, tracks 3,4), possibly reflecting the size difference between the translation products.

B. Translation products from CMV RNA 3.

CMV RNA 3 also proved to be an efficient monocistronic mRNA, under the *in vitro* translation conditions described. Attempts to obtain translation of the coat protein gene on RNA 3 by prior denaturation with CH_3HgOH as described by Payvar and Schimke (1979) were unsuccessful (data not shown). As shown in Fig. 2.3A, RNAs 3 from Q- and P-CMV yield *in vitro* translation products of almost identical mobility on SDS gels (track 1,2). The size of this translation product was calculated from the nucleotide sequence to be 36,700 (p 36.7) (Gould and Symons, 1982). The translation products of T- and M- CMV RNA 3 were both about M_r 3,000 larger (tracks 3,4; Table 2.1). The RNA 3 preparations from T- and M- CMV were contaminated with fragments from the larger RNAs, but the identities of their translation products were confirmed by subsequent peptide mapping.

The RNA 3 translation products were recovered from the SDS gels as described above and analysed for overall charge by non-equilibrium pH gradient electrophoresis (NEPHGE). As shown in Fig. 2.2B, all translation products were basic, with the smallest (from Q-CMV, track 5) being the most basic. That from P-CMV was only slightly less basic, whereas those from T- and M-CMV were similar (tracks 7,8), but both less basic than either of the Q- or P-CMV proteins. These observations suggest that the extra amino acids present in the larger translation products carry an overall acidic charge. All the translation products were resolved into several differently migrating forms, for reasons which are not known but possibly, due to differential modifications, since the proteins within each track appear to be related.

C. Partial proteolytic digestion of RNA 3 translation products.

The three proteases used to compare the RNA 3 translation products from the four strains of CMV were the *Staphylococcus aureus* V8 protease (which

FIGURE 2.3.

In vitro translation products of RNA 3 from Q (tracks 1, 5), P (tracks 2, 6), T (tracks 3, 7) and M (tracks 4,8) strains of CMV. The RNAs were purified and translated as described in Methods. The products were first analysed by 12% polyacrylamide SDS slab gel electrophoresis (A). Note that the T- and M-CMV RNA 3 preparations were contaminated with other RNAs. These translation products (arrowed) were then purified and analysed by NEPHGE (see text) (B).

RNA 3 TRANSLATION PRODUCTS

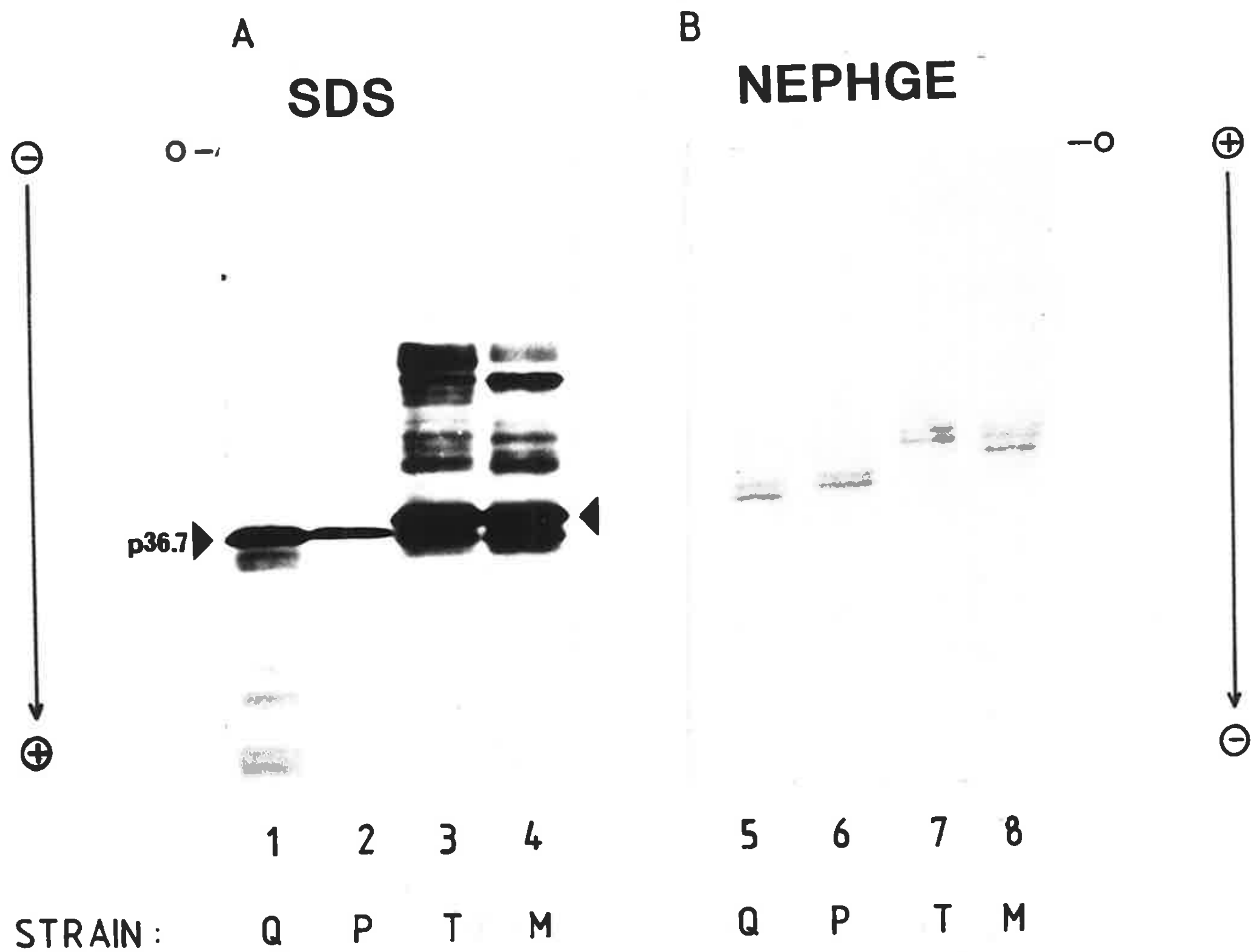
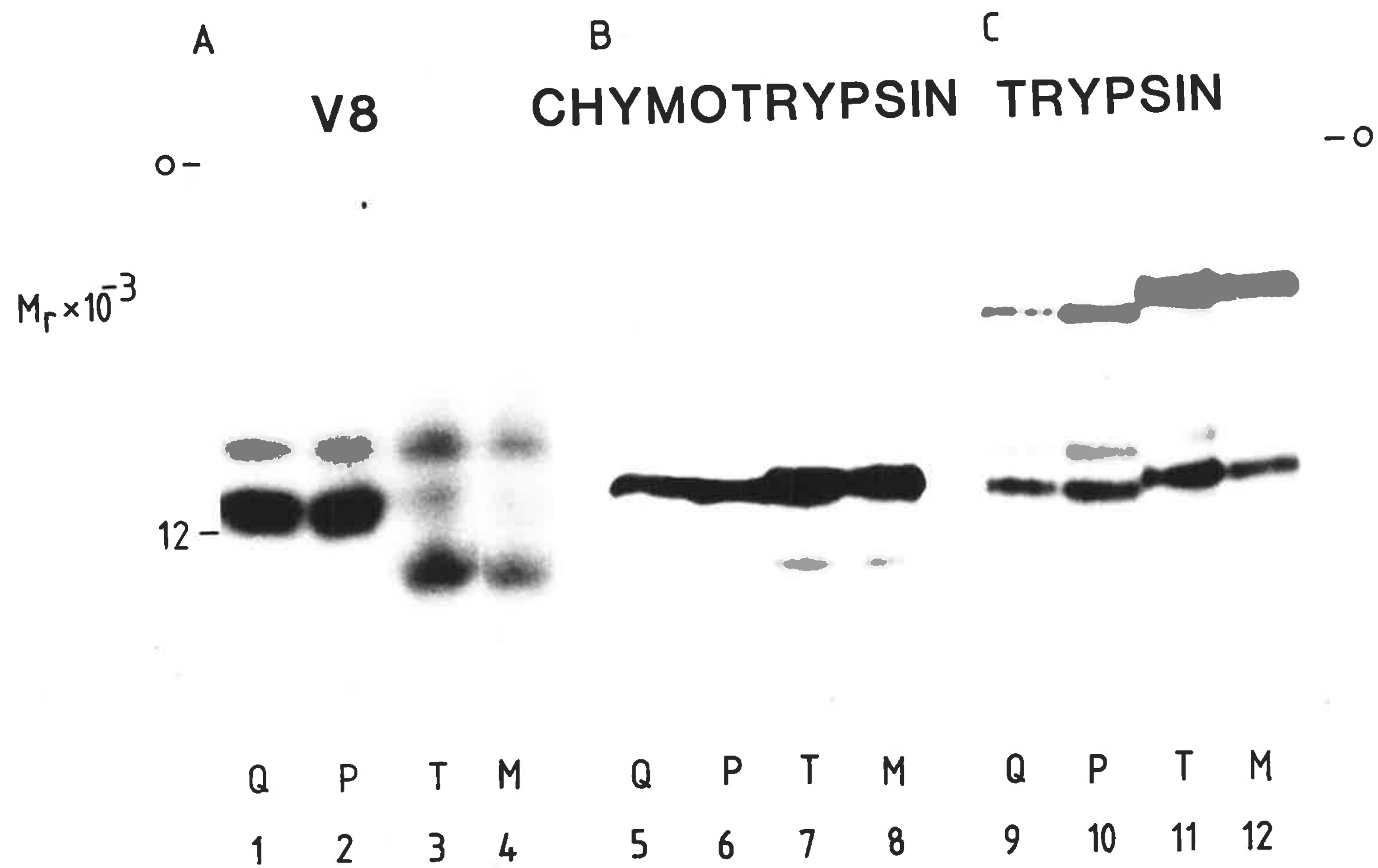


FIGURE 2.4

Proteolysis of RNA 3 translation products from four strains of CMV using V8 protease (tracks 1 to 4), chymotrypsin (tracks 5 to 8) or trypsin (tracks 9 to 12). Q-MCV: tracks 1, 5, 9; P-CMV: tracks 2, 6, 10; T-CMV: tracks 3, 7, 11; M-CMV: tracks 4, 8, 12. All translation products were purified and digested as described in Methods. Electrophoresis was on a 18% polyacrylamide SDS gel. Molecular weights were derived from coelectrophoresis of marker proteins.

PROTEOLYSIS OF RNA 3 TRANSLATION PRODUCTS



cleaves at glutamate residues), trypsin (cleaves at arginine and lysine) and chymotrypsin (cleaves at phenylalanine, tryptophan and tyrosine). All three proteases confirmed that the RNA 3 translation products from Q- and P-CMV, and from T- and M-CMV form two distinct, though related groups (Fig. 2.4). It was not possible to distinguish between the RNA 3 translation products from Q- and P-CMV, nor between those from T- and M-CMV, using any of these proteases.

The V8 protease cleaved the translation products from T- and M-CMV (Fig. 2.4, tracks 3,4) into some smaller fragments than those from Q- and P-CMV (tracks 1,2) whereas trypsin yields larger peptides from strains T- and M- (tracks 11,12) than from strains Q and P (tracks 9,10). The chymotrypsin peptides are less conclusive (tracks 5 to 8). The observations suggest that the additional amino acid residues in the larger translation products possibly contain acidic rather than basic residues, in line with the findings mentioned in section B above. In no case, however, could peptide fragments be definitely related to the amino acid sequence deduced from the RNA 3 nucleotide sequence. From their sizes, all major peptides appeared to be partial digestion products.

DISCUSSION

The genomic RNAs (1,2 and 3) of Q-CMV each yielded a single major *in vitro* translation product (of M_r 95,000, M_r 110,000 and M_r 35,000 respectively) whose size varied only slightly among the four CMV strains studied. This suggests that the genome organization of CMV is highly conserved. In fact, the genomic RNAs have been found to vary little in size (Mossop and Francki, 1977, 1979; Mossop, 1978).

The translation products of the three RNAs of two strains, Q- and P-CMV were compared by partial proteolytic digestion. Corresponding RNAs from each strain gave translation products with very similar peptide maps. The proteolytic digests of the RNA 3 translation products of the other two strains, T- and M-CMV, showed these to form a distinct group, which is closely related to the Q- and P-CMV strains.

The extra peptide sequences in the RNA 3 translation products from T- and

M-CMV (compared to Q-CMV) have shown a number of characteristics which suggest that they consist of more acidic than basic amino acid residues. Assuming that this insertion arose from a simple nucleotide sequence duplication, the most likely source of the extra sequence is the carboxy terminal region of this gene. The 27 codons here encode five acidic and only one basic amino acid (Gould and Symons, 1982) and would, if duplicated, increase the molecular weight of the protein by the required 3,000.

Gonda and Symons (1978) studied the sequence relationships among corresponding RNAs of Q-, P- and M-CMV by cDNA hybridization. They found that the four RNAs of P-CMV were indistinguishable from the corresponding Q-CMV RNAs. This is consistent with the present peptide mapping results. In contrast, only partial sequence homology was observed between the RNAs of Q-CMV and those of M-CMV. A further comparison of the RNAs of these four strains of CMV involved the determination of the nucleotide sequences of the 3'-terminal approximately 200 residues (Symons, 1979; Wilson, Barker and Symons, unpublished). All these sequences showed the conserved 3' tRNA-like structure, in which about half the nucleotides are conserved. Q- and P-CMV RNAs yielded almost identical sequences, as did those of T- and M-CMV, but the latter sequences differed considerably from those found for Q- and P-CMV. The four CMV strains thus fall into two groups.

The conservation of genome organisation in CMV is consistent with the finding that pseudo-recombinants could be constructed using purified genomic RNAs from several strains of CMV (Mossop and Francki, 1979; Rao and Francki, 1981), Mossop and Francki (1977) were able to construct pseudorecombinants between RNAs 1 and 2 (both from the same strain) and RNA 3 of Q- and M-CMV.

Pseudo recombinants have also been constructed using RNAs 1 and 2 from Q-CMV and RNA 3 of V-TAV. V8 protease mapping of the RNA3 translation products from N- and V-TAV showed that they were indistinguishable from those of Q- and P-CMV (data not given). The RNA 1 and 2 translation products from both strains of TAV co-migrate on SDS gels (at M_r 100,000) and so were not separated for peptide mapping, but V8 protease digests of

the mixed proteins showed very similar patterns for these two, closely related strains (data not shown). This is consistent with the finding of extensive 3'- terminal nucleotide sequence homology among the genomic RNAs of these two TAV strains (Wilson and Symons, 1981). Hybridization analysis with cDNA has shown that V-TAV RNAs 1, 2 and 3 have partial sequence homology with the corresponding RNAs of Q-CMV (Gonda and Symons, 1978), although to a lesser degree than M-CMV RNAs do.

The most important result from this work has been to establish that the translation products from RNAs 1, 2 and 3 do vary slightly in size, among the Q, P and T strains. Since these strains all replicate in cucumber, they could be used to investigate the possible involvement of viral gene products in viral RNA replication. This is described in Chapter 3.

CHAPTER 3

HIGHLY PURIFIED CMV-INDUCED RNA DEPENDENT RNA POLYMERASE DOES NOT CONTAIN ANY OF THE FULL LENGTH TRANSLATION PRODUCTS OF THE GENOMIC RNAs.

INTRODUCTION

As considered in Chapter 1, three polypeptides of M_r 110,000, 100,000 and 35,000 are unique to partially purified preparations of RNA replicase from Q-CMV-infected cucumber. They have approximately the same mobility on SDS-polyacrylamide gels as the full length translation products of Q-CMV RNAs 2, 1 and 3, respectively (Kumarasamy and Symons, 1979a; Gill et al., 1981). In order to test whether any of these RNA replicase components are actually viral RNA translation products, the first approach was to compare the mobilities of these polypeptides in partly purified RNA replicase preparations from cucumbers infected with different strains of CMV. Should any RNA replicase polypeptides be of viral origin, they would be expected to show variation in their mobilities on SDS gels similar to that found among the translation products of the RNAs from these different strains. Altogether, the three CMV strains Q, P and T were sufficient to test this proposition. The electrophoretic mobilities of T-CMV RNA 1 and 3 translation products differ from the corresponding Q-CMV products, as does the RNA 2 translation product of P-CMV from that of Q-CMV RNA 2. Since it was possible, if unlikely, that post-translational modifications might mask the mobility differences among the primary translation products, the enzyme polypeptides were next directly compared to the translation products by peptide mapping.

MATERIALS AND METHODS

A. Extraction and purification of RNA replicase

The RNA replicase used for comparison with the translation products of CMV RNAs was solubilized from particulate extracts of CMV-infected plants harvested 8 to 10 days after infection and purified by column chromatography on phospho-cellulose and poly (C)-cellulose as in Gill et al., (1981), and Gill (1983).

Protein in the enzyme fractions was labelled by reductive methylation with $^{3}H-KBH_4$ in the presence of 40% formamide, essentially as described by Kumarasamy and Symons (1979b). For some experiments, poly (C)-cellulose preparations of RNA replicase were labelled with ^{125}I (Amersham) essentially as described in "Radioiodination Techniques" (Amersham Review 18).

B. Peptide mapping CNBr cleavage

^{3}H-leucine labelled translation products of CMV RNAs and the ^{3}H-labelled polypeptides of the RNA replicase were electroeluted and precipitated as above. After drying, the protein samples (50,000 cpm of each) plus 50 µg BSA were dissolved in 20 µl of 75% formic acid. Digestions were started by adding 5 µl of CNBr (0.2 g/ml in 75% formic acid) and carried out under nitrogen in Eppendorf tubes for 60 h at 25°C in the dark. The tube contents were then diluted with 1 ml of water and lyophilized; this process was repeated four times. The digestion products were fractionated on 18% polyacrylamide SDS slab gels before fluorography.

RESULTS

A. Comparison of Polypeptides of CMV-induced particulate RNA replicase with in vitro translation products of CMV RNAs 1, 2 and 3

A direct comparison of electrophoretic mobilities of polypeptides of RNA replicase purified from Q-CMV-infected cucumber plants and of the *in vitro* translation products of the three largest RNAs of Q-CMV was made on SDS-polyacrylamide slab gels (Figs. 3.1A and B). A comparison of the gel pattern of poly(C)-cellulose step purified Q-CMV RNA replicase and the Q-CMV RNAs 1 and 2 translation products (Fig. 3.1A tracks 1, 2 and 3, respectively) with a mixture of RNA replicase with the translation products of Q-CMV RNA 1 (Fig. 3.1; track 4) and of Q-CMV RNA 2 (Fig. 3.1; track 5) shows clearly that the M_r110,000 polypeptide of CMV RNA replicase was identical in electrophoretic mobility to full length translation product of Q-CMV RNA 2. However, the M_r 100,000 polypeptide of RNA replicase was

FIGURE 3.1

Comparison of CMV RNA replicase polypeptides with *in vitro* translation products of CMV RNAs 1, 2 and 3 by SDS-polyacrylamide gel electrophoresis.

The *in vitro* translation products of three largest RNAs of Q-CMV and the ^{3}H-labelled protein samples of phospho-cellulose and poly(C)-cellulose step purified particulate Q-CMV RNA replicase were prepared and analysed by SDS-polyacrylamide slab gel electrophoresis as described in Chapter 2. The positions of marker proteins are given on the left-hand side.

A. 10% polyacrylamide-SDS slab gel:

(1) Poly(C)-cellulose step Q-CMV RNA replicase;

(2) Translation products of Q-CMV RNA 1;

(3) Translation products of Q-CMV RNA 2;

(4) Mixture of 1 & 2;

(5) Mixture of 1 & 3

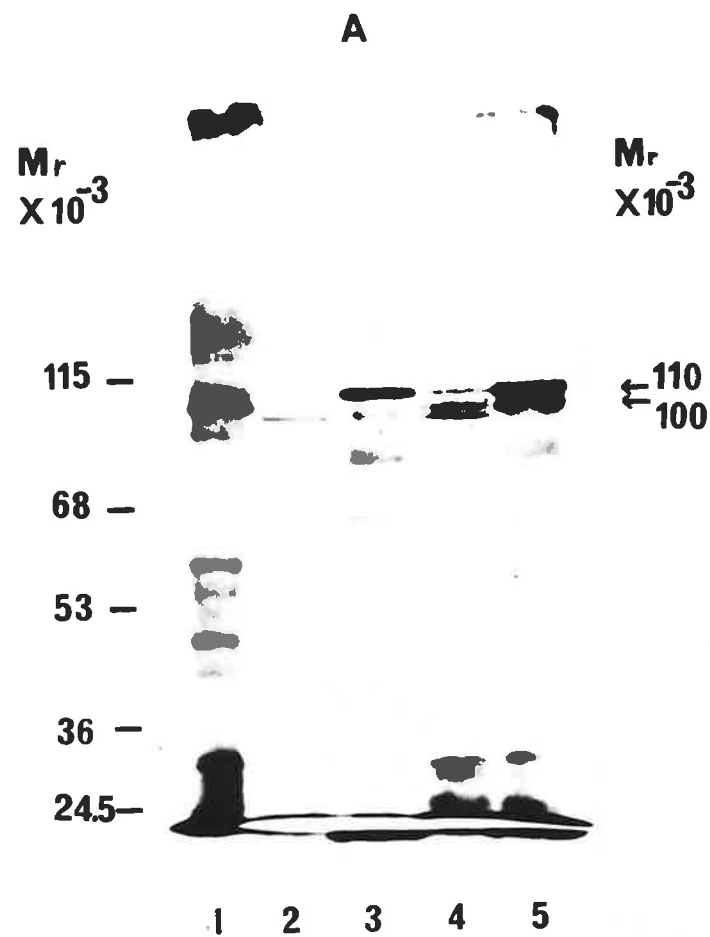
B. 12% polyacrylamide-SDS slab gel:

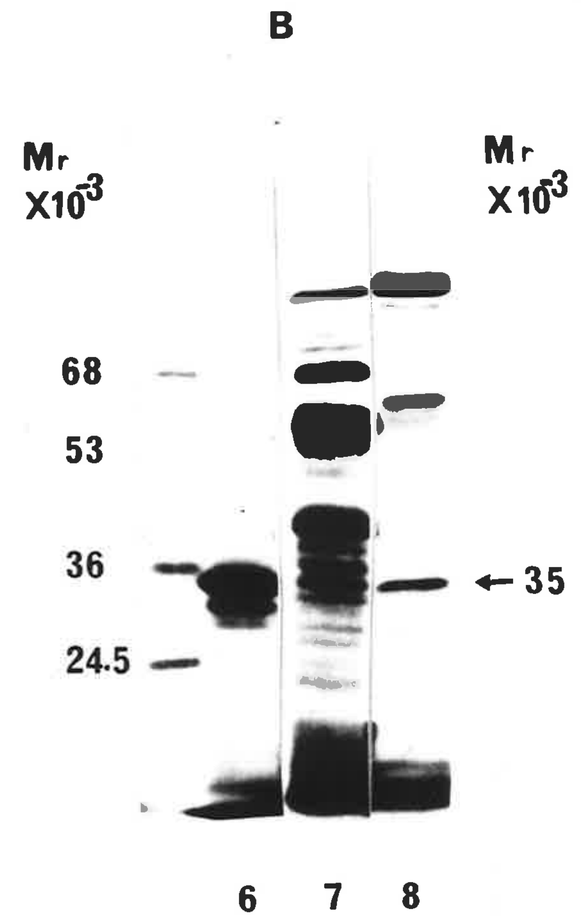
(6) Translation products of Q-CMV RNA 3;

(7) Phosphocellulose step Q-CMV RNA replicase;

(8) Poly(C)-cellulose step Q-CMV RNA replicase.

(From Gill et al., (1981)).


A
Mr
X10-3
115
68
53
36
24.5
Mr
X10-3
110
100
1
2
3
4
5


B
Mr
X10-3
68
53
36
24.5
Mr
X10-3
35
6
7
8

significantly larger (M_r about 5,000) than the full length translation product of Q-CMV RNA 1.

Similarly a comparison of the polypeptide pattern of phosphocellulose and poly (C)-cellulose step purified enzyme samples of Q-CMV RNA replicase (Fig. 3.1B; tracks 7 and 8, respectively) with the in vitro translation products of Q-CMV RNA 3 (Fig. 3.1B; track 6) shows that the electrophoretic mobility of the M_r 35,000 polypeptide component of RNA replicase, unique to CMV-infected plants, was similar (but not identical) to the in vitro full length translation product of Q-CMV RNA 3.

B. Different strains of CMV induce RNA replicase with identical polypeptide components.

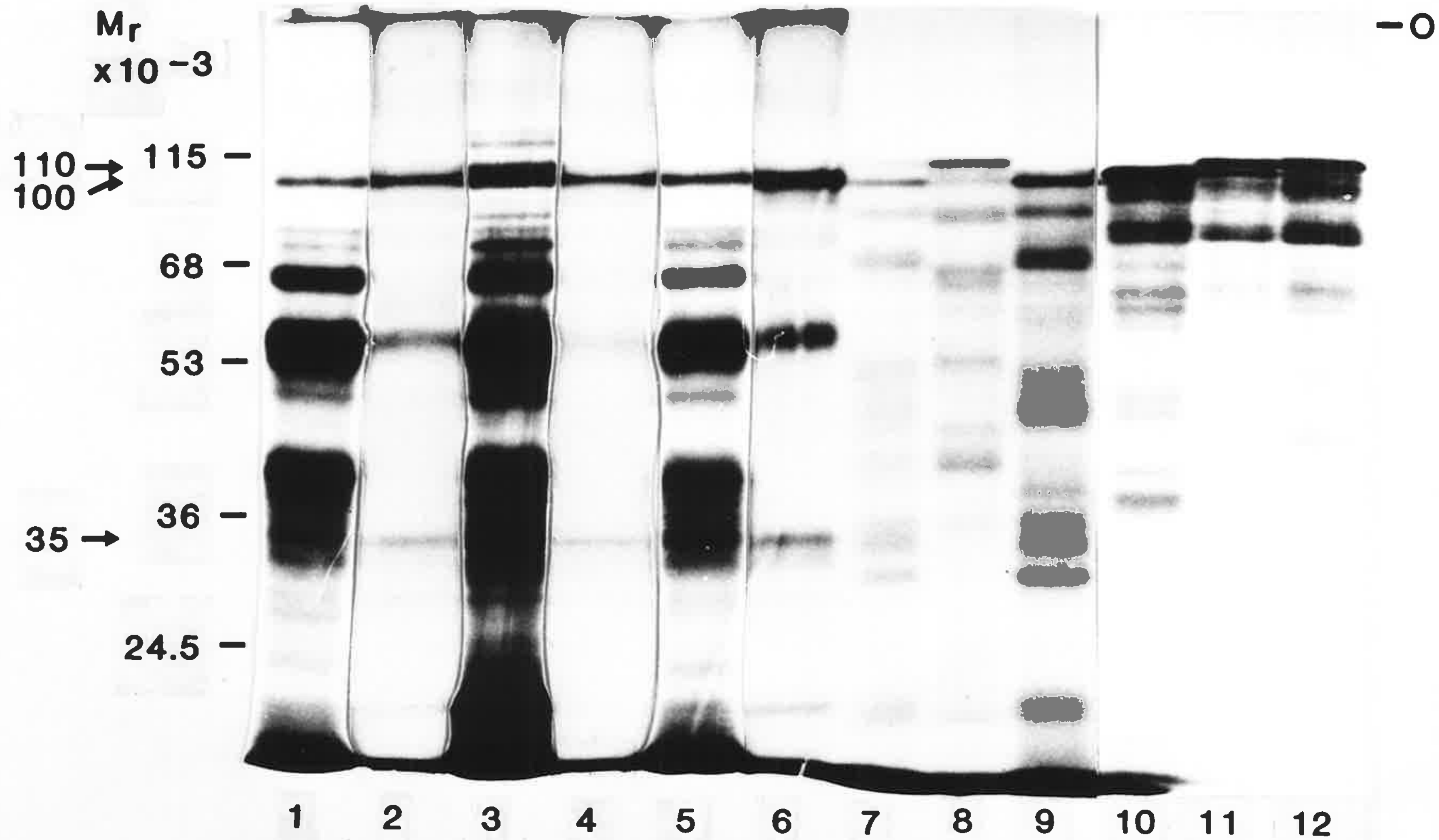
The translation products of P-CMV RNA 1 and T-CMV RNA 2 showed identical electrophoretic mobilities to the corresponding products of the same two Q-CMV RNAs (Fig. 3.2, tracks 7 and 9, 8 and 12). However, the translation product of P-CMV RNA 2, with an electrophoretic mobility equivalent to about M_r 105,000, was slightly smaller than that from Q-CMV RNA 2 (Fig. 3.2, tracks 8 and 10), whereas the translation product of T-CMV RNA 1, with an electrophoretic mobility equivalent to about M_r 100,000 was slightly larger than that from Q-CMV RNA 1 (Fig. 3.2, tracks 7 and 11).

In contrast, the phosphocellulose RNA replicase preparations from cucumber plants infected with either Q-CMV, P-CMV or T-CMV all contained the two high molecular weight polypeptides with electrophoretic mobilities equivalent to M_r 110,000 and 100,000 (Fig. 3.2, tracks 1, 3 and 5). The variation in the intensity of the M_r 110,000 polypeptide relative to the M_r 100,000 polypeptide between the three strains is not considered important here since we have observed a similar variation in different enzyme preparations from Q-CMV infected plants (data not given). In the more purified enzyme preparations after poly(C)-cellulose, no electrophoretic mobility differences were observed among the polypeptides present (Fig. 3.2, tracks 2, 4 and 6).

The translation products of RNA 3 from the three strains of CMV also

FIGURE 3.2.

