STRUCTURES OF VIROIDS VIRUSOIDS AND SATELLITES

Jim Haseloff B.Sc. (Hons)

Adelaide University Centre for Gene Technology,

South Australia.

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

May, 1983.

CONTENTS

STATEMENT

This thesis has not previously been submitted for an academic award at this or any other University, and is the original work of the author, except where due reference is made in the text.

JIM HASELOFF

ACKNOWLEDGEMENTS

I wish to thank Prof. W. H. Elliot for permission to undertake these studies in the Department. I also wish to thank my supervisor, Bob Symons for the advice and support provided to me during the course of this work. In addition, I wish to express my appreciation to the following people :

Dr. Peter Palukaitis, for purified CSV and arousing my interest in viroids;

Drs. Richard Francki and John Randles, for helpful discussions and interesting viruses;

Dr. Nizar Mohamed, Julita Imperial and Judith Rodriguez, for providing ample amounts of ccRNAs;

Dr. George Bruening, for unpublished results and stimulating discussion in the lab;

Dr. Detlev Riesner, for providing unpublished results and eintopf;

Dr. Allan Gould, for his help and dangerous sense of humour;

Karl Gordon, for his red wine and socio-politics (more like rosè), and scientific discussion;

My fellow viroid/virusoid infected co-workers, Dr. Peter Murphy, Jane Visvader and Paul Keese, as well as the other numerous laboratory and departmental inmates;

Jenny Rosey, Sharon Freund and Lisa Waters, for excellent technical assistance and preparation of the figures for this thesis;

Mrs. To, for her care in typing this thesis;

and finally my family, for putting up with this student.

SUMMARY

The work described in this thesis concerns the establishment and application of techniques for the rapid sequence determination of small circular RNAs such as those of viroids and virusoids. The determined sequences of chrysanthemum stunt viroid, the variant RNAs of coconut cadang-cadang viroid and the virusoids of velvet tobacco mottle virus, solanum nodiflorum mottle virus and subterranean clover mottle virus are presented. The overall conclusions from the work are outlined briefly below.

1. Viroids contain highly conserved sequences central to their rod-like native structures.

2. Virusoids also contain highly conserved sequences central to their rod-like native structures and share the pentanucleotide sequence GAAAC with that of viroids.

3. In addition the conserved sequences of virusoids are shared by the linear RNA of tobacco ringspot virus.

Presumably the common sequences of each class of RNAs reflect common function, and perhaps suggest some functional similarity between viroids and virusoids. Sequence homology between virusoids and the satellite RNA of tobacco ringspot virus allows prediction of sites for processing of these RNAs from multimeric RNA intermediates of replication.

CHAPTER 1

INTRODUCTION

A. Viroids

Viroids constitute a unique class of infectious plant pathogens, and as such are a fairly recent dicovery. The viroid concept was first recognized when the infectious agents of the spindle tuber disease of potato (Diener and Raymer, 1967) and the exocortis disease of citrus (Semancik and Weathers, 1968), which were thought to be viruses, were shown to possess unusual properties : (1) phenol or other organic solvents had no effect upon the infectivity of buffered extracts from infected plants; (2) no virus particles could be isolated or demonstrated by electron microscopy; (3) the infectious agents displayed a resistance to nucleases and an elution profile off cellulose columns similar to double-stranded RNA; and (4) the infectious agents were always present in high speed supernatants, possessing sedimentation coefficients of 10-15S. As a more detailed knowledge of the sizes and physico-chemical properties of these two disease agents became available (Raymer and Diener, 1969; Diener and Raymer, 1969; Semancik and Weathers, 1972a), it became apparent that these two agents were the first of a new class of infectious nucleic acids (Diener, 1971b; Semancik and Weathers, 1972b; Sänger, 1972); the term viroid was proposed (Diener, 1971b), and

the causal agents were renamed potato spindle tuber viroid (PSTV) and citrus exocortis viroid (CEV).

Since that time, viroids have been shown to be the causative agents of a further eight plant diseases, and are listed in Table 1-1. These viroids consist of infectious low molecular weight RNA species which are unencapsidated, single stranded, covalently-closed circular molecules with a high degree of intramolecular base-pairing (Diener, 1972; Semancik et al., 1975; Sänger et al., 1976). Physico-chemical studies of several viroids (Henco et al., 1977; Klump et al., 1978; Langowski et al., 1978; Domdey et al., 1978) culminated in a model for viroid structure in which the circular RNA molecules exist as extended rod-like structures, characterized by series of base-paired sections interspersed with single-stranded loop sections. Determination of the complete nucleotide sequence of PSTV (Gross et al., 1978), together with dye-binding experiments (Riesner et al., 1979) and tRNA-binding experiments (Wild et al., 1980) established the validity of this model.

Apart from their unique structures, the single feature which distinguishes viroids from viruses is their apparent lack of encoded polypeptide products. Viroids RNAs appear to be naked with no associated protein (Diener, 1971a; Semancik and Weathers, 1972a),

Table 1-1 Viroids that are presently known[1]

Viroid		References
1. potato spindle tuber viroid	(PSTV)	(Diener, 1977)
2. citrus exocortis viroid	(CEV)	(Semancik and Weathers, 1972; Sänger, 1972)
3. chrysanthemum stunt viroid	(CSV)	(Diener and Lawson, 1973)
4. chrysanthemum chlorotic mottle viroid	(ChCMV)	(Romaine and Horst, 1975)
5. cucumber pale fruit viroid	(CPFV)	(Van Dorst and Peters, 1974)
6. coconut cadang-cadang viroid	(CCCV)	(Randles, 1975)
7. hop stunt viroid	(HSV)	(Sasaki and Shikata, 1977)
8. avocado sunblotch viroid	(ASBV)	(Thomas and Mohamed, 1979)
9. tomato bunchy top viroid	(TBTV)	(Walter, 1981)
10. tomato planta macho viroid	(TPMV)	(Galindo *et al.*, 1982)

[1] Recent studies indicate that the causative agent of burdock stunt disease may possess properties atypical of those of other viroids (Chen Wei and Tien Po, personal communication), and has been tenatively omitted from the list.

and no viroid-coded translation products have been found either *in vivo* (Conjero and Semancik, 1977; Flores *et al.*, 1978; Conjero *et al.*, 1979; Camacho and Sänger, 1982a,b) or *in vitro* (Davies *et al.*, 1974; Semancik *et al.*, 1977). If viroids do not encode functional polypeptide translation products, they must rely entirely on plant host cell components for their replication.

Viroid replication has been shown to be inhibited by actinomycin D, inhibiting DNA-dependent RNA synthesis (Diener and Smith, 1975; Takahashi and Diener, 1975; Mühlbach and Sänger, 1979), and α-amanitin at concentrations which inhibit RNA polymerase II, and thus mRNA synthesis (Mühlbach and Sänger, 1979). While the effects of these drugs on viroid replication may be indirect, due to general effects on host cell metabolism, some evidence suggests that DNA-dependent RNA polymerase II, the target for α-amanitin, may play a direct role in viroid replication. Rackwitz *et al.* (1981) have shown that RNA polymerase II purified from healthy plant tissues is capable of the α-amanitin sensitive transcription of full-length linear complementary viroid RNAs from viroid template *in vitro*. However, the *in vitro* transcription of viroid templates has also been shown for RNA-dependent RNA polymerases isolated from healthy plant tissue (Boege *et al.*,1982)

and from cucumber mosaic virus infected plant tissue (D.S. Gill and R.H Symons, unpublished results).

So, while the exact nature of the enzymes involved in viroid replication *in vivo* remains unclear, the following details are known. (1) Viroids appear to replicate through complementary RNA intermediates (Grill and Semancik, 1978; Zaitlin *et al.*, 1980; Hadidi *et al.*, 1981; Zelcer *et al.*, 1982). (2) Longer than unit-length complementary viroid RNA intermediates have been detected in viroid infected tissue extracts (Branch *et al.*, 1981; Rohde and Sänger, 1981; Owens and Diener, 1982; Bruening *et al.*, 1982) (3) Oligomeric series of RNAs of avocado sunblotch viroid (ASBV) have been detected in infected avocado tissue (Bruening *et al.*, 1982). Various workers have therefore postulated rolling circle mechanisms for the transcription of larger than unit length viroid intermediates (Branch *et al.*, 1982; Owens and Diener, 1982; Bruening *et al.*, 1982). Such mechanisms require that unit-length linear viroid, produced by either specific transcription or cleavage of oligomeric RNAs, be ligated to produce the final circular product.

B. *Virusoids*

Four members of a new and unique group of plant viruses have been recently described (Randles *et*

al., 1981; Gould and Hatta, 1981; Tien-Po et al., 1981). These viruses, velvet tobacco mottle virus (VTMoV), solanum nodiflorum mottle virus (SNMV), lucerne transient streak virus (LTSV) and subterranean clover mottle virus (SCMoV) were isolated in Australasia (see Figure 1-1) and each consist of 30 nm polyhedral capsids containing two major single-strand RNA species. RNA 1 is a linear molecule of about 4,500 residues, whereas RNA 2 is a circular covalently closed molecule of 300-400 residues with a high degree of internal base-pairing. The RNA 2 molecules therefore possess physical properties similar to those of viroids and have been termed virusoids.

Gould et al. (1981) have shown that both RNA 1 and RNA 2 of VTMoV and SNMV are required for viral infection and that therefore these virusoid molecules contribute some function essential for replication. In contrast, Jones et al. (1983) have shown that the virusoid of LTSV is apparently not required for viral infection and that it therefore behaves as a satellite RNA. The nature of the relationship between the RNA components of SCMoV is unknown.

C. AIMS

Three unresolved questions stand behind the work described in this thesis.

Figure 1-1 Origins of viruses containing virusoids.

The geographic locations of isolates of VTMoV (Randles et al., 1981), SNMV (Gould and Hatta, 1981), LTSV from New Zealand (LTSV-NZ) and Australia (LTSV-A) (Tien-Po et al., 1981) and SCMoV (Francki et al., 1983) are show mapped.

ORIGIN OF VIRUSES CONTAINING VIRUSOIDS

(1) How do viroids replicate ?

(2) What function (if any) do virusoids contribute to virus replication ?

(3) Does the replication of viroids and virusoids share common features ?

In order to address these general questions of viroid and virusoid function, the primary and secondary structures of a number of these molecules were determined. It was hoped that the location of regions essential for function within these molecules would be mirrored by the presence of conserved nucleotide sequences and/or structures.

The various techniques required for viroid/virusoid sequence determination are described in the following chapter. Chapters 3, 4, 5 and 6 describe various applications of these techniques, and the final chapter deals with the overall conclusions from the work.

CHAPTER 2

RNA SEQUENCE DETERMINATION

INTRODUCTION

The techniques available for the rapid sequence determination of RNA rely on the presence of a fixed reference point within the RNA, eg. a unique site for primed synthesis of transcripts (Zimmern and Kaesberg, 1978; Symons, 1978) or a 5' or 3' radiolabelled terminus (Donis-Keller *et al.*, 1977), and therefore these techniques are ideally applied to linear RNAs. Modified approaches must be used for the sequence determination of circular RNAs, and the approaches used in this work fall into two classes.

First, by exploiting the base-paired nature of viroids and virusoids, specific linear RNA fragments may be produced by partial RNase cleavage of the circular molecules. The linear viroid or virusoid fragments may be radiolabelled, purified and sequenced using the partial enzymic digestion (Donis-Keller *et al.*, 1977) or dideoxynucleotide chain termination (Zimmern and Kaesberg, 1978; Symons, 1978) sequencing techniques. Second, double-strand cDNA can be transcribed from linearized viroid or virusoid RNA, inserted and propagated in a bacteriophage M13 vector, and the inserted DNA sequenced directly (Sanger *et al.*, 1980) or excised or transcribed and used to prime dideoxynucleotide chain termination sequencing of the

circular viroid or virusoid RNA (Zimmern and Kaesberg, 1978;Symons, 1978). The details of these techniques are given below.

MATERIALS

RNases A and T_1, calf intestinal alkaline phosphatase, E.coli tRNA, deoxynucleoside triphosphates and isopropylthiogalactoside (IPTG) were obtained from the Sigma Chemical Co.

T_4 polynucleotide kinase, T_4 DNA ligase and Klenow fragment of E.coli DNA polymerase were obtained from Boehringer.

T_4 RNA ligase, dideoxynucleoside triphosphates and $d(T_8C)$ were obtained from P.L. Biochemicals.

M13 specific 17-mer primer and all restriction endonucleases were obtained from New England Biolabs.

RNase U_2 was obtained from Sankyo.

5-bromo-4-chloro-3-indoyl-galactoside (BCIG) was obtained from Bethesda Research Laboratories.

Avian myeloblastosis virus reverse transcriptase was obtained from Life Science Inc., Petersburg, Florida.

$[\alpha\text{-}^{32}P]$dCTP and $[\alpha\text{-}^{32}P]$dATP, at specific activities of 1000 Ci/mmol, and $[\gamma\text{-}^{32}P]$ATP at a

specific activity of 2000 Ci/mmol were prepared by Dr. R.H. Symons as previously described (Symons, 1977; 1981).

E.coli poly(A) polymerase was purified according to Sippel (1973). Phy M RNase was prepared (Donis-Keller, 1980) from culture supernatants of Physarum polycephalum, the inoculum of which was kindly provided by the School of Biological Sciences, Flinders University of South Australia. The extracellular RNase of Bacillus cereus was prepared as described by Lockard et al. (1978) from a culture supplied by Dr. G. Brownlee.

METHODS

A. Isolation and sequence determination of linear viroid or virusoid fragments

A-1 Analytical RNase digests

Circular RNA was digested with a range of RNase concentrations to obtain optimal conditions for the production of specific linear fragments. Five aliquots, each of 0.1 µg to 0.5 µg purified circular RNA were resuspended in 9 µl of the appropriate high salt RNase digestion buffer (600mM NaCl, 10mM $MgCl_2$ with either 20mM sodium citrate pH 3.5 for RNase U_2, or 20mM Tris-HCl pH 7.5 for RNase T_1 or RNase A digestions) and placed on ice. 1 µl of 100,000 units/ml RNase T_1,

1 mg/ml RNase A or 100 units/ml RNase U_2 was added to one of the tubes, which was to contain the highest concentration of RNase. The tube contents were mixed and 1 µl removed to a second tube; this was followed subsequently by another two similar 10-fold RNase dilutions. No RNase was added to the fifth tube. For example, if RNase T_1 was used, the five tubes would contain 10,000, 1,000, 100, 10 and 0 units/ml RNase T_1 respectively. After incubation at 0°C for 60 minutes, the digestions were terminated by extraction with 100 µl water saturated phenol : chloroform (1 : 1) and 100 µl 0.2M NaCl, 0.1mM Na_2EDTA; the aqueous phases were removed, washed twice with ether and precipated with 3 volumes of ethanol. After 20 minutes at -80°C, the tubes were centrifuged at 10,000g for 15 minutes at 4°C, the supernatants were removed, and the precipitates were dried <u>in vacuo</u> with 10 µCi [γ-^{32}P]ATP.

The dried pellets containing RNA fragments and [γ-^{32}P]ATP were resuspended in 5 µl H_2O, heated at 80°C for 1 minute, snap cooled on ice, and 1 µl of 5x polynucleotide kinase buffer (125mM Tris-HCl pH 9.0, 50mM $MgCl_2$, 50mM DTT) added. 0.25 units of polynucleotide kinase was added, and the reaction was incubated at 37°C for 30 minutes. 5 µl of formamide loading buffer (95% deionized formamide, 10mM EDTA, 0.02% xylene cyanol FF, 0.02% bromophenol blue) was

added, the tubes were heated at 80°C for 1 minute, snap cooled on ice, and loaded onto a 20x40x0.05 cm 6% polyacrylamide gel containing TBE (90mM Tris-borate pH 8.3, 2mM Na_2EDTA) and 7M urea (Sanger and Coulson, 1978). After electrophoresis, the gel was autoradiographed.

A-2 Preparative RNase digests

Purified circular RNA (2 to 20 µg) was resuspended in 50 µl of RNase T_1 and or RNase U_2 high salt digestion buffer (see above) and cooled on ice. The appropriate amount of RNase T_1, RNase A or RNase U_2, as determined by analytical ribonuclease digests, was added and incubation continued for 60 minutes at 0°C. Digestions were terminated by extraction with 150 µl water saturated phenol : chloroform (1 : 1) and 100 µl H_2O. The aqueous phase was removed, washed twice with 1 ml ether, and 450 µl ethanol was added. After 20 minutes at -80°C the sample was centrifuged at 10,000g for 15 minutes at 4°C and the supernatant discarded. The precipitated RNA fragments could be either 5' or 3' radiolabelled.

A-3 5'-^{32}P-labelling of RNA fragments

The pellet, with 200 µCi of added [γ-^{32}P]ATP (2000 Ci/mmol), was dried *in vacuo*, resuspended in 8 µl

1.5mM spermidine, heated at 80°C for 1 minute and snap cooled on ice. 2 µl of 5x T_4 polynucleotide kinase buffer(see above) and 1 µl (4 units) T_4 polynucleotide kinase were added and the reaction incubated at 37°C for 30 minutes.

A-4 3'-^{32}P-labelling of RNA fragments

i) Synthesis of [5' -^{32}P]dpCp

500 µCi [γ-^{32}P]ATP (2000 Ci/mmol) was dried *in vacuo*, resuspended in 5 µl H_2O, and 2 µl 10 mg/ml 3' dCMP, 2 µl 5x T_4 polynucleotide kinase buffer and 1 µl (4 units) T_4 polynucleotide kinase were added. The reaction was incubated at 37°C for 30 minutes, heated at 90°C for 1 minute and stored at -15°C before use.

ii) 3'-^{32}P-labelling

The precipitated RNA fragments were dried *in vacuo*, resuspended in 20 µl 10mM Tris-HCl pH 9.0 containing 0.01 units calf intestinal alkaline phosphatase, and incubated at 50°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 fragments were 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 µM 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.

A-5 Polyacrylamide gel fractionation

10 µl of formamide loading buffer (95% deionized formamide, 10mM Na_2EDTA, 0.02% bromophenol blue, 0.02% xylene cyanol FF) was added to each reaction mixture containing 5'- or 3'- radiolabelled RNA fragments. The reaction mixtures were heated at 80°C for 30 seconds, snap cooled on ice and loaded onto an 80x20x0.05 cm 6% polyacrylamide gel containing 90mM Tris-borate pH 8.3, 2mM EDTA and 7M urea. After electrophoresis for 6 hours at 25 mA, the gel was autoradiographed at room temperature for 5-30 minutes and the resultant autoradiograph used as a template to locate and excise the ^{32}P-labelled fragments. Excised bands were eluted by soaking overnight at room temperature in 500 µl of 500mM ammonium acetate, 1mM Na_2EDTA, 0.1% SDS, which contained 60 µg *E.coli* tRNA as

carrier if the fragments were to be sequenced using the partial enzymic cleavage technique. After soaking, the elution buffer was removed and the RNA fragments were precipitated by the addition of 1 ml ethanol and storage at -80°C for at least 30 minutes. After centrifugation at 10,000g for 15 minutes at 4°C, the pelleted fragments were resuspended in 100 µl 0.2M NaCl, 0.1mM Na_2EDTA and re-ethanol precipitated with 300 µl ethanol. After centrifugation the pellets were dried *in vacuo*. Purified 5' or 3' radiolabelled RNA fragments were used for sequence determination by either the partial enzymic cleavage or dideoxynucleotide chain termination techniques.

A-6 *Sequence determination of RNA fragments using the partial enzymic cleavage technique*

5'- or 3'-^{32}P-labelled RNA fragments, with 60 µg *E.coli* tRNA, were resuspended in 12 µl H_2O and six aliquots of 1 µl dispensed. The following buffers were added to the six tubes.

Tube N (No enzyme)	9 µl	20mM Na citrate pH 5.0 1mM Na_2EDTA 7M urea
Tube T (RNase T_1)	9 µl	20mM Na citrate pH 5.0 1mM Na_2EDTA 7M urea

Tube U (RNase U_2)	9 µl	20mM Na citrate pH 3.5 1mM Na_2EDTA 7M urea
Tube L (alkali ladder)	5 µl	50mM $Na_2CO_3/NaHCO_3$ pH 9.0
Tube D (RNase PhyM)	9 µl	20mM Na citrate pH 5.0 1mM Na_2EDTA 7M urea
Tube B (Bacillus cereus RNase)	5 µl	20mM Na citrate pH 5.0 1mM Na_2EDTA

Tubes N, T, U, P, and B were heated at 80°C for 1 minute, snap cooled on ice, and the ribonucleases added.

Tube N	-	
Tube T	1 µl	10,000 units/ml RNase T_1
Tube U	1 µl	5 units/ml RNase U_2
Tube L	-	
Tube P	1 µl	RNase PhyM extract
Tube B	1 µl	B. cereus RNase extract

Tube L was heated at 100°C for 90 seconds while the remaining tubes were incubated at 50°C for 20 minutes. At completion of the sequencing reactions, the tubes were stored at -80°C or were kept on ice while being prepared immediately for electrophoresis. Before polyacrylamide gel electrophoresis, formamide loading buffer (95% deionized formamide, 10mM EDTA, 0.02% (w/v)

bromophenol blue, 0.02% (w/v) xylene cyanol FF) was added to the samples to a final volume of 12 µl. Samples were heated at 80°C for 1 minute and snap cooled on ice before electrophoresis.

A-7 Sequence determination of RNA fragments using the dideoxynucleotide chain termination technique

i) Phosphatase treatment

Purified 5'-^{32}P-labelled RNA fragments, obtained by RNase T_1 digestion, were each resuspended in 20 µl 50mM Tris-HCl pH 9.0, heated at 80°C for 1 minute and snap cooled on ice; 1 µl (0.01 units) calf intestinal alkaline phosphatase was added, and the reactions incubated at 50°C for 20 minutes. The reactions were extracted with 100 µl water saturated phenol : chloroform (1 : 1) and 100 µl 0.2M NaCl, 0.1mM EDTA; the aqueous phases were washed twice with 1 ml ether and the fragments precipitated with 300 µl ethanol at -80°C for 20 minutes. After centrifugation at 10,000g for 15 minutes at 4°C, the pellets were dried *in vacuo*.

ii) Polyadenylation

Phosphatase treated RNA fragments were resuspended in 10 µl water, heated at 80°C for 1 minute and snap cooled on ice. 2 µl of 2mM ATP, 4 µl of 5x *E.coli* poly(A) polymerase reaction mixture, comprising 105 µl H_2O, 50 µl 1M Tris-HCl pH 7.9, 25 µl

0.1M $MnCl_2$, 10 µl 1M $MgCl_2$ and 10 µl 0.1M DTT, and 6 µl of E.coli poly(A) polymerase extract were added, and the reaction incubated at 37°C for 60 minutes. The reactions were extracted with 100 µl water saturated phenol : chloroform (1 : 1) and 100 µl 0.2M NaCl, 0.1mM EDTA, washed twice with 1 ml ether and precipitated with 300 µl ethanol at -80°C for 20 minutes. After centrifugation at 10,000g for 15 minutes at 4°C, the pellets were dried in vacuo.

iii) Reverse transcription

Polyadenylated RNA fragments were resuspended in 10 µl H_2O with 1 µl 0.25 mg/ml dT_8C, heated at 80°C and allowed to cool to room temperature over 20 minutes. 2.5 µl aliquots were dispensed into four tubes to give reaction mixtures of 5 µl, containing 50mM Tris-HCl pH 8.3, 50mM KCl, 8mM $MgCl_2$, 10mM DTT, 2 units of avian myeloblastosis virus reverse transcriptase, 2 µCi [α-^{32}P]dATP or dCTP, 50 µM of the remaining deoxynucleoside triphosphates and a single dideoxynucleoside triphosphate species essentially as described by Symons (1978). Reactions were incubated at 37°C for 30 minutes; 5 µl of formamide loading buffer (95% deionized formamide, 10mM Na_2EDTA, 0.02% (w/v) bromophenol blue, 0.02%(w/v) xylene cyanol FF) was added to each reaction, and the tubes were heated at 100°C for 2 minutes and snap cooled on ice before

electrophoresis.

A-8 Polyacrylamide gel electrophoresis

Radiolabelled products of the partial enzymic cleavage and dideoxynucleotide chain termination techniques were fractionated by electrophoresis in 80x20x0.05 cm 8% polyacrylamide gels or 40x20x0.05 cm 20% polyacrylamide gels containing 90mM Tris-borate pH 8.3, 2mM Na_2EDTA and 7M urea (Sanger and Coulson, 1978).

The sequence determination of some RNA fragments was complicated by the presence of band compression artifacts (Kramer and Mills, 1978) arising from incomplete denaturation of RNA or cDNA fragments during electrophoresis. In order to eliminate band compression, fragments were fractionated in polyacrylamide gels containing 98% formamide.

Sequencing reaction mixtures were precipitated by adding 100 µl 0.2M NaCl, 0.1mM Na_2EDTA and 300 µl ethanol, and the samples were kept at -80°C for 20 minutes, centrifuged at 10,000g for 15 minutes at 4°C, and the pellets dried *in vacuo*. The samples were resuspended in 5 µl formamide loading buffer, heated at 80°C for 1 minute, cooled on ice and loaded onto a 40x20x0.05 cm 20% polyacrylamide gel containing 98% formamide buffered with 16mM Na_2HPO_4, 4mM NaH_2PO_4 as

described by Maniatis and Efstratiadis (1980). Following electrophoresis, gels were autoradiographed at -80°C, using calcium tungstate intensifying screens.

B. Sequence determination using cloned viroid or virusoid sequences

B-1 Synthesis and cloning of viroid and virusoid ds cDNA

i) Linearization and polyadenylation

Purified circular RNA (2-10 µg) in 10 µl distilled water was heated at 100°C for 30 minutes in a sealed capillary to generate randomly cleaved full-length linear molecules. As described above for radiolabelled RNA fragments in A-7(i), terminal 2'(3')-phosphate groups were removed from the cleaved molecules by the addition of Tris-HCl pH 9.0 to 50mM and 0.01 units calf intestinal alkaline phosphatase followed by incubation at 50°C for 20 minutes. Reactions were phenol-chloroform extracted, twice ether washed and ethanol precipitated.

Phosphatase treated RNA molecules were resuspended in 50 µl water, heated at 80°C for 1 minute and snap cooled on ice; 10 µl of 2mM ATP, 20 µl of 5x E.coli poly(A) polymerase reaction mixture (see A-7(ii)), and 30 µl E.coli poly(A) polymerase extract were added and the reaction incubated at 37°C for 60

minutes. Reactions were extracted with 100 µl water saturated phenol : chloroform, twice ether washed and precipitated with 3 volumes of ethanol.

ii) First strand cDNA synthesis

Polyadenylated RNA was resuspended in a 50 µl reaction mixture containing 50mM Tris-HCl pH 8.3, 50mM KCl, 10mM $MgCl_2$, 10mM DTT, 1mM each of dATP, dGTP and dTTP, 200 µM [α-^{32}P]dCTP (60 µCi), 0.6 µg $(dT)_{10}$ and 30 units avian myeloblastosis virus reverse transcriptase, and incubated at 42°C for 60 minutes. The reaction mixture was heated at 100°C for 1 minute, snap cooled on ice, 1 µl (10 µg) heat-treated RNase A added and incubation continued at 50°C for 20 minutes. Reactions were extracted with 100 µl 0.2M NaCl, 0.1mM Na_2EDTA and 100 µl water saturated phenol : chloroform (1 : 1), twice ether washed and precipitated with 3 volumes of ethanol.

iii) Second strand cDNA synthesis

Single strand cDNA was resuspended in 10 µl H_2O, boiled for 1 minute and snap cooled on ice, added to make up a 25 µl reaction mixture containing 50mM Tris-HCl pH 7.5, 10mM $MgCl_2$, 10mM DTT, 1mM dATP, dCTP, dGTP and dTTP and 2 units Klenow fragment of E.coli DNA polymerase 1, incubated for 4 hours at 37°C and kept at -20°C before use.

iv) Restriction enzyme cleavage and fractionation

Double strand cDNA was subjected to digestion by various restriction endonucleases under conditions as specified by the suppliers of the enzymes. The cleaved ds cDNA was fractionated by electrophoresis in a 20x40x0.05 cm 6% polyacrylamide gel (Sanger and Coulson, 1977) containing TBE buffer (see above) and 2M urea. Following electrophoresis, the gel was autoradiographed, the ds cDNA fragments were excised and each eluted in 400 µl 0.5M ammonium acetate, 0.1% SDS, 0.1mM Na_2EDTA at room temperature overnight, and ethanol precipiated twice.

v) M13 cloning

Purified ds cDNA fragments were ligated into an appropriate restriction site of the replicative form of the phage M13 mp7 using phage T_4 DNA ligase (Goodman and MacDonald, 1979; Messing, Crea and Seeburg, 1981). Further specific details are provided in following chapters. Ligated M13 RF and ds cDNA was used to transform competent E.coli JM101 and the cells were plated on agar media with BCIG (5-bromo-4-chloro -3-indoyl-β-D-galactoside) and IPTG (Isopropyl-β -D-thiogalactopyranoside). Recombinant M13 phage were screened by sequencing as described below.

B-2 Sequence determination of recombinant phage M13

M13 phage containing cloned ds cDNA inserts

were selected, as judged by insertional inactivation of β-galactosidase activity (Messing, Crea and Seeburg, 1981), and the inserted sequences were determined using the dideoxynucleotide chain termination technique as described by Sanger et al. (1980) with a 17-mer M13 specific oligonucleotide primer ($GTA_4CGACG_2C_2AGT$).

B-3 Sequence determination of RNAs using cloned DNA primers

i) Preparation of primer

The DNA primers used in this work were of two types, being either fragments restricted from recombinant phage M13 RF DNA or transcribed using recombinant phage M13 ss DNA as a template. In the former case, recombinant M13 RF was isolated using the method of Birnboim and Doly (1979), restricted with the appropriate enzyme, and the fragment(s) containing viroid or virusoid sequences purifed by electrophoresis through a 6% polyacrylamide gel containing 7M urea (Sanger and Coulson, 1978). In the latter case, recombinant M13 ss DNA containing viroid or virusoid sequences of the same polarity as the RNA sequence was transcribed using an M13 specific oligonucleotide primer and the Klenow fragment of E.coli DNA polymerase 1 with $[\alpha\text{-}^{32}P]$dCTP and $[\alpha\text{-}^{32}P]$dATP (specific activities of 1000 Ci/mmol) essentially as described by Bruening et

al. (1982). The resulting partially double-stranded DNA molecules were subjected to restriction enzyme digestion, and the labelled fragments fractionated by polyacrylamide gel electrophoresis. In both cases the purified DNA primers were eluted from polyacrylamide gels by soaking (Maxam and Gilbert, 1980).

ii) RNA-DNA hybridization

RNA-DNA hybrids were prepared as follows. The purified restriction fragments and 1 to 2 µg of the appropriate RNA were resuspended in 25 µl of 0.18M NaCl, 10mM Tris-HCl pH 7.0, 1mM EDTA, 0.05% SDS, heated at 100°C for 2 minutes, and incubated at 60°C for 2 hours. The RNA-DNA hybrids were twice ethanol precipitated and dried in vacuo.

iii) Reverse transcription

The RNA-DNA hybrids were reverse transcribed in the presence of dideoxynucleoside triphosphates essentially as described above in A-7(iii). However, if ^{32}P-radiolabelled transcripts of recombinant M13 ss DNA were used as primers, [α-^{32}P]dNTPs were omitted and replaced by the unlabelled dNTP species. Reverse transcripts were fractionated by polyacrylamide gel electrophoresis as described above in A-8.

All experiments involving the use of recombinant DNA molecules were performed within safety guidelines set out by the Australian Academy of Science

Committee on Recombinant DNA molecules (ASCORD).

RESULTS AND DISCUSSION

A. Techniques

The various rapid gel sequencing techniques used for the determination of viroid and virusoid sequences are outlined in Figure 2-1. These techniques share the advantage of requiring only small amounts of purified RNA for their use and, while there are disadvantages inherent in the use of each procedure, the combined use of different techniques allows rapid and reliable RNA sequence analysis. The various approaches are reviewed briefly below.

B. Partial enzymic cleavage of viroid and virusoid RNAs

Under conditions of high salt concentration (600mM NaCl, 10mM $MgCl_2$) and low temperature (0°C), the highly base-paired native structures of viroids and virusoids are stabilized and so cleavage by the single strand specific RNases T_1, U_2 and A is initially limited to relatively few accessible sites on the molecules. Thus, partial ribonuclease digestion of the native circular RNA molecules gives rise to relatively few, specific linear RNA fragments which may then be radiolabelled and fractionated by polyacrylamide gel electrophoresis (Figs 2-2,3). The main disadvantage in

Figure 2-1a

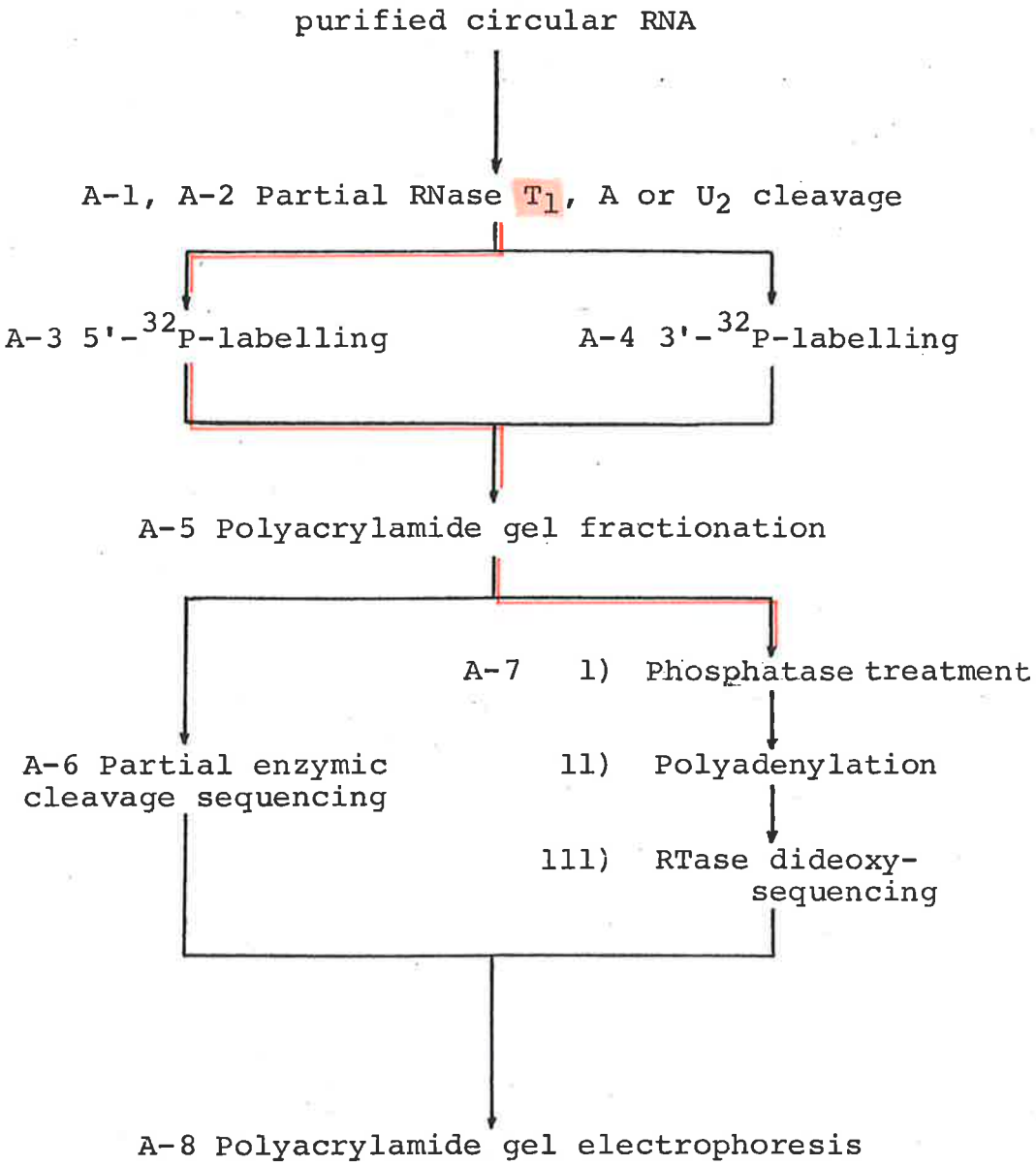

purified circular RNA
A-1, A-2 Partial RNase T_1, A or U_2 cleavage
A-3 5'-^{32}P-labelling
A-4 3'-^{32}P-labelling
A-5 Polyacrylamide gel fractionation
A-7 1) Phosphatase treatment
11) Polyadenylation
111) RTase dideoxy-sequencing
A-6 Partial enzymic cleavage sequencing
A-8 Polyacrylamide gel electrophoresis