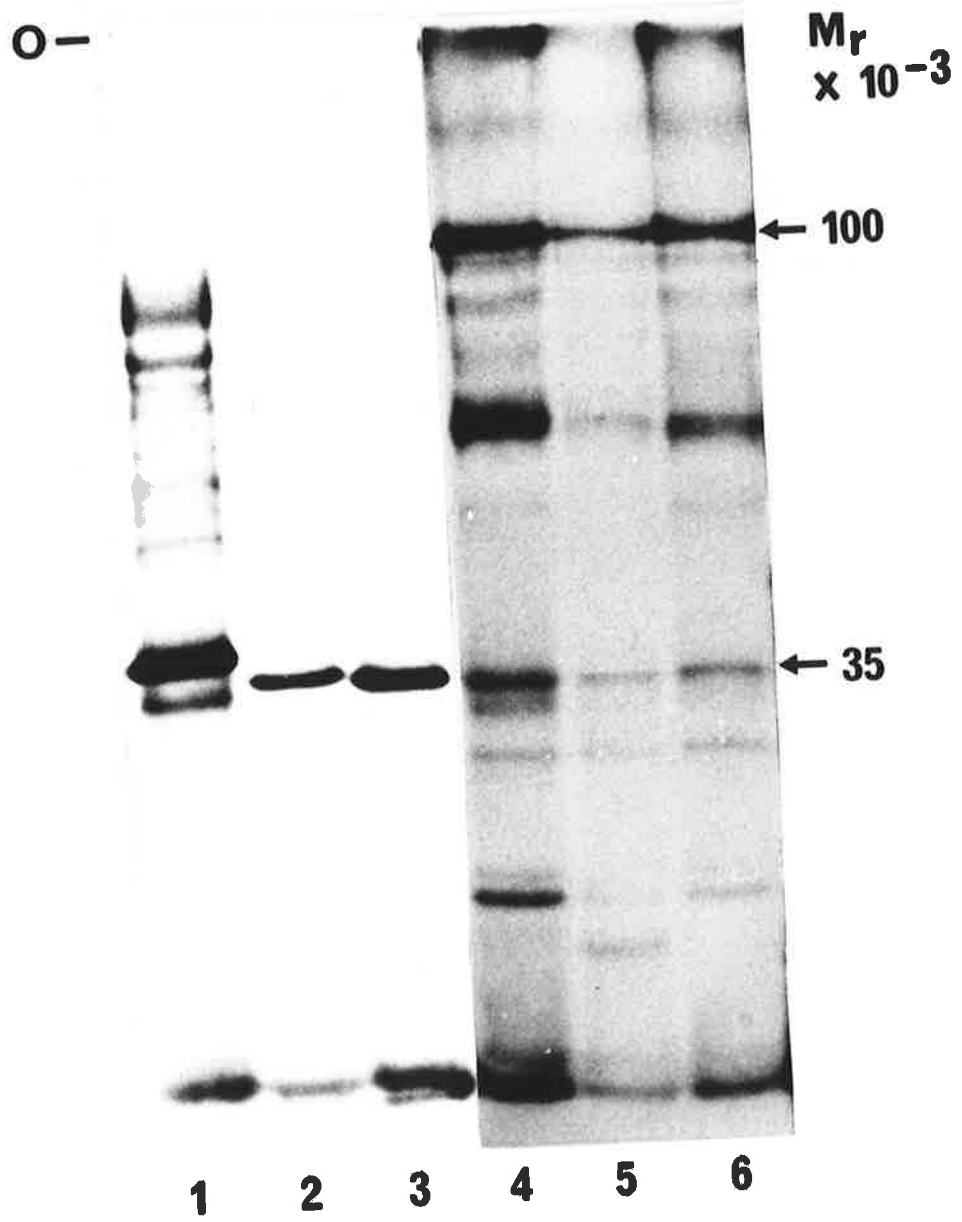
Electrophoresis on a SDS 12% polyacrylamide slab gel of RNA replicase preparations from cucumber plants infected with Q (tracks 1 and 2), P (tracks 3 and 4) and T (tracks 5 and 6) strains of CMV. Phosphocellulose step enzyme preparations were run in tracks 1, 3 and 5. These fractions were then further purified on poly(C)-cellulose columns and the resulting preparations run in tracks 2, 4 and 6. Protein in the fractions was labelled with $^{3}H-KBH_4$ as in Methods. Purified RNAs were translated at 20 µg/ml in the rabbit reticulocyte cell-free system in the presence of ^{3}H-leucine as described in Chapter 2. Translation products of Q-CMV RNAs 1 and 2 were run in tracks 7 and 8, those from P-CMV RNAs 1 and 2 in tracks 9 and 10, and those from T-CMV RNAs in tracks 11 and 12, respectively. Molecular weights were derived from the positions of ^{3}H-markers. Note that track 11 shows that the T-RNA 1 used for *in vitro* translation was contaminated with T-CMV RNA 2 (see track 12).



Mr
x10 -3
110
100
115
68
53
36
35
24.5
O
1
2
3
4
5
6
7
8
9
10
11
12

FIGURE 3.3.

Electrophoresis on a SDS 13% polyacrylamide gel of the *in vitro* translation products of RNA 3 from three strains of CMV (track 1: T-CMV; track 2: P-CMV; track 3: Q-CMV) and RNA replicase preparations from cucumber plants infected with T-CMV (track 4), P-CMV (track 5) and Q-CMV (track 6). The enzyme preparations were purified by chromatography on phosphocellulose and poly(C)-cellulose and the protein in the enzyme fractions labelled with ^{3}H-KBH_4 as described in Methods. Purified RNA samples were translated at 40 µg/ml in the presence of ^{3}H-leucine in the reticulocyte cell-free systems as in Chapter 2. Note that track 1 indicates that the T-CMV RNA 3 was contaminated with fragments of larger RNAs.


O–
M_r
x 10^{-3}
← 100
← 35
1
2
3
4
5
6

varied in their electrophoretic mobilities, with that from T-CMV being larger than those from P-CMV and Q-CMV which had similar mobilities (Fig. 3.3, tracks 1, 2 and 3); its mobility was equivalent to M_r 38,000 compared with M_r 35,000 for the other two products. On the other hand, there was no corresponding variation in the RNA replicase polypeptide of M_r35,000 in enzyme preparations from plants infected with the three strains of CMV (Fig. 3.3, tracks 4, 5 and 6).

These results therefore show that none of the full length translation products of CMV RNAs 1, 2 or 3 was present in the highly purified RNA replicase.

C. Peptide mapping confirms the absence of full length translation products from the highly purified RNA replicase.

The peptides generated by digestion with the *S. aureus* V8 protease of the largest RNA replicase polypeptides (M_r 110,000 and M_r 100,000) and of the translation products of Q-CMV RNAs 1 and 2 had electrophoretic mobilities from about M_r 10,000 to about M_r 50,000 (Fig. 3.4). The peptide patterns were sufficiently different to indicate that the two high molecular weight RNA replicase polypeptides were not translation products of CMV RNAs 1 and 2. Some bands in the peptide map of the RNA 1 translation product were like some from the enzyme polypeptides, but the larger ones were not.

It is of interest that the M_r 110,000 and M_r 100,000 polypeptides of the RNA replicase had several bands with similar mobilities in their peptide maps (Fig. 3.4, tracks 1 and 2). Further, peptide mapping by V8 protease digestion of two other enzyme polypeptides of M_r 75,000 and M_r 65,000 indicated that the smaller polypeptide was derived by partial breakdown of the larger one (data not given). These two polypeptides also gave V8 protease bands corresponding in mobility to some of those found in the peptide maps of the M_r 110,000 and the M_r 100,000 enzyme polypeptides. These results suggest that the enzyme polypeptides undergo breakdown during enzyme preparation, storage or labelling.

The polypeptides used for Fig. 3.4 were labelled differently; the

FIGURE 3.4.

Peptide mapping using S. aureus V8 protease of the M_r 110,000 and M_r 100,000 polypeptides of the Q-CMV RNA replicase (tracks 1 and 2, respectively) and of the full length translation products of Q-CMV RNAs 1 and 2 (tracks 3 and 4, respectively). The translation products and the enzyme polypeptides purified by chromatography on phosphocellulose and poly(C)-cellulose were all ^{3}H-labelled. After recovery from SDS polyacrylamide gels and digestion with protease as in Chapter 2, all samples were electrophoresed on a SDS 18% polyacrylamide gel. Molecular weights were derived from co-electrophoresis of marker proteins.

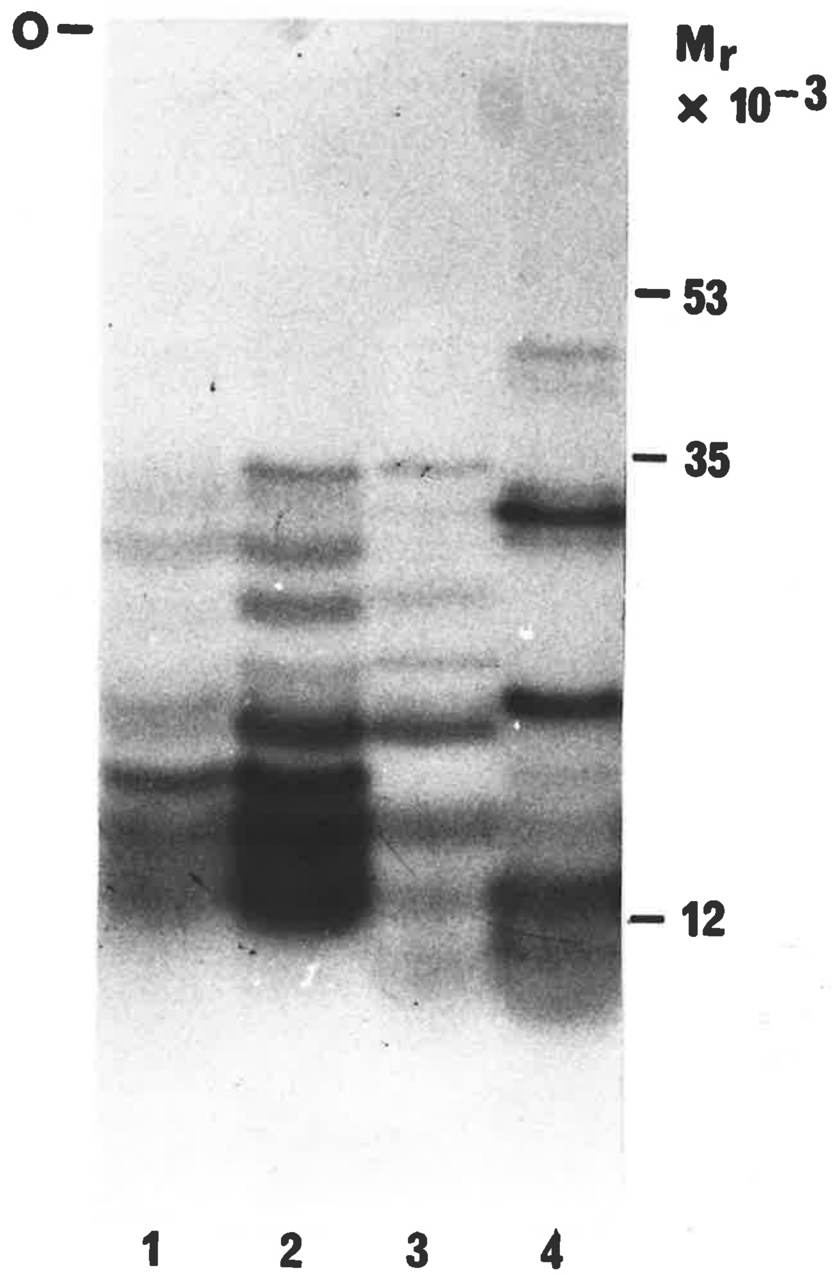
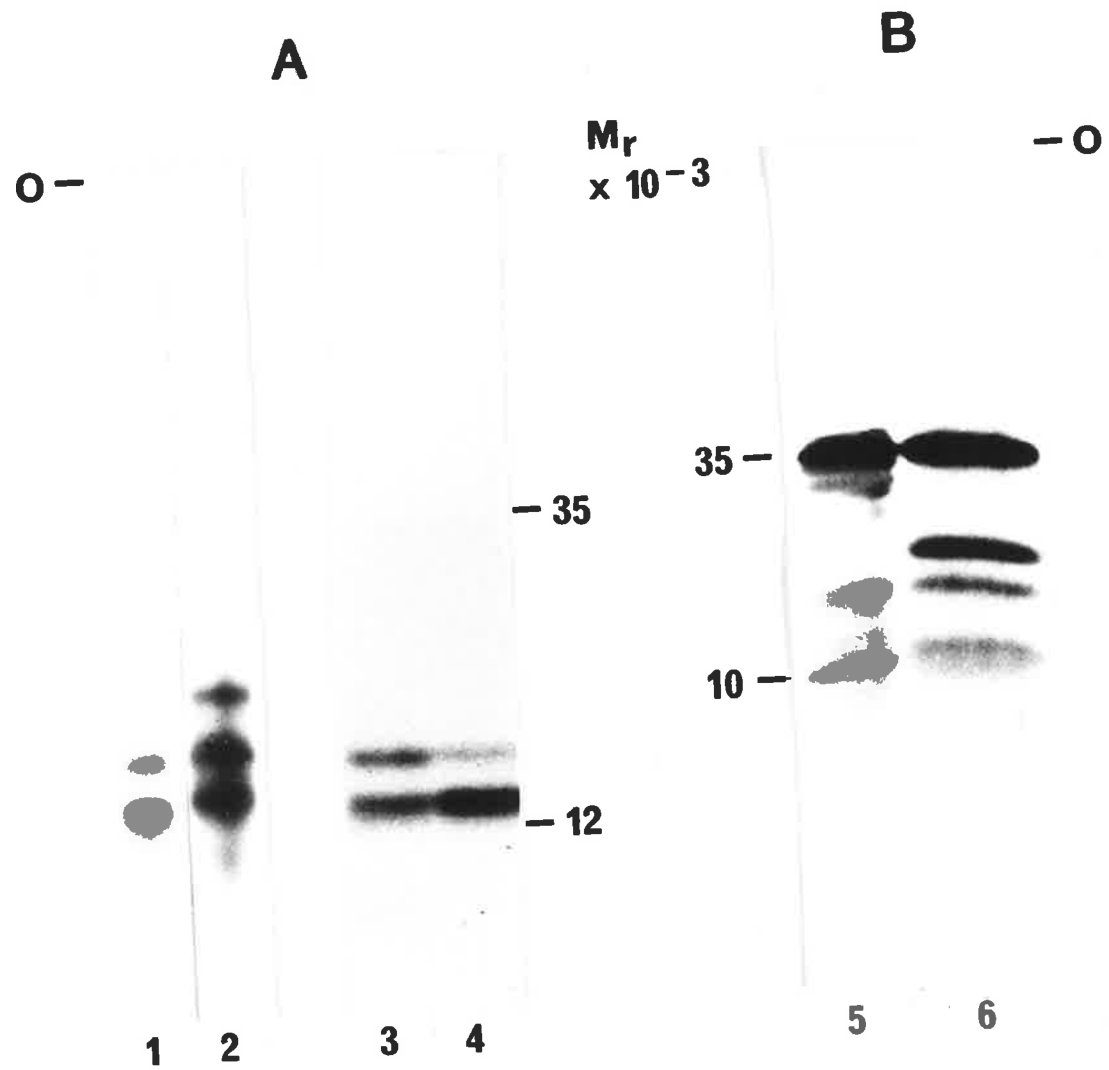

O
M_r
$\times 10^{-3}$
53
35
12
1
2
3
4

FIGURE 3.5.

Peptide mapping of the translation product of Q-CMV RNA 3 and the M_r 35,000 polypeptide of the Q-CMV RNA replicase. RNA 3 was translated in the reticulocyte cell-free system in the presence of either ^{3}H-leucine or ^{35}S-methionine. The RNA replicase preparations were purified by chromatography on phosphocellulose and poly (C)-cellulose, before labelling with ^{3}H-KBH_4 or ^{125}I as described in Methods.

A. Gel slices containing ^{35}S-labelled translation products (track 1) and the ^{125}I-labelled M_r 35,000 enzyme polypeptide (track 2) were placed in the wells of a SDS 18% polyacrylamide gel and digested with 5 µg/well of S. aureus V8 protease during electrophoresis, as described by Cleveland et al. (1977). In another experiment to show that digestion of the translation product gives the same peptide map irrespective of the labelled amino acid, ^{3}H- and ^{35}S-labelled products of RNA 3 (tracks 3 and 4, respectively) were electroeluted, digested with V8 protease and run on a SDS 18% polyacrylamide gel.

B. Cleavage with CNBr of the ^{3}H-labelled translation product of Q-CMV RNA 3 (track 5) and the M_r 35,000 polypeptide of Q-CMV RNA replicase purified by chromatography on phosphocellulose and poly(C)-cellulose (track 6). Protein samples were recovered by electroelution (Chapter 2), and cleaved with CNBr as described in Methods. The samples were then run on a SDS 18% polyacrylamide gel. Note that the digestions are incomplete as appreciable amounts remained of undigested M_r 35,000 protein. Molecular weights were derived from co-electrophoresis of ^{3}H-markers.


A
B
O
M_r
x 10^{-3}
O
35
12
35
10
1
2
3
4
5
6

translation products with ^{3}H-leucine and the enzyme polypeptides by reductive methylation with ^{3}H-KBH_4 to produce ϵ-N,N-^{3}H-dimethyl lysine (Rice and Means, 1971). However, comparison of the peptide maps of the translation products labelled with either ^{3}H-leucine or ^{35}S-methionine showed that there were no significant differences between them (data not shown). Hence, the different location of the radioactive label detected by fluorography, leucine or methionine for the translation products and lysine for the enzyme polypeptides, was not critical when comparing the peptide maps.

Further support for the non-identity of the two large enzyme polypeptides and the RNAs 1 and 2 translation products was obtained by peptide mapping of labelled polypeptides digested with CNBr (data not shown). However, these results did show that the M_r 110,000 and M_r 100,000 enzyme polypeptides had very similar CNBr peptide maps.

Peptide mapping of the M_r 35,000 RNA replicase polypeptide and the translation product of RNA 3 with the S. aureus V8 protease (Fig. 3.5A, tracks 2 and 1) confirmed that these two proteins were different. Although different amino acids were radiolabelled in the two samples (^{35}S-methionine in the translation products and ^{125}I-tyrosine in the enzyme polypeptide), the peptide maps of ^{3}H-leucine and ^{35}S-methinine labelled translation products of CMV RNA 3 were identical (Fig. 3.5A, tracks 3 and 4). Cleavage with CNBr of the ^{3}H-labelled M_r 35,000 polypeptide and of the ^{3}H-labelled RNA 3 translation product also showed different patterns (Fig. 3.5B, tracks 5 and 6).

DISCUSSION

Results presented in this chapter clearly show that none of the three RNA replicase polypeptides of M_r 110,000, M_r 100,000 and M_r 35,000 which are unique to enzyme preparations from CMV-infected plants were gene products of the CMV RNAs. Three different strains of CMV-induced RNA replicase showed no variation in the electrophoretic mobilities of these polypeptides although the in vitro translation products of each of the genomal RNAs of

CMV varied in mobility among the three strains. It was therefore concluded that these three enzyme polypeptides were host encoded, despite their similar electrophoretic mobilities on SDS polyacrylamide gels to the in vitro translation products of CMV RNAs. Peptide mapping studies using both S. aureus V8 protease and CNBr cleavage have confirmed these findings. In addition, polyacrylamide gel electrophoresis in a cationic acetic acid-urea-triton X-100 system (Bonner et al., 1980) showed that the translation products of CMV RNAs 1 and 2 did not comigrate with the M_r 100,000 and M_r 110,000 enzyme polypeptides, respectively (data not given). Under these conditions, polypeptide characteristics other than size, e.g. hydrophobicity and charge, affect mobility.

The copurification of the host-encoded M_r 100,000 polypeptide with RNA replicase activity indicated that this polypeptide is responsible for the enzyme activity. The M_r 110,000 polypeptide also copurified with the M_r 100,000 polypeptide but it was generally obtained in smaller amounts while its disappearance on storage of enzyme preparations did not affect enzyme activity significantly. The similar peptide maps of the two polypeptides suggested that they are structurally related. The M_r 110,000 polypeptide might therefore be the main, or even the only, in vivo form of this protein.

The finding that the full length translation products of the CMV genomal RNAs were not present in the highly purified RNA replicase raises the question of what roles these proteins play during the replication of viral RNAs since RNAs 1, 2 and 3 are all required for infection (Peden and Symons, 1973; Lot et al., 1974). The induction of very high levels of RNA replicase activity in extracts of CMV-infected plants, whereas healthy plants lack such activity altogether, indicates that the CMV gene products have an important role in the production of the enzyme which has long been considered to be multi-component (Clark et al., 1974; Kumarasamy and Symons, 1979a). The gene products of the CMV RNAs may therefore be responsible for the induction or regulation of host proteins involved in viral RNA replication

or in the assembly of the membrane-bound RNA replication structure. Another possibility, which must still be considered, is that the viral coded proteins are processed or modified before incorporation in the enzyme complex.

Our inability to recover the viral gene products using the solubilization and purification procedures described here suggests that they might be firmly membrane-bound within the RNA replication complex. If these gene products play a role in the regulation of template specificity, then the highly purified enzyme would lack this specificity due to the absence of the viral coded proteins. Such a lack of template specificity was indeed observed for the highly purified soluble form of the enzyme (Kumarasamy and Symons, 1979a).

The evidence for the in vivo form of the CMV-induced RNA replicase being a membrane-bound complex goes back to the work of May et al. (1970). Comparison of the crude particulate enzymes isolated from cucumber seedlings infected with either CMV or tobacco ringspot virus (TRSV) provided some evidence for virus specific differences in the attachment of RNA replicase activity to membranes (Peden et al., 1972). In the TRSV-infected plants, such activity was much more difficult to solubilize with $MgSO_4$ than was the case with CMV. Furthermore, no solubilization of active enzyme occurred during the RNA polymerase assay of the particulate fraction, again in marked contrast to the CMV-induced enzyme. Freezing and thawing the particulate fraction from TRSV-infected plants gave low and variable release of RNA replicase activity whereas 50% of the CMV-induced enzyme was released under the same conditions. Hence, properties of the particulate enzyme were markedly influenced by the infecting virus. It would be of considerable interest to see if the M_r 100,000 host protein induced by CMV infection is also associated with the RNA replicase activity in TRSV-infected cucumber seedlings.

CHAPTER 4

SUBGENOMIC RNAs FROM CMV RNA 2

INTRODUCTION

In order to determine the nucleotide sequence of Q-CMV RNA 2, double stranded cDNA was cut with restriction endonucleases to yield fragments for cloning in the bacteriophage vector M13mp7. These clones have not only been used in the actual sequencing studies, but provided the means for a detailed study of the expression of RNA 2 genetic information. The positive clones allowed synthesis of specific probes for (+) RNAs derived from RNA 2, whereas the (-) clones could be used to select by hybridization such RNAs for further analysis e.g. *in vitro* translation. This chapter describes subgenomic mRNAs derived from RNA 2, and analyses their *in vitro* function in the expression of RNA 2 genetic information.

MATERIALS AND METHODS

A. Materials.

Klenow fragment of *E. coli* DNA polymerase was obtained from Boehringer (Mannheim) or BRESA (Adelaide). AMV reverse transcriptase was obtained from Life Science Inc., (Florida). Deoxy- and dideoxynucleoside triphosphates were from P. L. Biochemicals (Wisconsin). M13-specific 17-mer and all restriction endonucleases were from New England Biolabs. Low melting point agarose was obtained from BRL, (Maryland). (α-^{32}P)dATP and (α-^{32}P) dCTP (specific activities of 2 kCi/mmol) were from BRESA. Cellulose nitrate and "Genescreen" membranes were from Sartorius and New England Nuclear, respectively.

B. Virus and RNA.

Q-CMV was grown in either cucumber or tobacco (*Nicotiana clevelandii* A. Gray) and viral RNA prepared as described in Chapter 2. For some experiments, RNA was obtained from a crude preparation of RNA-dependent RNA polymerase, the particulate fraction (Gill et al., 1981; Gill, 1983). Ten ml of particulate fraction were made 1% SDS and extracted with an

equal volume of water saturated phenol, before addition of 10 ml of chloroform. After an ether extraction, the aqueousphase was made 0.25 M Na acetate and the RNA precipitated with 25 ml of ethanol.

C. Cloning and sequence analysis of CMV RNA 2.

Double stranded cDNA to RNA 2 was made as described in Gould and Symons (1982), digested with restriction endonucleases Taq I or Sau 3A I and the resulting fragments cloned in the bacteriophage vector M13mp7 (Messing et al.,1981) by Dr. A. Gould. The nucleotide sequences of the inserts were obtained using the dideoxynucleotide chain termination technique (Sanger et al; 1980) and the specific M13 17-mer primer (5'-GTAAAACGACGGCCAGT-3') by Drs. A. Gould, A. Rezian and R. Williams.

D. Direct sequence analysis of RNA 2.

(i) Preparation of DNA restriction fragments for use as primers.

Recombinant single stranded M13 DNA, containing RNA 2 sequences of the same polarity as the RNA, was transcribed using the M13 specific 17-mer (see above) with (α-^{32}P)dATP and (α-^{32}P)dCTP, essentially as described by Bruening et al., (1982). Appropriate restriction endonucleases were used, as recommended by the manufacturers, to excise the fragments desired as primers. These were then fractionated on 6% polyacrylamide, 7M urea, gels in 90 mM Tris-boric acid, pH 8.3, 1mM EDTA (Sanger and Coulson, 1978). After detection by autoradiography, the DNA primers were excised from the gel and recovered by soaking in 0.5M ammonium acetate, 0.1% SDS, 1 mM EDTA (Maxam and Gilbert, 1980).

(ii) RNA - DNA hybridization.

Purified DNA primers and 2 μg of appropriate RNA were redissolved in 30 μl of 0.18 M Nacl, 10 mM Tris-HCL, pH 7.0, 1 mM EDTA, 0.05% SDS, heated at 100°C for 2 min and incubated at 65°C for 2 h. The hybrids were twice ethanol precipitated, twice washed with cold 70% ethanol and dried *in vacuo*.

(iii) Reverse transcriptase sequencing.

The RNA - DNA hybrids were redissolved in 4.5 µl of 0.1 mM EDTA and aliquots (1µl) dispensed into four tubes. Reactions contained, in a final volume of 4 µl, 50 mM Tris-HCL, pH 8.3, 50 mM KCl, 8 mM $MgCl_2$, 10 mM DTT, 2 units of AMV reverse transcriptase, 100 µM of the four dNTPs and a single ddNTP. The appropriate ddNTP concentrations used were: 5 µM ddATP; 1 µM ddCTP; 1 µM ddGTP; 5 µM ddTTP (Zimmern and Kaesberg, 1978; Symons, 1979). Reactions were incubated at 37°C for 30 min, then 15 µl of 0.3 M Na acetate were added and the hybrids precipitated with 50 µl of cold ethanol. The pellets were dried in vacuo, resuspended in 3 µl of 95% deionised formamide, 10 mM EDTA, 0.03% xylene cyanol FF, 0.03% bromophenol blue, heated to 100°C for 2 min, and electrophoresed on 40x40x0.025 cm, 6% polyacrylamide gels containing 7M urea (Sanger and Coulsen, 1978).

In some experiments, designed to determine the 5'-termini of subgenomic RNAs, some of the DNA primer was hybridised to RNA samples containing the desired subgenomic RNAs. Reverse transcription of these samples was carried out in the absence of ddNTPs. These samples were coelectrophoresed with normal sequencing reactions of full-length RNAs, as markers.

E. RNA blotting hybridizations.

(i) ^{32}P-DNA probes.

Recombinant, single stranded M13 DNA, containing inserts of (+) polarity, was used to prepare ^{32}P-DNA probes of high specific activity, as described above in section D (i), and by Bruening et al., (1982). Restriction endonuclease Eco RI was used to excise the probes before gel electrophoresis.

(ii) Gel electrophoresis and blotting of RNA.

Nucleic acid samples were denatured with 1M glyoxal in a total volume of 10 µl of 10 mM Na phosphate, pH 6.5, 0.1 mM EDTA, at 50°C for 15 min (McMaster and Carmichael, 1977), before electrophoresis on 1.5% agarose gels (15x14x0.3 cm) in 10mM Na phosphate, pH 6.5, 0.1 mM EDTA. After staining with ethidium bromide (5 µg/ml) and destaining (both in electrophoresis buffer), the RNA was transferred to nitrocellulose or "Genescreen"

filters by blotting and baked _in vacuo_ at 80°C (Thomas, 1980). The ^{32}P-DNA probes were hybridised to the bound RNA essentially as described by Thomas (1980). The washing conditions employed were 5x5 min in 1xSSC, 0.1% SDS at room temperature, followed by 30 min in 0.1xSSC, 0.1% SDS, at 50°C. The filters were wrapped in plastic (Glad-wrap) and exposed to X-ray film at room temperature.

F. Hybrid-selection of RNA.

The appropriate recombinant single stranded M13 DNAs containing inserts complementary to viral RNA were grown in 50 or 500 ml cultures (Messing et al., 1981) to yield sufficient DNA. DNA was bound to nitrocellulose filters, of pore size .45 µm or .22 µm (which has a higher binding capacity - Bill Kalionis, personal communication) and the hybrid-selection performed as described (Parnes et al., 1981; Maniatis et al., 1982). Washing conditions were of higher stringency than described by Maniatis et al., (1982) and comprised 6x5 min washes in 1xSSC, 0.5% SDS, at 65°C, followed by 6x5 min in 0.1xSSC, 0.5% SDS, at 50°C (the last two minus SDS). After elution, the RNA was ethanol precipitated, washed once with cold 70% ethanol and dried _in vacuo_, before being resuspended in 0.1 mM EDTA. Aliquots were analysed by _in vitro_ translation (see Chapter 2) or by RNA blotting (section E, above).

G. Other methods for RNA fractionation.

Initially, viral RNA was fractionated on 10-40% sucrose gradients as described by Schwinghamer and Symons (1975). Appropriate regions of the gradients were collected in 0.2 to 0.5 µl fractions. In later experiments, virus RNA was electrophoresed on 1.5% low melting point agarose gels, in 20 mM Tris-acetic acid, pH 7.5, 1 mM EDTA. The RNA bands were detected by staining with 5 µg/ml ethidium bromide, and excised. The gel slices were melted at 65°C for 15 min and the RNA recovered by 3 phenol extractions as described by Kuhn et al., (1979).

FIGURE 4.1

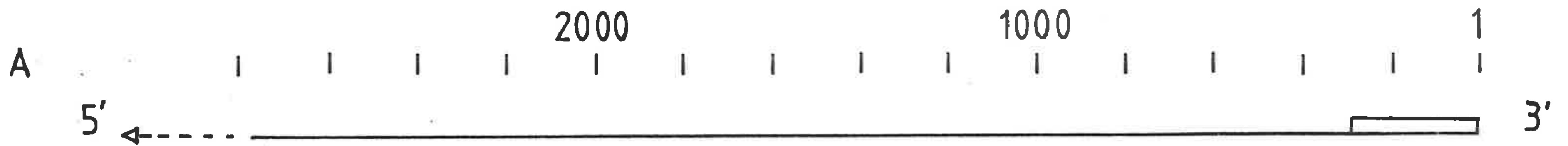
Diagram showing the extent of sequence information and cDNA clones of Q-CMV RNA 2.

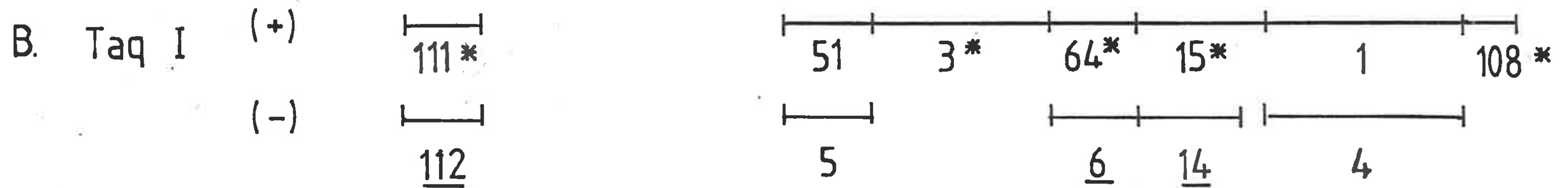
A. The sequence is numbered from the 3' end, with the 3' conserved region (Symons, 1979; Gould and Symons, 1982) indicated by a box.