Figure 2-1b

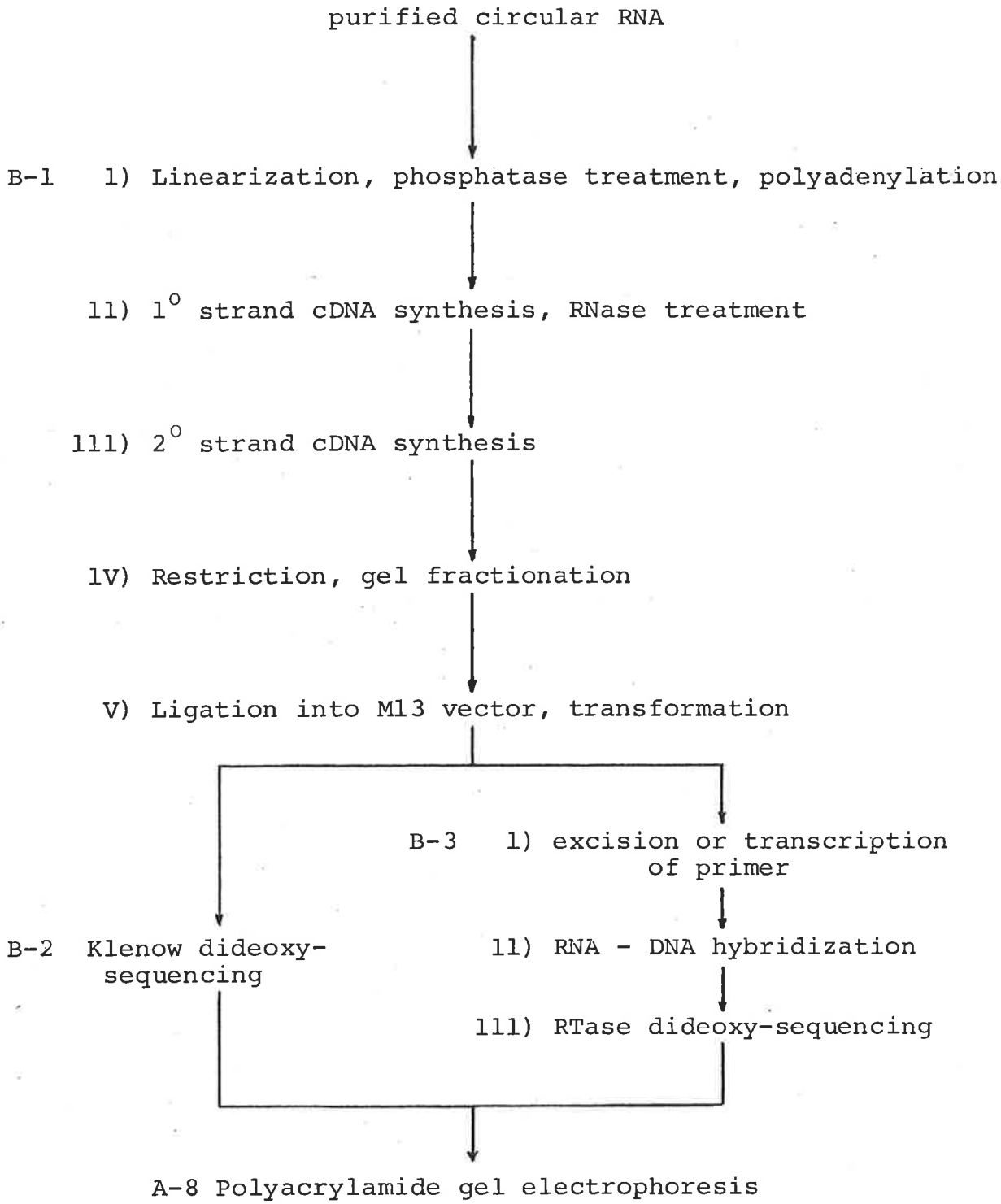

purified circular RNA
B-1 1) Linearization, phosphatase treatment, polyadenylation
11) 1° strand cDNA synthesis, RNase treatment
111) 2° strand cDNA synthesis
1V) Restriction, gel fractionation
V) Ligation into M13 vector, transformation
B-3 1) excision or transcription of primer
B-2 Klenow dideoxy-sequencing
11) RNA - DNA hybridization
111) RTase dideoxy-sequencing
A-8 Polyacrylamide gel electrophoresis

Figure 2-2 Analytical RNase A digestions of SNMV RNA 2.

As described in the text, 0.5 µg purified SNMV RNA 2 was variously digested with 0, 0.1, 1, 10 or 100 µg/ml RNase A under conditions of high salt and low temperature, and the resulting linear RNA fragments 5'-^{32}P-labelled and fractionated by denaturing polyacrylamide gel electrophoresis. The presence of a band in the track containing SNMV RNA 2 untreated with RNase A corresponds to a small amount of full-length linear breakdown product (377 residues in size) present with the intact circular RNA. Concentrations of between 0.1 and 1 µg/ml RNase A were used to obtain fragments suitable for sequencing.

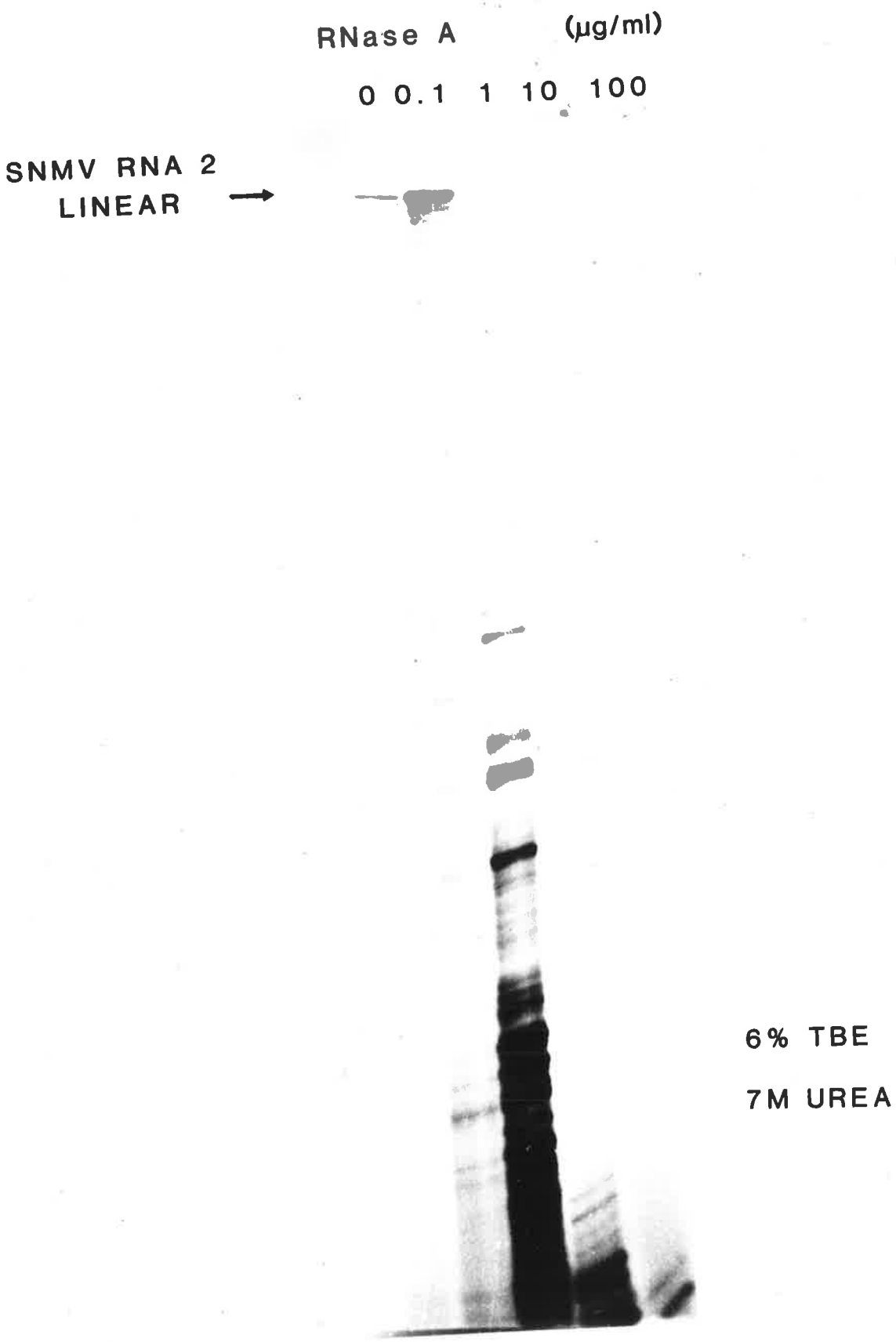
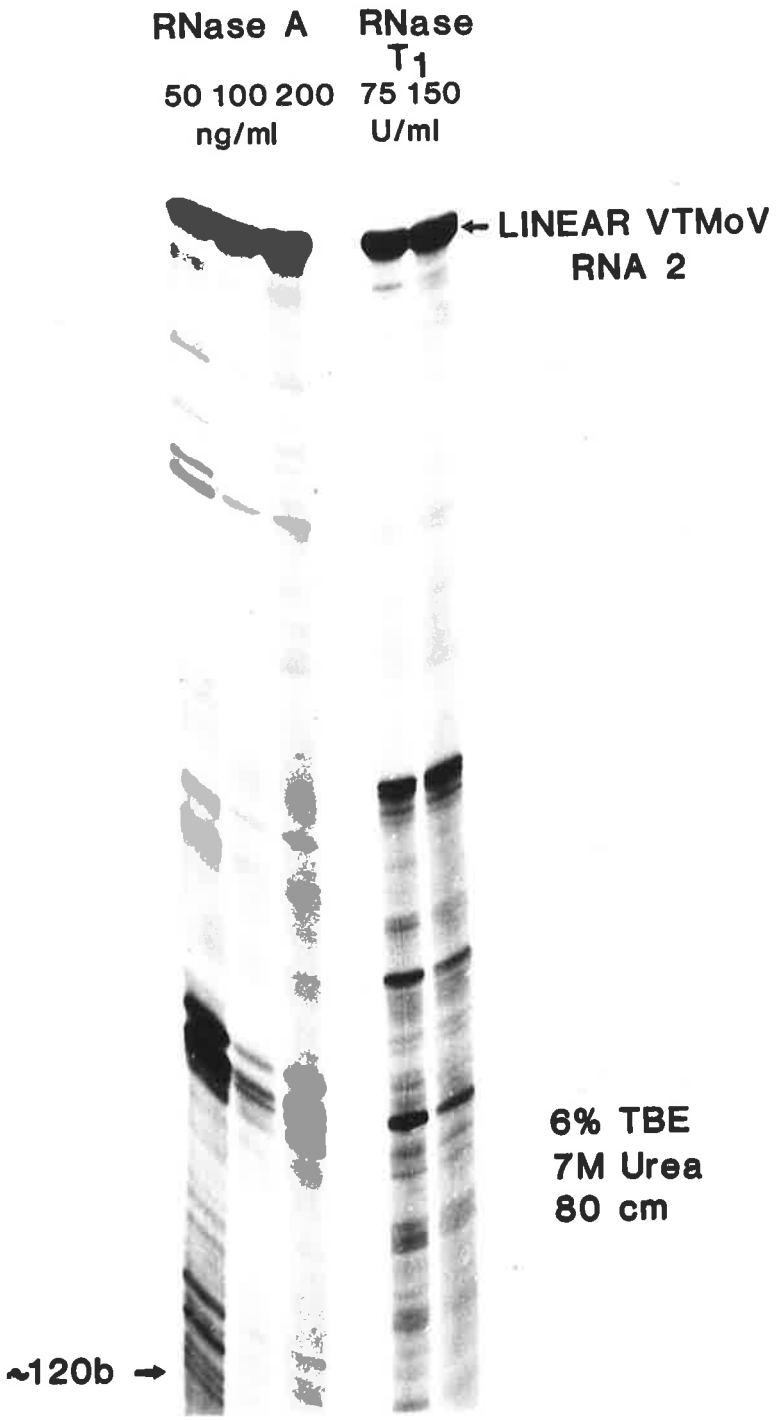

RNase A (µg/ml)
0 0.1 1 10 100
SNMV RNA 2
LINEAR
6% TBE
7M UREA

Figure 2-3 Preparative RNase A and RNase T_1 digestions of VTMoV RNA 2.

5 µg amounts of purified VTMoV RNA 2 were digested with 0.2, 0.1 or 0.05 µg/ml RNase A and 150 or 75 units/ml RNase T_1 under conditions of high salt and low temperature. 5'-radiolabelled products are shown fractionated on an 80 cm long 6% poly -acrylamide gel containing 7M urea. Only the bottom portion of the gel was autoradiographed and the band corresponding to full-length linear VTMoV RNA 2 (365/366 residues) migrated about 30 cm from the origin. Following a 5 minute autoradiographic exposure, bands were excised and eluted for sequence determination.


RNase A
50 100 200
ng/ml
RNase
T1
75 150
U/ml
← LINEAR VTMoV
RNA 2
6% TBE
7M Urea
80 cm
~120b →

using this technique is that, if the native circular RNA molecule possesses an exposed single-strand region (such as a terminal hairpin loop) containing accessible sites for all ribonucleases, it is difficult to obtain RNA fragments spanning such an exposed region. This disadvantage is not shared by those sequencing methods which rely on cloned viroid or virusoid fragments.

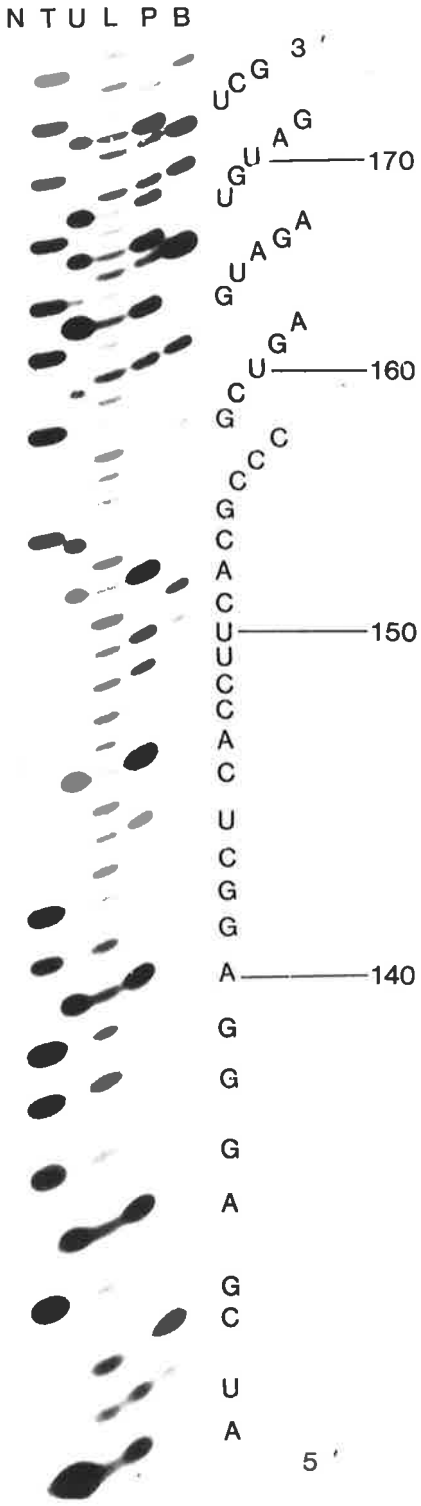
C. Partial enzymic cleavage of radiolabelled RNA fragments

The purified 3'- or 5'- radiolabelled RNA fragments obtained after partial RNase cleavage of intact viroid or virusoid molecules were sequenced using the partial enzymic cleavage method (Donis-Keller, 1977; Lockard et al., 1978; Krupp and Gross, 1979; Donis-Keller, 1980). An example of one sequencing gel is given in Figure 2-4. In particular, the use of fragments labelled separately at either the 5' or 3' terminii allowed the sequence determination of long RNA fragments and, with shorter fragments, resolved gel compression artefacts (Kramer and Mills, 1978) when the relevant nucleotide sequences were determined from both directions.

D. Dideoxynucleotide chain termination sequencing of RNA fragments

Figure 2-4 Partial enzymic digestion sequencing technique.

A 5'-^{32}P-labelled RNA fragment, obtained by partial RNase T_1 digestion of SNMV RNA 2 under non-denaturing conditions, was subjected to partial digestion under denaturing conditions with various RNases as described in the text (N, no enzyme; T, RNase T_1; U, RNase U_2; L, alkali ladder; P, RNase PhyM; B, Bacillus cereus RNase). The resulting fragments were separated by 80 cm, 8% polyacrylamide gel electrophoresis and the resulting autoradiograph is shown with residues 132 to 175 of SNMV RNA 2.


N T U L P B
3'
5'
170
160
150
140

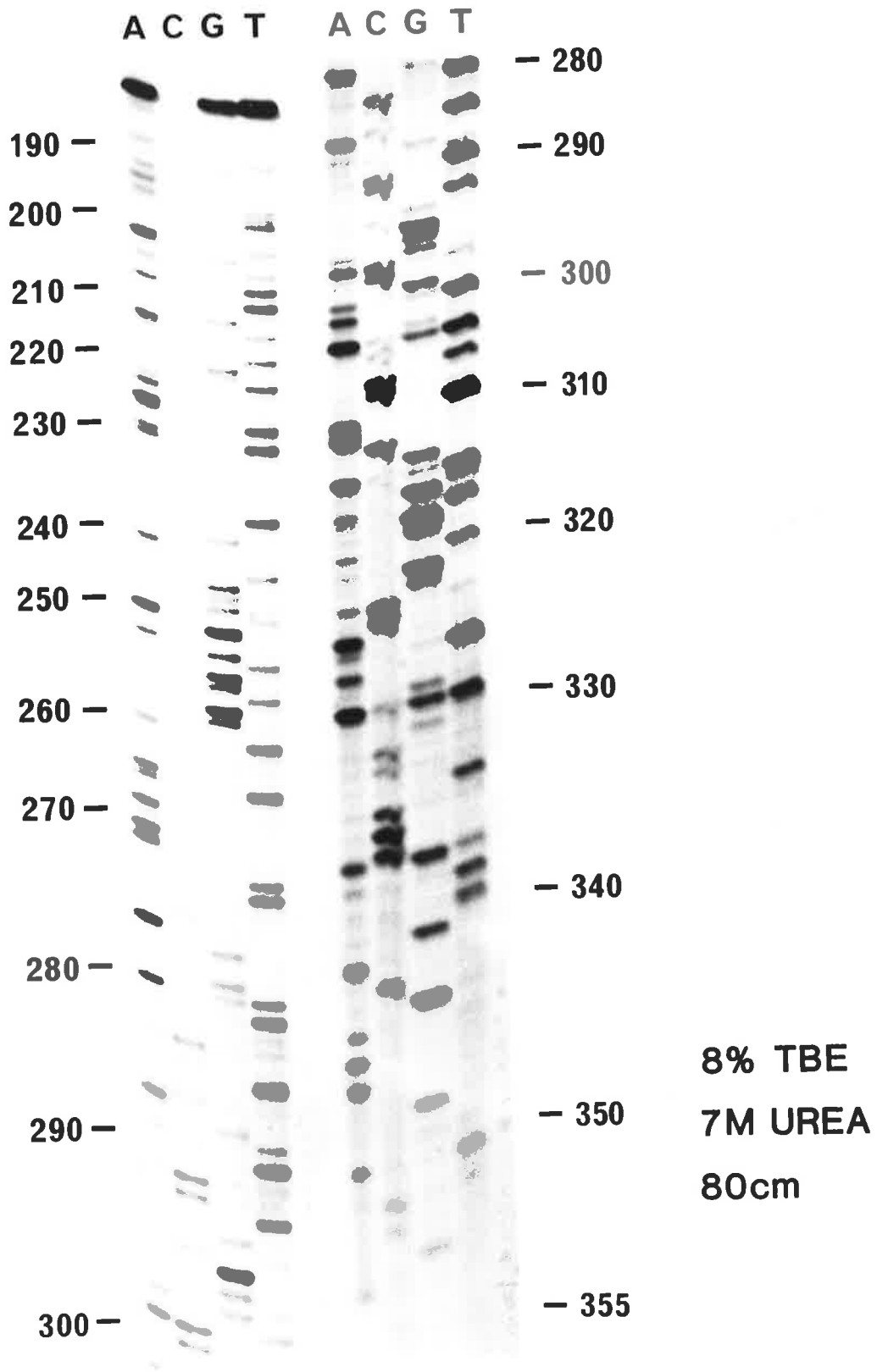
All 5'-^{32}P-radiolabelled fragments produced by RNase T_1 cleavage of viroid or virusoid RNA possess a 3' proximal guanine residue, together with a 2'(3')-phosphate group. Treatment of such fragments with calf intestinal phosphatase (Efstratiadis et al., 1977) removed both the radiolabelled 5' phosphate and 2'(3')-phosphate groups, and the RNA fragments could then be used as templates for polyadenylation using E.coli poly(A) polymerase (Sippel, 1973; Gould et al., 1978). The synthetic oligonucleotide d(T_8C) was then used as a specific primer for reverse transcription in the presence of dideoxynucleoside triphosphates. The polyacrylamide gel fractionated products of such a transcription reaction are shown in Figure 2-5. The dideoxynucleotide chain termination and partial enzymic digestion techniques are complementary, allowing further confirmation of sequences and the resolution of occasional band compressions.

E. Cloning of viroid and virusoid sequences

Purified circular RNAs were hydrolysed by prolonged heating at 100°C to generate randomly nicked full-length linear molecules. Terminal 2'(3')-phosphates were removed by treatment with calf intestinal phosphatase and the RNA molecules polyadenylated with E.coli poly(A) polymerase. First

Figure 2-5 Dideoxynucleotide chain termination sequencing of RNA fragments

An RNase T_1 cleaved fragment of VTMoV RNA 2 was phosphatase-treated, polyadenylated and reverse transcribed in the presence of dideoxynucleoside triphosphates as described in the text. Two loadings of the resultant radiolabelled transcripts are shown fractionated on an 80 cm 8% polyacrylamide gel containing 7M urea. Residues 355 to 190 (3'→5') of VTMoV RNA 2 are shown.


A C G T
A C G T
190
200
210
220
230
240
250
260
270
280
290
300
280
290
300
310
320
330
340
350
355
8% TBE
7M UREA
80cm

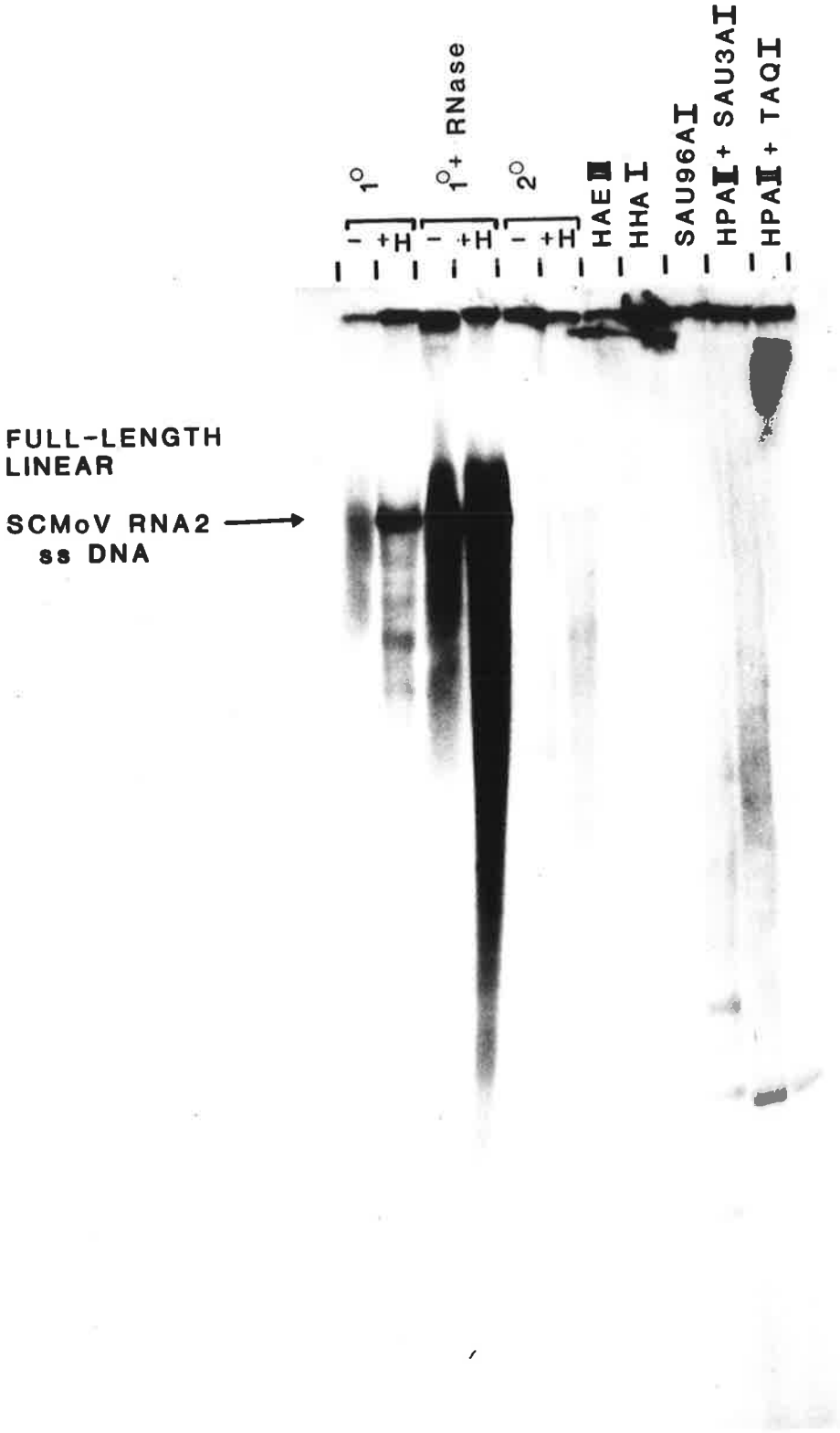
strand ^{32}P-cDNA was synthesized using reverse transcriptase and oligo-$(dT)_{10}$ as primer (Gould and Symons, 1982). The RNA.DNA hybrids were heat denatured and the template was removed by RNase A treatment. Immediately prior to the synthesis of the second DNA strand, the cDNA was heat-denatured; this step was found necessary to avoid the formation of anomalous ds cDNAs with large apparent molecular weights. The self-primed second DNA strand was transcribed using the Klenow fragment of E.coli DNA polymerase 1, and the resulting population of circularly permuted ds cDNA molecules was then digested with an appropriate restriction enzyme(s) to produce specific DNA fragments for cloning (see Figure 2-6). The restriction fragments were then purified by non-denaturing polyacrylamide gel electrophoresis and ligated into a bacteriophage M13 vector, transformed and propagated in an E.coli host (Messing, Crea and Seeburg, 1981; Sanger et al., 1980).

F. Sequence determination using cloned viroid or virusoid sequences

Viroid or virusoid sequences were determined using recombinant M13 phage in either of two ways. First, recombinant phage ss DNA was sequenced using the dideoxynucleotide chain termination technique with an M13 specific primer (Sanger et al., 1980) (see Figure

Figure 2-6 Synthesis and restriction endonuclease cleavage of SCMoV RNA 2 ds cDNA.

As described in the text, purified SCMoV RNA 2 was linearized, phosphatase treated, polyadenylated and used as a template for first strand cDNA synthesis (1°). The ss cDNA was denatured and treated with RNase (1° + RNase) and second strand cDNA was synthesised by self priming (2°). Double-strand cDNA was digested with Hae III, Hha I, Sau96 I, Hpa II + Sau3A I or Hpa II + Taq I. Samples taken at each step were fractionated by 6% polyacrylamide gel electrophoresis in the presence of 2M urea. Samples 1°, 1° + RNase and 2° were loaded with (+H) and without (-) heating at 100°C for 2 minutes. Specific bands produced by restriction endonuclease cleavage were suitable for cloning.


1°
1°+ RNase
2°
- +H
- +H
- +H
HAE III
HHA I
SAU96A I
HPA II + SAU3A I
HPA II + TAQ I
FULL-LENGTH
LINEAR
SCMoV RNA2
ss DNA
6% TBE
2M Urea

2-7). In this way, cloned sequences corresponding to either viroid/virusoid or its complement, depending on orientation of the cloned fragments, could be determined. Second, cloned sequences were excised or transcribed for use as primers for reverse transcribed dideoxynucleotide chain termination sequencing of intact viroid or virusoid RNAs (Zimmern and Kaesberg, 1978; Symons, 1978). The primed transcripts obtained in this way are representative of the whole population of RNA templates, and thus this sequencing technique has been used to estimate the relative proportions of RNA sub-species within heterogeneous populations (see Figure 2-8).

G. Determination of complete viroid or virusoid primary structures

Construction of the complete primary structures of each circular molecule depended on the obtaining of numerous overlapping sequences by one or more of the techniques described above. Once the entire base sequence of a molecule had been determined by merging the various overlapping sequences, secondary structure models were constructed using the methods of Tinoco et al. (1971). The following chapters present the determined structures of several viroid and virusoid RNAs and outline some of the interesting features of these unique molecules.

Figure 2-7 M13 dideoxynucleotide chain termination sequencing.

Single-strand DNAs isolated from recombinant bacteriophage M13 were sequenced by the dideoxynucleotide chain termination technique using a 17-mer specific primer. The sequence of a cloned Sau3A I fragment of VTMoV RNA 1 ds cDNA is shown determined from both individually cloned DNA strands.

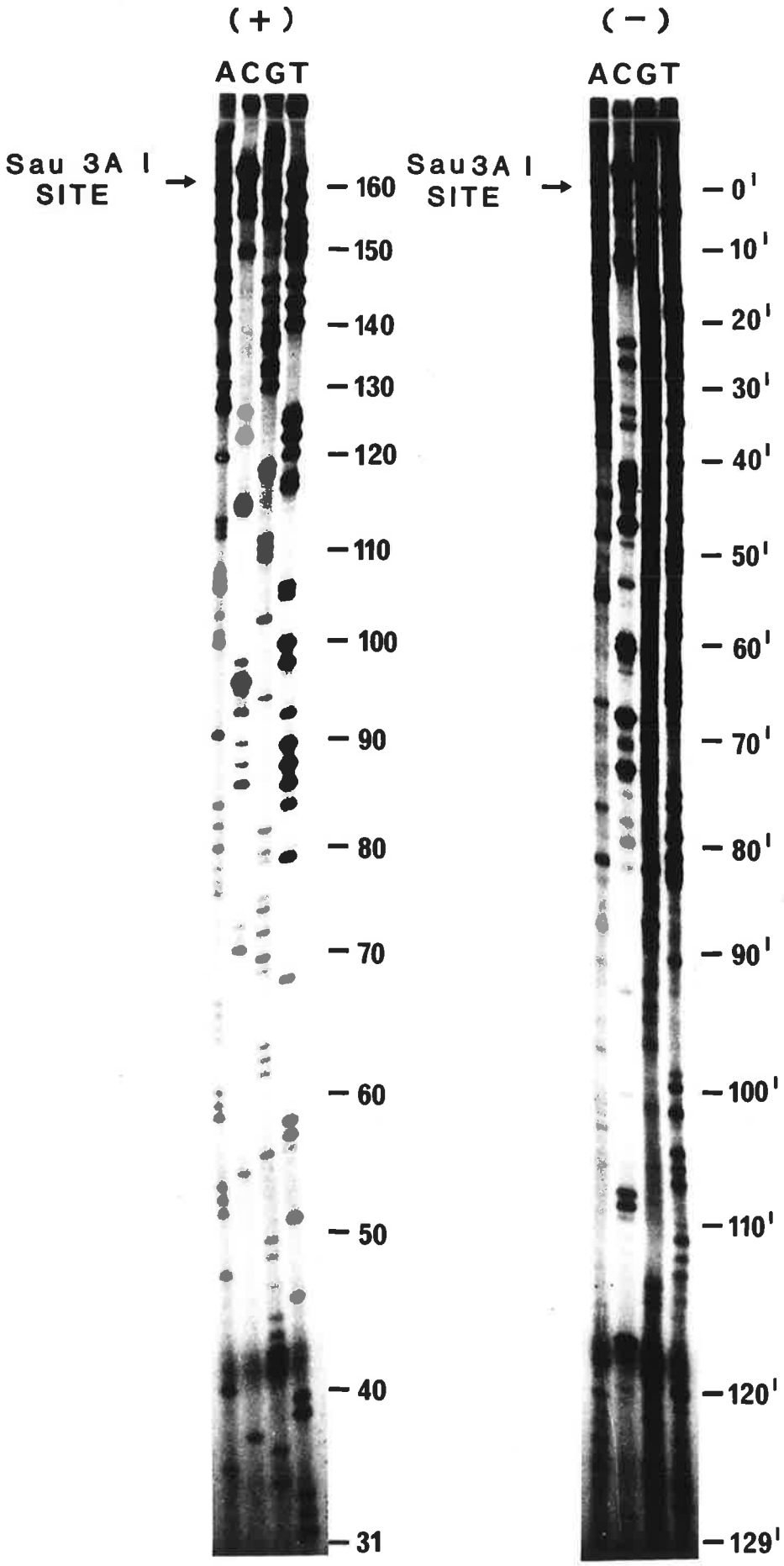
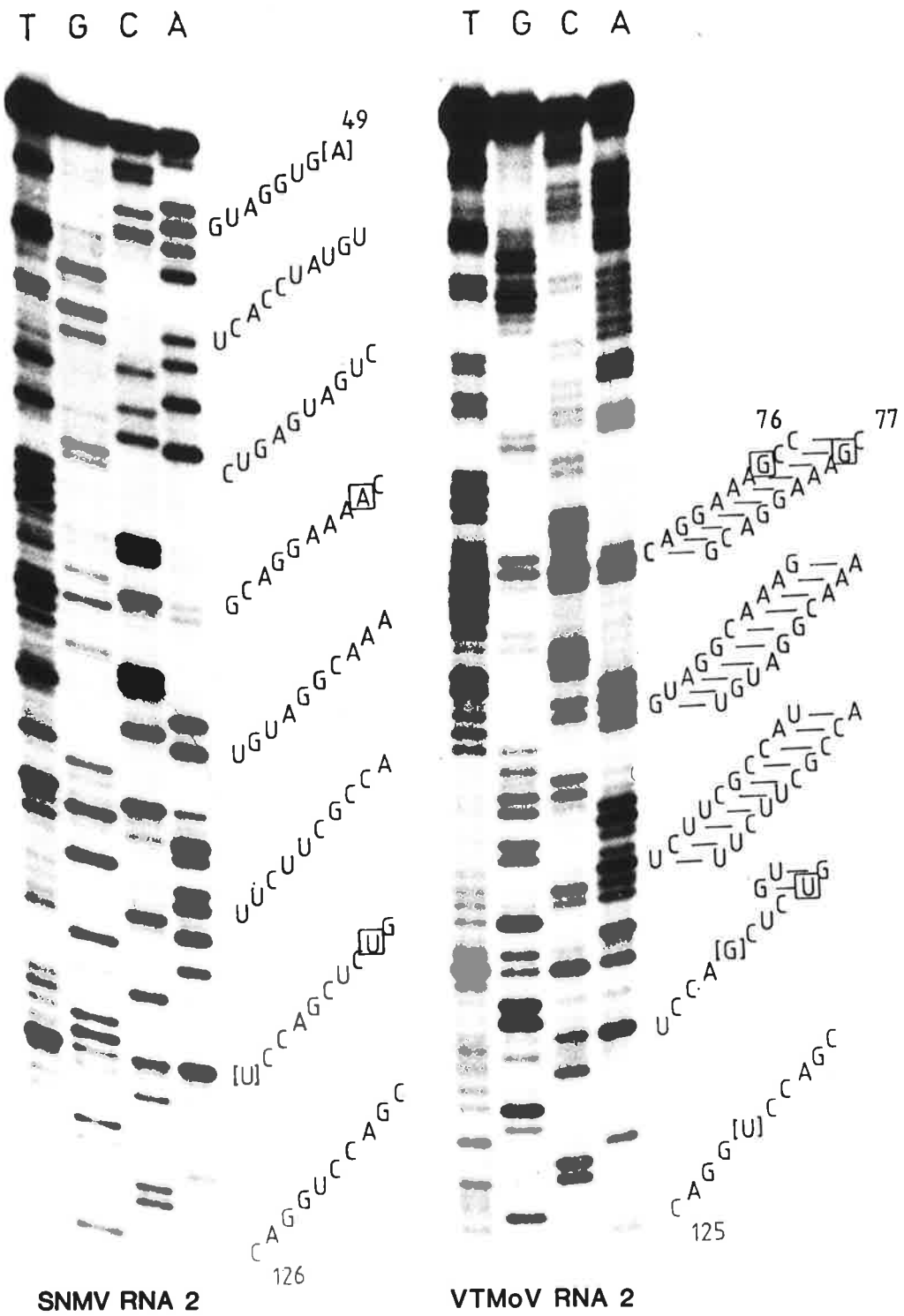