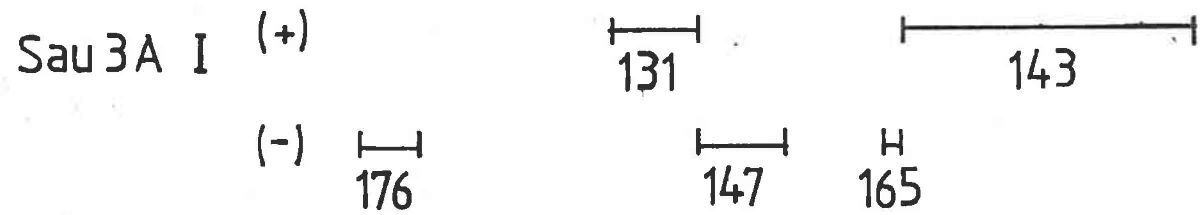
B. M13mp7 clones of restriction fragments from cDNA cut with Taq I (upper series) or Sau 3A I (lower series). (+) or (-) indicates their orientation with respect to RNA 2. Clones marked with an asterisk have been used to synthesise ^{32}P-cDNA probes for RNA blotting; underlined clones have been used in hybrid-selection experiments.

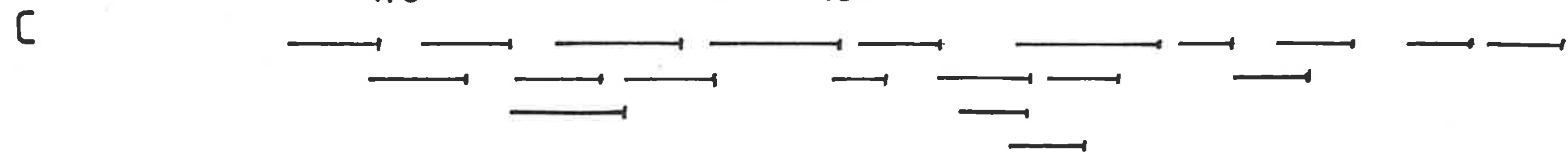
C. RNA sequences determined by dideoxynucleotide chain termination using DNA fragments as primers on RNA 2.

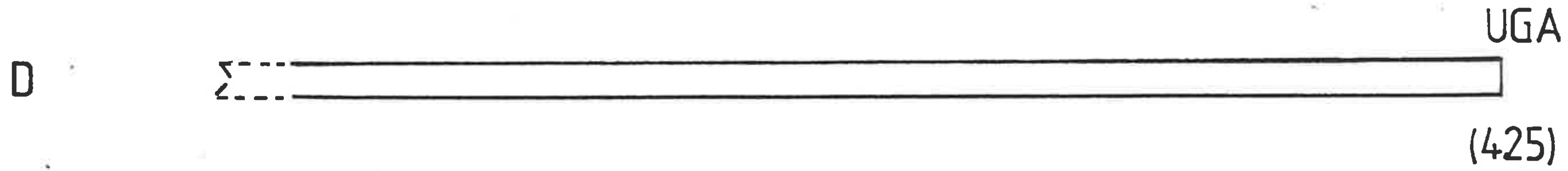
D. Open reading frame traversing most of RNA 2.


A
2000
1000
1
5′
3′


B. Taq I
(+)
111*
51
3*
64*
15*
1
108*
(−)
112
5
6
14
4


Sau3A I
(+)
131
143
(−)
176
147
165


C


D
UGA
(425)

RESULTS

A. Nucleotide sequence of CMV RNA 2.

At present, the sequence of 2754 nucleotides from the 3' end of Q-CMV RNA 2 has been determined, in collaboration with Drs. A. Gould, A. Rezian and R. Williams. This sequence is equivalent to 80% of RNA 2. Most of the sequence is included in two partially overlapping series of clones in M13mp7. The two series consist of Taq I and Sau 3A I cDNA fragments, respectively. Direct dideoxynucleotide chain termination sequencing of RNA with reverse transcriptase and DNA primers from M13 clones was used to sequence uncloned sections of RNA 2, to link clones where necessary and to check sequences in most of the cloned regions (Fig. 4.1).

A long open reading frame traverses the known RNA 2 sequence, to terminate at a UGA codon 425 nucleotides from the 3' end. Neither the (+) nor (-) strands contain any other open reading frame longer than 100 amino acids. Details of the sequence will not be presented in this thesis.

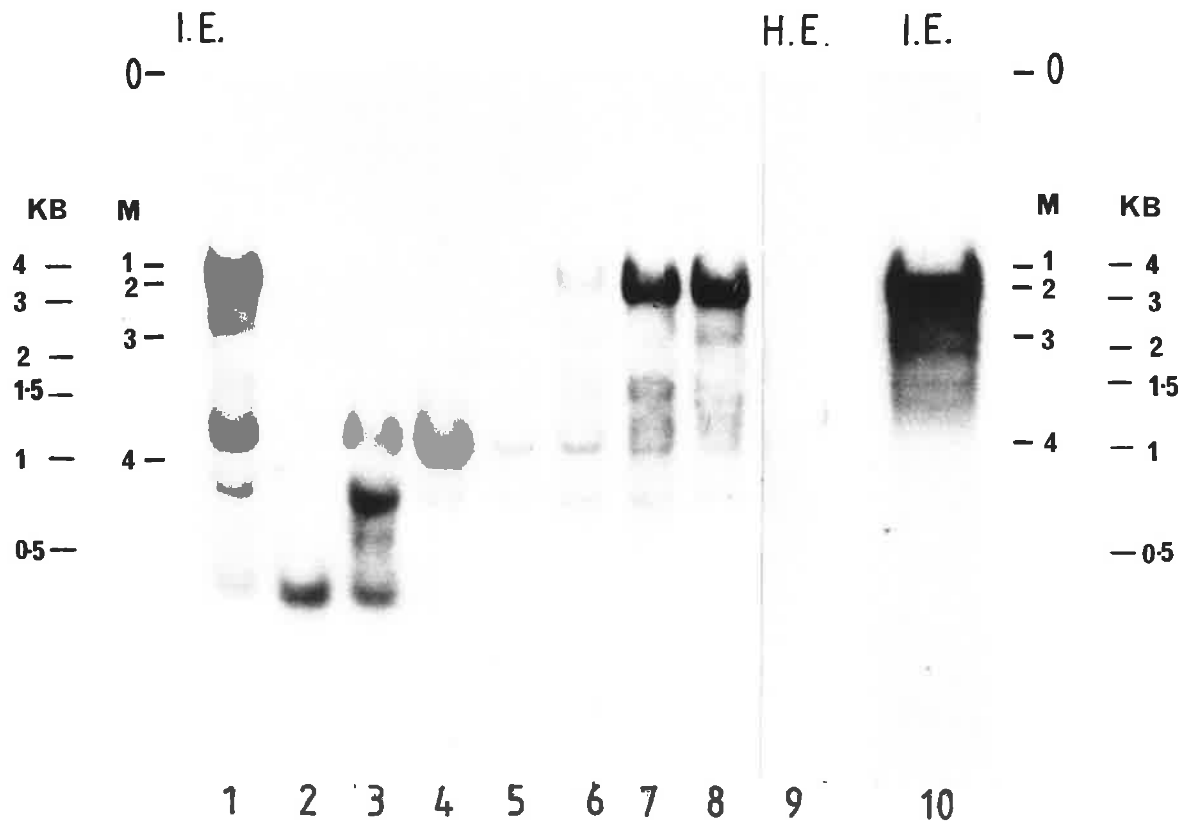
B. RNA blotting of subgenomic RNAs.

This section describes the discovery of over 20 subgenomic RNAs derived from RNA 2. RNA blotting experiments showed that these RNAs were of discrete sizes and that most were 3'-coterminal with genomic RNA. This suggested that the subgenomic RNAs possessed specific, 5' termini, which was later confirmed using primer extension with reverse transcriptase. Other, smaller, RNAs correspond to specific internal sequences of RNA 2. All of these RNAs are distinct from a low background of randomly degraded RNA.

In an initial experiment to determine whether RNA 2 generated any subgenomic RNAs, RNA extracted from the particulate fraction of CMV-infected cucumber leaves was fractionated on sucrose gradients, followed by agarose gel electrophoresis and RNA blotting. The plus probe was ^{32}P-cDNA to clone 64^{+}, whose insert is a Taq I fragment located between 813 and 1,015 nucleotides from the 3' end of RNA 2. As shown in Fig. 4.2, a number of RNA

FIGURE 4.2

Possible subgenomic RNAs from CMV RNA 2. RNA was extracted from the particulate fraction of CMV infected plants and fractionated on 10-40% sucrose gradients. Aliquots (10 µg) of each 0.5 ml fraction were denatured with glyoxal before electrophoresis on a 1.6% agarose gel. Probe was from clone 64^{+} (tracks 1-9), as described in Methods. The sucrose gradient fractions are in tracks 2-8; the arrow indicates the direction of sedimentation. Tracks 1 and 9 contain unfractionated RNA (10µg) extracted from CMV-infected (I.E.) and uninfected (H.E.) plants, respectively. Track 10 (from a different gel) shows I. E. RNA (10µg) probed with cDNA to clone 3^{+}. The positions of marker CMV RNAs 1-4 are shown for each gel (M), as is a scale in Kilobases (KB).


PROBE:
64+
3+
I.E.
H.E.
I.E.
0
0
KB
M
M
KB
4
3
2
1·5
1
0·5
1
2
3
4
1
2
3
4
4
3
2
1·5
1
0·5
1
2
3
4
5
6
7
8
9
10
SEDIMENTATION

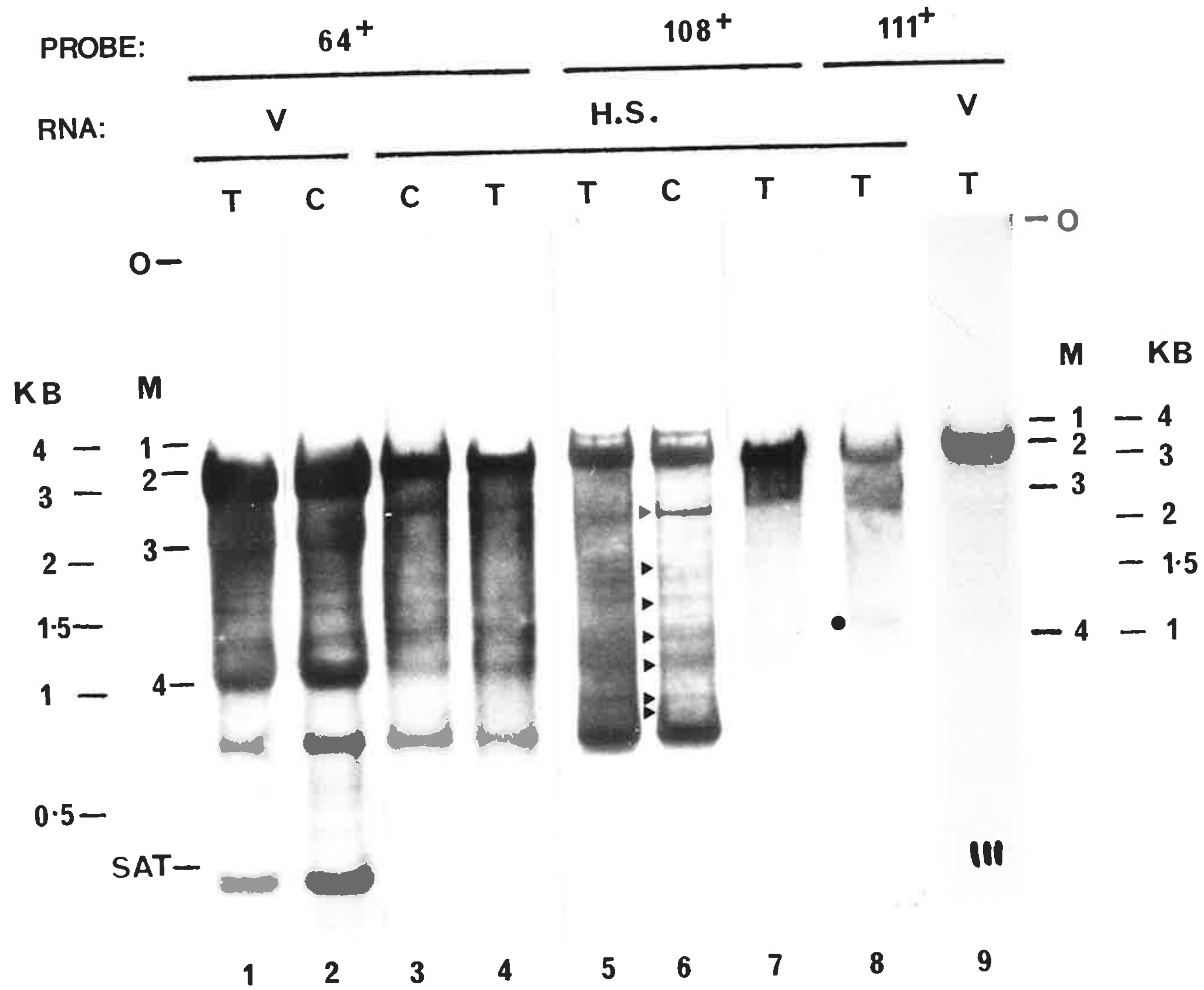
species smaller than RNA 2 were detected. Among these were three prominent species of approximately 400, 900 and 1,100 nucleotides (tracks 2-4). At least five other RNAs, whose sizes ranged from 1,400 to 3,000 nucleotides, were detected in smaller amounts (tracks 5 to 8). These larger RNAs could also be detected using a probe from clone 3 (located 5' to clone 64 - see Fig. 4.1) (Fig. 4.2, track 10). None of the probes hybridized to RNA from uninfected plants (Fig. 4.2, track 9).

When the total RNA encapsidated in virus obtained from either cucumber or tobacco was analysed by agarose gel electrophoresis and blotting, a probe from clone 64^+ disclosed a similar, large range of subgenomic RNAs (Fig. 4.3, tracks 1,2). Three of the strongest subgenomic bands, at 2,200, 1,100 and 750 nucleotides, were, however, found to correspond to three prominent bands detected by staining with ethidium bromide, of which the two largest are RNAs 3 and 4. Some of these subgenomic RNAs may therefore be due to artefacts of electrophoresis. In order to eliminate this possibility, the RNA 2 derived species present in the total viral RNA were purified by hybrid-selection using a (-) clone, 14^-, located from 600 to 816 nucleotides from the 3' end of RNA 2 (see Fig. 4.1). These RNA samples were run on the same gel (Fig. 4.3, tracks 3,4). The probe from clone 64^+ disclosed an extensive range of genuine subgenomic RNAs, down to 750 nucleotides long (Fig. 4.3, tracks 3,4). There was no significant difference between RNAs obtained from the two different hosts.

Since the subgenomic RNAs were of discrete sizes and hybridised to a clone located near the 3' end of RNA 2, it was surmised that they were in fact 3'-coterminal with genomic RNA. To test this, RNA was selected from total virus RNA by hybridization to clone 14^- DNA. A probe from clone 108^+, covering the 145 nucleotides at the extreme 3' end of RNA 2 showed (Fig. 4.3 tracks 5,6) that all the RNA bands already detected in tracks 3 and 4 were also detected with the 3' probe. Furthermore, two small RNAs of 800 and 880 nucleotides were detected with clone 108^+ but not

FIGURE 4.3

Further analysis of subgenomic RNAs from CMV RNA 2. Either 10 µg of total encapsidated (V) or 1 µg of RNA (H.S.) selected from encapsidated RNA by hybridization to clone 14^- (tracks 3 - 6, 8) or 112^- (track 7) was denatured by glyoxal and electrophoresed on 1.6% agarose gels. Probes were from clone 64^+ (tracks 1 - 4), clone 108^+ (tracks 5, 6, 7) or clone 111^+ (tracks 8, 9) as described in Methods. The virus was from either cucumber (C) or tobacco (T) plants. Separate gels were run for tracks 1 - 4, tracks 5 - 8 and track 9. The markers (CMV RNAs) on the left refer to tracks 1 - 8, those on the right to track 9. The scales shown are in Kilobases (KB).



PROBE:
64+
108+
111+
RNA:
V
H.S.
V
T
C
C
T
T
C
T
T
T
O
KB
M
4
3
2
1·5
1
0·5
1
2
3
4
SAT
M
KB
1
2
3
4
4
3
2
1·5
1
1
2
3
4
5
6
7
8
9

clone 64^+ (see below). Note that the hybrid-selected RNAs were contaminated with a small amount of RNA 1 (tracks 5,6), which could only be detected with a probe for the 3' conserved sequences. Comparison of these RNAs with those shown in Fig. 5.2 (tracks 2,3) confirmed that the subgenomic RNAs from RNA 1 were present in amounts too low to be detected in this experiment.

These observations led to the conclusion that the 20 subgenomic RNAs above 750 nucleotides in size were indeed 3'-coterminal with RNA 2. An observation consistent with this conclusion is that, when total virus or hybrid-selected RNA was probed with clone 111^+, located at the 5' end of the known RNA 2 sequence (over 2,300 nucleotides from the 3' end - see Fig. 4.1) only RNAs over 2,400 nucleotides long were detected (Fig. 4.3, tracks 8,9). In the hybrid-selected RNA, an exception appears to be a band at 1,600 nucleotides (track 8). Furthermore, when RNA was selected by hybridization to clone 112^-, which is complementary to clone 111^+, only the RNA species longer than 2,400 nucleotides were detected by the clone 108^+ probe (Fig. 4.3, track 7). Table 4.1 lists all the RNA species detected in the RNA blotting experiments.

The subgenomic RNAs under 1,000 nucleotides long present some interesting puzzles. Total virus RNA contains many small RNAs, of which only those of 750 nucleotides or longer were selected by hybridization to clone 14^- (Fig. 4.3, tracks 1 to 4). The major band at 750 nucleotides hybridized to probes from (+) clones 64^+ and 108^+, although a continuous RNA transcript of this length would not cover both these clones, i.e. reach from the 3' end of RNA 2 to clone 64^+. This observation suggests that the 750 nucleotide band may contain a mixture of RNAs, or be a non-continuous transcript.

Two slightly larger RNAs, of 800 and 880 nucleotides, were not detected with cDNA to clone 64^+ (Fig. 4.3, tracks 1 - 4). This is consistent with their being 3'-coterminal with RNA 2, for their lengths would allow them to cover all or most of clone 14^- but very little of clone 64^+. All the RNA species below 750 nucleotides in length and detected by probes from clone 64^+, must

TABLE 4.1

SUBGENOMIC RNAs AND THEIR 5' TERMINI

RNAs (in nucleotides) detected by[a]

Blotting[b]	Primer Extension[c]	Others in Primer Extensions[c]	Primers Used[d]
3,300	n.d.		
3,200	n.d.		
2,900	n.d.		
2,700	2,630		147.100
2,400	2,380	(2,251 ; 2,281)	147.00
2,200	2,204		147.100
2,100	2,112		147.100
2,050	2,051	(2,046) (1,959)	51.89 ; 147.100
1,850	1,815	(1,800)	51.89
1,750	1,765	(1,694)	51.89
1,650	1,626	(1,615)	3.110
1,550	1,536		3.110
1,500	1,490		3.110
1,450	1,440		3.110
1,250	1,259		3.69
1,200	1,215		3.69
1,050	n.d.		
1,000	n.d.		
880	n.d.		
800	n.d.		
750	n.d.		
550 *	n.d.		
500 *	n.d.		
450 *	n.d.		
320 *	n.d.		

a - all 3'-coterminal except *

b - from data in Fig. 4.3

c - from data in Fig. 4.5B

d - see Fig. 4.5A

n.d. - not determined

be specific internal fragments, since they are too short to reach the $3'$-end of RNA 2. Furthermore, since they were not selected by hybridization to clone 14^-, they cannot extend far, if at all beyond the sequences corresponding to the $3'$ terminus of clone 64^+.

C. Primer extension mapping.

The aim of the primer extension experiments was to determine whether distinct $5'$-termini could be identified for each of the subgenomic RNAs. The primers were cut with restriction endonucleases from appropriately located clones, and extended as described in Methods. In initial experiments primer extension of RNAs in sucrose gradient fractions yielded very long transcripts, among which it was very difficult to distinguish specific stop points. Total virus RNA (200 µg) was therefore first fractionated on a non-denaturing low melting point agarose gel and the RNA recovered from gel slices. Before choosing the primers, the RNA species present in each fraction had to be identified. RNA 2 subgenomic RNAs were therefore hybrid - selected with clone 14^- (20µg/reaction) and electrophoresed on a 1.6% agarose gel after glyoxal denaturation. A probe from clone 64^+ (Fig. 4.4) showed that only few species were present in each gel fraction, with most RNAs being present in several tracks.

Three fractions from this gel, containing subgenomic RNAs from 1,400 to 2,500 nucleotides long, were chosen for primer extension analysis, using primers from clones 3, 51 and 147, as shown in Fig. 4.5A. Aliquots (2µg) of the RNA recovered from each gel slice (not the hybrid-selected RNA) were hybridised to appropriate primers. Primer 3.69 (Fig. 4.5A), was used to analyse fractions from a sucrose gradient and from a preparative 2.8% polyacrylamide gel, both enriched for RNAs from 1,200 to 1,700 nucleotides long. These extensions (from 3.69) were designed to detect only the subgenomic RNAs around 1,200 nucleotides (Fig. 4.5B). Reactions with the other primers achieved long transcripts and specific stopping. They were carried out with or without 4 mM Na pyrophosphate, which prevents

FIGURE 4.4

Fractionation of subgenomic RNAs on agarose gels. Encapsidated RNA (200 μg) was electrophoresed on a 1.5% low melting point agarose gel (sample track shown), which was sliced as indicated and the RNA recovered by phenol extraction (Methods). RNA 2 subgenomic RNAs were selected by hybridization with clone 14^- (10 μg/fraction) and 30% of each hybrid-selected RNA sample denatured with glyoxal and analysed by electrophoresis on a 1.6% agarose gel. The probe was from clone 64^+ (Methods). The positions of CMV RNAs 1 - 4 run as markers, are indicated (M), as is a scale in Kilobases (KB).

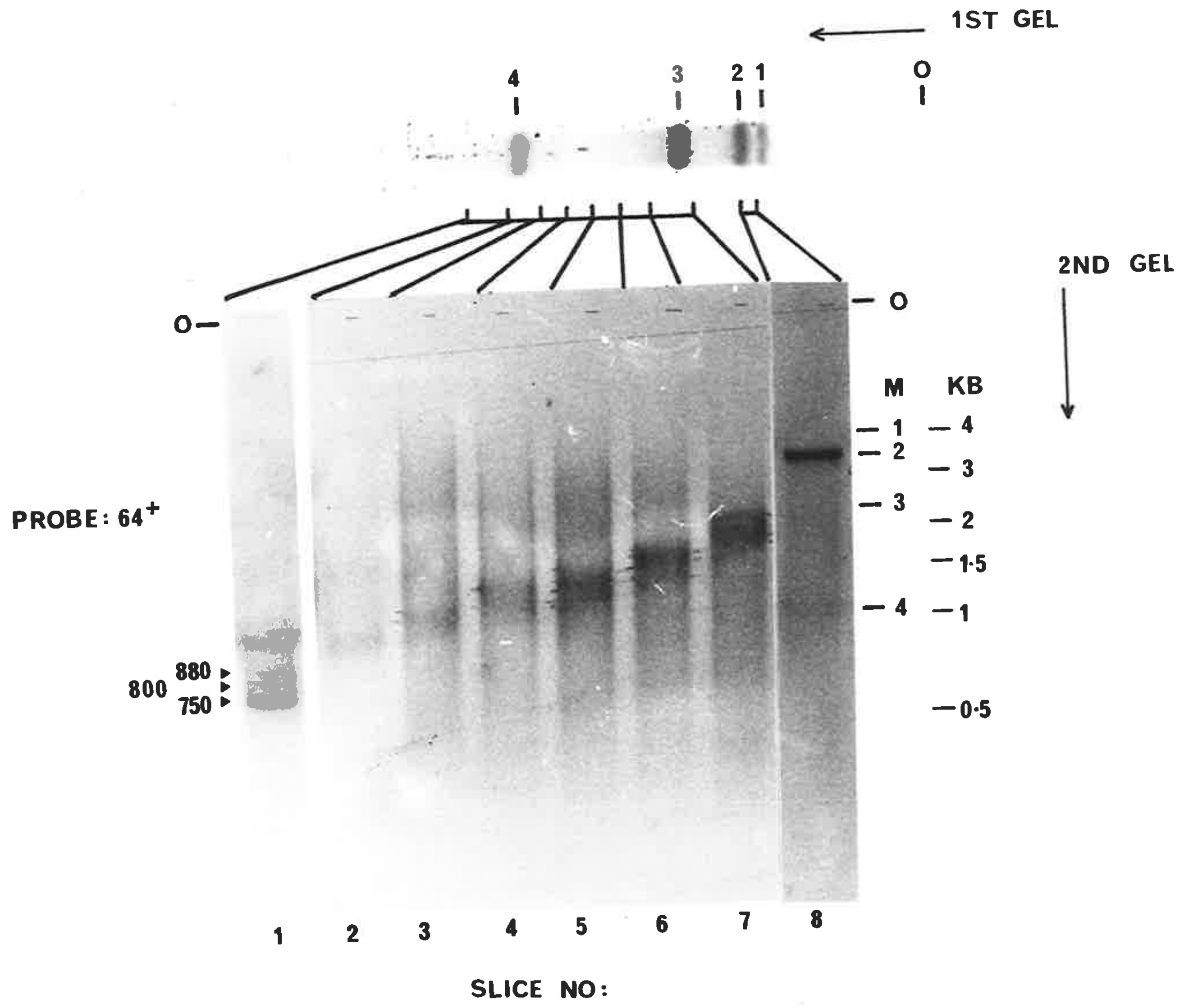
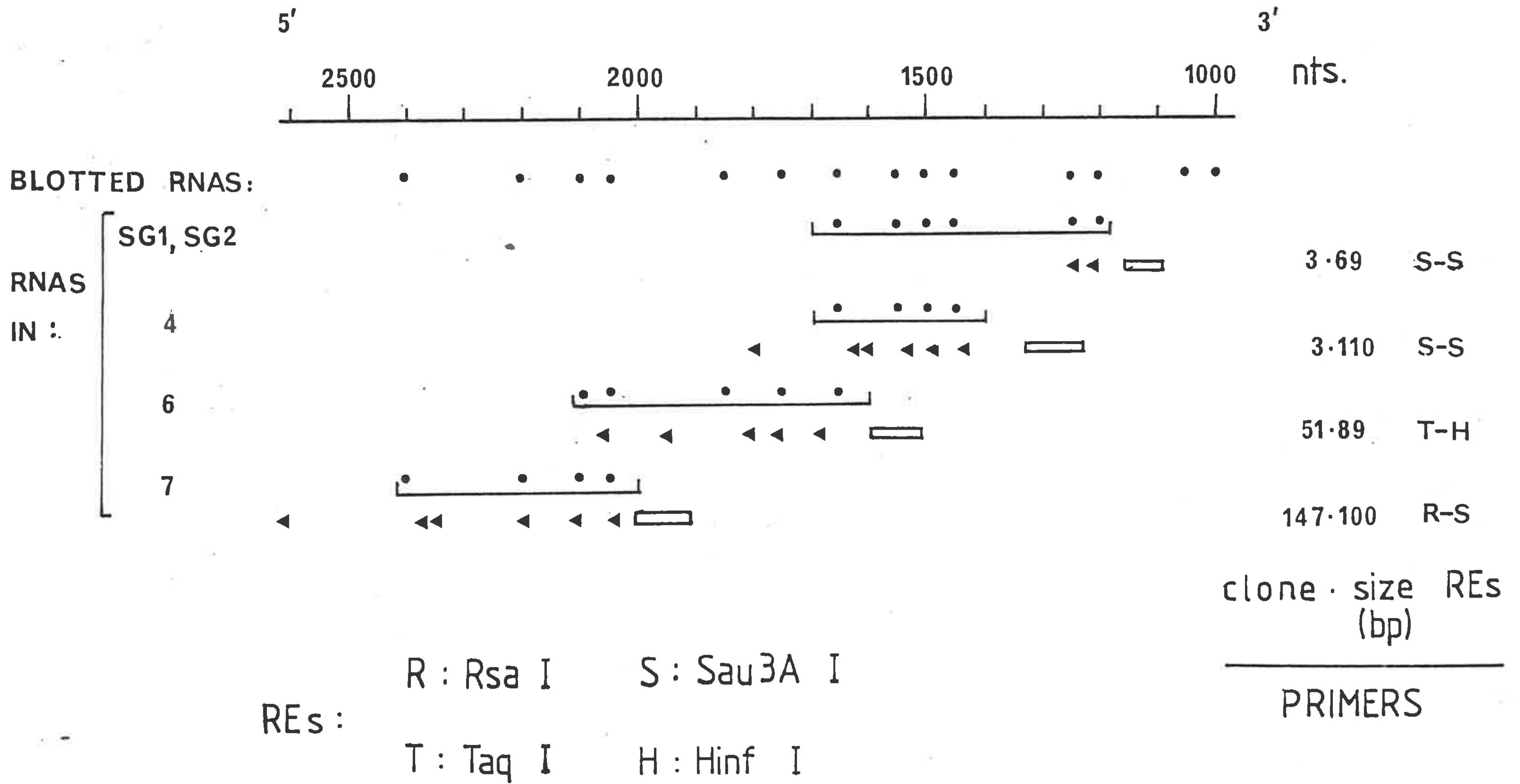

1ST GEL
2ND GEL
4
3
2 1
O
O
O
M
KB
1
4
2
3
3
2
1·5
4
1
0·5
PROBE: 64+
880
800
750
1
2
3
4
5
6
7
8
SLICE NO:

FIGURE 4.5

Primer extension 5' mapping of subgenomic RNAs.

A. Diagram showing restriction fragments (from cDNA clones) and the RNAs present in each fraction to which the primers were hybridized. The scale depicts part of RNA 2, numbered from the 3' end. The dots represent 5' ends of subgenomic RNAs, based on sizes determined by RNA blotting. Five RNA fractions were analysed: SG1, SG2, 4, 6, 7. SG1 was a fraction from a 10-40% sucrose gradient, on which total virus RNA (200 µg/gradient) was run; this fraction was shown by RNA blotting to be enriched with RNAs from 1,200 to 1,700 nucleotides, as indicated by the horizontal line. SG2 contained RNAs from 1,200 to 1,700 nucleotides, eluted from a preparative 2.8% polyacrylamide gel (Symons, 1978), and checked by RNA blotting (not shown). Fractions 4, 6 and 7 contained total RNA from slices 4, 6 and 7, respectively, of the preparative agarose gel in Fig. 4.4. They contained the RNAs shown by the dots and horizontal line. Each primer is identified by first the clone from which it was derived, followed by its size in bp. The primers were cut with the restriction endonucleases (REs) indicated and purified as described in Methods. Each primer was then hybridised to a different RNA fraction (as shown) and extended, as described in Methods. 5' termini detected by this approach, and which were near those identified by blotting, are represented by arrowheads. The sizes of RNAs given by this experiment are compared to those determined by RNA blotting in Table 4.1, which also notes the primer from which each RNA was determined.

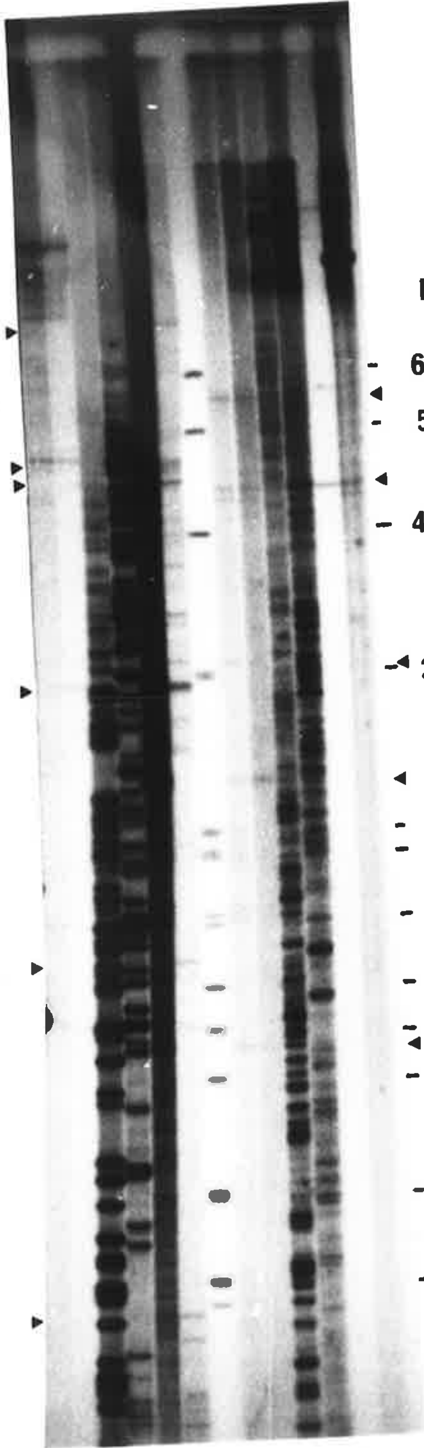


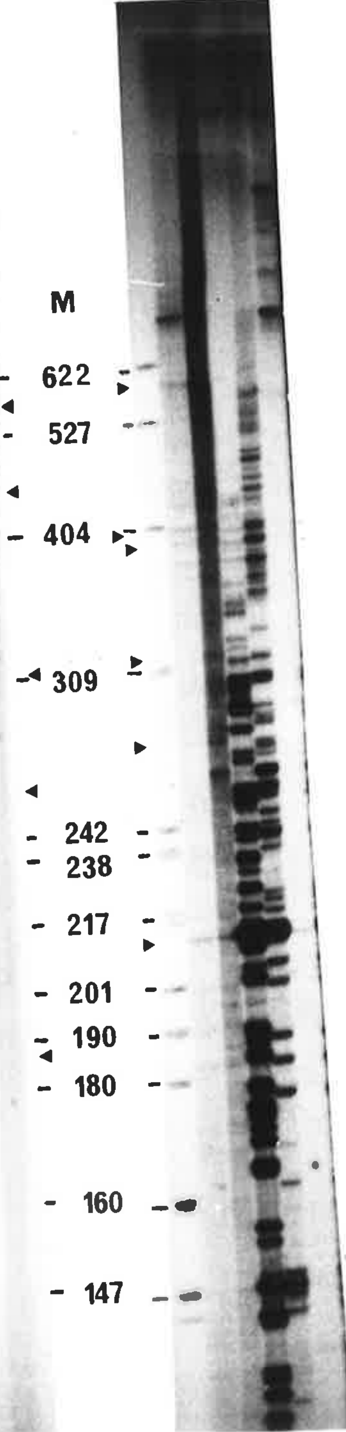
5'
3'
2500
2000
1500
1000
nts.
BLOTTED RNAS:
RNAS IN:
SG1, SG2
4
6
7
3·69
S–S
3·110
S–S
51·89
T–H
147·100
R–S
clone · size (bp)
REs
PRIMERS
REs:
R : Rsa I
S : Sau3A I
T : Taq I
H : Hinf I

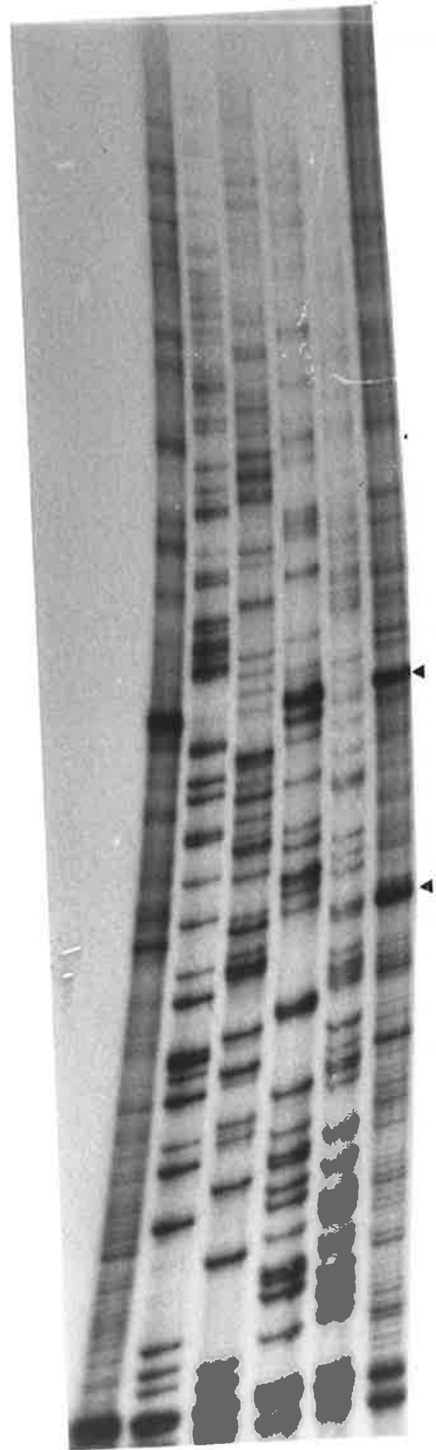
FIGURE 4.5

B. Gel electrophoresis of the primer extension reactions shown in A. The four primers used (see part A) are given at the top of the Figure. SG denotes RNA fractions 4, 6 or 7 (see part A) containing subgenomic RNAs. These primer extensions (E) were carried out in the presence of (α-^{32}P)dCTP, as described in Methods. Primer 3.69 was hybridised to SG1 or SG2 (see part A for details) and extended in the presence of unlabelled dNTPs. Part of each primer was hybridised to sucrose gradient fractions containing full length (FL) RNA 2, and subjected to the dideoxynucleotide chain termination sequencing reactions, to identify A, C, G or U residues, as indicated. Control primer extension (E) reactions on FL RNA 2 omitted all ddNTPs. Tracks marked with stars were reactions in the presence of 4 mM Na pyrophosphate; stop points in these tracks are marked with arrowheads, where they gave RNA sizes to those determined by RNA blotting (see part A). All reactions were analysed on 6% polyacrylamide, 7M urea gels (see Methods). Markers (M) were ^{32}P-end-filled Hpa II fragments of pBR322 (sizes given in nucleotides). Primers were electrophoresed to the bottom of each gel. The precise stop points derived from each primer extension, and which corresponded closely to RNA sizes determined by RNA blotting, (as denoted by arrowheads) are listed, with primers, in Table 4.1.

B







anomalous elongation through second strand synthesis (Myers and Spiegelman, 1978); the 5' termini shown in Fig. 4.5A and Table 4.1 are those from reactions in the presence of Na pyrophosphate.

As shown in Fig. 4.5, more 5'-termini were detected by primer extension than the number of RNAs detected by blotting in each gel fraction. This probably reflects the greater resolving power of the primer extension technique. Nonetheless, there is good overall correspondence between the RNAs detected by primer extension (Fig. 4.5A, B) and by blotting (Fig.4.3) as summarized in Table 4.1. Some of the extra bands detected by primer extension may correspond to non-3'-coterminal subgenomic RNAs, which would therefore have sizes quite unrelated to their 5'-termini. Furthermore, the reverse transcription observed, in some cases, beyond the largest RNAs detected by blotting probably arose through contamination of the RNA fractions with randomly degraded RNA derived from 5' to the distinct species.

Fig. 4.6 illustrates the 3'-coterminal and smaller internal subgenomic RNAs from RNA 2. The RNA sizes shown are those obtained from the blotting experiments, but there may be even more subgenomic RNAs, as discussed above.

D. In Vitro translation products of the subgenomic RNAs.

The discovery of the subgenomic RNAs from RNA 2 immediately raised the question of whether they could function as mRNAs *in vitro*. To test this, aliquots of the unfractionated RNA hybrid-selected with clone 14, and shown in Fig. 4.3, in addition to RNA hybrid-selected from CMV-infected plant extracts, were translated in the reticulocyte lysate, as detailed in Chapter 2. Fig. 4.7 shows the resulting translation products. RNA 2, purified on the basis of size (Chapter 2) yielded the major full-length product of M_r 110,000 (track 1), which was also one of the major translation products of unfractionated virus RNA (track 2). Hybrid-selected RNA from CMV-infected plants (track 3) yielded the full-length M_r 110,000 polypeptide, in addition to other products of M_r 17,000, M_r 30,000, M_r 37,000, M_r 60,000 and M_r 85,000. The hybrid-selected RNAs from virus RNA,

FIGURE 4.6

Diagram showing 3'-coterminal and internal subgenomic RNAs derived from CMV RNA 2. The genomic RNA 2 is shown at the top; the dashed portion represents the 5' unknown sequence. The positions of the UGA termination codon for the long open reading frame and of the 3' conserved sequences (boxed) are shown. Lines represent the subgenomic RNAs, whose sizes from RNA blotting are given on the right. Small boxes indicate the likely initiation codon for each major translation product (as shown), below the smallest RNA capable of encoding it. The hatched box at the bottom locates clone 14, used in the hybrid-selection experiments. The likely positions of small internal RNAs are shown at the bottom (RNAs from 320 to 750 nucleotides).

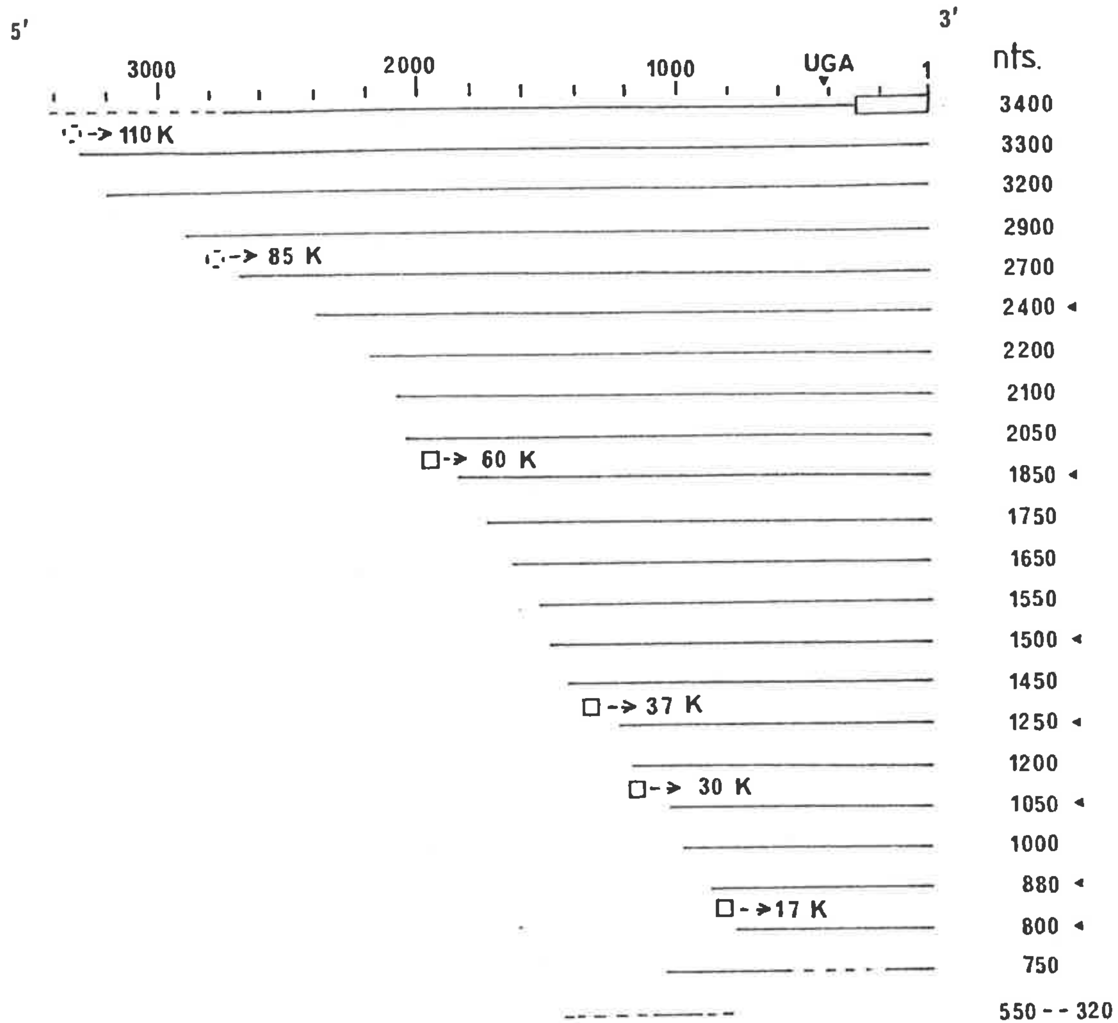
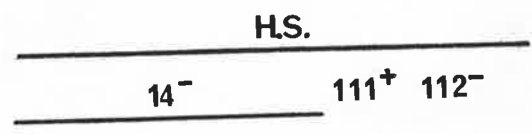

5'
3'
3000
2000
1000
UGA
1
nts.
3400
110 K
3300
3200
2900
85 K
2700
2400
2200
2100
2050
60 K
1850
1750
1650
1550
1500
1450
37 K
1250
1200
30 K
1050
1000
880
17 K
800
750
550 - - 320
H. S. done 14⁻

FIGURE 4.7

In vitro translation products of RNA 2 and its subgenomic RNAs. Track 1 shows the product from RNA 2 which had been purified by 2.8% polyacrylamide gel electrophoresis (see Chapter 2); track 2 shows the translation products of total encapsidated RNA (V). The other tracks show translation products of hybrid-selected (H.S.) RNAs from CMV-infected cucumber extracts (I.E.C., track 3) or encapsidated RNA (tracks 4-7). The RNAs were selected with clones 14^-(tracks 3-5), 111^+ (track 6) or 112^-(track 7). Molecular weights of the translation products from CMV RNAs 1-4 are indicated on the left, and those of the translation products in tracks 3-5, on the right. F is the dye front for track 1 and E an endogenous band (tracks 2-7) due to ^{35}S-methionine. *In vitro* translations (RNA at 50-150 μg/ml) and 13% polyacrylamide SDS gel electrophoresis were as described in Chapter 2.

H.S. CLONE :



shown in tracks 3-6 of Fig. 4.3, also yielded these translation products Fig. 4.7, tracks 4,5). The M_r 110,000 polypeptide appeared in variable amounts, but never as the major product. The most efficiently translated product was clearly that of M_r 17,000, possibly with some related, smaller proteins. When total virus RNA was hybrid-selected with clone 112^-, which is located about 2,400 nucleotides from the 3'-end, a different pattern of translation products was obtained (Fig. 4.7, track 7). This RNA, (shown in Fig. 4.3, track 7) yielded the M_r 110,000 polypeptide as its major product, together with several other peptides over M_r 40,000. Peptides of M_r 37,000 and one possibly corresponding to M_r 20,000 were very minor products. Clone 111^+, the positive insert complementary to 112^-, selected no translateable RNA (Fig. 4.7, track 6). Note that in all cases, the hybrid-selected RNAs have been very efficiently depleted of the other CMV RNAs (especially RNAs 3 and 4), since their full length translation products (p 36.7 and p 26.2) were clearly absent.

Once it was clear that the complete mixture of CMV RNA 2 and its subgenomic RNAs encoded several proteins, the next question to be asked was which protein was produced by which specific subgenomic RNA(s). This was investigated using the hybrid-selected RNA fractions shown in Fig.4.4. Their *in vitro* translation products are shown in Fig. 4.8A, where the track numbers refer to the same RNAs as those tracks in Fig. 4.4. The smallest RNAs present, (from 750 to 1000 nucleotides long, track 1) yielded enormous amounts of products around M_r 17,000 in size. RNAs from 1,000 to 1,400 nucleotides (tracks 2,3) gave no significant translation products, whereas the longer RNAs (tracks 4 to 8) produced polypeptides of increasing size, from M_r 30,000 to M_r 85,000 and the full-length polypeptide (M_r 110,000). Interestingly, these larger RNAs, from 1,400 to 2,500 nucleotides, each appear to encode polypeptides from M_r 30,000 up to the largest products consistent with the RNA size (see Discussion).

FIGURE 4.8

In vitro translation products of subgenomic RNAs.

A. Aliquots of the RNAs selected by hybridization to clone 14^- and shown in Fig. 4.4 were translated and the products analysed by 13% polyacrylamide SDS gel electrophoresis (see Chapter 2). The track numbers (1-8) refer to the same RNA samples as those in Fig. 4.4. Molecular weights of the major translation products are shown. 'E' indicates the endogenous band due to ^{35}S-methionine.

B. Viral RNA which cosedimented with RNA4 on sucrose gradients was subjected to hybrid-selection with 20 µg of DNA from clone 6^- (track 9) or 14^- (track 10). The selected RNAs were translated *in vitro* and the products analysed by SDS gel electrophoresis as described in Chapter 2. Molecular weights of the major translation products are indicated.

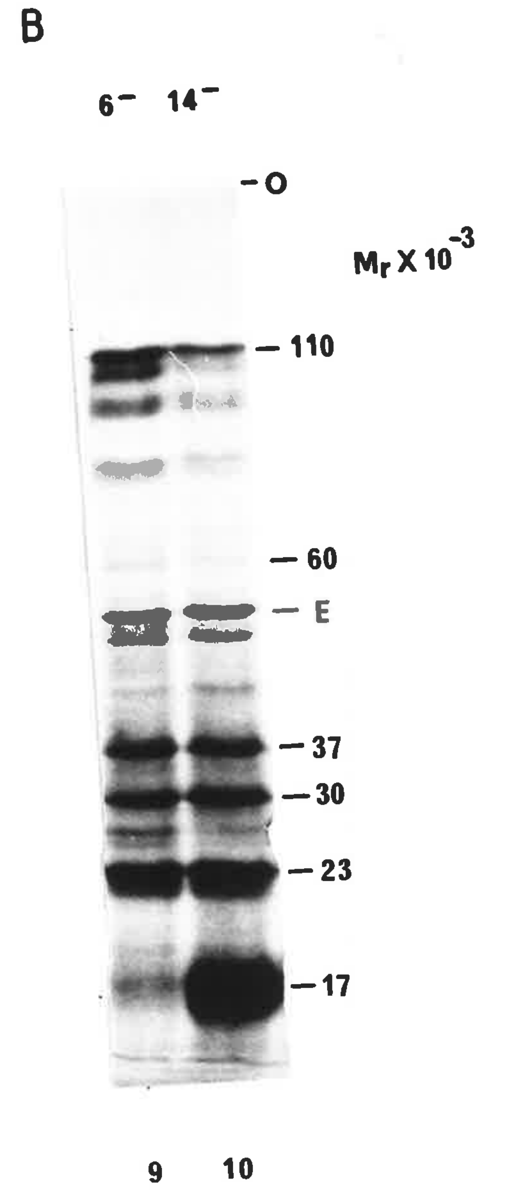
A

H.S. CLONE : 14⁻

E —

— O
— 110
— 85
— 60
— E
— 37
— 30
— 17

1 2 3 4 5 6 7 8



The variation in production of the M_r 60,000 and M_r 23,000 proteins may be due to competition between the subgenomic RNAs (tracks 7, 9,10).

Possible mRNAs for the M_r 17,000 translation product (Fig. 4.8A, track 1) include several RNAs from 750 to 880 nucleotides (Fig. 4.4, track 1). This mRNA for the M_r 17,000 protein was selected by hybridization to DNA from clone 14^-, located 600 to 816 nucleotides from the 3' end of RNA 2 (Fig. 4.8A, track 1 and B, track 10), but not from clone 6^- (813 to 1,015 nucleotides from the 3' end) (Fig. 4.8B, track 9). The only open reading frame available for the M_r 17,000 protein corresponds to the carboxy-terminal part of the long open reading frame traversing RNA 2, and runs from an AUG codon 853 nucleotides from the 3' end to the UGA codon 425 nucleotides from this end. This makes the protein (calculated M_r 17,000) 145 amino acids long. It would be encoded by the 880 nucleotide, but not the 800 nucleotide, RNA. Clone 6^- is unlikely to hybrid-select the 880 nucleotide RNA, with an overlap of about 70 nucleotides, which did not allow a probe from clone 64^+ (complementary to clone 6^-) to hybridize to this RNA (Fig. 4.3, tracks 1 - 4). All these observations are consistent with the conclusion that the 880 nucleotide RNA is the mRNA for the M_r 17,000 protein. The next largest open reading frame on RNA 2 could encode a polypeptide of 100 amino acids (calculated M_r 11,000) starting 627 nucleotides from the 3' end. This could be translated from the 750 nucleotide RNA (i.e. from the 3'-coterminal one if there is a mixture in this band) and may thus be present in the lower part of the translation product smear below M_r 17,000 (Fig.4.8A, track 1).

DISCUSSION

The work described in this chapter has shown that over 20 subgenomic RNAs are generated from CMV RNA 2. Most of these are 3'-coterminal with RNA 2, but the smallest ones are specific internal fragments located near,

although not reaching, the 3' end. These RNAs have been reproducibly detected in several samples of encapsidated virus RNA and in RNA extracts from CMV-infected plant fractions known to include a CMV RNA replication complex. They were detected both before and after hybrid-selection using cloned DNAs, and RNA blotting experiments using probes for specific sequences along RNA 2 confirmed that they did not arise from random degradation. A more thorough study using clones from all sections of the complete RNA 2 sequence may, however, disclose other subgenomic RNAs from e.g. the 5' part of RNA 2.

Many of the subgenomic RNAs were found to have considerable *in vitro* messenger activity. Although these *in vitro* translation products may be artefactual, it is significant that they were produced in considerable amounts even in the presence of RNA 2 or, in some cases, total virus RNA. Their efficiency at competing with RNA 2, during translation of the hybrid-selected RNAs, presents a puzzle, although it is possible that hybrid-selection has favoured shorter RNAs over the longer ones. Other possible reasons will be discussed in Chapter 6. In view of the specific sizes of the subgenomic RNAs, their 3'-cotermination with RNA 2 and their efficient mRNA activity, it would be hard to distinguish between these RNAs and functional subgenomic mRNAs. No evidence is, however, available to show whether or not they have any *in vivo* mRNA function.

A single long open reading frame traverses RNA 2 and therefore also each of the subgenomic RNAs. Since these are 3'-coterminal with RNA 2, their *in vitro* translation products must correspond to the appropriately-sized carboxyterminal parts of the M_r 110,000 translation product(Fig.4.9). The larger translation products from RNAs of increasing size must therefore differ at their amino termini. An alternative explanation, early termination of translation to produce polypeptides sharing their aminotermini, can be excluded because the smallest 3'-coterminal RNAs have been found to make the smallest proteins. Similar observations of a family of carboxy-

FIGURE 4.9

RNA 2 encoded polypeptides corresponding to the major *in vitro* translation products from subgenomic RNAs. The polypeptides are shown as derived from the carboxyterminal region of the long open reading frame (see text), whose dashed region shows the unknown nucleotide sequence.
The calculated sizes of the polypeptides, together with those observed for the translation products, are given on the right. No initiation codon has been identified for the M_r 85,000 product. The position of clone $\overline{14}$, used in the hybrid-selection, is shown. The highly basic carboxyterminal region is coloured blue and the stretches of amino acid sequence showing homology to AlMV RNA 2 (Cornelissen et al., 1983) and BMV RNA 2 (Ahlquist et al., 1983) are shown in yellow. Within these areas, hatching indicates regions of greatest homology, extending to TMV (Goelet et al., 1982).

PROTEIN $M_r \times 10^{-3}$

calc.	obs.
n.d.	110
n.d.	85
60	60
39	37
30	30
17	17

H.S. clone 14^-

816 600

terminal related proteins encoded by 3'-coterminal subgenomic RNAs have been made for the M_r 30,000 gene of TMV (Hunter et al., 1983).

Sixteen initiation codons were found in phase with the long open reading frame traversing the 3' 2,300 nucleotides of RNA 2. Only 7 of these have a purine in the -3 position and/or a G in the +4 position, thus satisfying Kozak's (1981a) proposed requirements for active initiation codons. Of these 7, 3 would yield translation products of 275, 340 and 528 amino acids (Fig. 4.9), whose sizes (M_r 31,000, M_r 39,000 and M_r 60,000, respectively) correspond to the observed _in vitro_ translation products of M_r 30,000, M_r 37,000 and M_r 60,000 respectively. As already mentioned, the initiation codons nearest the 3' end would produce polypeptides corresponding to the M_r 17,000 product. The minor translation products (M_r 35,000, M_r 47,000, M_r 49,000 and M_r 56,000) seen in some tracks of Figs. 4.7 and 4.8 may arise from other initiation codons (in the same reading frame), which lack the surrounding purines. The sequence from 800 to 2300 nucleotides from the 3' end shows 18 AUG codons in the other two reading frames. Ten of these codons show the surrounding purines required for initiation, but these 10 would produce polypeptides no longer than 46 amino acids. Such small proteins would not have been detected in these experiments.

Each of the major translation products appears to be translated from all RNAs long enough to encode it and up to at least 2500 nucleotides in size, rather than only from those RNAs whose 5' termini lie between a specific AUG initiation codon and the next one towards the 5' end (see Fig. 4.6). It is unlikely that this is due to contamination with smaller RNAs, since these would have been visible in Fig. 4.4. A further possibility is that protein initiation is a leaky process, as far as these internal AUG codons are concerned. RNA sequences and secondary structure may allow some ribosomes to pass 5' AUG codons to reach others e.g. for the M_r 30,000 protein, before initiating peptide synthesis. Other viral mRNAs have been observed to initiate translation _in vivo_ or _in vitro_ at

two or more AUG codons. These include the RNAs from the Rous sarcoma virus src gene (Mardon and Varmus, 1983) and from the E1B genes of adenovirus (Bos et al., 1981; Saito et al., 1983). Lomedico and McAndrew (1982) have shown that the introduction of extra initiation codons into the 5' leader sequence of a preproinsulin mRNA does not significantly affect correct protein initiation. Kozak (1981b) has, however, pointed out that nuclease degradation during *in vitro* translation (Lawrence, 1980) cannot be excluded as the cause of apparently internal initiation.

Computer analysis of the available RNA 2 sequence showed that no specific sequences could be correlated with the start points of more that 2 or 3 of the subgenomic RNAs, although secondary structure analysis of the complete sequence, when available, may disclose features of interest. It is worth noting that immediately 5' to the proposed AUG codon for the M_r 17,000 protein is the sequence

```
                  ↓
RNA 2:            G U U U U G U U G A - C A G G C U G A A G U U U    AUG
                  ● ● ● ●   ● ● ● ●     ● ●     ● ● ● ●
RNA 4:   m7GpppG  U U U A G U U G U U C A - C C U G A  (35 nts)      AUG
                  ⇑
```

which, as shown, is very similar to the 5' sequence of RNA 4, (Gould and Symons, 1982), whose 5' end is marked with the double arrow. The arrowed G from the RNA 2 sequence is residue 885 from the 3' end, but primer extension experiments are needed to determine if this actually represents the start of the 880 nucleotide mRNA.

Both this mRNA and RNA 4 are very efficient mRNAs. The 880 nucleotide mRNA appears to have a very short 5' untranslated region (23 nucleotides if the RNA is indeed 885 nucleotides long) which cannot form any very stable secondary structure. Short, largely single stranded 5' untranslated regions are thought to favour efficient translation of plant viral mRNAs (Ravelonandro et al., 1983; Gehrke et al., 1983).

The amino acid sequence of the long open reading frame traversing CMV RNA 2 shows extensive homology with that from BMV RNA 2 (Ahlquist et al., 1983). The overall homology is 45% for the 400 amino acids from residues 150 to 549 from the carboxyterminus of the CMV protein, when these are aligned for maximum homology with residues 183 to 582 from the carboxyterminus of the BMV gene. The corresponding amino acid sequence from AlMV RNA 2 (Cornelissen et al., 1983) spans 419 residues and shows 30% sequence homology to CMV RNA 2. In the read-through portion of the TMV M_r 183,000 gene (Goelet et al., 1982) 414 amino acids show 15% sequence homology to the 400 residues of the CMV RNA 2 gene. This homology between CMV and TMV is concentrated in two blocks, as shown in Fig. 4.9. The CMV 3'-coterminal subgenomic RNAs selectively express the conserved regions, especially those homologous to TMV and CMV, in the M_r 30,000 and M_r 37,000 translation products.

The carboxyterminal 145 amino acids of the CMV RNA 2 gene, which are thought to correspond to the M_r 17,000 translation product, show very little sequence homology to the BMV RNA 2 gene and none to those of AlMV. This 145 amino acid sequence from CMV encodes a basic protein, with 18% arginine or lysine residues, and 10% acidic residues. The preceeding 400 amino acids, covering the conserved regions have 12% basic and 11% acidic residues. Therefore the CMV RNA 2 carboxyterminal M_r 17,000 protein appears to be a basic, CMV-specific product, whose gene is adjacent to more neutral, conserved domains of the full length reading frame of RNA 2.

CHAPTER 5

SUBGENOMIC RNAS FROM RNA 1

INTRODUCTION

As suggested in Chapter 1, CMV RNA 1 is unique in that it requires only about 65% of its sequence to encode the full-length *in vitro* translation product of M_r 95,000. The remaining sequence is about 1000 nucleotides long, excluding approximately 300 conserved nucleotides at the 3' end and a putative 5' leader sequence of 100 nucleotides. These 1000 nucleotides could encode a protein of M_r 35,000 and are assumed to lie at the 3' end of RNA 1. The nucleotide sequence of RNA 1 thus becomes of especial interest.

The work described in Chapter 4 raises the further question of whether RNA 1 also generates any subgenomic mRNAs. In this chapter I report the initial characterization of subgenomic mRNAs from RNA 1 and of their *in vitro* translation products.