(+)
ACGT
Sau 3A I SITE
160
150
140
130
120
110
100
90
80
70
60
50
40
31
(−)
ACGT
Sau3A I SITE
0'
10'
20'
30'
40'
50'
60'
70'
80'
90'
100'
110'
120'
129'

Figure 2-8 Dideoxynucleotide chain termination sequencing of intact RNA using cloned DNA primers. Recombinant M13 ssDNA containing sequences corresponding to SNMV RNA 2 residues 131 to 216 was transcribed to produce a labelled complementary strand as described in the text. The cloned insert was excised, purified and the labelled transcript hybridized to intact VTMoV and SNMV RNA 2; this primer was elongated using reverse transcriptase in the presence of dideoxynucleotides. The transcripts were fractionated by 80 cm 6% polyacrylamide gel electophoresis and the tracks A, C, G and T correspond to the dideoxynucleotide species present. Note the sequence heterogeneity evident in VTMoV RNA 2 at residue 108 which results in band doubling, and in the termination of reverse transcription at residue 49.


T G C A
T G C A
49
76
77
125
126
SNMV RNA 2
VTMoV RNA 2

CHAPTER 3

CHRYSANTHEMUM STUNT VIROID

INTRODUCTION

Chrysanthemum stunt disease was first described by Dimock (1947) and was epidemic among cultivated Chrysanthemum morifolium varieties during 1945-1947 in the U.S.A. (Brierly and Smith, 1949). At that time the production and distribution of chrysanthemum varieties in the U.S.A. were highly centralized, allowing the rapid spread of the disease (Keller, 1953). Transmission of the chrysanthemum stunt disease by both grafting and sap inoculation (Brierly and Smith, 1949; Olson, 1949; Brierly and Smith, 1951) demonstrated its infectious nature and allowed the development of control measures to prevent accidental spread of disease (Keller, 1953). Diener and Lawson (1973) demonstrated that the causative agent of chrysanthemum stunt disease was a viroid, a low molecular weight RNA species with properties similar to, but distinct from potato spindle tuber viroid (PSTV). Chrysanthemum stunt viroid (CSV) was shown to possess a greater electrophoretic mobility than PSTV, and the host range of CSV was shown to be mostly confined to some plant species in the compositae family whereas PSTV will replicate in species from a number of plant families (Diener, 1979; Hollings and Stone, 1973). Ribonuclease fingerprinting has also shown that the primary sequence of CSV differs significantly from that of PSTV (Gross,

Domdey and Sänger, 1977). However, cDNA-RNA hybridization techniques have indicated that at least 20% of the PSTV primary sequence is common to that of CSV (Owens, Smith and Diener, 1978). RNA sequence determination studies were undertaken to obtain primary and predicted secondary structures of CSV and to compare these with the structures previously determined for PSTV (Gross et al., 1978).

MATERIALS

The CSV isolate used in this work was originally obtained from infected Chrysanthemum cultivars, Charm type, and was kindly provided by T.C. Lee, Adelaide Botanic Gardens, via Dr. R.I.B. Francki, University of Adelaide. CSV, purified from infected chrysanthemums as described by Palukaitis and Symons (1980), was kindly provided by Dr. P. Palukaitis.

METHODS

A. Primary structure determination

The primary sequence of CSV was determined using techniques described in Chapter 2. Linear viroid fragments were obtained by partial digestion of purified CSV under conditions of high salt and low temperature using 3750 units/ml RNase T_1, 2 µg/ml RNase A or 2 units/ml RNase U_2. The linear RNA fragments were

5 -^{32}P-radiolabelled using [γ-^{32}P]ATP and T_4 polynucleotide kinase, fractionated by denaturing polyacrylamide gel electrophoresis and sequenced using only the partial enzymic cleavage method, as described in Chapter 2. Some partial enzymic digests of viroid fragments were fractionated in polyacrylamide gels containing 98% formamide, as described in Chapter 2 A-8, in order to eliminate band compression artifacts (Kramer and Mills, 1978) arising from incomplete denaturation of the fragments during electrophoresis. The complete base sequence of CSV was assembled from sequence data of a large number of RNA fragments obtained by partial RNase digestion.

B. Secondary structure determination

A possible secondary structure model was constructed from the complete CSV sequence using the matrix method of Tinoco et al. (1971, 1973) and the predicted thermodynamic stability of the model was calculated using values provided by Dr. D. Reisner (Steger, Gross, Randles, Sanger and Reisner, in preparation).

RESULTS

A. Sequence determination

Purified circular CSV was subjected to partial

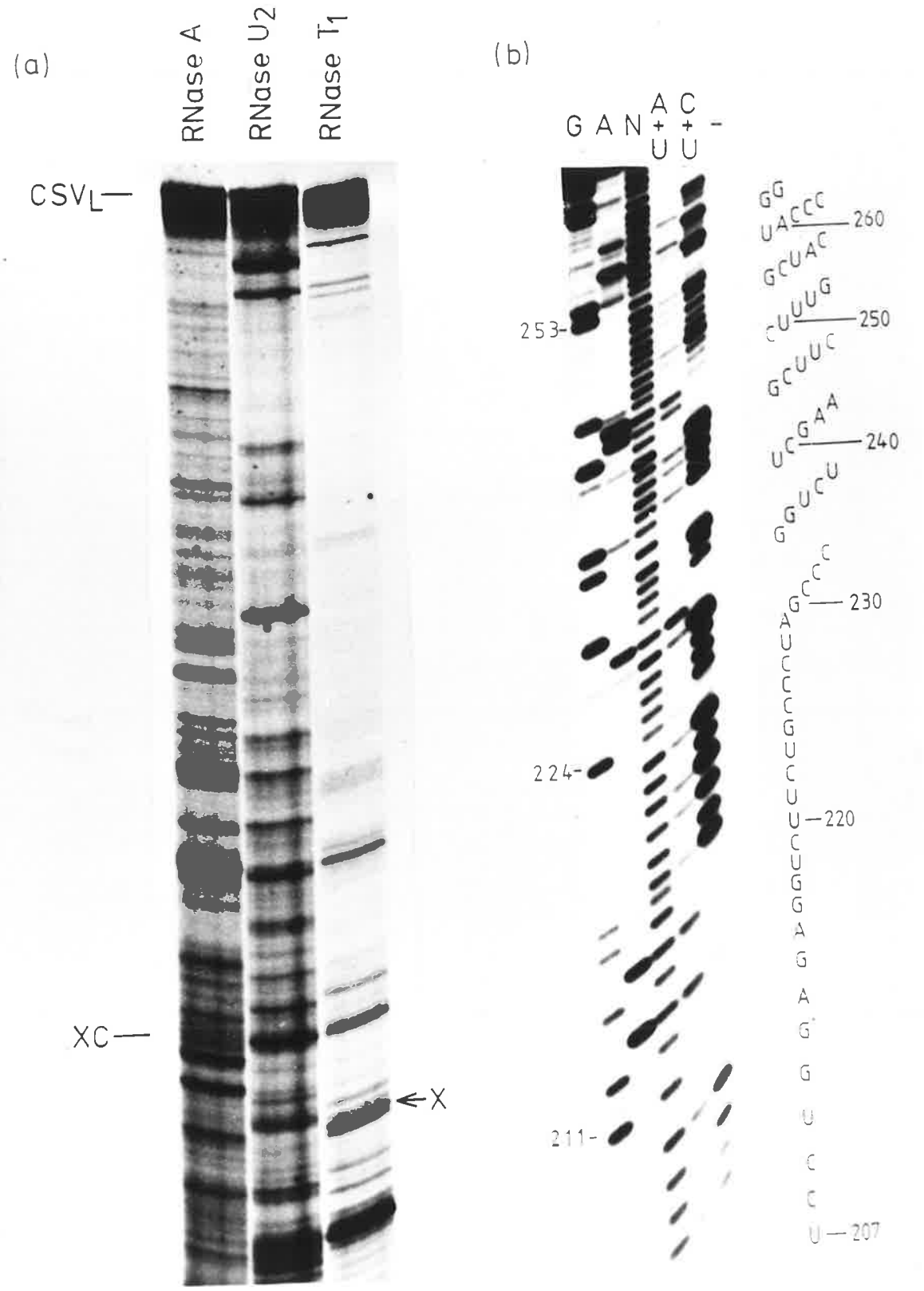
digestion with RNase T_1 U_2 or A under conditions of high salt concentration (600mM NaCl, 10mM $MgCl_2$ and at 0°C in order to limit cleavage by the single-strand specific RNases to relatively few accessible sites on the highly base-paired RNA molecule. The resulting viroid fragments were 5'-^{32}P-labelled *in vitro* using T_4 polynucleotide kinase and [γ-^{32}P]ATP and fractionated by size on a denaturing polyacrylamide gel. Figure 3-1a shows the gel patterns obtained for partial digestions of CSV with RNases T_1, U_2 and A. Digestion with either of the single base specific RNases T_1 (G specific) or U_2 (A specific) gave rise to fewer fragments than digestion with the C and U specific RNase A.

The gel fractionated 5'-labelled fragments obtained by partial RNase digestion of CSV were excised, eluted and sequenced using the partial enzymic cleavage method as described in Chapter 2 A-6. An example of one sequencing gel is given in Figure 3-1b. For some regions of the viroid molecule, sequencing was complicated by band compression (Kramer and Mills, 1978) due to the presence of stable base-paired hairpin structures. However, these band compressions could be eliminated by the use of sequencing gels which contained 98% formamide rather than 7M urea in order to ensure complete denaturation of the RNA fragments.

Figure 3-1 Purification and nucleotide sequence determination of CSV fragments.

(a) Autoradiogram of the 5'-^{32}P-labelled products of the partial digestion of CSV by RNases A, U_2 and T_1 after fractionation by electrophoresis on a polyacrylamide slab gel as described in Chapter 2 A-5. The largest radiolabelled fragment is full length linear CSV (CSV_L) which migrated about 30 cm from the origin. XC is the position of the xylene cyanol FF dye marker which corresponds to fragments about 80 residues long. A number of the shorter CSV fragments (including band X) were excised and eluted for sequencing by partial enzymic digestion.

(b) Autoradiogram of part of a sequencing gel (8% polyacrylamide) containing the various partial enzymic digests of fragment X. Digestions, as described in Chapter 2 A-6, were with RNase T_1 (G), RNase U_2 (A) alkali (N) to produce the reference ladder, RNase Phy M (A+U) and Bacillus cereus RNase (C+U). Part of the nucleotide sequence of fragment X from residues 207 to 265 is given.


(a)
RNase A
RNase U2
RNase T1
CSVL
XC
X
(b)
G A N A C
+ +
U U
253-
224-
211-
260
250
240
230
220
207

B. Primary Sequence and Secondary Structure of CSV

The complete base sequence of CSV was assembled from sequence data of a large number of RNA fragments obtained using the partial RNase digestion technique (Figure 3-2). The 356 residues are numbered according to the scheme of Gross et al. (1978) for PSTV and the main overlapping sequences used for the primary structure determination are shown in Figure 3-2.

A possible secondary structure model was constructed from the CSV sequence using the methods of Tinoco et al. (1971; 1973) and is compared with the published structure for PSTV (Gross et al., 1978) (Figure 3-3). The relative number of G.C base pairs in the predicted CSV structure (64 G.C, 44 A.U, 16 G.U) is lower than that of PSTV (73 G.C, 37 A.U, 16 G.U) and, using values kindly provided by Dr. D. Reisner (Steger, Gross, Randles, Sanger and Reisner, in preparation), the thermodynamic stabilities of the proposed models for CSV and PSTV were calculated to be ΔG (25°C, 1M NaCl) = -540 KJ/mol and -610 KJ/mol respectively.

DISCUSSION

A. Homology between CSV and PSTV

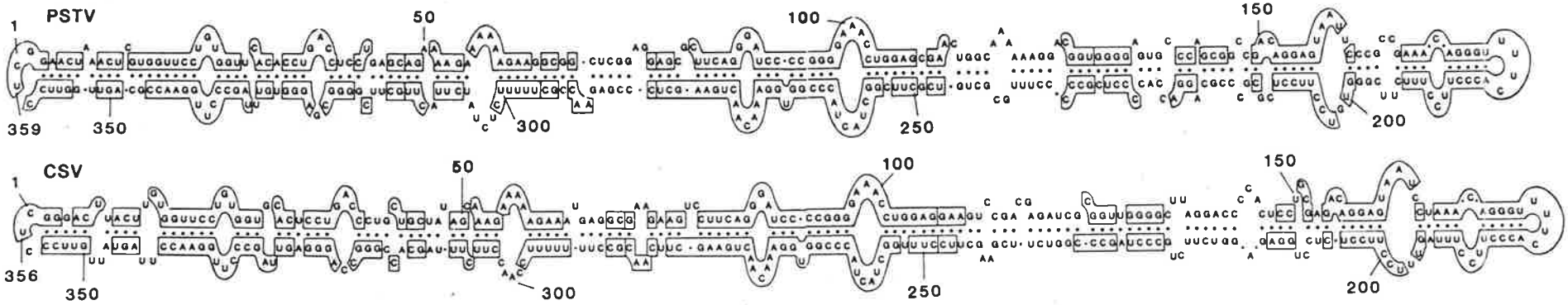
The striking feature of both the primary and postulated secondary structures of CSV is the extent of homology with the previously sequenced PSTV molecule

Figure 3-2 Primary sequence of CSV.

The sequence of the 356 residues of CSV is given and the residues numbered according to the published sequence of PSTV (Gross et al., 1978). The 247 residues homologous with PSTV are boxed. Within the circular sequence are given the locations of overlapping sequences obtained from RNase fragments of CSV; these sequences do not represent the entire length of these fragments. Each sequence is labelled with the RNase (A, T_1, or U_2) which gave the fragment from which that sequence was derived.

Figure 3-3 The predicted secondary Structures of CSV and PSTV (Gross et al., 1978).

The boxed areas contain residues homologous between the two viroids.


PSTV
1
50
100
150
359
350
300
250
200
CSV
1
50
100
150
356
350
300
250
200

(Gross et al., 1978). Of the 356 residues of CSV, 247 residues (69%) are homologous with those of PSTV, and occur in two main areas in the primary structure (Figure 3-2) extending from residues 247 to 110 and 148 to 206. These areas are separated by two regions of about 40 residues each containing only two small areas of homology. The postulated secondary structure model for CSV (Figure 3-3) shows that the two main areas of homology each correspond to one base-paired end of the native molecule. These are separated by the two regions of lesser homology which are positioned almost exactly opposite each other in the native molecule and are predominantly base-paired. Thus, the conservative arrangement and base-pairing of such non-conserved regions in the primary sequence allows the CSV molecule to form a stable secondary structure similar to that of PSTV.

B. Replication of CSV and PSTV

Although the host ranges of PSTV and our isolate of CSV differ significantly (see Introduction), they do overlap in such plant hosts as the composite Gynura aurantiaca (Diener,1979; Palukaitis and Symons, 1980; Niblett et al., 1980). It is feasible that replication of the two viroids in these plants will occur by similar mechanisms in view of their similarities in size and

sequence.

Assuming the existence of translatable linear forms of the RNAs (Kozak, 1979; Konarska *et al.*, 1981), the possible polypeptide products of both the viroid and putative complementary RNA strands of CSV and PSTV can be predicted from their known primary structures. Major differences are found between the possible polypeptide translation products of CSV and PSTV (see Chapter 7), suggesting that neither viroid codes for proteins involved in their replication. This is consistent with the lack of evidence for any viroid translation *in vivo* (Conjero and Semancik, 1977) and *in vitro* (Davies *et al.*, 1974; Semancik *et al.*, 1977).

In contrast, the overall secondary structures of CSV and PSTV are conserved despite differences in sequence. Given the lack of evidence for functional viroid-coded translation products, the replication of CSV and PSTV may involve recognition by host enzymes which are capable of RNA-dependent RNA synthesis. Thus, the sequence and structural features common to both CSV and PSTV may play a role in such recognition processes. An example of such a conserved feature is situated at the centre of the native molecules (Fig. 3-3) (CSV residues 74-110, 247-284; PSTV residues 76-112, 247-284) and includes two relatively large single-stranded regions which are completely conserved between CSV and PSTV.

C. Relationship of this isolate of CSV to other viroid isolates

Both Owens et al. (1978) and Gross et al. (1977; 1982) have used isolates of CSV, of different origins from ours, for comparative studies with a PSTV isolate obtained from Dr. T. O. Diener. The sequence of this PSTV isolate has been determined (Gross et al., 1978).