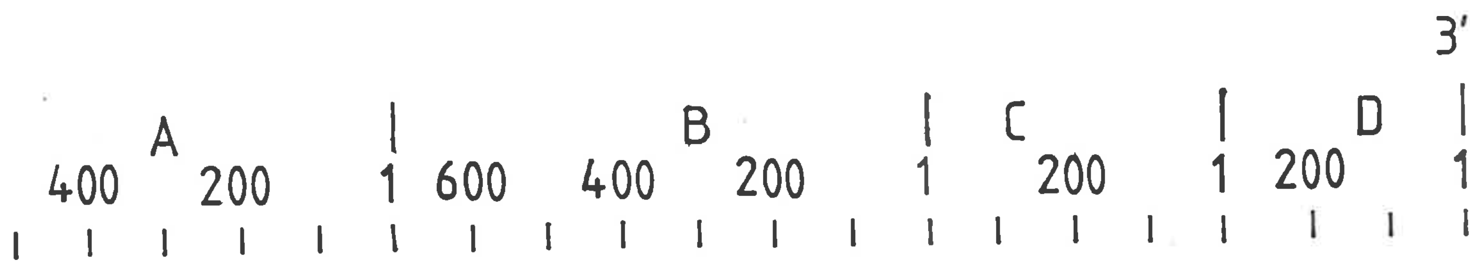
RESULTS

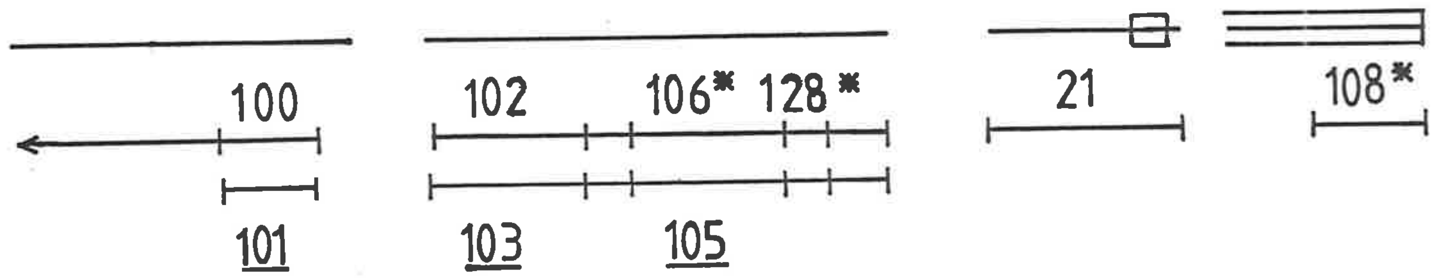
A. Clones from CMV RNA 1

Several Taq I and Sau 3A I restriction fragments, from cDNA to the 3' third of Q-CMV RNA 1, were cloned into M13mp7 and sequenced, as described in Chapter 4 (A. Gould, A. Rezian, R. Williams, unpublished). Direct RNA sequencing with cloned DNA primers, by the dideoxynucleotide chain termination technique (Chapter 4) was used to check the sequences of the clones and to link some of them (R. Williams, unpublished). As shown in Fig. 5.1, the clones form 4 unconnected blocks of sequence, with most being in blocks A and B. At least one (+) clone from each of blocks A, B and D was checked by RNA blotting (results not shown). Three (-) clones 101⁻, 103⁻ and 105⁻, were shown to contain sequences complementary to RNA 1, by hybridization-selection experiments (see below). These (+) and (-) clones from unique RNA 1 sequences, were used in a preliminary analysis of RNA 1 subgenomic RNAs, despite the lack of contiguous sequence information.

FIGURE 5.1

The extent of sequence information at the 3' end of RNA 1 (A. Gould, A. Rezian, R. Williams, unpublished). This consists of 4 unlinked blocks, each containing 1 (C,D) or more (A,B) clones. Clones marked with asterisks were identified as RNA 1 clones by RNA blotting(not shown); underlined clones were identified by hybrid-selection (see Fig. 5.3). The clones in B have been linked by direct dideoxy sequencing of RNA (R. Williams, unpublished). All methods were as in Chapter 4. Clone 108^{+} is within the 3' conserved sequence (Symons, 1979) indicated by a box (the 5' limit has not been determined). Clone 21 has not been confirmed as from RNA 1, but contains 1 stretch (boxed) of sequence homology with RNA 2 (333 to 386 nucleotides from the 3' end) and none with RNA 3. The relative order shown for A and B is based on hybrid-selection data (see text).





(+)

Taq I

(−)

CLONES

(+)

Sau3A I

RNA

B. Subgenomic RNAs from RNA 1

When total encapsidated RNA from virus grown in tobacco was analysed by RNA blotting and a probe from clone 128^+(located near the 3' end of RNA 1 - see Fig. 5.1), many subgenomic RNAs down to below 1000 nucleotides were detected (Fig. 5.2, track 1). As already observed in Fig. 4.3, however, some of the strongest apparent subgenomic bands corresponded to the other genomic RNAs (in this case, 2, 3 and 4). Similar gel artefacts have been observed by Palukaitis et al. (1983). Total, encapsidated, virus RNA, from either tobacco or cucumber, was therefore subjected to hybridization-selection with a mixture of clones 101^-, 103^- and 105^-, since the relative order of these clones was not known for certain. The hybrid-selected RNAs were electrophoresed on a 1.6% agarose gel. Probe was from clone 108^+, located in the 3' conserved region of the CMV RNAs. Tracks 2 and 3 (Fig. 5.2) show the over 20 subgenomic RNAs detected by this experiment. This RNA pattern was clearly different to that from RNA 2 (Fig. 4.3, tracks 5,6), analysed on the same agarose gel. The sizes of the subgenomic RNAs from RNA 1 are listed in Table 5.1. Note that the 900 and 800 nucleotide RNAs are apparently not 3'-coterminal (Fig. 5.2, tracks 2,3) although they were detected with clone 128^+ (track 1).

In order to determine whether these subgenomic RNAs could be detected in plants, RNA extracted from the particulate fraction of CMV-infected plants was subjected to hybridization-selection with clones 101^- and 103^-. The hybrid-selected RNAs (Fig. 5.2, tracks 4,5) were electrophoresed on the same 1.6% agarose gel as used from tracks 2,3; the probe was from clone 108^+. Very similar subgenomic RNA patterns were observed to those present in the encapsidated RNA, at least for RNAs over 1400 nucleotides in size. The smaller RNAs (down to 1000 nucleotides) were presumably too short to be selected by clones 101^- and 103^-, although they were observed in RNA blots of total RNA extracted from the plant fractions (results not shown). These observations suggest that the order of clone blocks A and B is as shown in Fig. 5.1.

FIGURE 5.2

Subgenomic RNAs from RNA 1. Total virus RNA (V) (10ug) was denatured with glyoxal and electrophoresed on a 1.6% agarose gel. The probe was cDNA to clone 128^{+}(track 1) Tracks 2 and 3 show RNAs (H.S.) selected from virus RNA by hybridization to a mixture of clones 101^{-}, 103^{-}, and 105^{-}. The probes was from clone 108^{+}, after 1.6% agarose gel electrophoresis. Tracks 4 and 5 show RNAs selected from extracts of CMV-infected plants (I.E.) by hybridization to a mixture of clones 101^{-} and 103^{-}. Probe was cDNA to clone 108^{+}. Separate gels were used for tracks 1, and 2 - 5. The positions of CMV marker RNAs 1 - 4, and Sat-RNA, are indicated (M). For Methods, see Chapter 4. The RNAs marked with arrowheads in track 3 correspond to similarly marked RNAs in Table 5.1, which lists all the subgenomic RNAs. A size scale in Kilobases (KB) is given.

H.S.

V | V | I.E.

T | C | T | C | T

O – | – O

KB | M | M | KB

4 – 1 – | – 1 – 4
3 – 2 – | – 2 – 3
3 – | – 3
2 – | – 2
1·5 – | – 1·5
1 – 4 – | – 4 – 1
0·5 – | – 0·5

SAT –

1 2 3 4 5

TABLE 5.1

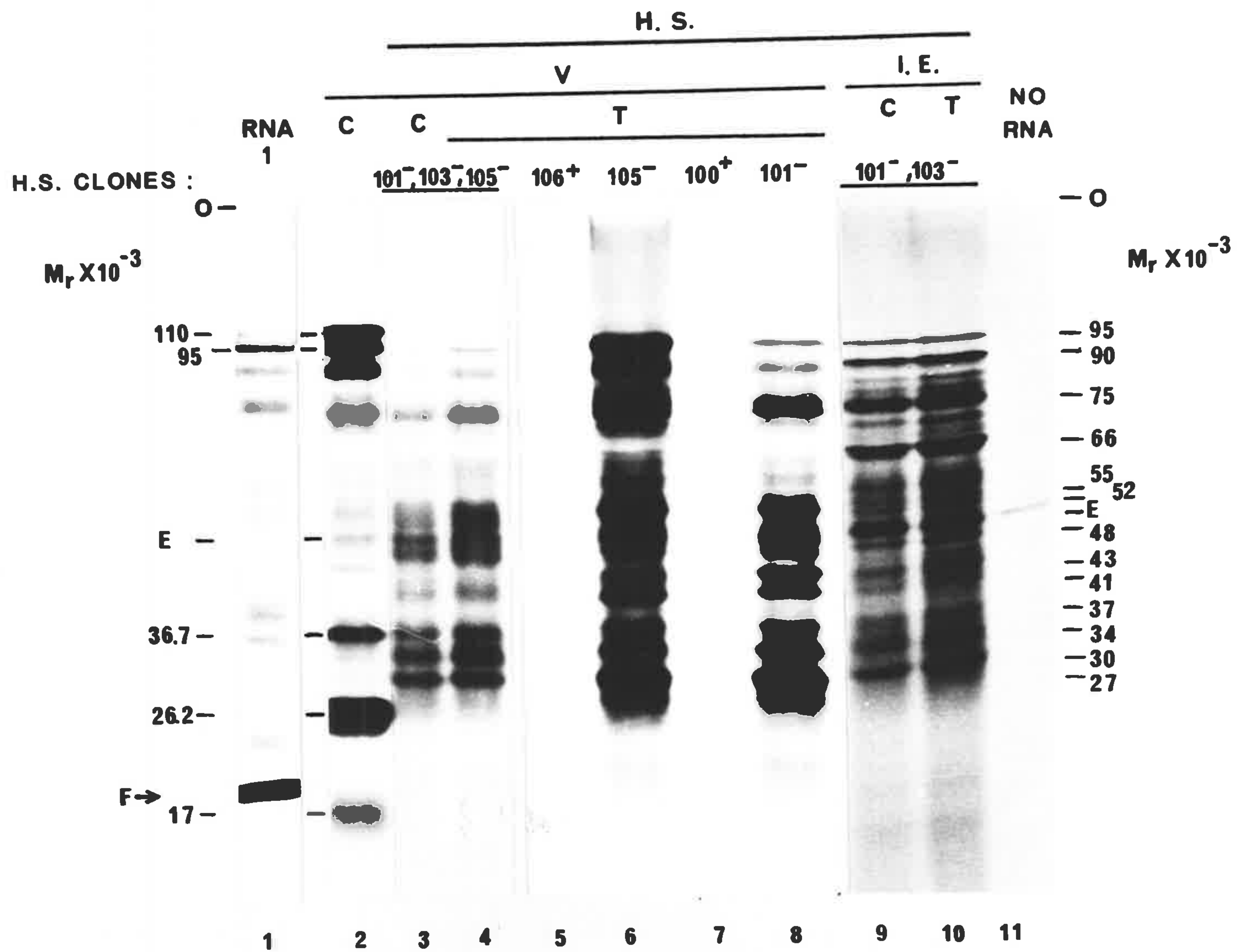
SUBGENOMIC RNAs FROM RNA 1 AND IN VITRO TRANSLATION PRODUCTS [a]

RNAs (nucleotides) [b]	Polypeptides (M_rx10^{-3}) [c]
3,500	90
3,300	75
3,150	66
3,050	55
2,700 ◄	52
2,450	48
2,400	43
2,350 ◄	41
2,200	37
1,850 ◄	34
1,800	30
1,650	27
1,600	
1,520	
1,450	
1,400 ◄	
1,250	
1,150	
1,030 ◄	
900 *	
800 *	
550 *	

a - for possible relationships between RNAs and some translation products, see Fig. 5.5
b - all 3'-coterminal except *; from data in Fig 5.2
c - from data in Fig. 5.3
◄ - refer to RNAs marked (◄) in Fig.5.2, track 3

FIGURE 5.3

In vitro translation products of RNA 1 and its subgenomic RNAs. Track 1 shows the product from RNA 1 which had been purified by 2.8% polyacrylamide gel electrophoresis (Chapter 2); track 2 shows the translation products from (150 μg/ml) total encapsidated RNA (V). Tracks 3 to 10 show the translation products of RNAs hybrid-selected by clones 101^{-}, 103^{-}, 105^{-}, (mixed) (tracks 3,4), 106^{+} (track 5), 105^{-} (track 6), 100^{+} (track 7), 101^{-} (track 8) or 101^{-}, 103^{-} (tracks 9,10) from encapsidated RNA (tracks3-8) or extracts (I.E.) of CMV-infected plants (tracks 9,10). RNAs were derived from either cucumber (C) or tobacco (T). Track 11 is a control translation (no RNA). Molecular weights of the translation products from CMV RNAs 1 - 4 are shown on the left, and from the subgenomic RNAs on the right. 'E' is the endogenous band due to ^{35}S-methionine. 'F' is the dye front in track 1. All products were analysed on 13% polyacrylamide SDS gels (as in Chapter 2).


H. S.
V
I. E.
RNA
1
C
C
T
C
T
NO
RNA
H.S. CLONES :
101⁻,103⁻,105⁻
106⁺
105⁻
100⁺
101⁻
101⁻,103⁻
O
M_r X10^-3
110
95
E
36.7
26.2
F
17
95
90
75
66
55
52
E
48
43
41
37
34
30
27
1
2
3
4
5
6
7
8
9
10
11

C. In vitro translation of subgenomic RNAs

RNA 1, when size-purified on 2.8% polyacrylamide gels (Chapter 2) yielded only one significant translation product, of M_r 95,000 (Fig. 5.3, track 1). This polypeptide was one of the four major products of in vitro translation of total encapsidated CMV RNA (Fig. 5.3, track 2).

In vitro translation of the virus RNAs selected by hybridization to a mixture of clones 101^-, 103^- and 105^- (Fig. 5.2, tracks 2,3) showed many other products, down to M_r 30,000 in size (Fig. 5.3, tracks 3,4). These products were observed after in vitro translation of RNA selected from total encapsidated RNA using clones 105^- (track 6) or 101^- (track 8) alone. In these latter experiments, some higher molecular weight products, up to M_r 90,000 and the full-length M_r 95,000 polypeptide, were observed in greater amounts. In vitro translation of the mixture of RNA 1 and its subgenomic RNAs thus yielded over ten products (as listed in Table 5.1), some of which appeared in variable amounts. The full length M_r 95,000 protein was the longest protein observed, in variable amounts, but never as the major product. In control experiments, the positive-insert clones 106^+ and 100^+ did not select any translateable RNA from total virus RNA (Fig. 5.3, tracks 5, 7, 11).

The RNAs purified by hybrid-selection from RNA extracted from plant fractions, and shown in Fig. 5.2, tracks 4 and 5, also yielded many in vitro translation products. As shown in Fig. 5.3, (tracks 9, 10), these polypeptides correspond to the many products already noted (especially in track 6) with an additional major polypeptide of M_r 66,000. Again, the M_r 95,000 polypeptide was the longest, but not the major product.

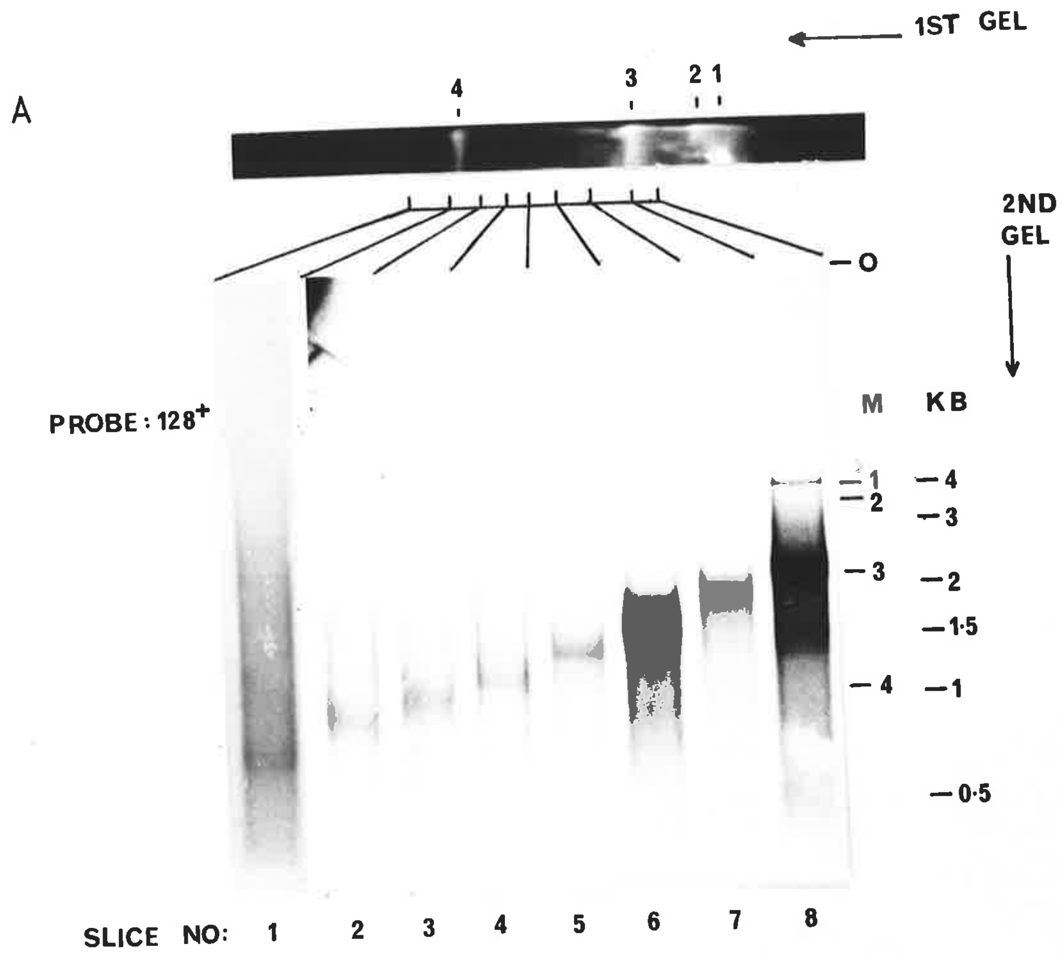
In order to confirm that these translation products were indeed derived from the subgenomic RNAs, and to allocate them to specific RNAs, if possible, total encapsidated RNA was fractionated by electrophoresis on a low melting point agarose gel (see Chapter 4) prior to hybrid-selection with clone 105^-.

FIGURE 5.4

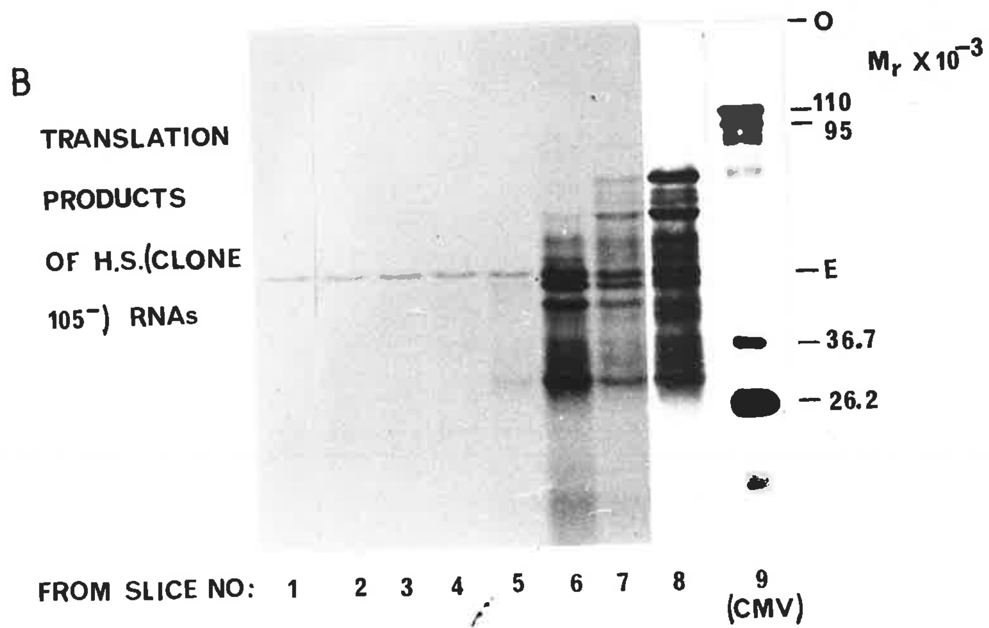
Fractionation and translation of subgenomic RNAs.

A. Total encapsidated RNA (200 μg) was fractionated on a 1.5% low melting point agarose gel as in Chapter 4. RNA 1 subgenomic RNAs from each fraction were selected by hybridization to clone 105^- (10 μg/fraction) and 30% of the selected RNA analysed on a 1.6% agarose gel by cDNA to clone 128^+ (Methods). The positions of CMV RNAs 1-4, run as markers, are shown.

B. Aliquots (30%) of the RNAs shown in A, tracks 1-8, were translated *in vitro* and the products analysed by 13% polyacrylamide SDS gel electrophoresis (Chapter 2). Tracks 1-8 refer to the same samples as in A. Track 9 shows the translation products of total CMV RNA (molecular weights shown on right).



A
1ST GEL
4
3
2 1
2ND
GEL
— O
PROBE : 128⁺
M
KB
— 1
— 2
— 3
— 4
— 4
— 3
— 2
— 1·5
— 1
— 0·5
SLICE NO: 1 2 3 4 5 6 7 8



B
TRANSLATION
PRODUCTS
OF H.S.(CLONE
105⁻) RNAs
— O
Mr X10⁻³
— 110
— 95
— E
— 36.7
— 26.2
FROM SLICE NO: 1 2 3 4 5 6 7 8 9
(CMV)

Aliquots of the hybrid-selected RNA from each fraction shown were analysed by RNA blotting, with a probe from clone 128^{+}(Fig. 5.4A). All RNAs from 800 to 2500 nucleotides were recovered. Note that the initial gel fractionation was incomplete, with RNAs of approximately 1400 to 1500 nucleotides present in tracks 5 to 8. This was probably due to aggregation on the non-denaturing agarose gel.

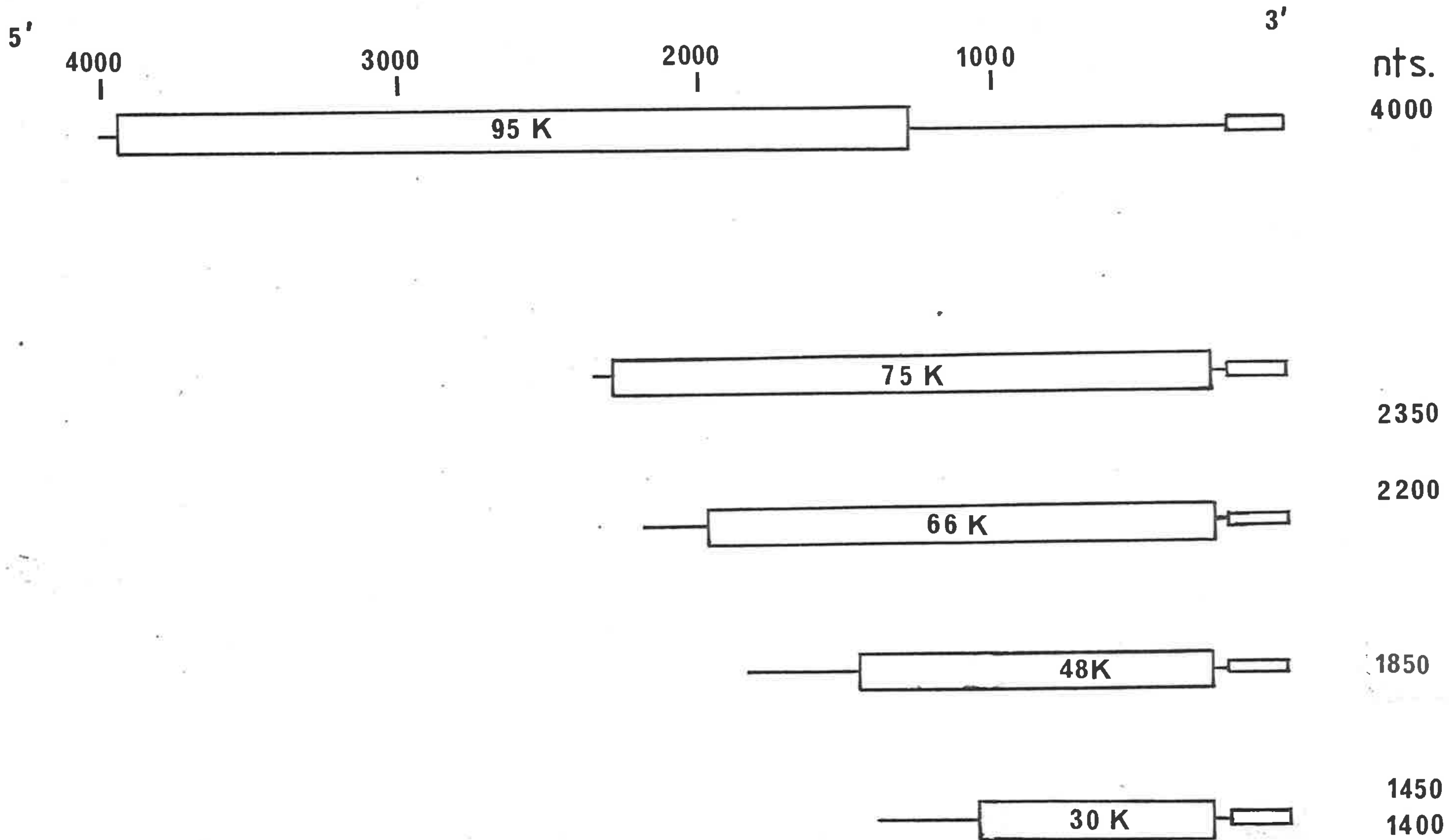
In vitro translation of the RNA fractions of Fig. 5.4A yielded many translation products (Fig. 5.4B) with similar sizes to those observed in Fig. 5.3. Virtually all of the mRNA activity was concentrated in fractions with RNAs over 1500 nucleotides. One fraction (in track 8) had by far the greatest activity and alone produced almost all of the polypeptides from M_r 75,000 to M_r 27,000. As shown in Fig. 5.4A (track 8), this fraction has been enriched for RNAs from 2000 to 2500 nucleotides. The smallest RNAs (tracks 1 to 4) lacked almost any detectable mRNA activity.

DISCUSSION

The results presented in this chapter indicate that RNA 1 possesses a complex genetic organization. This RNA generates many subgenomic RNAs which, like those from RNA 2, are discrete in size, 3'-coterminal with RNA 1 (in most cases) and reproducibly observed in encapsidated virus RNA and in RNA from infected plants. These subgenomic RNAs yield many in vitro translation products - up to twice as many as from RNA 2. Some experimental variability was observed in the occurrence of some of these polypeptides, but the reasons for this are not known. The results shown in Fig. 5.4A, B (especially track 8) suggest that all these proteins arise from genuine subgenomic RNAs; most of these proteins (in track 8 of Fig. 5.4B) were translated very efficiently from this RNA fraction before hybrid-selection (not shown). Hybrid-selection artefacts may be limited to variably efficient hybridization of larger vs. smaller RNAs, when using different M13 clones e.g. tracks 6 and 8 in Fig. 5.3.

FIGURE 5.5

Hypothetical protein coding arrangement of RNA 1. RNA 1 is shown at the top, with the reading frame for the M_r 95,000 translation product in the proposed 5' position (see text). The 3' box represents the about 300 conserved nucleotides. Possible coding regions are indicated for four translation products of some of the subgenomic RNAs. The locations of these frames were on the basis of evidence in Fig. 5.4 A,B: track 8 (M_r 75,000, 66,000), track 6 (M_r 48,000) and track 5 (M_r 30,000).


5'
3'
4000
3000
2000
1000
nts.
95 K
4000
75 K
2350
66 K
2200
48K
1850
30 K
1450
1400

Obviously a complete unravelling of the genetic organization of RNA 1 will have to await the complete determination of its nucleotide sequence.

The observation that the translateable, subgenomic mRNAs (over 1,400 nucleotides long) are 3'-coterminal with RNA 1 suggests that they may indeed play a role in expressing a 3' open reading frame as proposed in the introduction. Clearly, however, no one single translation product can be allocated to such a reading frame. The observations with RNA 2 suggest that a 3' reading frame could encode several polypeptides, with differing aminotermini. The longer polypeptides could derive from a reading frame on RNA 1 overlapping with, but different to, that for the M_r 95,000 polypeptide. This hypothesis is depicted in Fig. 5.5, for some of the major polypeptides whose presence appears to be correlated with specific subgenomic RNAs. Note that these major subgenomic RNAs (e.g. 2,400 to 2,200 nucleotides) are only just long enough to encode the M_r 75,000 or M_r 66,000 products, respectively. It is also possible that a single RNA yields translation products of different sizes, as appears to happen with RNA 2 subgenomic RNAs.

The nucleotide sequences of AlMV RNA 1 (Cornelissen et al., 1983a) and BMV RNA 1 (Ahlquist et al., 1983) have recently been determined. Both these RNAs contain a single open reading frame traversing their full length. It will be of great interest to elucidate any sequence homologies between their translation products and the proposed open reading frames of CMV RNA 1, or between these (CMV RNA 1) and the TMV gene products (Goelet et al., 1982).

CHAPTER 6

CMV SUBGENOMIC RNAS : SYNTHESIS

AND

GENERAL DISCUSSION

INTRODUCTION

Many subgenomic RNAs have been found to be derived from CMV RNAs 1 and 2. These subgenomic RNAs were of discrete, reproducible sizes, generally 3'-coterminal with their genomic RNA and, in many cases, capable of efficient, *in vitro* translation. These observations suggest that the subgenomic RNAs are specific products of RNA replication. In order to determine whether they play a significant role during CMV gene expression *in vivo*, it is necessary to search for their translation products.

The only previously characterised subgenomic RNAs among the tripartite plant RNA viruses were those expressing the coat protein. If subgenomic RNAs prove to have a much wider role in viral gene expression, then the question of how they are synthesised becomes more pressing. Three basic mechanisms have been proposed (e.g. for the replication of AlMV RNA 4; Smit and Jaspers, 1982):

(i) specific cleavage of the full-length (+) RNA,

(ii) internal initiation of transcription on a full-length (-) RNA or

(iii) replication via a subgenomic (-) strand.

Although the elucidation of the details of this process will be a difficult task, a simple question which can be asked now is whether (-) strands corresponding to the CMV subgenomic RNAs exist in a double stranded RNA fraction enriched for possible replicative forms. Such RNA should be present in the particulate fraction, which contains the viral RNA replication complex (Symons et al., 1982; Gill, 1983). In this chapter I present data showing the existence of at least some subgenomic (-) RNAs. Also discussed are the significance of the subgenomic RNAs for viral gene expression, and the roles they may play in the evolution of RNA viruses.

RESULTS

The approach used in this preliminary study of the subgenomic (-) RNAs from RNAs 2 and 1 was to fractionate the total RNA extracted from the

particulate fraction into single and double stranded components, using LiCl precipitation (see Chapter 7). The RNA samples were then glyoxalated and analysed in duplicate on a 1.6% agarose gel. After transfer to nitrocellulose, each sample was probed with equal amounts (in cpm) of single stranded ^{32}P-cDNA from M13 clones containing (+) or (-) inserts of about equal length and sequence. As shown in Fig. 6.1, small amounts of double stranded RNA were obtained, for each of tracks 4, 6, 7, 8 contains the 4M LiCl-insoluble RNA from 200 µg of original extract. High stringency washing conditions were employed after hybridization (see legend to Fig. 6.1), so that this experiment was designed to unambiguously distinguish (+) and (-) sequences present in each sample.

RNA 2 (-) sequences were easily detected (Fig. 6.1A). Surprisingly, total encapsidated RNA was found to contain (-) RNA (track 2), although in much smaller amounts than the (+) RNA present (track 1). The single stranded RNA obtained (by 2M LiCl precipitation) from RNA extracted from two different CMV hosts also included considerable amounts of (-) RNA (tracks 3,5) in similar proportion to the (+) sequences (not shown) as observed for encapsidated RNA (tracks 1,2). The double stranded samples (obtained by 4 M LiCl precipitation) hybridized about equally to the (-) and (+) probes (tracks 4,6 and 7,8, respectively), with no significant difference between samples from the two hosts. In all cases, the major (-) RNA corresponded in size to full-length RNA 2. The next strongest bands in the double stranded samples were RNAs of 1050 and 750 nucleotides (tracks 4, 6-8). Single stranded RNA showed bands of these sizes (tracks 3, 5), in addition to a number of strong bands which corresponded to major RNAs detected by staining with ethidium bromide (not shown), and thus probably were due to gel artefacts (see Chapters 4,5). Other, minor (+) and (-) bands in the double stranded RNAs corresponded to some of the subgenomic RNAs detected in Chapter 4.

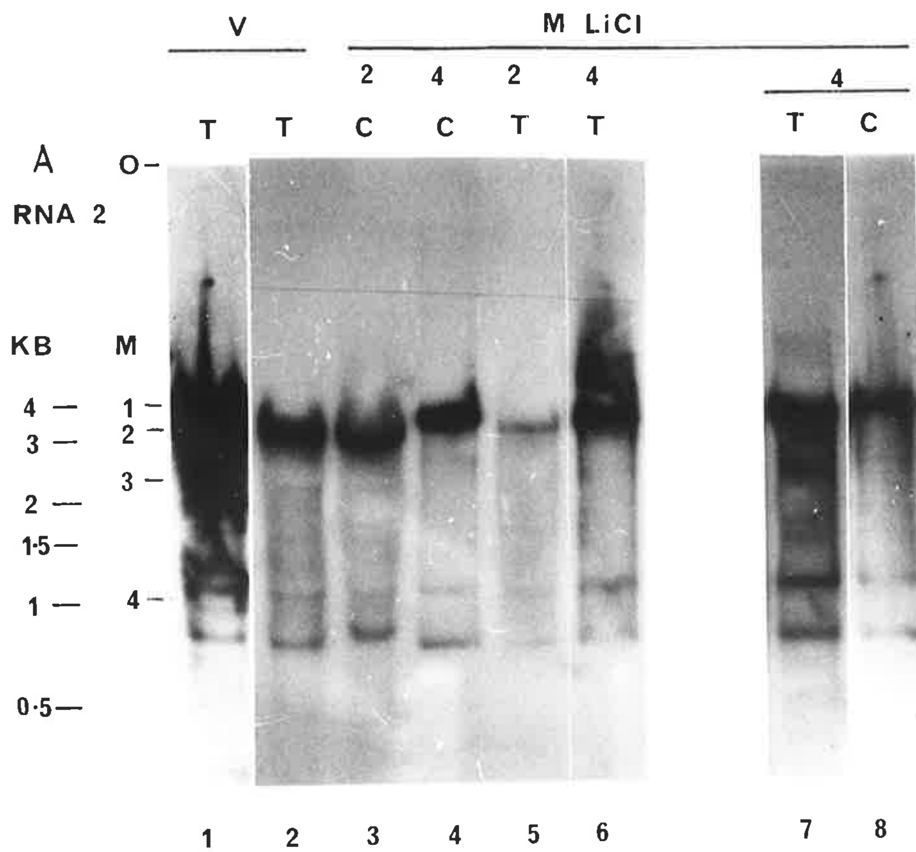
Similar observations were made for RNA 1 (Fig. 6.1B). Again, encapsidated RNA contained (-) RNA, in amounts greater than for RNA 2, although

FIGURE 6.1

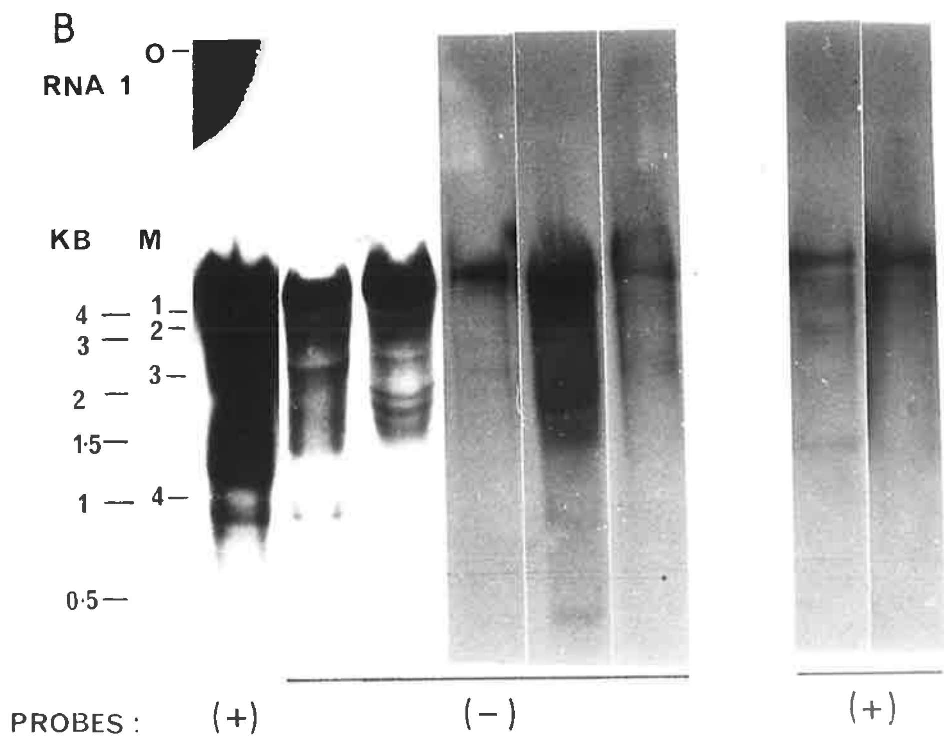
Detection of (+) and (-) sequences in double stranded RNAs from CMV RNA 2 (A) and RNA 1 (B). Total nucleic acid was extracted from the particulate fraction of CMV infected plants and double stranded RNA prepared by LiCl fractionation (see Chapter 7), before glyoxalation and 1.6% agarose gel electrophoresis (see Chapter 4). Total virus RNA (V) (20μg) was run as markers (tracks 1 and 2); the positions of RNAs 1 - 4 are shown (M), as is a scale in Kilobases (KB). The single stranded RNA obtained by 2 M LiCl precipitation (20μg) is shown in tracks 3 and 5, and the double stranded RNA, obtained by 4 M LiCl precipitation, in tracks 4, 6, 7 and 8. CMV was grown in either cucumber (C) or tobacco (T). Specific (+) or (-) cDNA probes were synthesised as described in Chapter 4. Equal amounts (in cpm) of (+) and (-) probes of the same specific activity, from the following clones were used:

RNA 2	(+)	15	A, tracks 1, 7, 8
	(-)	14	A, tracks 2 - 6
RNA 1	(+)	106	B, tracks 1, 7, 8
	(-)	105	B, tracks 2 - 6

Note that the full length (+) and (-) RNA 1 and 2 bands in tracks 4, 6, 7, 8 (A and B) migrated more slowly than the full length markers in tracks 1 and 2, or the CMV RNAs detected by blotting in tracks 3 and 5. The reason for this is not known, but may be due to the much lower overall amounts of RNAs loaded in tracks 4, 6, 7 and 8. A similar effect has been observed in Fig. 4.3 (tracks 1 - 4). The prominent subgenomic RNAs visible in tracks 2 and 3 (A and B) comigrated with host RNAs detected by staining; the latter were present in RNA extracted from uninfected plants, but did not hybridize to the specific probes used (not shown).



V
M LiCl
2
4
2
4
4
T
T
C
C
T
T
T
C
A
O–
RNA 2
KB
M
4 —
3 —
2 —
1·5 —
1 —
0·5 —
1 –
2 –
3 –
4 –
1
2
3
4
5
6
7
8



B
O–
RNA 1
KB
M
4 —
3 —
2 —
1·5 —
1 —
0·5 —
1 –
2 –
3 –
4 –
PROBES :
(+)
(–)
(+)

still in minor proportion to the (+) RNA (tracks 2,1). Since the same filters were used as in Fig. 6.1A (after boiling in water for 5 min to remove the RNA 2 probes), this observation may represent a real difference in encapsidation. The single stranded RNAs contained considerable (-) RNA (tracks 3,5) again showing many apparent subgenomic species, of which some are probably gel artefacts. The double stranded RNA fractions contained equal amounts of (+) and (-) full-length RNA 1 (tracks 4, 6, 7, 8), but no other major bands. Many faint (+) and (-) bands between 1500 and 3500 nucleotides long were detected, corresponding to the size range of most of the larger subgenomic RNAs. These subgenomic RNAs are not shown very clearly by the reproductions in Fig. 6.1B.

DISCUSSION

A. Synthesis of subgenomic RNAs

This preliminary study of the double stranded RNAs present at low levels in CMV-infected plants has shown that at least some of the many subgenomic RNAs identified for CMV RNAs 1 and 2 possess corresponding (-) RNAs in totally or partially double stranded form, which may be involved in replication. It is not known, however, whether the other subgenomic RNAs lack (-) forms, or whether their levels are below detection. Obviously a more detailed analysis, e.g. using hybridization-selection, is required to determine the complete (-) RNA complement of the replication complex. Furthermore, it is possible that the double stranded RNAs have arisen through nuclease degradation of hybrids formed from replication intermediates by full-length (-) and subgenomic (+) RNAs or through copying of (+) subgenomic RNAs to give dead-end double stranded RNAs.

Other reports have been made of double stranded RNAs corresponding to the subgenomic RNA 4 of CMV, in CMV-infected plants (Diaz-Ruiz and Kaper, 1978) and in CMV-infected tobacco protoplasts (Takanami et al., 1977). It should be noted that the 1050 bp RNA from RNA2 is almost exactly the same size as double stranded RNA 4, so that any putative replicative form

of RNA 4 would have to be confirmed e.g. with a specific probe. Double stranded RNAs corresponding, in size, to RNA 4 of AlMV (Pinck and Hirth, 1972; Bol et al., 1975) and BMV (Bastin and Kaesberg, 1976; Loesch-Fries and Hall, 1980) have also been observed, but the former may be degradation products (Mohier et al., 1974) or not specific for RNA 4 (Bol et al., 1976a). In fact, a more comprehensive study of AlMV replication in protoplasts has suggested that AlMV RNA 4 is produced by internal initiation of transcription on full-length (-) RNA 3, since no specific (-) RNA 4 could be detected (Nassuth et al., 1983). In contrast double stranded subgenomic RNAs of TMV have frequently been detected in TMV-infected plants (Zelcer et al., 1981; Dawson and Dodds, 1982; Goelet and Karn, 1982; Palukaitis et al., 1983) although they seem to be fewer in number than the many subgenomic RNAs detected (Hunter et al., 1983).

The cytoplasmically replicating RNA viruses in which the problem of subgenomic RNA synthesis has been most closely studied are the alphaviruses. Evidence here has implicated internal initiation of transcription on a full-length (-) strand, in the synthesis of a subgenomic RNA (Sawicki et al., 1978; Petterson et al., 1980).

Families of 3'-related subgenomic mRNAs have been observed for a number of animal viruses. Among the best studied examples are those from adenovirus (Ziff, 1980), especially the E1B region (Bos et al., 1981; Saito et al., 1982) and the subgenomic mRNAs expressing retroviral _env_ and _src_ genes (Hayward, 1977; Mellon and Duesberg, 1977; Weiss et al., 1977; Krzyzek et al., 1978). Other examples include the subgenomic mRNAs derived from RNAs 7 and 8 of influenza virus (Lamb and Choppin, 1983) and from coronavirus RNA (Siddell et al., 1982).

In all these cases, the 3'-coterminal subgenomic mRNAs possess spliced 5'-leader sequences, which, for the RNA viruses, are derived from the 5'-terminal genomic RNA sequence. The splicing of 5'-leader sequences onto the coronavirus subgenomic mRNAs (Lai et al., 1982) represents the first

such example in a cytoplasmically replicating RNA virus. In this case, the apparent splicing is thought to be due to a mechanism, in which the RNA dependent RNA polymerase, attached to the leader sequence, re-initiates transcription at internal sites on the full-length (-) RNA (Mahy, 1983). This raises the question of whether the CMV subgenomic RNAs have spliced 5'-leader sequences. The data do not exclude this possibility, but correlation of the primer extension and RNA blotting data indicates that any putative spliced segments would have to be quite short (e.g. under 100 nucleotides long).

B. CMV subgenomic RNAs and gene expression

In vitro translation of either CMV RNA 1 or 2 and its subgenomic RNAs yields a diverse population of polypeptides, probably covering all available reading frames. The genomic RNAs produce full-length polypeptides, including the aminoterminal regions. The aminoterminal sequences are also present in any early termination products, e.g. the less abundant and shorter polypeptides observed upon translation of size-purified genomic RNAs (see Chapter 2). In contrast, the 3' subgenomic RNAs provide the carboxyterminal regions of these polypeptides, in addition to expressing any silent genes at the 3' end.

The proper test of the significance of these subgenomic RNAs is to determine whether their translation products can be detected in CMV-infected plants. Although the subgenomic RNAs available for translation may vary during the course of the infection, or at different intracellular locations, it seems unlikely that only the genomic RNAs are translateable even if their translation products are the major ones detected upon *in vitro* translation of total encapsidated RNA - a somewhat artificial situation.

A remaining paradox is the efficiency with which the subgenomic RNAs of CMV RNAs 1 and 2 compete with their respective genomic RNA during *in vitro* translation of hybrid-selected RNAs. Since the full-length RNA is still the most abundant single species, it seems unlikely that the hybrid-

selection process is significantly biased e.g. towards short RNAs. Possibly, gene products from several viral genomic RNAs combine during *in vitro* translation of total encapsidated RNA to influence the translation product pattern.

The efficiency with which the subgenomic RNAs appear to compete with genomic RNAs for translation *in vitro* may be related to another intriguing dilemma: the inability to detect the full-length translation products of CMV RNAs 1, 2 and 3 *in vivo*. They have been hunted for in CMV-infected cowpea protoplasts, cucumber cells, cucumber leaf slices at early and later stages of infection and cucumber seedlings, using radiolabelled amino acids and cell fractionation techniques (Gonda and Symons, 1979; Haseloff, 1979; Gordon et al., 1982; and unpublished data). In contrast, the coat protein has easily been detected in CMV-infected protoplasts (Gonda and Symons, 1979) and even by staining following SDS-gel electrophoresis of particulate fraction proteins from cucumber (Gordon and Gill, unpublished data).

Possible explanations which have been advanced for the failure to find these proteins *in vivo* include their synthesis at low levels, rapid processing or the unsuitability of the extraction procedures used (Symons et al., 1982). The reason may lie, in part, in the case with which the subgenomic mRNAs can be translated, so that their products become the preferred means by which CMV expresses its genetic information. They could easily have hitherto escaped detection among the many proteins between M_r 10,000 and M_r 100,000 in the various extracts analysed. Antibodies raised against synthetic peptides corresponding to defined portions of each protein would be a powerful tool in the study of these proteins (Sheppard, 1983; Walter and Doolittle, 1983).

The CMV subgenomic RNAs may solve a further problem faced by CMV during its infection cycle. The (+) genomic RNAs participate in three competing processes - replication, translation and encapsidation. It is likely that encapsidation is linked to replication, so that the effective competition is between translation and replication. In the absence of a

regulatory mechanism, each process could initiate at opposite ends of one (+) RNA molecule, leading to conflict when the macromolecular complexes collide. Viral RNAs appear to have solved this dilemma by various strategies, involving modifications to their termini (to distinguish between encapsidated and messenger RNAs) or sequestration of one (e.g. the 5') end through intermolecular interactions (reviewed in Kozak, 1981b).

While none of these mechanisms appear to apply to the CMV RNAs, the subgenomic RNAs may allow CMV to tolerate the interruption or displacement of protein synthesis from any one RNA by a progressing RNA replicase. Such displaced protein synthesis would still yield polypeptides corresponding to the aminoterminal part of any long open reading frames. Translation of the subgenomic mRNAs would provide complementary peptides corresponding to the carboxyterminal region. An assumption behind this mechanism is that the CMV gene products contain different functional domains, which do not need to be linked in a single polypeptide molecule. This mechanism represents a more direct response to this biological problem than hypothetical interactions involving protein translation factors (Hall and Wepprich, 1976; Ahlquist et al., 1983).

C. Other plant viral subgenomic mRNAs

3'-coterminal subgenomic mRNAs may represent a common strategy for gene expression adopted by plant RNA viruses. Apart from the coat protein mRNAs, the best characterised such subgenomic mRNAs are derived from the M_r 30,000 gene of TMV (Hunter et al., 1983), where a family of RNAs from 1900 to 1500 nucleotides long encodes carboxyterminal-overlapping proteins of M_r 30,000, M_r 28,000 and M_r 23,000. Other TMV proteins which overlap with the longer (M_r 126,000 and M_r 183,000) gene products have been reported (Zelcer et al., 1981; Goelet and Karn, 1982; Joshi et al., 1983; Wilson and Glover, 1983). The full length translation products of TMV genomic RNA have been detected *in vivo* (Scalla et al., 1978).

Although there is no evidence about the *in vivo* function ofthe subgenomic RNAs from CMV RNAs 1 and 2, the comparison of these CMV RNAs to those from

the TMV M_r 30,000 gene is significant, because the latter subgenomic RNAs appear to be the only way by which this TMV gene can be expressed *in vivo* (Hunter et al., 1983). This cistron is known to be expressed *in vivo*, from genetic evidence (Leonard and Zaitlin, 1982). It is intriguing that crude fractionation of the encapsidated RNA from several other plant viruses, e.g. lucerne transient streak virus, tobacco etch virus and turnip crinckle virus, followed by *in vitro* translation, has disclosed polypeptides (other than the viral coat protein) whose mRNAs were smaller than the genomic RNA (Dougherty and Kaesberg, 1981; Morris-Krsinich and Foster, 1983; Otal and Hari, 1983).

Although no systematic study has been made of subgenomic mRNAs produced by any other tripartite virus, genetic studies of AlMV have identified two complementation groups on its RNA 1 (reviewed in Lane, 1979). *In vitro* translation studies of this RNA (purified by density gradients) have identified several smaller than full-length products (van Tol and van Vloten-Doting, 1979; Neeleman and van Vloten-Doting, 1982) although the sequence shows no internal termination codons (Cornelissen et al., 1983a). It is tempting to speculate that AlMV RNA 1 may also generate subgenomic RNAs. Possible subgenomic RNAs from AlMV have been observed, but not further characterised (Bol et al., 1976b).

D. Subgenomic RNAs and the evolution of CMV

If the CMV subgenomic RNAs are indeed important for *in vivo* gene expression, then they may also play a major role in the evolution of plant RNA viruses. The subgenomic RNAs may allow the virus to regulate the selective expression of different portions of each gene. This would afford greater flexibility in the virus's response to different environments and, especially for RNA 1, allow greater versatility in expressing possibly overlapping genes.

By expressing the carboxyterminal regions of viral genes, the subgenomic RNAs may serve to make the virus more tolerant of changes or differing mutation rates along its genomic RNAs. A region in the carboxy-

terminal half of the gene on CMV RNA 2 is conserved (see Chapter 4) among plant RNA viruses; this is the region expressed by the subgenomic mRNAs. If this finding becomes generally applicable, then the evolution of new viruses may result chiefly from alterations in the aminoterminal regions of the major genes. Possibly such changes or mutations affect those regions of the protein involved in host interactions, symptom expression etc., whereas the carboxyterminal regions perform more fundamental functions e.g. RNA replication.

Indeed, a totally new viral RNA, encoding a new gene product could even be derived from the linking of a 3'-coterminal subgenomic RNA from one virus to another viral RNA, thus yielding a new gene with the same fundamental capabilities but altered host-specificity. Such a rearrangement is most likely to be mediated by a viral induced RNA replicase after co-infection of a single host with several viruses. Reports of recombination in RNA viruses have been made (King et al., 1982; Fields and Winter, 1982) and RNA viruses have long been known to undergo rapid evolution (Holland et al., 1982; Reanney, 1982). Zimmern (1982) has, however, proposed the existence of such RNA-fusing mechanisms in virus-free plants.

CHAPTER 7

CMV SATELLITE RNA : SECONDARY STRUCTURE

AND

REPLICATIVE FORMS.

INTRODUCTION

Sequence and structural studies on satellite RNAs of CMV are important for several reasons. Not only might such studies provide information about their possible origins, but also, when compared to the known sequence and structural features of the viral RNAs, about minimum requirements for viral RNA replication. Very little direct sequence homology exists between the Adelaide Sat-RNA described in Chapter 1 and the known genomic RNA sequences. The longest stretches of homology are only 11 nucleotides long, and occur in isolation.

The RNA replicase is likely to recognize and interact with RNA structures rather than just a linear nucleotide sequence. Secondary structure formed by Sat-RNA was therefore investigated using nuclease digestion under non-denaturing conditions. The data allowed a complete secondary structure model to be derived. The sequences of other isolates of satellite RNAs (Richards et al., 1978; Collmer et al., 1983) also fit this model, which shows striking sequence and structural homology between its 3'-terminal structure and that formed by the genomic RNAs of CMV.

CMV satellite RNAs are replicated very efficiently *in vivo* and large amounts of double stranded CARNA 5 have been isolated from infected leaves (Diaz-Ruiz and Kaper, 1978). These double stranded forms apparently exist as unencapsidated end products *in vivo*, and are not derived from replicative forms (Habili and Kaper, 1981). In order to further analyse these RNAs, and to search for possible higher molecular weight replication intermediates, as have been found for the satellite RNA of tobacco ring-spot virus (Kiefer et al., 1982; Sogo and Schneider, 1982), total RNA was extracted from the particulate fraction (see Chapters 3,4) of cucumbers infected with Q-CMV and Sat-RNA. This work has confirmed the double stranded monomeric forms, but failed to find extensive oligomeric forms, using probes derived from M13 clones.

MATERIALS AND METHODS

A. Virus and RNA

Cucumber seedlings were inoculated with Q-CMV and Sat-RNA, virions purified 12-14 days after infection and virus RNA extracted, as described in Chapter 2. Sat-RNA was purified on 5% polyacrylamide, non-denaturing gels in 90 mM Tris-acetate, pH 7.5, 1 mM EDTA (Symons, 1978). The particulate fraction was prepared and RNA obtained as described in Chapter 4.

B. Materials

T4 polynucleotide kinase and nuclease S_1 were from Boehringer; T4 RNA ligase was from P. L. Biochemicals; ribonucleases A and T_1, and calf intestinal phosphatase were from Sigma; ribonuclease U_2 was from Sankyo. (α-^{32}P)dNTPs and (γ-^{32}P)ATP (specific activity 2kCi/mmol were from BRESA.

C. cDNA clones in M13mp9

Restriction endonuclease Hae III cleaves double stranded cDNA to Sat-RNA twice, at nucleotides 122 and 196. The resulting 74 bp fragment was cloned (by blunt end ligation), into the Sma 1 site of the bacteriophage vector M13mp9, by P. Keese and Dr. G. Bruening. Clones in both orientations were obtained, with the (+) insert (clone A9) and the (-) insert (clone A6) checked by dideoxy sequencing (in collaboration with P. Keese).

D. Secondary structure analysis of Sat-RNA

(i) 5'-^{32}P-labelling of Sat-RNA.

The m^{7}G cap was removed by periodate oxidation, followed by β-elimination in the presence of aniline, essentially as described by Fraenkel-Conrat and Steinschneider (1968). After ethanol precipitation, the RNA pellet was dried *in vacuo* with 200 µCi of (γ-^{32}P) ATP, redissolved in 10 µl of 25 mM Tris-HCl, pH 9.0, 10 mM $MgCl_2$, 10 mM DTT, 1 mM spermidine (Haseloff and Symons, 1981) and incubated with 4 units of polynucleotide kinase at 37°C for 30 min. Ten µl of loading buffer (95% deionised formamide, 10 mM EDTA, 0.02% bromophenol blue, 0.02% xylene cyanol FF) was added to each

reaction mixture and the end-labelled RNA purified by electrophoresis on 40x20x0.05 cm 3% polyacrylamide, 7 M urea gels, followed by elution at room temperature in 0.5 M ammonium acetate, 0.1% SDS, 1 mM EDTA, and ethanol precipitation.

(ii) 3'-^{32}P-labelling of Sat-RNA.

To synthesise (5'-^{32}P) dpCp, 500 µCi of (γ-^{32}P)ATP was dried _in vacuo_, and reacted with 20 µg of 3'-dCMP and 4 units of polynucleotide kinase in 10 µl of the reaction mixture described in (i) above, at 37°C for 30 min. The mixture was then heated to 90°C, for 1 min and stored at -15°C until used.

The precipitated Sat-RNA was dried _in vacuo_, resuspended in 20 µl 10mM Tris-HCl pH 9.0 containing 0.01 units calf intestinal alkaline phosphatase, and incubated at 37°C for 20 minutes. The reactions were then extracted with 100 µl water saturated phenol : chloroform (1 : 1) and 100 µl 0.2M NaCl, 0.1mM Na_2EDTA. The aqueous phase was removed, washed twice with 1 ml ether and the RNA precipitated with 450 µl ethanol at -80°C for 20 minutes. The reaction tubes were centrifuged at 10,000g for 15 minutes at 4°C and the supernatant was discarded. The precipitated, phosphatase treated RNA was dried _in vacuo_, resuspended in 5 µl H_2O, heated at 80°C for one minute and snap cooled on ice. 1 µl of (5'-^{32}P)dpCp (50 µCi), 6 µl of 2x T_4 RNA ligase buffer (100mM HEPES pH 7.5, 6.6mM DTT, 30mM $MgCl_2$, 20% (v/v) redistilled DMSO, 100 uM ATP) and 1 µl T_4 RNA ligase (4.6 units, 1.5 µg) were added, and the reaction was incubated at 4°C for 16 hours (England and Uhlenbeck, 1978). The 3'-^{32}P-Sat-RNA was purified as described in (i) above.

(iii) Nuclease digestions of Sat-RNA

For digestion with nuclease S_1, end-labelled Sat-RNA was preincubated for 10 min at 37°C in 40 mM sodium acetate, pH 5.0, 1.0 mM $ZnSO_4$, 0.2 M NaCl, 330 µg/ml carrier _E. coli._ tRNA, and then for 10 min at 37°C in the presence of varying concentrations of nuclease S_1 (see figure legends); final reaction volume was 6 ul for each assay. Digestions were stopped

by addition of 2 µl 1% SDS, 50 mM EDTA. For digestions with RNase T_1, end-labelled Sat-RNA was incubated for 10 min at 37°C with either 0.01 or 0.1 units RNase T_1 in 6 µl 20mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 0.2 M NaCl, 330 µg/ml carrier E. coli tRNA and terminated as for nuclease S_1 digestions.

To enable the precise location of nuclease sensitive bonds, partial RNase T_1 and RNase U_2 digests of RNA plus random alkali cleaved RNA were prepared as for the sequencing procedure using specific partial enzymic cleavage and electrophoresed as reference tracks (Haseloff and Symons, 1981). Aliquots (1 µl) of Sat-RNA, each with 10 µg of E. coli tRNA were digested in the following buffers:

RNase T_1 (total volume 9 µl): 20 mM Na-citrate, pH 5.0, 1 mM EDTA, 7 M urea;
RNase U_2 (total 9 µl): 20 mM Na-citrate, pH 3.5, 1 mM EDTA, 7 M urea.
After heating to 80°C for 1 min and snap-cooling on ice, 1 µl of RNase T_1 (10 units) or RNase U_2 (0.005 units) were added and the RNA digested at 50°C for 15 min. For random alkali cleavage, 5 µl of 50 mM $Na_2CO_3/NaHCO_3$, pH 9.0, were added to 1 µl of RNA and the tube heated at 100°C for 90 sec. All reactions were cooled on ice upon completion.

After addition of 5 µl loading buffer (95% deionised formamide, 10 mM EDTA, 0.03% xylene cyanol FF, 0.03% bromophenol blue) to each reaction mix, all samples were denatured by heating at 80°C for 1 min followed by snap cooling prior to analysis by electrophoresis on 40 x 20 x 0.05 cm 20% acrylamide gels or 80 x 20 x 0.05 cm 8% acrylamide gels in 90 mM Tris-borate, pH 8.3, 1 mM EDTA, 7 M urea.

(iv) Generation of base-paired fragments with RNase T_1

Unlabelled Sat-RNA (10 µg) in 100 µl 20 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 0.5 M NaCl, was digested with RNase T_1 at 1,000 units/ml at 0°C for 1.0 h. After phenol-chloroform extraction and ethanol precipitation, the RNA fragments were 5'-end labelled with γ-^{32}P-ATP (200 µCi) and polynucleotide kinase as described above.

After another phenol-chloroform extraction and ethanol precipitation,

the labelled fragments were dissolved in 20 µl of 10% (v/v) glycerol, 1 mM EDTA, and fractionated by electrophoresis on a 40 x 20 x 0.05 cm non-denaturing 12% polyacrylamide gel in 90 mM Tris-borate, pH 8.3, 1 mM EDTA, for 6 h at 12 mA. After detection by autoradiography, labelled bands were excised and recovered by elution as above, and fractionated into individual, end-labelled fragments on a denaturing 12% polyacrylamide gel as above but in the presence of 7 M urea. After elution, each band was identified using partial enzyme digestion with RNases T_1 and U_2 and random alkali cleavage followed by gel electrophoresis, as described above.

E. Detection of (+) and (-) Sat-RNA

(i) LiCl fractionation of RNA extracts

Up to 2 mg of RNA extracted from the particulate fraction was made 2 M LiCl and kept at 0-4°C for 8 h, then centrifuged at 10,000g for 10 min (Diaz-Ruiz and Kaper, 1976). The precipitate was redissolved in 0.3 M Na acetate and recovered by ethanol precipitation. The supernatant was made 4 M LiCl and kept at 0-4°C for 12 h, then the precipitate recovered as for the 2 M LiCl step. The remaining supernatant, containing DNA and tRNA, was discarded.