Owens et al. (1978) used DNA complementary to PSTV to show that by hybridization analysis, their isolate of CSV contained about 20% sequence homology with PSTV, whereas a viroid isolated from Columnea erythrophae contained about 40% sequence homology. In addition, their CSV isolate and the Columnea viroid had electrophoretic mobilities in a non-denaturing polyacrylamide gel which were markedly faster than that of PSTV, indicating appreciably different sizes and/or secondary structures. In contrast, our isolate of CSV shares 69% sequence homology with PSTV, is only 3 residues shorter, and possesses a similar secondary structure. The data suggest that, even allowing for some errors in estimates of sequence homology determined by cDNA-RNA hybridization analysis, our isolate of CSV may be more closely related to PSTV in size and sequence than are the Owens et al. (1978) isolates of CSV and

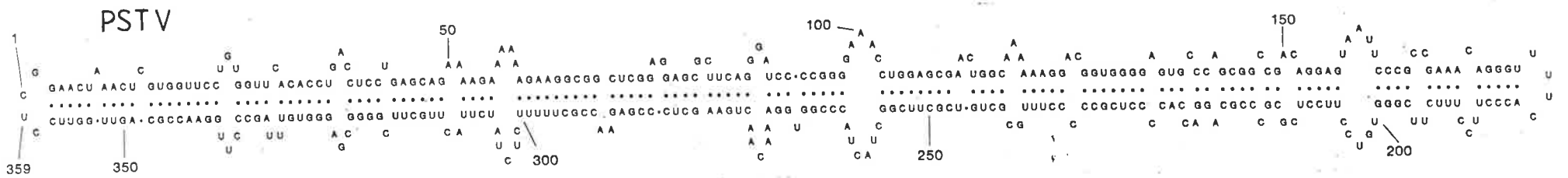
Columnea viroid.

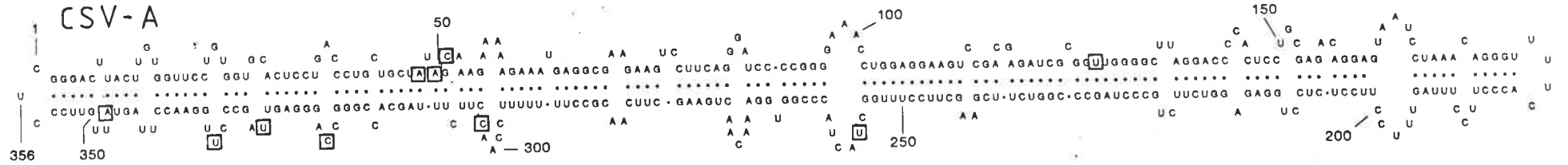
Gross *et al*. (1977) have compared the oligonucleotide fingerprints of an isolate of CSV, obtained from Dr. M. Hollings, with those of PSTV and obtained distinctly different patterns. It was concluded that CSV and PSTV differ significantly in sequence. Subsequent sequence determination of that isolate of CSV, while confirming its non-identity with PSTV, showed 73% sequence homology with PSTV (Gross *et al*., 1982). Figure 3-4 shows the predicted structures of the two isolates of CSV, together with that of PSTV. It can be seen that both isolates of CSV are closely related, sharing 346 residues in common, while the CSV isolate of Gross *et al*. (1978) is only 2 residues smaller in size. The few sequence differences between the two isolates are mainly located in the left hand sides (as they are drawn) of the native molecules. It is unknown whether the differences in structure between these CSV isolates corresponds to differences in biological properties.

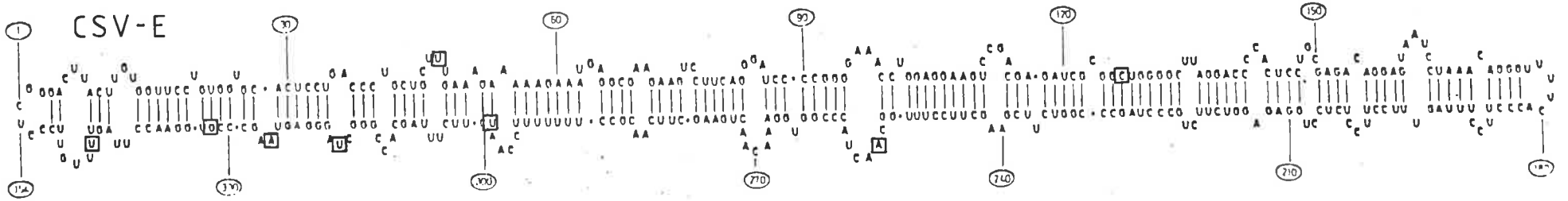
Three independent isolates of CSV have been characterized and it seems likely that all differ. These three CSV isolates, together with the *Columnea* viroid, are closely related to PSTV. It therefore appears likely that there exists a group of viroids,

Figure 3-4 Comparison of the secondary structures of two sequenced isolates of CSV.

The secondary structure models for the Australian CSV isolate (CSV-A; this work) and English CSV isolate (CSV-E; Gross et al., 1982) are presented. The two viroid isolates share over 97% base sequence homology, and the few residues not shared by both isolates are shown boxed.


PSTV
1
50
100
150
200
250
300
350
359


CSV-A
1
50
100
150
200
250
300
350
356


CSV-E

including PSTV, which share common sequences and possibly secondary structures, and which may be derived from a common ancestral viroid.

CHAPTER 4

COCONUT CADANG-CADANG VIROID

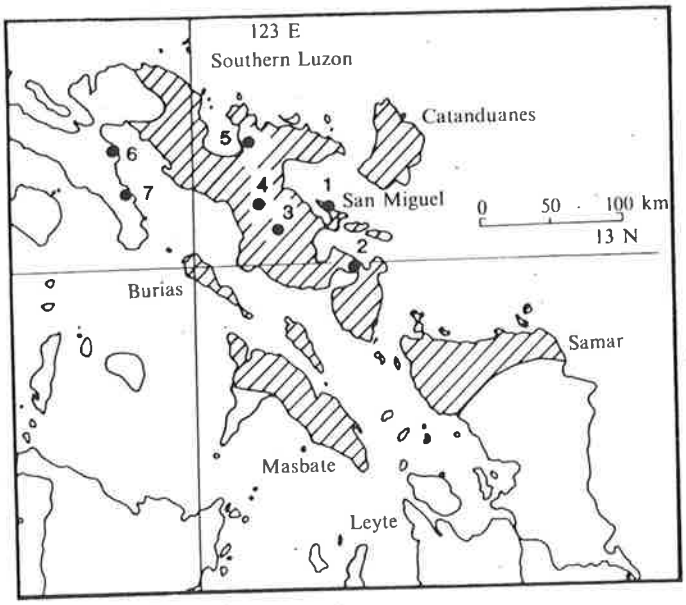
INTRODUCTION

Cadang-cadang is a serious and economically important disease of coconut palms which was first reported in 1927 on San Miguel Island in the Philippines (Ocfemia, 1937). In the following years, incidence of cadang-cadang disease was reported in surrounding areas of the Philippines and now, 56 years after its first known occurrence, the disease is found widespread over the south-east part of Luzon and many neighbouring islands (Figure 4-1). By 1962, only 100 of the 250,000 coconut palms on San Miguel Island had survived the disease (Bigornia, 1977). It is estimated that cadang-cadang disease is still responsible for the death of 500,000 palms each year in the Philippines (B. Zelazny, personal communication). Recent work has indicated that tinangaja disease of coconuts on Guam, an island 1,500 miles east of the Philippines, has the same aetiology as cadang-cadang disease (Boccardo *et al.*, 1981). However, cadang-cadang disease has not been found in any other coconut growing area.

The first symptom of the cadang-cadang disease is the development on the affected palm of small, rounded coconuts which are distinctively scarified. Later, the fronds develop characteristic yellow spots and as the disease advances the crown of

Figure 4-1 Incidence of cadang-cadang disease of coconuts.

Regions of the Philippine islands where the disease is widespread are shown cross-hatched, while isolated incidences of the disease occur in surrounding areas (Zelazny, 1979). The different ccRNA isolates used in this work were each obtained from separate diseased coconut palms in one of the following locations : 1, San Miguel Island; 2, Sorsogon; 3, Ligao; 4, Lake Baao; 5, Tinambac; 6, Guinayangan; and 7, San Nasciso.


123 E
Southern Luzon
Catanduanes
San Miguel
0 50 100 km
13 N
Burias
Samar
Masbate
Leyte
1
2
3
4
5
6
7

the palm is reduced to a tuft of short yellow fronds. The course of the disease invariably ends with the death of the infected palm, which occurs from 3 years to more than 15 years (usually about 10 years) after the appearance of the first symptoms (Zelazny and Niven, 1980).

Little was known of the nature of the pathogenic agent responsible for the cadang-cadang disease until Randles (1975) showed the existence of two RNA species that were present only in diseased palms. The RNA species were named ccRNA 1 and ccRNA 2 in order of increasing size, and were shown to share thermal denaturation, nuclease sensitivity, centrifugation and electrophoretic properties with viroids (Randles, 1975; Randles et al., 1976). The ccRNAs were also subsequently shown to be circular (Randles and Hatta, 1979), like viroids. Randles and Palukaitis (1979), using cDNA-RNA hybridization techniques, demonstrated that ccRNA 1 and ccRNA 2 shared common sequences and that the sequences of the ccRNAs were not found in healthy palms.

Recently it has been shown that ccRNA 1 and ccRNA 2 occur as fast and slow electrophoretic variants and that the occurrence of the variants is related to the stage of disease development in the coconut palms (Imperial et al., 1981). The fast electrophoretic

variants, ccRNA 1 fast and ccRNA 2 fast are present in infected palms at early stages of the disease, and as the disease progresses over a period of years the ccRNA 1 slow and ccRNA 2 slow variants first appear and then predominate (Imperial et al., 1981; Mohamed et al., 1982).

All four ccRNA species have been recently shown to be infectious (Mohamed and Imperial, unpublished results) and, as the ccRNAs are single-strand covalently closed circular RNAs with high degrees of secondary structure (Randles et al., 1976; Randles and Hatta, 1979), they possess both biological and physical properties similar to those of viroids. In order to further investigate the intriguing relationships between the cadang-cadang disease, the variant ccRNAs and viroids, the sequences and structures of the different ccRNAs were determined and compared.

METHODS

A. Isolation of the ccRNAs

Purified ccRNAs were kindly provided by Dr. Nizar A. Mohamed. Fronds were harvested from naturally infected coconut palms from a number of sites in the Philippines. Nucleic acids were extracted from the leaf tissue as described by Imperial et al. (1981) using their Method 1. Individual ccRNAs were purified by 3

cycles of polyacrylamide gel electrophoresis (Imperial et al., 1981).

B. Sizing of the ccRNAs

Sizes of the ccRNAs were estimated by electrophoresis in 6% polyacrylamide gels (40x20x0.05 cm) containing 98% formamide (Maniatis and Efstradiatis, 1980) The following were used as molecular weight markers - solanum nodiflorum mottle virus (SNMV) RNA 2, 377 residues (Haseloff and Symons, 1982); velvet tobacco mottle virus (VTMoV) RNA 2, 365 residues (Haseloff and Symons, 1982); chrysanthemum stunt viroid (CSV), 356 residues (Haseloff and Symons, 1981); Q strain of cucumber mosaic virus (CMV) RNA 4, 1027 residues (Gould and Symons,1982); CMV satellite RNA, 336 residues (Gordon and Symons, 1983); chicken 18S rRNA, 1800 residues (Spohr et al., 1976); alfalfa mosaic virus (AMV) RNA 4, 881 residues (Brederode et al., 1980); yeast 5.8S RNA, 158 residues (Rubin, 1973); yeast 5S RNA, 121 residues (Miyazaki, 1974); Escherichia coli phenylalanine tRNA, 76 nucleotides (Barrell and Sanger, 1969).

The circular RNAs (SNMV RNA 2, VTMoV RNA 2, CSV and the ccRNAs were boiled for 15 minutes in distilled water before electrophoresis to produce the linear forms. After electrophoresis, gels were stained

with 0.01 % toluidine blue and destained with water.

C. Fingerprinting of the ccRNAs

Purified ccRNAs (0.5 µg) were dried down, resuspended in 5 µl 5mM Tris-HCl pH 7.5, and digested with 0.1 µg RNase A at 37°C for 1 hour or with 20 units RNase T_1 at 56°C for 30 minutes. The resultant oligonucleotide fragments were transferred to another tube containing 5 µCi dried down [γ-^{32}P]ATP (2000 Ci/mmol) and 1.5 µl of 5x T_4 polynucleotide kinase buffer (250mM Tris-HCl pH 9.0, 50mM $MgCl_2$, 50mM dithiothreitol), and 0.5 µl (0.5 units) T_4 polynucleotide kinase added. The reaction was incubated at 37°C for 20 minutes and 5 µl formamide loading buffer (95% (v/v) deionized formamide, 10mM Na_2EDTA, 0.02% (w/v) bromophenol blue, 0.02% (w/v) xylene cyanol FF) added.

Radiolabelled oligonucleotides were fractionated by two dimensional polyacrylamide gel electrophoresis. For the first dimension, preparations were electrophoresed in 40x20x0.05 cm 10% polyacrylamide gels containing 25mM sodium citrate pH 3.5 (DeWachter and Fiers, 1972; Frisby, 1977). After the xylene cyanol FF dye marker had migrated 14 cm, electrophoresis was stopped, and gel strips were excised and embedded at the bottom of 40x20x0.05 cm 25% polyacrylamide gels

containing 89mM Tris-borate pH 8.3, 2mM Na_2EDTA (Frisby, 1977). Polymerization of the second dimension gels was catalysed by the addition of 300 µl 10% (w/v) ammonium persulphate, 30 µl TEMED, 50 µl 10% (w/v) ascorbic acid and 70 µl 30% (w/v) H_2O_2 per 50 ml of gel solution to ensure complete polymerization in the region of the first dimension gel strip. Samples were electrophoresed upwards until the bromophenol blue dye marker had migrated 18 cm, and the fractionated ^{32}P-labeled oligonucleotides were detected by autoradiography.

D. Sequence and structure determination of the ccRNAs

Purified ccRNAs were sequenced essentially as described in Chapter 2. Purified ccRNAs (5 µg) isolated from single infected coconut palms were subjected to limited digestion by 3 units/ml RNase U_2, 5 ng/ml RNase A or 2000 units/ml RNase T_1 under non-denaturing conditions (600mM NaCl, 10mM $MgCl_2$ at 0°C). The resulting linear RNA fragments were 5'-radiolabelled using [γ-^{32}P]ATP and T_4 polynucleotide kinase or, after treatment with calf intestinal phosphatase, 3'-radiolabelled using [5'-^{32}P]dpCp and T_4 RNA ligase, and fractionated by denaturing polyacrylamide gel electrophoresis as detailed in Chapter 2 A-3,4,5. Radiolabelled fragments were located by autoradiography, excised, eluted, and sequenced using partial enzymic

cleavage methods. The sequences of numerous overlapping fragments were assembled to give the complete primary structure of each circular molecule.

Secondary structures of the ccRNAs were mapped using S_1 nuclease (Wurst et al., 1978). Full length 5'- or 3'-radiolabelled ccRNAs, obtained as described above by RNase T_1 digestion, were suspended in 20 µl 200mM NaCl, 0.05mM $ZnSO_4$, 50mM sodium acetate pH 4.6, containing 5 µg Escherichia coli tRNA carrier, and incubated at 37°C for 10 minutes with 0.1, 1 or 10 units of S_1 nuclease (Boehringer). The reaction mixtures were extracted with phenol, precipitated with ethanol and fractionated by electrophoresis in a polyacrylamide gel containing 8M urea and TBE buffer (90mM Tris-borate pH 8.3, 2mM Na_2EDTA). Products of partial enzymic sequencing reactions of the same ccRNA species were run as markers, thus allowing sites of S1 nuclease sensitivity to be located. Data so derived were used during construction of the secondary structure models of ccRNA 1 fast and ccRNA 1 slow.

Secondary structure models for the native ccRNAs were constructed using the base pairing matrix procedure of Tinoco et al. (1971; 1973) and the thermodynamic stabilities of the predicted RNA structures were calculated using values kindly provided by Dr. D. Reisner (G. Steger, H.J. Gross, J.W. Randles,

H.L. Sanger and D. Reisner, unpublished results).

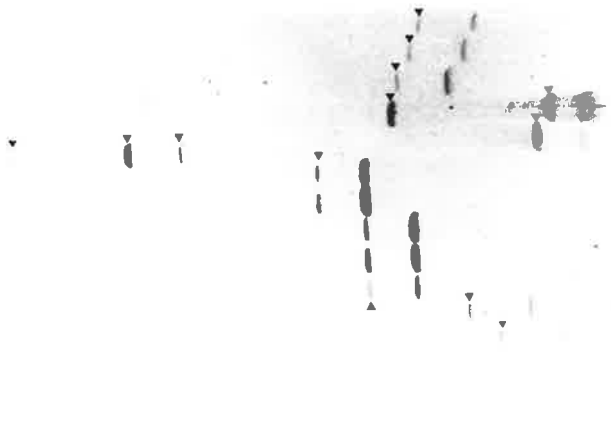
RESULTS and DISCUSSION

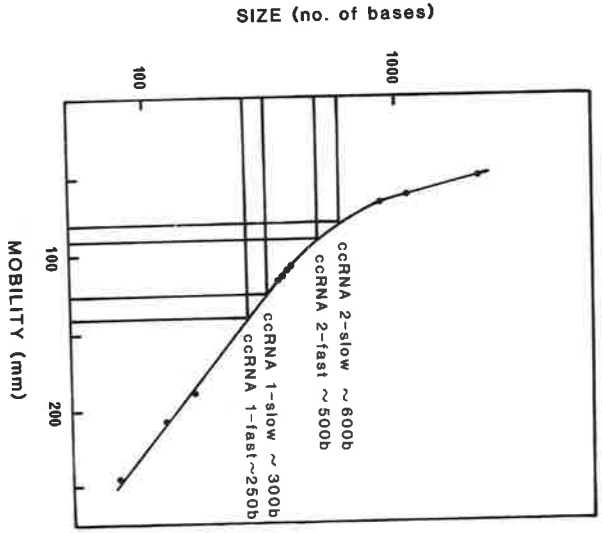
A. Sizing of the ccRNAs

The sizes of the linear fast and slow variants of ccRNA 1 and ccRNA 2 were estimated by electrophoresis in polyacrylamide gels containing 98% formamide, using linear RNAs of known molecular weight as markers (Figure 4-2). No differences in mobility were observed between the different ccRNA isolates. The sizes of the RNAs were estimated to be : ccRNA 1 fast, 250 residues; ccRNA 1 slow, 300 residues; ccRNA 2 fast, 500 residues; ccRNA 2 slow, 600 residues. Therefore the ccRNA 2 fast and ccRNA 2 slow forms are approximately twice the size of the corresponding ccRNA 1 forms and the slow forms of both ccRNA 1 and ccRNA 2 are 20% larger than the corresponding fast forms. These estimates of size differ from those of ccRNA 1 (310±3 residues) and ccRNA 2 (438±5 residues) as obtained by length measurements made after electron microscopy (Randles and Hatta, 1979). However, Randles and Hatta (1979) did not differentiate between the fast and slow ccRNA variants which may, together with perhaps under-estimated experimental error, have resulted in the differences observed.