(ii) RNA blotting

To test for single stranded RNAs contaminating the 4 M LiCl precipitated RNA, part of each 4 M LiCl precipitate was digested with RNases T_1, and A before RNA blotting. Digestions were carried out, in a total volume of 100 µl of 2 x SSC, with 30 units/ml RNase T_1 and 0.5 µg/ml RNase A, at 37°C for 30 min. The reactions were terminated by extraction with an equal volume of (1 : 1) phenol/chloroform and the RNA recovered by ethanol precipitation. RNAs were denatured by glyoxalation and analysed by electrophoresis on 1.6% agarose gels, as described in Chapter 4. After staining in 0.0005% ethidium bromide and destaining, RNA was transferred to nitrocellulose filters and probed with ^{32}P-cDNA probes from clones A6 (-) or A 9 (+), as described in Chapter 4.

RESULTS

A. Secondary Structure Analysis of Sat-RNA

Possible secondary structure models for the Sat-RNA sequence were constructed using the matrix procedure of Tinoco et al., (1971). In view of the many possible base-paired structures found, experimental studies were undertaken to determine which structures actually occur in solution. The approach used was to determine the residues present in single-strand and double-strand regions by digestion of Sat-RNA with specific nucleases.

When Sat-RNA was digested under non-denaturing conditions with the single-strand specific nuclease S_1, only about 10% of the residues were cleaved under the conditions used (Fig. 7.1). Thus, digestion of 5'-^{32}p-Sat-RNA for 10 min at 37°C with two levels of nuclease S_1 showed cleavage at residues 3 to 9 and 30 to 33 only when the digests were analysed on 40 cm long 20% polyacrylamide gels in 7 M urea (Fig. 7.1A). However, analysis on longer 80 cm 8% polyacrylamide gels disclosed further cleavage, predominantly at sites between residues 133 and 162 (not shown). Similarly, cleavage of 3'-^{32}P-Sat-RNA by nuclease S_1 was observed at residues 314 to 316, and at residue 297 (Fig. 7.1C). Again, analysis of digests on the longer 80 cm 8% acrylamide gels disclosed further cleavage at sites between residues 220 and 236.

RNase T_1 was found to cleave Sat-RNA at the same regions as nuclease S_1 under the same conditions as shown using both 5'- and 3'-^{32}P-Sat-RNA (Figs. 7.1B and D, respectively). All the bonds cleaved by both enzymes are indicated on the secondary structure model in Fig. 7.3.

These nuclease cleavage data indicated extensive base-pairing in Sat-RNA but it was necessary to obtain positive information on these base-paired regions in order to derive an overall secondary structure model. Hence, Sat-RNA was digested with a high concentration of RNase T_1 (1,000 units/ml) under conditions designed to maximize base-pairing (0.5 M NaCl,

FIGURE 7.1

Polyacrylamide gel analysis of partial nuclease digests of 5'- or 3'-^{32}P-Sat-RNA. Sat-RNA was end-labelled and digested under non-denaturing and denaturing conditions as described in Materials and Methods. N, no digestion (control); L, random alkali digestion to produce ladder; U_2, partial RNase U_2 digestion under denaturing conditions used for sequencing. Marker dyes: XC, xylene cyanol FF; BPB, bromophenol blue. 5'-^{32}P-Sat-RNA: Gel A. Nuclease S_1 cleavage under non-denaturing conditions at (a) 70 units/ml and (b) 120 units/ml. T_1, partial RNase T_1 digestion under denaturing conditions used for sequencing. Gel B. RNase T_1 cleavage under non-denaturing conditions at (b) 1.7 units/ml and (c) 17 units/ml. (a), partial RNase T_1 digest under denaturing conditions used for sequencing. 3'-^{32}P-Sat-RNA: Gel C. Nuclease S_1 cleavage under non-denaturing conditions at (a) 40 units/ml, (b) 120 units/ml, and (c) 170 units/ml. T_1, partial RNase T_1 digestion under denaturing conditions used for sequencing. Gel D. RNase T_1 cleavage under non-denaturing conditions at (b) 1.7 units/ml and (c) 17 units/ml. (a), partial RNase T_1 digest under denaturing conditions used for sequencing. All digests were analysed by electrophoresis on 40 x 20 x 0.05 cm 20% polyacrylamide gels in the presence of 7M urea.

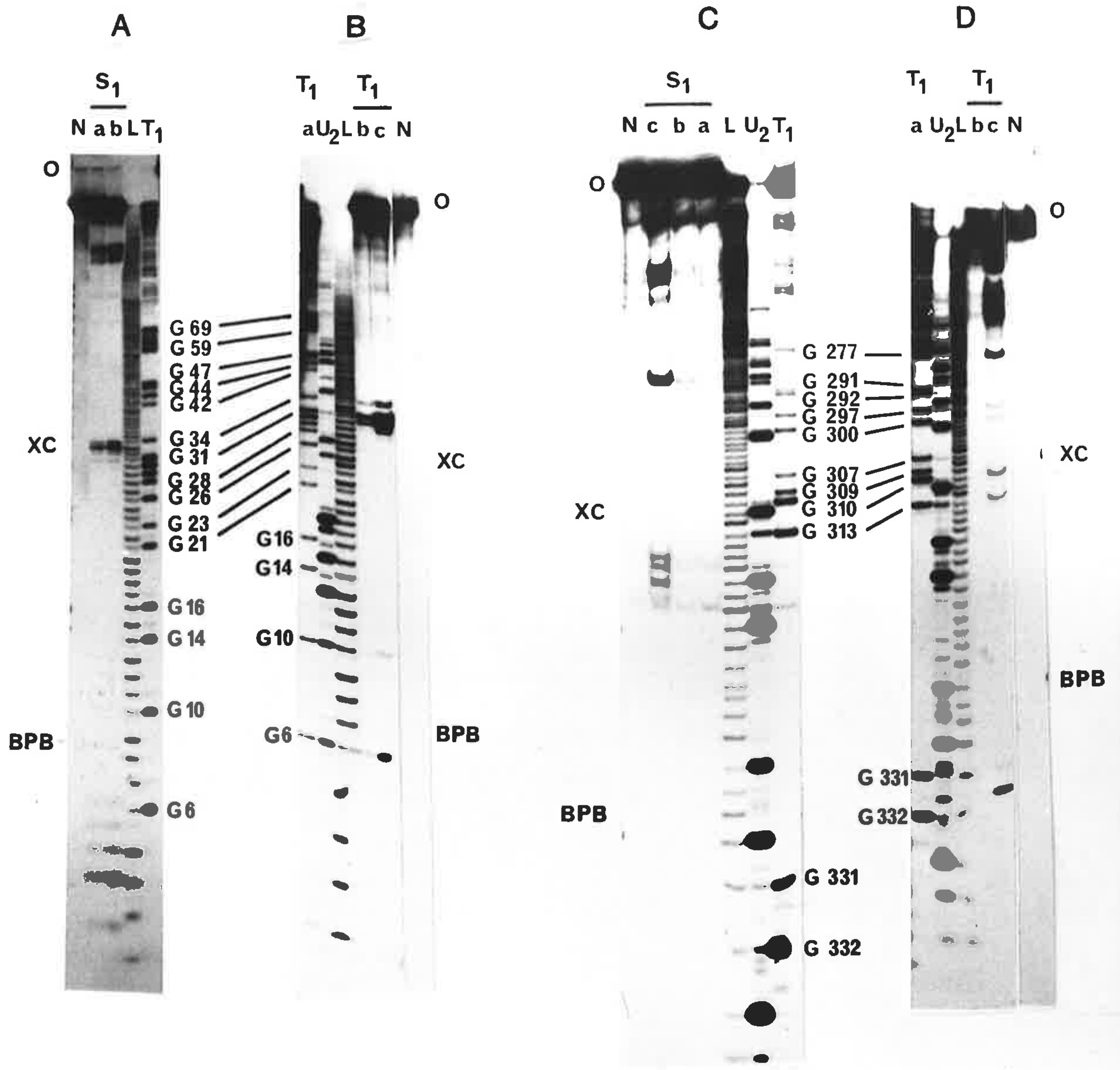
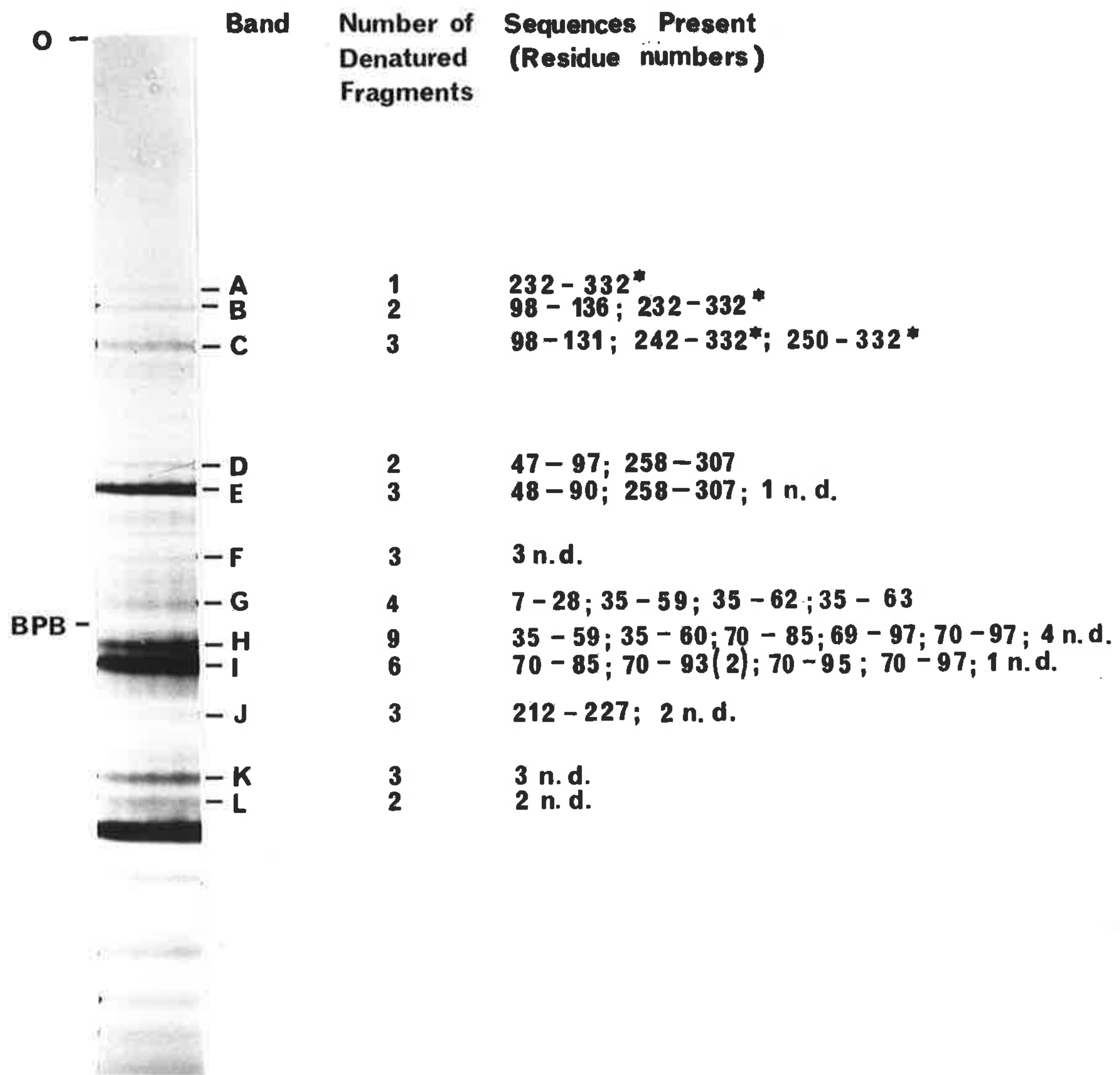

A
B
C
D
S1
N a b L T1
T1 T1
a U2 L b c N
N c b a L U2 T1
a U2 L b c N
O
XC
BPB
G 69
G 59
G 47
G 44
G 42
G 34
G 31
G 28
G 26
G 23
G 21
G 16
G 14
G 10
G 6
G 277
G 291
G 292
G 297
G 300
G 307
G 309
G 310
G 313
G 331
G 332

FIGURE 7.2

Polyacrylamide gel electrophoresis of base-paired fragments generated by RNase T_1 digestion of Sat-RNA under non-denaturing conditions followed by 5'-^{32}P-labelling as described in Materials and Methods. Bands A to L on the non-denaturing 12% polyacrylamide gel were excised, the RNA eluted and fractionated on a 12% polyacrylamide denaturing (7 M urea) gel. The number of labelled denatured fragments recovered from each band is given together with their residue numbers as determined by sequencing. n.d.; Sequence not determined due to cross contamination. Residues with asterisk; exact termination of fragment could not be accurately determined. Marker dye: BPB, bromophenol blue.



O
BPB
Band
Number of Denatured Fragments
Sequences Present (Residue numbers)
A 1 232 - 332*
B 2 98 - 136 ; 232 - 332*
C 3 98 - 131 ; 242 - 332*; 250 - 332*
D 2 47 - 97; 258 - 307
E 3 48 - 90; 258 - 307; 1 n. d.
F 3 3 n.d.
G 4 7 - 28; 35 - 59; 35 - 62 ; 35 - 63
H 9 35 - 59; 35 - 60; 70 - 85; 69 - 97; 70 - 97; 4 n.d.
I 6 70 - 85; 70 - 93(2); 70 - 95 ; 70 - 97; 1 n.d.
J 3 212 - 227; 2 n. d.
K 3 3 n. d.
L 2 2 n. d.

$0^{o}C$), followed by 5'-^{32}P-labelling, without prior denaturation, of the double-strand RNA fragments which were then fractionated on a non-denaturing 12% polyacrylamide gel (Fig. 7.2). Each of the bands A to L was eluted and rerun under denaturing conditions on a 12% polyacrylamide, 7 M urea gel to give from 1 to 9 single-strand fragments (not shown); 25 of the 41 fragments were pure enough to be sequenced and the residue numbers of each fragment are given in Fig. 7.2.

This sequence data showed that the bands on the native gel were of three types. The first type, represented by band A, contained only a single, large 5'-labelled fragment which either could have been present as a single piece of RNA with extensive internal base-pairing or had other RNA fragments base-paired to it that were not 5'-labelled with γ-^{32}P-ATP and polynucleotide kinase under the conditions used. Each of the second type of band (bands B, C, D, E, G and H) contained two or more fragments derived from two separate regions of Sat-RNA and presumably base-paired over at least some of their sequences.

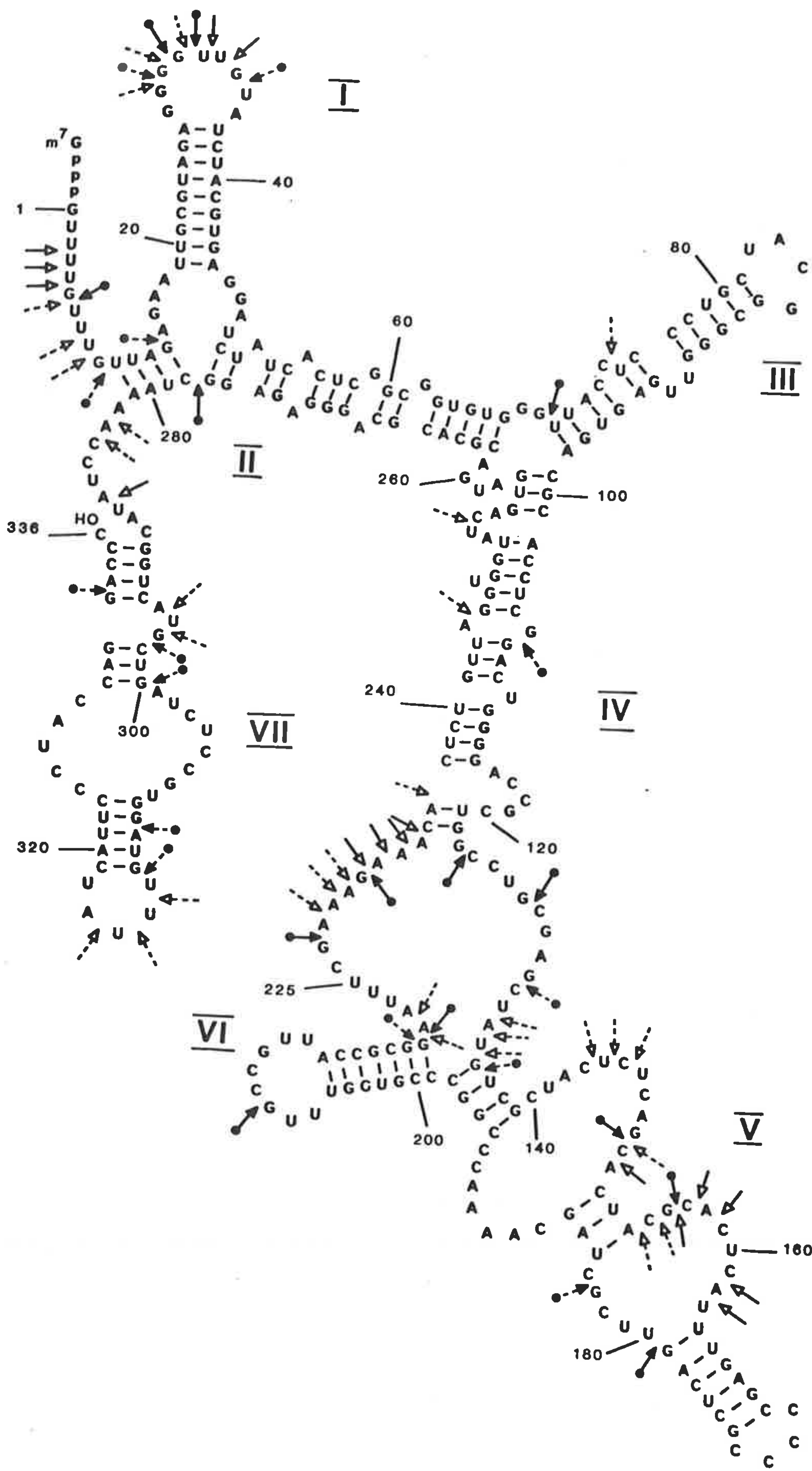
The third type is represented by band I which contained 6 fragments. Five of these each covered all or most of the sequence between residues 70 and 97 while the sequence of the remaining one could not be determined. These results indicated a prominent hairpin formed between residues 70 and 98. All the RNA molecules containing this sequence (band I, Fig. 7.2) migrated together in the non-denaturing gel despite the variation in lengths.

B. Secondary Structure Model of Sat-RNA

The secondary structure model for Sat-RNA in Fig. 7.3 is consistent with the nuclease cleavage date described above. It contains 87 base-pairs, which involve 52% of the 336 residues, in five major hairpin structures and two other base-paired regions. Within the central third of the molecule, from residues 123 to 234, many nuclease sensitive sites were found; apart from the hairpin structures V and VI (Fig. 7.3), little secondary structure is possible in this region. There are limited alternative base-pairing structures for those shown between residues 136 to 154 and residues 185 to

FIGURE 7.3

Secondary structure model of Sat-RNA. Residues are numbered from 5'-end as in Fig. 1. Major cleavage sites for nuclease S_1 and RNase T_1 are shown by —▷ and ●→, respectively. Minor cleavage sites are indicated by dashed arrows. Stems and other base-paired regions are identified by Roman numerals.



I
II
III
IV
V
VI
VII
m7Gppp
HO
1
20
40
60
80
100
120
140
160
180
200
225
240
260
280
300
320
336

198, but these do not affect the overall model significantly. The sites of digestion by nuclease S_1 are in single-strand regions in Fig. 7.3; these regions are presumably exposed in the tertiary structure of Sat-RNA in solution since nuclease S_1 is unable to cleave single base mismatches (Silber and Loeb, 1981) and cleaves short single-strand regions poorly (Dodgson and Wells, 1977; Gonda and Symons,(1978). The large internal loop in region VII of Fig. 7.3 contained no nuclease S_1 sensitive bonds which indicates that the single strands are protected in some way in the native molecule. Of a total of 11 single-strand stretches longer than 5 residues, only 4 are not cleaved by nuclease S_1.

The overall free energy of the structure is approximately -340 kJ/mol (25^oC, 1 M NaCl) calculated according to the rules of G. Steger, H. Gross, J. W. Randles, H. Sanger, and D. Riesner (personal communication). The presence of several branch junctions and most likely of tertiary interactions means that this value is not very accurate. However, it does indicate that the Sat-RNA structure proposed is considerably less stable than the structures proposed for viroids of similar size; e.g., the structure of the 356 residues of chrysanthemum stunt viroid gives a ΔG of -540 kJ/mol (Haseloff and Symons, 1981). The viroid-like RNAs of velvet tobacco mottle virus (365 residues) and of solanum nodiflorum mottle virus (377 residues), however, have proposed secondary structures with ΔG values more similar to that of Sat-RNA, namely -350 kJ/mol and -450 kJ/mol, respectively (Haseloff and Symons, 1982).

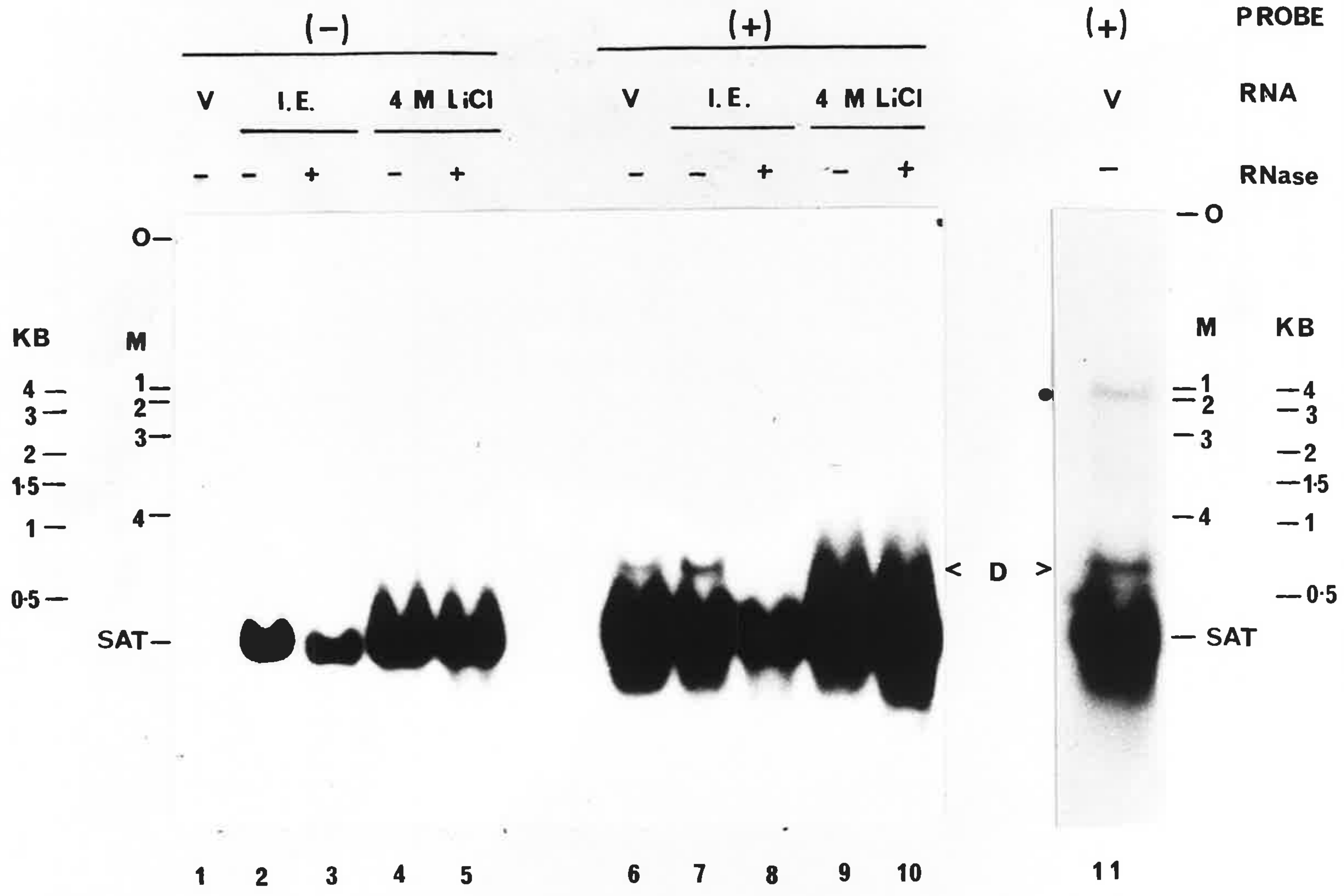
C. <u>Double stranded and oligomeric Sat-RNA</u>

Total encapsidated RNA from plants inoculated with CMV and Sat-RNA contained three readily detectable forms of Sat-RNA. The probe showed a (+) monomer, a small amount of a dimeric form (as determined from its relative mobility) and even less of a 3,500 nucleotide form (co-migrating with RNA 2) (Fig. 7.4, tracks 6, 11). All these forms were single stranded, as shown by their sensitivity to nuclease digestion (not shown). Virtually no (-) Sat-RNA was detected (Fig. 7.4, track 1).

FIGURE 7.4

(+) and (-) forms of Sat-RNA

RNA samples were glyoxalated, electrophoresed on a 1.6% agarose gel and probed with cDNA to clone A6 for (-) sequences (tracks 1 - 5) or to clone A9 for (+) sequences (tracks 6 - 10), as described in Methods. Equal amounts (in cpm) of (+) and (-) probes with the same specific activity were used. Track 11, from another gel, was also probed for (+) sequences. Tracks 1, 6, 11 : total encapsidated CMV RNA with Sat-RNA (20µg); the positions of the stained marker RNAs (M) are shown.
Tracks 2, 7: 20µg of total RNA extracted from the particulate fraction of plants infected with CMV and Sat-RNA (see Methods).
Tracks 3, 8: as for tracks 2 and 7, but with nuclease digestion, as described in Methods.
Tracks 4, 9 : RNA precipitated by 4 M LiCl from 200µg of total RNA extracted from the particulate fraction, as described in Methods.
Tracks 5, 10: as for tracks 4, 9, but with nuclease digestion.
A scale in Kilobases (KB) is provided.
"D" indicates the dimer of Sat-RNA (in tracks 6, 7 and 11); and the dot by track 11 the high molecular weight form (not detected in CMV RNA free of Sat-RNA).


(−)
(+)
(+)
PROBE
V
I.E.
4 M LiCl
V
I.E.
4 M LiCl
V
RNA
− − + − + − − + − + −
RNase
O
KB
M
M
KB
4
3
2
1·5
1
0·5
1
2
3
4
SAT
< D >
1 2 3 4 5 6 7 8 9 10 11

In contrast, the RNA extracted from the particulate fraction contained considerable amounts of (-) monomer Sat-RNA (track 2) much of which was refractory to nuclease digestion (track 3) and therefore double stranded. Similar results were obtained for (+) sequences in this fraction (track 7, 8) except that the (+) dimeric form did not survive nuclease digestion (track 8) and was presumably all single stranded. The double stranded RNA obtained by 4 M LiCl precipitation (tracks, 4, 9) was completely resistant to nuclease digestion (tracks 5, 10). It contained none of the higher molecular weight forms.

DISCUSSION

A. Nucleotide Sequence and Secondary Structure of Sat-RNA and CARNA 5

The secondary structure model proposed for Sat-RNA (Fig. 7.3) provides a basis for further studies on its structure and function and for the comparison of different strains of satellite RNAs of CMV. The model can be applied to CARNA 5 since most of the sequence differences in CARNA 5, relative to Sat-RNA, occur either in single-strand regions or maintain base-pairing. Since (L) CARNA 5 shows 92% sequence homology with Sat-RNA (Chapter 1), only two of the base pairs shown in the structure model are disrupted. These are U(11)-A(281) and C(199)-G(220). A further 4 base pairs are disrupted by sequence changes in (N) CARNA 5, although, however, a small structural rearrangement at the top of base-paired region $\overline{IV}$ (Fig. 7.3) means little decrease in overall stability. It will be of interest to determine if the other satellite RNAs of CMV (Takanami, 1981) or of the related peanut stunt virus (Kaper et al., 1978) can form a secondary structure similar to that proposed here.

Using computer analysis to maximise base-pairing, Dr. D. Riesner and coworkers (personal communication) have shown that the Sat-RNA sequence can be folded into an alternative secondary structure. This structure is more highly base-paired and slightly more stable than that shown in Fig. 7.3, but contradicts much of the nuclease cleavage data.

Sat-RNA and (N) CARNA 5 differ greatly in their modification of CMV symptoms in certain host plants (Mossop and Francki, 1979; Waterworth et al., 1979; Murant and Mayo, 1982). Given the extensive homology in nucleotide sequence and possibly secondary structure, it is feasible that the completely different potential translation products are responsible for these effects. Sat-RNA lacks detectable messenger activity in rabbit reticulocyte lysates but stimulates synthesis of a protein of M_r about 3,000 in the wheat germ system (P. Palukaitis and K.H.J. Gordon, unpublished) which could correspond to the 17 amino acid open reading frame starting at nucleotide 134 (Fig. 7.3). The (L) CARNA 5 sequence (Collmer et al., 1983) also predicts different translation products to those of either (N) CARNA 5 or Sat-RNA.

B. Sequence and Structural Homology between Sat-RNA and CMV Genomal RNAs

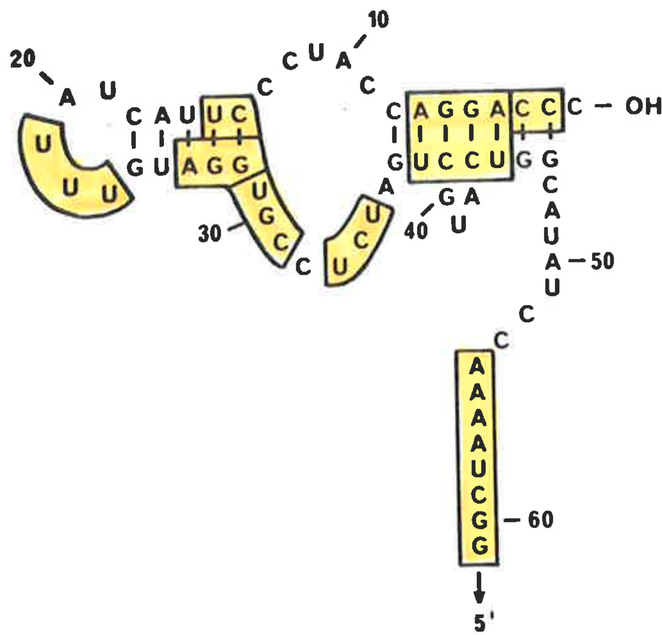
There are only short scattered stretches of direct nucleotide sequence homology between Sat-RNA and the known sequences of Q-CMV genomal RNAs. Ten of the first 12 5'-terminal residues of Sat-RNA are homologous to 10 of the 5'-terminal 11 residues of the subgenomic Q-CMV RNA 4 (Gould and Symons, 1982). This homology in a G and U-rich stretch is unlikely to be important for translation of Sat-RNA in view of its poor mRNA activity and, since these G and U-rich stretches are characteristic of the 5'-terminal sequences of some subgenomic mRNAs of plant viruses (Guilley et al., 1979; Koper-Zwarthoff et al., 1979), they may represent a signal for the generation of such RNAs from genomal RNAs during replication. Computer analysis has shown other short stretches of sequence homology between the Q-CMV RNAs and Sat-RNA; the longest is 11 residues corresponding to residues 63 to 73 in Sat-RNA and residues 1577 to 1587 in RNA 3 (Gould and Symons, 1982).

The most prominent structural feature hitherto identified on CMV genomal RNAs is the tRNA-like structure formed by the 3'-terminal 125 residues of the four RNAs (Symons, 1979), the biological significance of which is considered below. In Fig. 7.5, the 3'-terminal hairpin of Sat-RNA

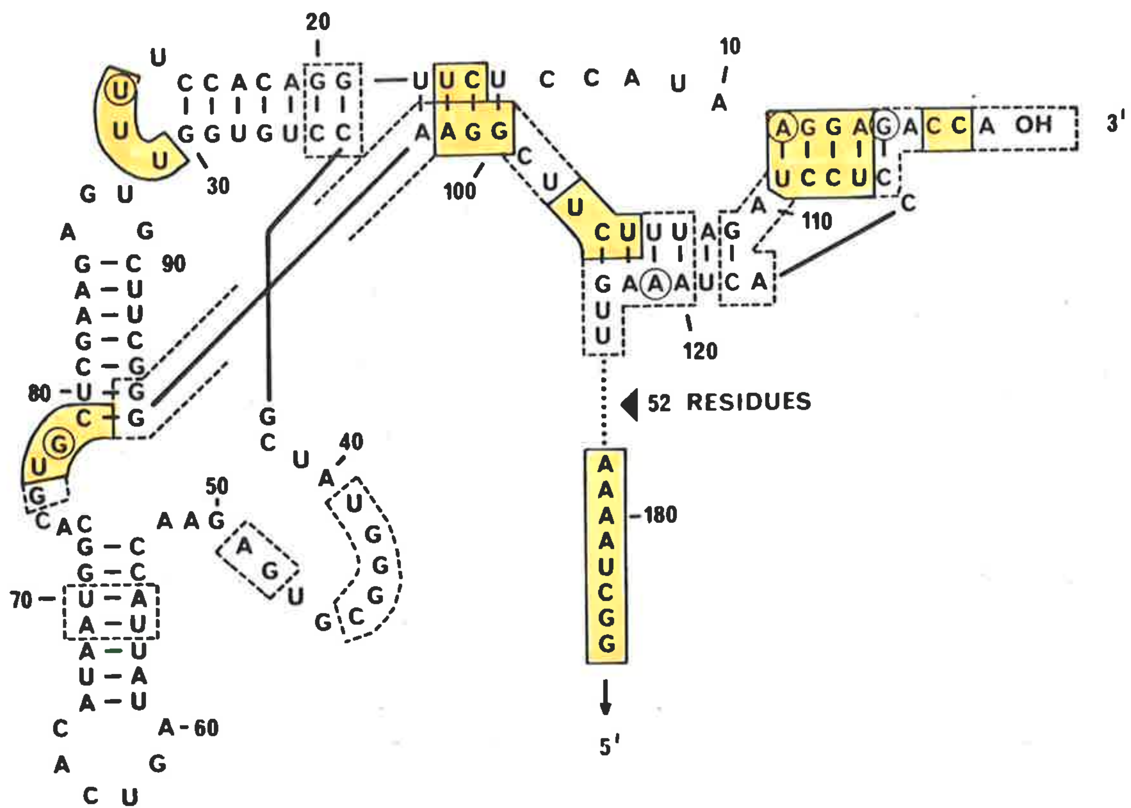
FIGURE 7.5

Sequence and structural homology at the 3'-termini of Sat-RNA and CMV genomal RNAs. Residues are numbered from the 3'-terminus in both structures. The Sat-RNA structure (A) is the 3'-terminal hairpin VII in Fig. 7.3. In B, the model proposed for CMV RNA 3 (Symons, 1979) has been re-arranged to show the possible derivation of the three-dimensional tRNA-like structure (Kim, 1976). Residues within dashed and solid boxes in B are conserved in all genomal RNAs of CMV strains Q, P, T and M (Symons, 1979) (Wilson, Barker and Symons, unpublished); of these, the circled residues vary in one to two of the 12 genomal RNAs studied. The solid boxed sequences in the Sat-RNA structure (A) and the CMV RNAs structure (B) are homologous.

A SAT - RNA



B CMV RNAs



is compared to the tRNA-like structure formed by Q-CMV RNA 3 which has been drawn to indicate the possible tertiary structure in a similar way to that employed for tRNA (Kim, 1976). About one-half of the 125 residues of Q-CMV forming this structure are conserved in the genomal RNAs of four strains of CMV studied so far; the Q, P, T and M strains (P.A. Wilson, J.M. Barker and R.H. Symons, unpublished). These conserved residues (shown in dashed plus solid boxes in Fig. 7.5) are heavily concentrated in that part of the structure which corresponds to the aminoacyl and TψC stems of canonical tRNAs.

The sequences homologous between Sat-RNA and the CMV RNAs are shown in solid boxes. It is striking that the proposed 3'-terminal structure of Sat-RNA shows extensive sequence and structural homology resulting in Sat-RNA possessing a truncated tRNA-like structure. In addition, most of the homologous residues in the CMV structure are also present in the 3'-terminal structure of the genomal RNAs of tomato aspermy virus (Wilson and Symons, 1981), another cucumovirus which can also support Sat-RNA replication (Gould et al., 1978). The hairpin formed between the 3'-terminal residues 14 to 28 of Sat-RNA represents a shortened version of the CMV loop formed by residues 15 to 36 plus 97 to 100 (Fig. 7.5), with conserved sequences in the loop and at the stem of the hairpin. A conserved double-strand region is also found near the 3'-terminus of both Sat-RNA and CMV RNAs while there is an intriguing conserved sequence of 8 residues between 3'-terminal residues 54 to 61 of Sat-RNA and 178 to 185 of CMV RNAs.

Only 10 of the 44 residues in the hairpins formed by CMV RNAs residues 53 to 96 (Fig. 7.5) are conserved in the four strains of CMV (see above). Of these 10, only 3 (UGC) show homology within the 3'-terminal sequence of Sat-RNA but they are not in the same order as the other homologous sequences between the two RNAs. In spite of this, tertiary folding of the CMV RNAs may bring this UGC sequence nearer to residue 100 and hence in an equivalent position to that of Sat-RNA.

C. Possible Function of 3'-terminal Structure and Replication of Sat-RNA

Although the four RNAs of CMV can be aminoacylated *in vitro* with tyrosine (Kohl and Hall, 1974), Sat-RNA cannot be aminoacylated (P. Palukaitis, personal communication) which is consistent with the lack of a complete tRNA-like structure. It has been suggested that the 3'-terminal aminoacylation of CMV and some other viral RNAs is important in the initiation of viral RNA replication, possibly through facilitating RNA-protein interactions (Hall, 1979; Gordon et al., 1982; Symons et al., 1982).

However, Sat-RNA would be incapable of replicating in exactly the same way but may parasitize a replication complex which has been assembled using aminoacylated viral RNAs and viral gene products. Some feature of Sat-RNA replication must explain the rapidity with which a very low level of Sat-RNA in virus used for inoculating plants can increase to up to 50% of the total encapsidated RNA (Gould et al., 1978); for example, the more compact shape of Sat-RNA's truncated tRNA-like structure may play a role. The complete dependence of Sat-RNA on CMV RNAs for replication plus the observation that Sat-RNA is replicated in the same particulate extract of infected plants in which CMV RNAs are replicated (D.S. Gill and K.H.J. Gordon, unpublished) suggest that Sat-RNA must possess sufficient sequence and structural homology with CMV RNAs to allow it to use the CMV-induced replication machinery.

Any model proposed for the replication of satellite RNA must explain the observation of double stranded monomeric, single stranded (+) dimeric and the higher molecular weight forms. The occurrence of large amounts of oligomeric (+) and (-) forms of satellite RNA of tobacco ringspot virus led to the suggestion that this satellite RNA replicates via a rolling circle mechanism (Kiefer et al., 1982; Sogo and Schneider, 1982). Similar observations of oligomeric (+) and (-) forms have been made for a number of viroids and cicular satellite RNAs (Symons et al., 1983), so that the proposed rolling circle mechanisms may have wider applicability. CMV Sat-RNA, however, has not shown these extensive, oligomeric (+) and (-)

forms; this difference may be related to the fact that it is capped, and therefore unlikely to be formed through processing of oligomeric forms.

REFERENCES

Ahlquist, P., Dasgupta, R. and Kaesberg, P. (1983).
J. Mol. Biol., in press.

Atabekov, J. G. and Morozov, S. Yu. (1979).
Advances in Virus Research 25, 1 - 91.

Bastin, M. and Kaesberg, P. (1976).
Virology 72, 536 - 539.

Beachy, R. N. and Zaitlin, M. (1977).
Virology 81, 160 - 169.

Benicourt, C., and A. - L. Haenni, (1978).
Biochem. Biophys. Res. Commun. 84, 831 - 839.

Bishop, D. H. L., Calisher, C. H., Casals, J., Chumakov, M. T., Gaidamovich, S. Y. A., Hannon, C., Lvov, D. K., Marshall, I. D., Oker-Blom, N., Petterson, R. F., Porterfield, J. S., Russell, P. K., Shope, R. E and Westaway, E. G. (1980).
Intervirology 14, 125 - 143.

Blumenthal, T. (1979).
Methods in Enzymology 60, 628 - 638.

Blumenthal, T. and Carmichael, G. G. (1979).
Ann. Rev. Biochem. 48, 525 - 548.

Bol, J. F., Brederode, F. T., Janze, G. C. and Rauh, D. K. (1975).
Virology 65, 1 - 15.

Boege, F. and Saenger, H. L. (1980).
FEBS Lett. 121, 91 - 96.

Bol, J. F., Clerx-van Haaster, C. M. and Weening, C. J. (1976a).
Ann. Microbiol. (Inst. Pasteur) 127A, 183 - 192.

Bol, J. F., Bakhuizen, C. E. G. C. and Rutgers, T. (1976b)
Virology 75, 1 - 17.

Bonner, W. M. and Laskey, R. A. (1974).
Eur. J. Biochem. 46, 83 - 88.

Bos, J. L., Polder, L. J., Bernards, R., Schrier, P. U., van den Elsen, P. J., van der Eb, A. J. and van Ormondt, H. (1981).
Cell 27, 121 - 131.

Bruening, G. (1981)
in "The Biochemistry of Plants" (P. K. Stumpf, E. E. Conn, eds.) Vol 6, 571 - 631.

Bruening, G., Beachy, R. N., Scalla, R. and Zaitlin, M. (1976).
Virology 71, 498 - 517.

Bruening, G., Gould, A. R., Murphy, P. J. and Symons, R. H. (1982).
FEBS. Lett. 148, 71 - 78.

Chifflot, S., Sommer, P., Hartmann D., Stussi-Garaud, C. and Hirth, L. (1980).
Virology 100, 91 - 100.

Clark, G. L., Peden, K. W. C. and Symons, R. H. (1974).
Virology 62, 434 - 443.

Clerx, C. M. and Bol, J. F. (1978).
Virology 91, 453 - 463.

Cleveland, D. W., Fischer, S. G., Kirschner, M. W. and Laemmli, U. K. (1977).
J. Biol. Chem. 252, 1102 - 1106.

Collmer, C. W., Tousignant, M. E. and Kaper, J. M. (1983.
Virology 127, 230 - 234.

Cornelissen, B. J. C., Brederode, F. Th., Moorman, R. J. B. and Bol, J. F.(1983a)
Nucl. Acids Res. 11, 1253 - 1265.

Cornelissen, B. J. C., Brederode, F. Th., Veeneman, G. H., van Boom, J. H. and Bol, J. F. (1983).
Nucl. Acids Res. 11, 3019 - 3025.

Dasgupta, A., Zabel, P. and Baltimore, D. (1980).
Cell 19, 423 - 429.

Davies, J. W. (1979).
In "Nucleic Acids in Plants" (T. C. Hall and J. W. Davies, eds.) II, 111 - 149. CRC Press Boca Raton, Fla.

Davies, J. W. and Hull, R. (1982).
J. Gen. Virol. 61, 1 - 14.

Dawson, W. O. and Dodds, J. A. (1982).
Biochem. Biophys. Res. Commun. 107, 1230 - 1235.

Diaz-Ruiz, J. R. and Kaper, J. M. (1977).
Virology 80, 204 - 213.

Diaz-Ruiz, J. R. and Kaper, J. M. (1978).
Preparative Biochem. 8, 1 - 17.

Dodgson, J. B. and Wells, R. D. (1977).
Biochemistry 16, 2374 - 2379.

Dorssers, L. (1983)
Ph.D. Thesis, Agricultural University of Wageningen.

Dorssers, L., vander Meer, J., van Kammen, A. and Zabel, P. (1983).
Virology 125, 155 - 174.

Dorssers, L., Zabel, P., van der Meer, J. and van Kammen, A. (1982).
Virology 116, 236 - 249.

Dougherty, W. G. and Kaesberg, P. (1981).
Virology 115, 45 - 56.

Duda, C. T. (1979).
Virology 92, 180 - 189.

England, T. E. and Uhlenbeck, O. C., (1978).
Nature 275, 560 - 561.

Federoff, N. (1975).
In "RNA Phages" (Zinder, N., ed.) 235 - 238. Cold Spring Harbour Lab. New York.

Fields, S. and Winter, G. (1982).
Cell 28, 303 - 313.

Fraenkel-Conrat, H. and Steinschneider, A. (1968).
Methods in Enzymology 12B, 243 - 246.

Francki, R. I. B., Mossop, D. W. and Hatta, T. (1979).
In "CMI/AAB Description of Plant Viruses" No. 213. Commonwealth Mycological Institute and Association of Applied Biologists.

Francki, R. I. B., Randles, J. W., Chambers, T. C and Wilson, S. B. (1966).
Virology 28, 729 - 741.

Friesen, P. D. and Rueckert, R. R. (1981).
J. Virol. 37, 876 - 886.

Gehrke, L., Auron, P. E., Quigley, G. J., Rich, A. and Sonenberg, N. (1983).
Biochemistry 22, 5157 - 5164.

Gerlinger, P., Mohier, E., Le Meur, M. A. and Hirth, L. (1977).
Nucl. Acids Res. 4, 813 - 826.

Gill, D. S. (1983).
Ph.D. Thesis, University of Adelaide.

Gill, D. S., Kumarasamy, R. and Symons, R. H. (1981).
Virology 113, 1 - 8.

Goelet, P. and Karn, J. (1982).
J. Mol. Biol. 154, 541 - 550.

Goelet, P., Lomonossoff, G. P., Butler, G. P., Akam, M. E., Gait, M. J. and Karn, J. (1982).
Proc. Natl. Acad. Sci. USA 79, 5818 - 5822.

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

Gonda, T. J. and Symons, R. H. (1978).
Virology 88, 361 - 370.

Gonda, T. J. and Symons, R. H. (1979).
J. Gen. Virol. 45, 723 - 736.

Gordon, K. H. J., Gill, D. S. and Symons, R. H. (1982).
Virology 123, 284 - 295.

Gordon, K. H. J. and Symons, R. H. (1983).
Nucl. Acids Res. 11, 947 - 960.

Gould, A. R., Palukaitis, P., Symons, R. H. and Mossop, D. W. (1978).
Virology 84, 443 - 455.

Gould, A. R. and Symons, R. H. (1977).
Nucl. Acids Res. 4, 3787 - 3802.

Gould, A. R. and Symons, R. H. (1982).
Eur. J. Biochem. 126, 217 - 226.

Guilley, H., Jonard, G., Kukla, B. and Richards, K. E. (1979).
Nucl. Acids Res. 6, 1287 - 1308.

Gunn, M. R. and Symons, R. H. (1980).
FEBS Lett. 109, 145 - 150.

Habili, N. and Francki, R. I. B. (1974)
Virology 57, 392 - 401.

Habili, N. and Kaper, J. M. (1981).
Virology 112, 250 - 261.

Hackett, P. B., Swanstrom, R., Varmus, H. E. and Bishop, J. M. (1982).
J. Virol. 41, 527 - 534.

Haenni, A.-L., Joshi, S. and Chapeville, F. (1982).
Progr. Nucl. Acids Res. and Mol. Biol. 27, 85 - 104.

Hall, T. C. (1979).
Int. Rev. Cytol. 60, 1 - 26.

Hall, T. C., Miller, W. A. and Bujarski, J. J. (1982).
In "Advances in Plant Pathology" (Ingram, D. and Williams, P.H., eds.)
Vol I, pp. 179 - 211. Academic Press, New York.

Hall, T. C. and Wepprich, R. (1976).
Ann. Microbiol. (Inst. Pasteur) 127A, 143 - 152

Haseloff, J. (1979).
B.Sc. (Hons) Thesis, University of Adelaide.

Haseloff, J. and Symons, R. H. (1981).
Nucl. Acids Res. 9, 2741 - 2752.

Haseloff, J. and Symons, R. H. (1982).
Nucl. Acids. Res. 10, 3681 - 3691.

Hayward, W. S. (1977).
J. Virol., 24, 47 - 63.

Hirth, L. (1982).
In "Nucleic Acids and Proteins in Plants". (B. Parthier and
D. Boulter, eds.) Encycl. of Plant Physiol. vol. 14B. pp. 302-338.
(Springer, Berlin).

Holland, J., Spindler, K., Horodyski, F., Grabau, E., Nichol, S. and VandePol,S.
(1982).
Science, 215, 1577 - 1585.

Hollings, M. and Stone, O. M. (1971).
In "CMI/AAB Descriptions of Plant Viruses" No. 79. Commonwealth
Mycological Inst., Kew, England.

Hunter, T., Jackson, R. and Zimmern, D. (1983).
Nucl. Acids Res. 11, 801 - 821.

Ikegami, M. and Fraenkel - Conrat, H. (1978a).
Proc. Natl. Acad. Sci. USA 75, 2122 - 2124.

Ikegami, M. and Fraenkel-Conrat, H. (1978b).
FEBS Lett. 96, 197 - 200.

Joshi, S., Pleij, C. W. A., Haenni, A. L., Chapeville, F. and Bosch, L. (1983).
Virology 127, 100 - 111.

Kamen, R. I. (1975).
In "RNA Phages" (Zinder, N., ed) pp. 203 - 234. Cold Spring Harbour Laboratory, New York.

Kaper, J. M., Tousignant, J. E. and Lot, H. (1976).
Biochem. Biophys. Res. Commun. 72, 1237 - 1243.

Kaper, J. M., Tousignant, M. E., Diaz-Ruiz, J. R. and Tolin, S.A. (1978).
Virology 88, 166 - 170.

Kiefer, M. C., Daubert, S. D., Schneider, I. R. and Bruening, G. (1982).
Virology 121, 262 - 273.

Kim, S. (1976).
Prog. Nucl. Acids Res. and Mol. Biol. 17, 181 - 216.

King, A. M. Q., McCahon, D., Slade, W. R. and Newman, J. W. I. (1982).
J. Virol. 41, 66 - 77.

Kitamura, N., Semler, B. L., Rothberg, P. G., Larsey, G. R., Adler, C. J., Dorner, A. J., Emini, E. A., Hanacak, R., Lee, J. J., van der Werf, S., Anderson, C. W. and Wimmer, E. (1981).
Nature 291, 547 - 553.

Kohl, R. J. and Hall, T. C. (1974).
J. Gen. Virol. 25, 257 - 261.

Koper-Zwarthoff, E. C., Brederode, F. Th., Walstra, P. and Bol, J. F. (1979).
Nucl. Acids Res. 7, 1887 - 1900.

Kozak, M. (1981a).
Nucl. Acids Res. 9, 5233 - 5252.

Kozak, M. (1981b).
Curr. Topics in Microbiol. and Immunol. 93, 81 - 123.

Krzyzek, R. A., Collett, M. S., Lau, A. F., Perdue, M. L., Leis, J. P. and Faras, A. J. (1978).
Proc. Natl. Acad. Sci. USA 75, 1284 - 1288.

Kuhn, S., Anitz, H. J. and Starlinger, P. (1979).
Mol. Gen. Genet. 167, 235 - 241.

Kumarasamy, R. and Symons, R. H. (1979a).
Virology 96, 622 - 632.

Kumarasamy, R. and Symons, R. H. (1979b).
Anal. Biochem. 95, 359 - 363.

Laemmli, U. K. (1970).
Nature 227, 680 - 685.

Lai, M. M. C. Patton, C. D. and Stohlman, S.A. (1982).
J. Virol. 41, 557 - 565.

Lamb, R. A. and Choppin, P. W. (1983).
Ann. Rev. Biochem. 52, 467 - 506.

Lane, L. C. (1979).
In "Nucleic Acids in Plants" (T. C. Hall and J. W. Davies, eds.)
Vol. 2, pp. 111 - 149. CRC Press, Boca Raton, Fla.

Laskey, R. A. and Mills, A. D. (1975).
Eur. J. Biochem. 56, 335 - 341.

Lawrence, C. B. (1980).
Nucl. Acids Res. 8, 1307 - 1317.

Loenard, D. A. and Zaitlin, M. (1982).
Virology 117, 416 - 424.

Loesch-Fries, L. S. and Hall, T. C. (1980).
J. Gen. Virol. 47, 323 - 332.

Loesch-Fries, L. S. and Hall, T. C. (1982).
Nature 298, 771 - 773.

Lomedico, P. T. and McAndrew, S. J. (1982).
Nature 299, 221 - 226.

Lot, H., Marchoux, G., Marron, J., Kaper, J. M., West, C. K., van Vloten-Doting, L. and Hull, R. (1974).
J. Gen. Virol. 22, 81 - 93.

Mahy, B. W. J. (1983).
Nature 305, 474 - 475.

Maniatis, T. Fritsch, E. F and Sambrook, J. (1982).
"Molecular Cloning." Cold Spring Harbour Press, New York.

Mardon, G. and Varmus, H. E. (1983).
Cell 32, 871 - 879.

Maxam, A. and Gilbert, W. (1981).
Methods Enzymol. 65, 499 - 560.

May, J. T., Gilliland, J. M. and Symons, R. H. (1970).
Virology 41, 653 - 664.

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

Mellon, P. and Duesberg, P. H. (1977).
Nature 270, 631 - 634.

Messing, J., Crea, P. and Seeburg, P. H. (1981).
Nucl. Acids Res. 9, 309 - 321.

Mohier, E., Pinck, L. and Hirth, L. (1974).
Virology 58, 9 - 15.

Morris-Krsinich, B. A. M. and Forster, R. L. S. (1983).
Virology 128, 176 - 185.

Mossop, D. W. (1978).
Ph.D. Thesis, University of Adelaide.

Mossop, D. W. and Francki, R. I. B. (1977).
Virology 81, 177 - 181.

Mossop, D. W. and Francki, R. I. B. (1979).
Virology 95, 395 - 404.

Mossop, D. W., Francki, R. I. B. and Grivell, C. J. (1976).
Virology 74, 544 - 546.

Murant, A. F. and Mayo, M.O. (1982).
Ann. Rev. Phytopathol. 20, 49 - 70.

Myers, J. and Spiegelman, S. (1978).
Proc. Natl. Acad. Sci. USA 75, 5329 - 5333

Nassuth, A. and Bol, J. F. (1982).
Virology 124, 75 - 85.

Nassuth, A., ten Bruggencate, G. and Bol, J. F. (1983).
Virology 125, 75 - 84.

Newman, J. F. E., Cartwright, B., Doel, T. R. and Brown, F. (1979).
J. Gen. Virol. 45, 497 - 507.

O'Farrell, P. Z., Goodman, H. M. and O'Farrell, P. H. (1977).
Cell, 12, 1133 - 1142.

Otal, T. and Hari, V. (1983).
Virology 125, 118 - 126.

Pallansch, M. A., Kew, O. M., Palmenberg, A. C., Golini, F., Wimmer, E. and Rueckert, R. R. (1980).
J. Virology 35, 414 - 419.

Palukaitis, P., Garcia-Arenal, F., Sulzinski, M. A. and Zaitlin, M. (1983).
Virology 131, in press.

Parnes, J. R., Velan, B., Felsenfeld, A., Ramanathan, L., Ferrini, U., Appella, E. and Sidman, J. G. (1981).
Proc. Natl. Acad. Sci. USA 78, 2253 - 2257.

Payvar, F. and Schimke, R. T. (1979).
J. Biol. Chem. 254, 7636 - 7642.

Peden, K. W. C., May, J. T. and Symons, R. H. (1972).
Virology 47, 498 - 501.

Peden, K. W. C. and Symons, R. H. (1973).
Virology 53, 487 - 492.

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

Pelham, H. R. B. and Jackson, R. J. (1976).
Eur. J. Biochem. 67, 247 - 256.

Petterson, R. F., Söderlund, H. and Kääriäinen, L. (1980).
Eur. J. Biochem. 105, 435 - 443.

Pinck, L. and Hirth, L. (1972).
Virology, 49, 413 - 425.

Pinck, L. and Pinck, M. (1979).
FEBS Lett. 107, 61 - 65.

Polatnick, J. and Wool, S. (1981).
J. Virology 40, 881 - 889.

Ranu, R. S. and London, I. M. (1976).
Methods in Enzymology 60, 459 - 484.

Rao, A. L. N. and Francki, R. I. B. (1981).
Virology 114, 573 - 575.

Rao, A. L. N., Hatta, T. and Francki, R. I. B. (1982).
Virology 116, 318 - 326.

Ravelonandro, M., Godefroy-Colburn, T. and Pinck, L. (1983).
Nucl. Acids Res. 11, 2815 - 2826.

Reanney, D. C. (1982).
Ann. Rev. Microbiol 36, 47 - 73.

Rice, R. H. and Means, G. E. (1971).
J. Biol. Chem. 246, 831 - 832.

Richards, K. E., Jonard, G., Jacquemond, M. and Lot, H. (1978).
Virology, 89, 395 - 408.

Romaine, C. P. and Zaitlin, M. (1978).
Virology 86, 241 - 243.

Rutgers, A. S. (1977).
Ph.D. Thesis, University of Leiden.

Saito, I., Shiroki, K. and Shimojo, H. (1983).
Virology 127, 272 - 289.

Sanger, F. and Coulsen, A. R. (1978).
FEBS Lett. 87, 107 - 110.

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

Sawicki, D. L., Kääriäinen, L., Lambek, C. and Gomatas, P. J. (1978).
J. Virol. 25, 19 - 27.

Scalla, R., Romaine, P., Asselin, A., Rigaud, J. and Zaitlin, M. (1978).
Virology 91, 182 - 193.

Schwinghamer, M. W. and Symons, R. H. (1975).
Virology 63, 252 - 262.

Schwinghamer, M. W. and Symons, R. H. (1977).
Virology 79, 88 - 108.

Sheppard, R. C. (1983).
Chemistry in Britain, May, 402 - 414.

Siddell, St., Wege, H. and ter Meulen, V. (1982).
Curr. Topics in Microbiol. and Immunol. 99, 135 - 163.

Silber, J. R. and Loeb, L. A. (1981).
Biochim. Biophys. Acta 656, 256 - 264.

Smit, J. M. and Jaspers, E. M. J. (1982).
Virology, 117, 271 - 274.

Sogo, J. M. and Schneider, I. R. (1982).
Virology 117, 401 - 415.

Symons, R. H. (1975).
Mol. Biol. Reports 2, 277 - 285.

Symons, R. H. (1978).
Aust. J. Biol. Sci. 31, 25 - 37.

Symons, R. H. (1979).
Nucl. Acids Res. 7, 825 - 837.

Symons, R. H. Gill, D. S., Gordon, K. H. J. and Gould, A. R. (1982)
In "Manipulation and Expression of Genes in Eukaryotes" (P. Nagley, A. W. Linnane, W. J. Peacock and J. A. Pateman, eds.). pp. 373 - 380.

Symons, R. H., Haseloff, J., Visvader, J. E., Keese, P., Murphy, P. J., Gill, D. S., Gordon, K. H. J. and Bruening, G. (1983).
In "Subviral Pathogens of Plants and Animals: Viroids and Prions" (K. Maramosch, ed.) in press (Academic).

Takanami, Y. (1981).
Virology 109, 120 - 126.

Takanami, Y., Kubo, S. and Imazumi, S. (1977).
Virology 80, 376 - 389.

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

Tinoco, I., Uhlenbeck, O. C. and Levine, M. D. (1971).
Nature 230, 362 - 367.

Traub, A., Diskin, B., Rosenberg, H. and Kalmar, H. (1976).
J. Virology 18, 375 - 382.

Van Dyke, T. A. and Flanegan, J. B. (1980).
J. Virology 35, 732 - 740.

Van Tol, R. G. L. and van Vloten-Doting, L. (1979).
Eur. J. Biochem. 93, 461 - 468.

Van Vloten-Doting, L. and Jaspers, E. M. J. (1977).
In Comprehensive Virology (H. Fraenkel-Conrat and R. R. Wagner, eds.). Vol. 11, pp. 1 - 47. Plenum Press, N.Y.

Van Vloten-Doting, L. and Neeleman, L. (1982).
in "Nucleic Acids and Proteins in Plants" (B, Parthier and D. Boulter, eds.). Encycl. of Plant Physiol. vol. 14B, pp. 337 - 367.

Walter, G. and Doolittle, R. F. (1983).
In "Genetic Engineering; Principles and Methods". (J. K. Setlow and A. Hollaender, eds.) vol. 5, pp. 61 - 91.

Waterworth, H. E., Kaper, J. M. and Tousignant, M. E. (1979).
Science 204, 845 - 847.

Weiss, S. R. Varmus, H. E. and Bishop, J. M. (1977).
Cell 12, 983 - 992.

White, J. L. and Dawson, W. O. (1978).
Virology 88, 33 - 43.

Wilson, P. A. and Symons, R. H. (1981).
Virology 112, 342 - 345.

Wilson, T. M. A. and Glover, J. F. (1983).
Biochim. Biophys. Acta. 739, 35 - 41.

Zagorski, W. (1978).
Eur. J. Biochem. 86, 465 - 472.

Zelcer, A., Weaber, K. F., Balazs, E. and Zaitlin, M. (1981).
Virology, 113, 417 - 427.

Ziff, E. B. (1980).
Nature 287, 491 - 499.

Zimmern, D. (1982).
Trends in Biochem. Sci. 7, 205 - 207.

Zimmern, D. and Kaesberg, P. (1978).
Proc. Natl. Acad. Sci. USA 75, 4257 - 4261.

DOKTRIN 1842

Schlage die Trommel und fürchte dich nicht,
Und küsse die Marketenderin!
Das ist die ganze Wissenschaft,
Das ist der Bücher tiefster Sinn.

Trommle die Leute aus dem Schlaf,
Trommle Reveille mit Jugendkraft,
Marschiere trommelnd immer voran.
Das ist die ganze Wissenschaft.

Das ist die Hegelsche Philosophie,
Das ist der der Bücher tiefster Sinn!
Ich hab' sie begriffen, weil ich gescheit,
Und weil ich ein guter Tambour bin.

(Heinrich Heine)

For Ottoveccio, Trautchen and the Wombat