Figure 4-2 Size estimation of the fast and slow forms of ccRNA 1 and ccRNA 2 by electrophoresis in a 6% polyacrylamide gel containing 98% formamide.

14B, 620C and X2 were different isolates of ccRNA obtained from Ligao, Philippines. Sizes of the ccRNAs were determined from a standard curve of mobilities of the RNA markers (described in Methods) plotted against their known sizes on a logarithmic scale. Only the linear RNAs, not circular RNAs were used for size estimation. For SNMV RNA 2, VTMoV RNA 2, CSV, and the ccRNAs, the linear forms are indicated by arrows. For CMV plus satellite (sat) RNA, the bands are (in order from the top) RNAs 1, 2, and 3 running as a broad band, RNA 4 and satellite RNA. Marker RNAs, in order from the top, are chicken 18S, AMV RNA 4, yeast 5.8S RNA, yeast 5S RNA and E.coli phenylalanine tRNA.


SNMV RNA 2
VTMoV RNA 2
CSV
CMV + sat
MARKERS
14B
620C
ccRNA 1-fast
14B
620C
X2
ccRNA 1-slow
620C
ccRNA 2-fast
X2
ccRNA 2-slow
– ORIGIN


SIZE (no. of bases)
1000
100
100
200
MOBILITY (mm)
ccRNA 2-slow ~ 600b
ccRNA 2-fast ~ 500b
ccRNA 1-slow ~ 300b
ccRNA 1-fast ~ 250b

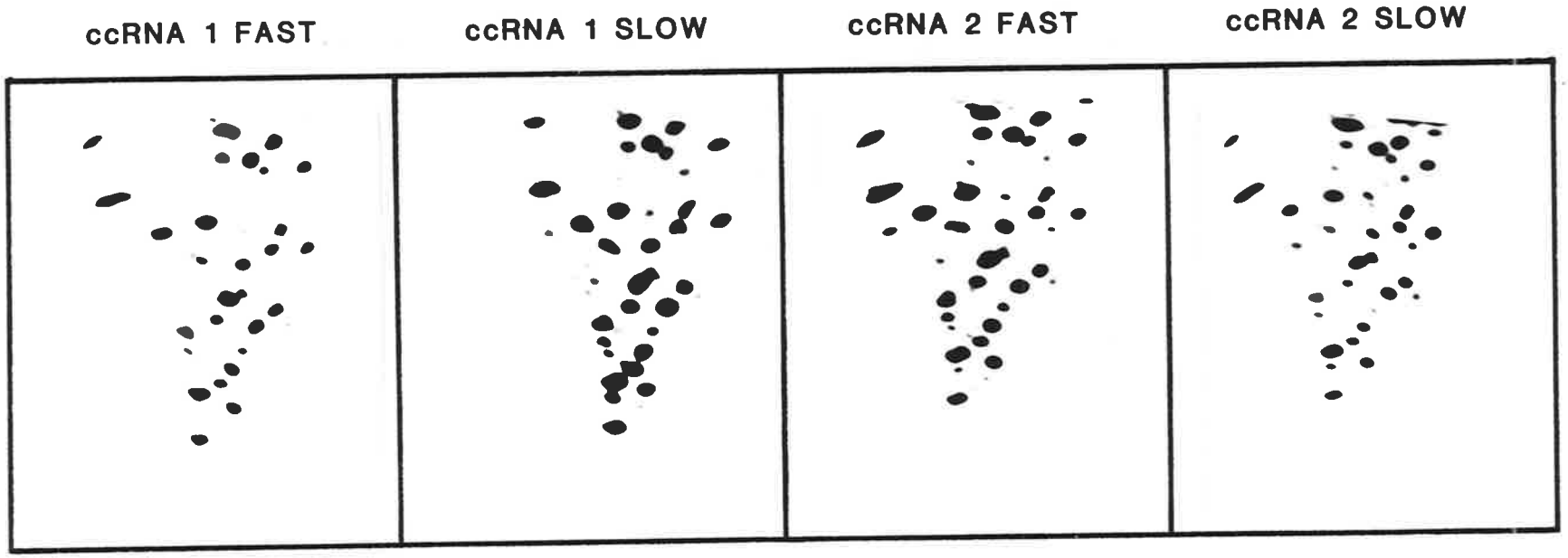
B. Fingerprinting of the ccRNAs

The sequence relationships between the four ccRNAs were further investigated using RNase A and RNase T_1 fingerprints. The individual purified circular ccRNAs isolated from a single infected palm were digested to completion with RNase A or RNase T_1, 5'-^{32}P radiolabelled, and the resultant oligonucleotides fractionated by two-dimensional gel electrophoresis. The RNase A fingerprints of the four forms of ccRNA extracted from the same tree (isolate Ligao T_1) show essentially identical patterns of labelled oligonucleotides (Figure 4-3). This indicates that the three larger ccRNAs contain the same sequences as the smallest ccRNA 1 fast. Similarly, essentially identical patterns of radiolabelled oligonucleotides were obtained after digestion with RNase T_1 (data not shown), however three extra oligonucleotides were found in the RNase T_1 fingerprints of the fast ccRNA forms, which were not in the fingerprints of the slow ccRNA forms; the significance of this is not known, but may be related either to sequence heterogeneity observed in ccRNA fast forms (see below) or to difficulty experienced in ensuring the RNase T_1, in contrast to RNase A digestions, always go to completion.

The ccRNA 2 fast and slow forms are estimated to be twice the sizes of their respective ccRNA 1 forms

Figure 4-3 RNase A fingerprints of ccRNAs.

The fast and slow forms of ccRNA 1 and ccRNA 2 were digested with RNase A, 5'-^{32}P-labelled and fractionated by two dimensional gel electrophoresis. Migration in the first dimension is from left to right and in the second dimension from bottom to top.


ccRNA 1 FAST
ccRNA 1 SLOW
ccRNA 2 FAST
ccRNA 2 SLOW

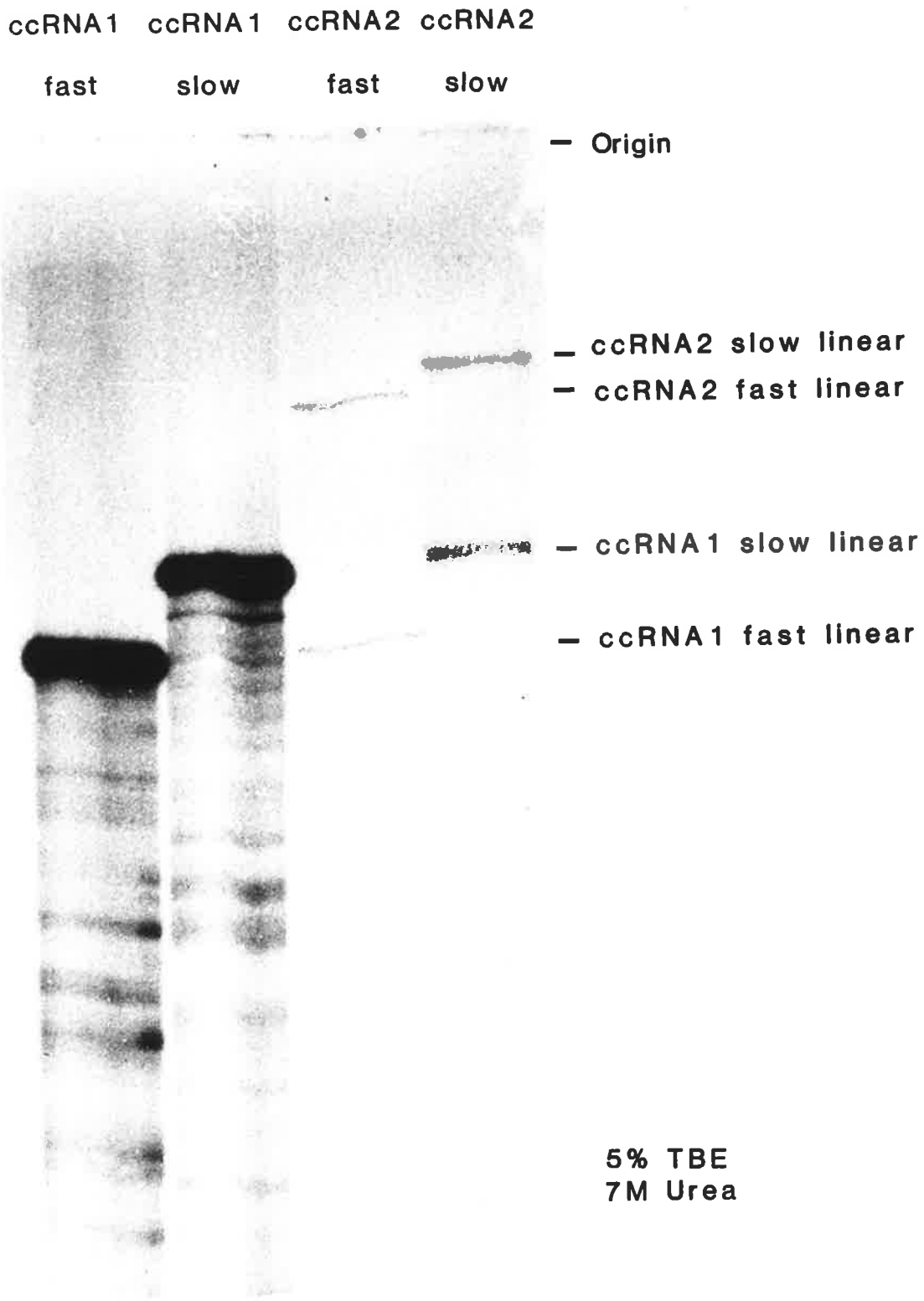
(Figure 4-2) and, since the RNase fingerprints show that ccRNA 1 and ccRNA 2 possess similar sequence complexity, it is possible that the ccRNA 2 fast and slow forms are dimers of the ccRNA 1 fast and slow forms respectively. Furthermore, it is likely that each ccRNA slow speices contains only repeated sequences of the respective ccRNA fast species. In order to extend these observations, the sequences of the ccRNAs were determined.

C. Sequences and structure determination of the ccRNAs

Native circular ccRNAs were subjected to limited digestion either by RNase T_1, which catalysed cleavage of ccRNA 1 species at single sites and ccRNA 2 species at either or both of two sites to produce specific full length linear ccRNAs (Figure 4-4), or by ribonucleases A or U_2, which produced smaller overlapping RNA fragments. These linear RNA molecules were 5'- or 3'- radiolabelled and then purified by polyacrylamide gel electrophoresis. The sequences of these 5'- or 3'- labelled fragments were determined by the partial enzymic digestion technique. The use of fragments labelled separately at both the 5'- and 3'- ends allowed the sequence determination of long RNA fragments up to 574 residues long and, with shorter fragments, resolved gel compression artefacts (Kramer and Mills, 1978) when the relevant nucleotide sequences

Figure 4-4 Partial RNase T_1 digest of ccRNAs.

Purified ccRNAs (isolate Ligao T_1) were digested with 2000 units/ml RNase T_1 under conditions of high salt and low temperature as described in the text. Resulting linear RNAs were 5'-^{32}P-labelled and fractionated on a 5% polyacrylamide gel containing 7M urea. Digestion of ccRNA 2 species gave rise to linear RNAs corresponding in sizes to those of the full-length ccRNAs 2 and ccRNAs 1.

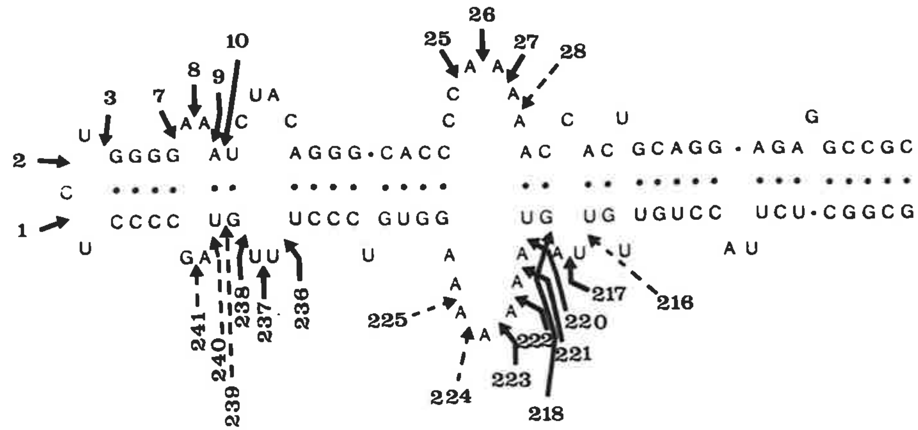

ccRNA1 fast
ccRNA1 slow
ccRNA2 fast
ccRNA2 slow
Origin
ccRNA2 slow linear
ccRNA2 fast linear
ccRNA1 slow linear
ccRNA1 fast linear
5% TBE
7M Urea

were determined from both directions. The sequences of overlapping fragments were assembled to construct the complete primary structures of the circular RNA molecules.

Secondary structure models for the native ccRNAs were constructed using the base-pairing matrix procedure of Tinoco et al. (1971), values for the thermodynamic stability of the predicted RNA structures (G. Steger, H.J. Gross, J.W. Randles, H.L. Sanger and D. Reisner, unpublished results), and experimental evidence for the location of ribonuclease sensitive single stranded regions on the native molecules. In addition, specific full-length linear ccRNAs, produced by limited ribonuclease T_1 cleavage, were either 5'- or 3'-^{32}P labelled and the susceptible single strand regions in the native structures located by the S1 nuclease mapping procedure (Wurst et al., 1978). The sites of cleavage were determined by co-electrophoresis of the radiolabelled fragments of the S1 nuclease digest with products of sequencing reactions using the partial enzymic digestion procedure (Figure 4-5). Thus the possibility of specifically linearizing the circular ccRNAs by limited cleavage with RNase T_1 both facilitated sequence determination of the molecules and allowed S1 nuclease mapping of the renatured linearized molecules.

Figure 4-5 Partial enzymic digestion sequencing and S1 nuclease mapping of the ccRNAs.

Full-length linear isolate Baao 54 ccRNA 1 fast was produced by partial RNase T_1 digestion under non-denaturing conditions, 5'-^{32}P-labelled and purified by polyacrylamide gel eletrophoresis as described in the text. Purified radiolabelled RNA was subjected to treatment by no enzyme (N), RNase T_1 (T), RNase U_2 (U), alkali (L), RNase PhyM (P), Bacillus cereus RNase (B) and 10 units nuclease S1 (S). The products were fractionated by 80 cm 6% polyacrylamide gel electroporesis, and cleavage sites for nuclease S1 are shown arrowed on the predicted secondary structure model of ccRNA 1 fast.


GGGG AU AGGG·CACC AC AC GCAGG·AGA GCCGC
CCCC UG UCCC GUGG UG UG UGUCC UCU·CGGCG
1 2 3 7 8 9 10 25 26 27 28
216 217 218 220 221 222 223 224 225 236 237 238 239 240 241


N T U L P B S1
28 27 26 25
10 9 8 7 3 2 1
241 240 239 238 237 236
225 224 223 222 221 220
218 217 216

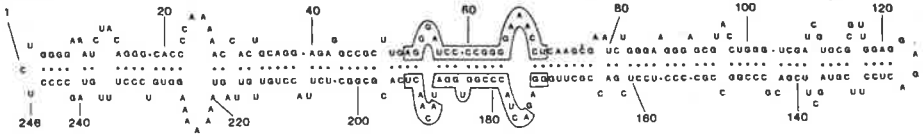
D. ccRNAs differ in size but not sequence complexity.

The sequences and predicted structures of the ccRNA 1 fast and slow forms isolated from a single infected coconut palm (isolate Baao 54, Figure 4-1) are shown in Figure 4-6 together with the known structures of four viroids, PSTV (Gross et al., 1978), CSV (Haseloff and Symons, 1981), CEV (Visvader et al., 1982) and ASBV (Symons, 1981). The two native ccRNAs 1 possess extensive regions of intramolecular base pairing and can form rod-like native structures similar to other viroids. The ccRNA 1 fast and ccRNA 1 slow possess 246 and 287 residues respectively, and have calculated thermodynamic stabilities of -320 and -360 $KJmol^{-1}$ respectively. The ccRNA 1 slow contains the entire sequence and structure of the smaller ccRNA 1 fast but differs by an additional duplicated sequence and structure of 41 residues (residues 103-143 in ccRNA 1 fast) which is added at the right-hand end of the native molecule between residues 123 and 124 of ccRNA 1 fast (Figure 4-6). Thus , the rod-like, base-paired native structure is maintained in the larger molecule.

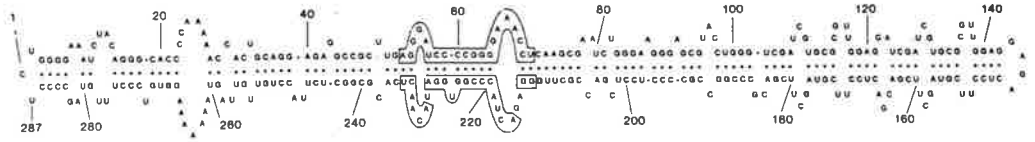
The nucleotide sequences of ccRNA 2 fast and ccRNA 2 slow, consisting of 492 and 574 residues respectively, are perfect dimers of the respective ccRNA 1 forms. A schematic summary of the relationships

Figure 4-6 Sequences and predicted secondary structures of the Baao 54 isolate of ccRNA 1 fast and ccRNA 1 slow are shown with those of PSTV, CSV, CEV and ASBV. The structures are aligned under the central conserved regions of these viroids (boxed).

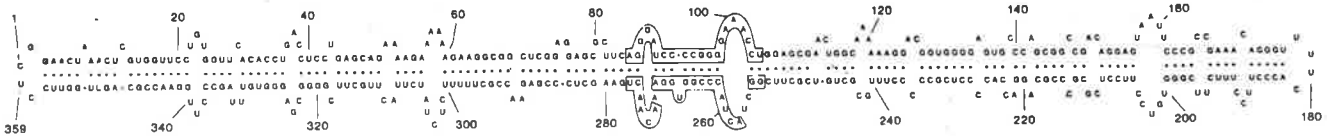
Cadang–cadang RNA 1 fast



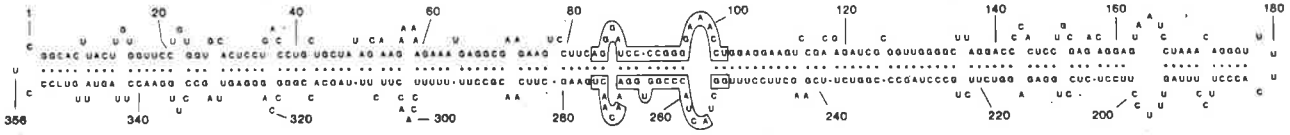
Cadang–cadang RNA 1 slow



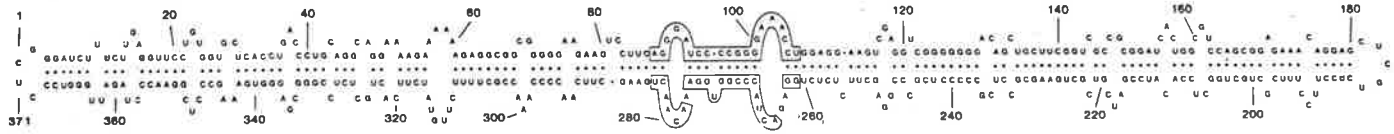
PSTV



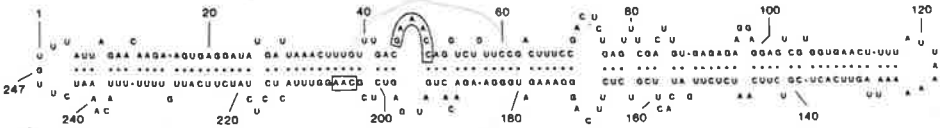
CSV



CEV



ASBV

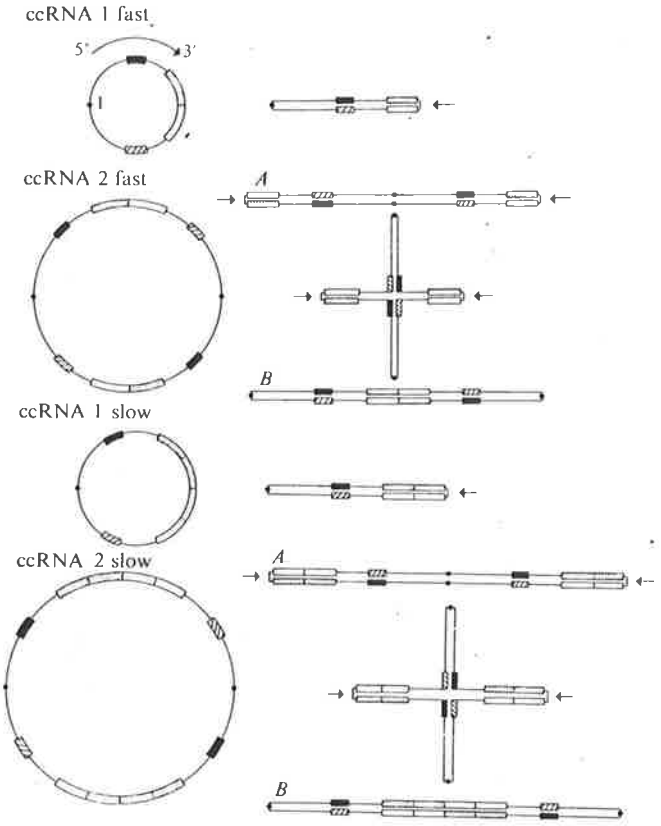


between the primary structures of all four ccRNAs and their predicted secondary structures is given in Figure 4-7. While each of the monomeric ccRNA 1 forms can base pair intramolecularly to form a single rod-like conformer, the ccRNA 2 forms, due to their dimeric nature, can each form either of two rod-like conformers (A or B, Figure 4-7) and a large number of intermediate cruciform-shaped structures, one of which is given in Figure 4-7.

The ccRNA 1 fast and ccRNA 1 slow molecules each possess, under experiemental conditions, a highly accessible site for cleavage by RNase T_1 at the right-hand terminal hairpin loop of the predicted native structure (between residues 124 and 125 in Baao 54 ccRNA 1 fast and between residues 145 and 146 in Baao 54 ccRNA 1 slow). Limited RNase T_1 digestion of each ccRNA 2 species produced specific linear RNA fragments corresponding to both the respective full-length linear ccRNA 2 and ccRNA 1 molecules. Sequence determination of these fragments showed that cleavage of the ccRNA 2 molecules occurred at two sites located at the same sequences as for the two ccRNA 1 molecules. This suggests that either the predicted conformer A of the two ccRNA 2 molecules or possible cruciform intermediates exist in solution, whereby the appropriate terminal hairpin loops are exposed. However, the

Figure 4-7 Schematic representation of the sequences and predicted structure relationships between the ccRNAs.

The circular sequences of the four ccRNAs are shown with black, and cross-hatched boxed regions representing the sequences highly conserved between ccRNAs, PSTV, CSV and CEV. The white, and stippled boxed regions represent those sequences duplicated within the ccRNA 1 slow species. Positions corresponding to residue 1 of ccRNA fast or ccRNA 1 slow are indicated by black dots. Both ccRNAs 2 are dimers of the respective ccRNA 1 forms and can potentially form either of two rod-like conformers A or B as well as a large number of cruciform-shaped intermediates, of which one is shown. Each ccRNA 1 species possesses a single, highly accessible site for RNase T_1 cleavage located on a terminal hairpin loop; these sites are indicated by arrows. Each ccRNA 2 species possesses two such accessible sites for RNase T_1 cleavage and arrows indicate where these sites also occur on hairpin loops in the different ccRNA 2 conformers.


ccRNA 1 fast
5'
3'
1
ccRNA 2 fast
A
B
ccRNA 1 slow
ccRNA 2 slow
A
B

existence of type B conformers cannot be precluded.

E. Variation in sequence between different ccRNA isolates

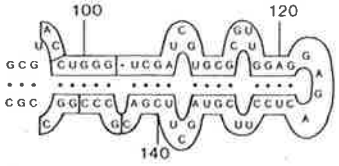
Different isolates of cadang-cadang RNAs were each obtained from single infected coconut palms from different localities in the Philippines (figure 4-1). Sequence differences between the isolates consist of two types. First, the sequences of the ccRNA 1 slow forms can differ. While all ccRNA 1 fast forms are essentially identical (see below), ccRNA 1 slow forms can differ in the length of the repeated sequence inserted between residues 123 and 124 of ccRNA 1 fast. Three different repeated sequences found in nine sequenced isolates of ccRNA 1 slow are given in Figure 4-8; these vary in length from 41 to 55 residues but are all internally base-paired to produce duplicated structures as well as sequences at the right-hand ends of the native molecules. Interestingly, the right-hand ends of the molecules of PSTV, CSV, CEV and ASBV (Figure 4-6) are similarly distanced from the central conserved regions of these molecules. In contrast, the right-hand side of the ccRNA 1 fast molecule is shorter while those of the elongated ccRNA 1 slow molecules are closer in size to those of PSTV, CSV, CEV and ASBV.

Second, four of the six isolates of ccRNA 1

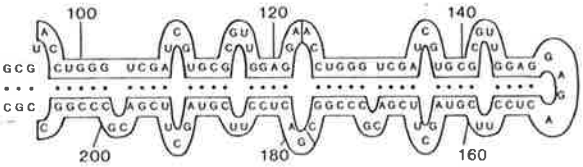
Figure 4-8 Sequence variation between ccRNA 1 slow of three ccRNA isolates.

The sequences and structures of various ccRNA 1 fast and ccRNA 1 slow isolates were determined as described in the text. As essentially all sequence variation occurred at the right hand end of the ccRNA 1 slow molecules, only this region is shown. Boxed regions represent those sequences which are duplicated in the ccRNA 1 slow molecules and which are 41 (isolate Baao 54), 50 (isolate Ligao 14B) or 55 residues (isolate Ligao T1) long. All sequenced ccRNA 1 slow isolates correspond to one of these forms (Table 4-1).

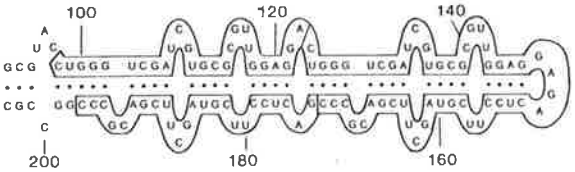
ccRNA 1 fast



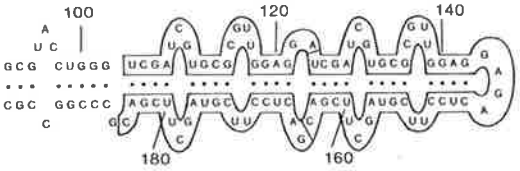
ccRNA 1 slow Isolate Ligao T_1



ccRNA 1 slow Isolate Ligao 14B



ccRNA 1 slow Isolate Baao 54



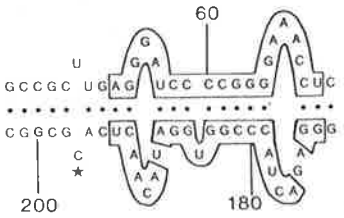
fast sequenced each consist of two populations of molecules, one of 246 residues and the other of 247 residues, which differ in the presence or absence of a C at residue 198 (Figure 4-9, Table 4-1). Similar sequence heterogeneities have also been reported for CEV (Gross et al., 1982) and the viroid-like RNA of VTMoV (Haseloff and Symons, 1982). The relative proportions of the two ccRNA 1 fast subspecies vary between different isolates as listed in Table 4-1. For the two ccRNA 2 fast isolates sequenced, the relative proportions of the two forms are the same as those of the corresponding ccRNA 1 fast. In contrast to the ccRNA 1 and 2 fast species, no similar sequence heterogeneity has been observed in nine isolates of ccRNA 1 slow and the one sequenced isolate of ccRNA 2 slow (Table 4-1). Each isolate of the ccRNA slow species thus consists entirely of either one subspecies or the other; in all except one case, the C at the position corresponding to ccRNA 1 fast residue 198 was absent. The various sequence differences between the ccRNA isolates do not seem to correlate with differences in geographic location.

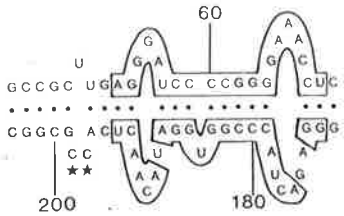
F. Structural similarities between ccRNAs and viroids

The ccRNAs share two regions of sequence homology, each of about 20 nucleotides, with the viroids

Figure 4-9 Sequence heterogeneity within the ccRNAs.

Purified ccRNAs were found to consist of either one or a mixture of two RNA species which differed in the presence or absence of an extra C residue at the position corresponding to ccRNA 1 fast residue 197 or 198. Portions of the predicted native structures of these two species are shown with stars to indicate the sequence differences. Boxed regions indicate those sequences common to PSTV, CSV and CEV.


60
GCCGC UGAG UCC CCGGG CUC
CGGCG ACUC AGG GGCCC GGG
200
180


60
GCCGC UGAG UCC CCGGG CUC
CGGCG ACUC AGG GGCCC GGG
200
180

* ccRNA isolates were purified from nucleic acid extracts of single, infected coconut palms.

† Relative proportions of sequence variants were determined by sequence analysis of RNase T_1-digested, [5'-^{32}P]radiolabelled full-length linear ccRNAs. If a ccRNA consisted of a mixture of variants, band doubling was observed on sequencing gels after ccRNA 1 fast residue 197. Relative proportions of the two sets of band doublets were taken as estimates of the molar proportions of the two variant ccRNA species.

‡ These ccRNA 1 fast species consisted of a mixture of two species, one of 246 and the other of 247 residues.

** Ligao T_1 ccRNA 2 fast species contained sequence heterogeneity at the positions corresponding to ccRNA 1 fast residue 198. Due to limitations of sequencing technique, it was not determined whether dimeric ccRNA 2 fast consisted of a 492 (2 x 246) residue species together with a 494 (2 x 247) residue species or only a 493 (246 + 247) residue species.

Table 4-1 Properties of RNAs of various ccRNA isolates.

ccRNA Isolate*	Relative proportions of sequence variants†		Total length of ccRNA (residues)	Length of sequence duplication (residues)
	C	CC		
ccRNA 1 fast				
Baao 54	1.0	0	246	
Tinambac	1.0	0	246	
Ligao 14B	0.8	0.2	246/247‡	
Ligao 620C	0.6	0.4	246/247	
Ligao 191D	0.4	0.6	246/247	
Ligao T1	0.2	0.8	246/247	
ccRNA 2 fast				
Baao 54	1.0	0	492	
Ligao T1	0.2	0.8	492-494**	
ccRNA 1 slow				
Baao 54	1.0	0	287	41
Ligao 14B	1.0	0	296	50
Ligao 620C	1.0	0	296	50
Ligao 191D	1.0	0	296	50
Ligao T1	1.0	0	301	55
Ligao 5	1.0	0	296	50
Guinayangan	1.0	0	296	50
San Miguel	1.0	0	296	50
San Nasciso	0	1.0	297	50
ccRNA 2 slow				
Baao 54	1.0	0	574	41

PSTV, CSV and CEV (Figure 4-6). The latter three viroids are closely related, sharing about 50% sequence homology. The two conserved regions are base-paired in the predicted structures of the native molecules to form highly conserved secondary structures.

The conserved regions shared by the ccRNAs correspond to regions of PSTV,CSV and CEV postulated to be involved in base-pairing of viroid complementary RNAs with a plant small nuclear RNA (snRNA) in a manner analogous to that proposed for the interaction of mRNA intron-exon splice junctions with mammalian U1a snRNA (Gross et al., 1982; Diener, 1981), and in the formation of a stabilizing stem-loop structure in the viroid complements (Gross et al., 1982). The proposed interaction between viroid complements and snRNA is postulated to reflect the origin of viroids from an intron ancestor (Diener, 1981) or as a basis for pathogenesis (Gross et al., 1982; Diener, 1981), but has not been proposed to be directly involved in viroid replication. However, the RNA or complementary RNA of at least one viroid, ASBV, is incapable of base-pairing with a U1a-like snRNA or of formation of the conserved stem-loop structure, despite up to 18% sequence homology betwen ASBV and other viroids (Symons, 1981). It is possible that the central conserved regions of viroids, including ASBV, reflect functional similarities related

to viroid replication rather than to the postulated snRNA binding.

G. Replication of ccRNAs

As PSTV,CSV,CEV and the ccRNAs are capable of autonomous replication, the enzymes involved are of considerable interest. However, no viroid-encoded translation products have been found *in vitro* (Davies *et al.*, 1974; Semancik *et al.*, 1977) or *in vivo* (Conjero and Semancik, 1977). Although PSTV, CSV and CEV share around 50% sequence homology, none of these viroids nor their putative complementary RNAs can theoretically encode similar translation products (Haseloff and Symons, 1981; Visvader *et al.*, 1982), even assuming the existence of translatable linear viroid RNAs *in vivo* (Kozak, 1979; Konarska *et al.*, 1981). Possible protein-coding regions similar to those of other viroids are not found in the ccRNAs or their complements nor are there any AUG initiation codons present. It therefore seems highly unlikely that the ccRNAs can code for any functional polypeptide product. All evidence indicates that ccRNAs and other viroids must rely entirely on host components for their replication.

Larger that unit-length complementary (-) RNA intermediates have been detected in PSTV, CEV and ASBV infected tissues (Branch *et al.*, 1981; Rohde and Sänger,

1981; Owens and Diener, 1982; Bruening et al., 1982). In addition, an oligomeric series of RNAs of ASBV (+) have been detected in infected avocado tissue; the dimer of ASBV has been purified and characterized as a single-strand, circular molecule similar to the ccRNA 2 molecules. Rolling circle mechanisms have been postulated for the synthesis of oligomeric complementary RNAs from circular viroid templates (Branch et al., 1981; Owens and Diener, 1982; Bruening et al., 1982), and oligomeric viroid (+) sequences could be simply generated by transcription of multimeric (-) strand templates. Unit length viroid produced by either specific transcription or cleavage of oligomeric viroid RNAs must be ligated to produce the final circular product. Such a model for viroid replication, involving oligomeric RNA intermediates, could readily account for the formation of the dimeric forms of both ccRNA 1 fast and ccRNA 1 slow; that is, ccRNA 2 fast and ccRNA 2 slow respectively. Rate-limiting steps during the transcription or the possible processing of viroid transcripts would allow the dimeric ccRNAs 2 to accumulate over the monomeric ccRNA 1 species.

H. ccRNA slow variants and the time course of infection

In the initial stages of cadang-cadang disease, only the fast forms of ccRNA 1 and ccRNA 2 are

present in infected palms and it is only after a further 24-30 months that the slow variants of ccRNA 1 and ccRNA 2 first appear and in the following years predominate (Mohamed et al., 1982). These data, plus preliminary evidence that the ccRNA fast species are more infectious that the ccRNA slow species (Imperial et al., 1981), are consistent with the de novo generation of the ccRNA slow variants during each cadang-cadang disease infection. This proposition is supported by the following sequence data.

1) The ccRNA 1 slow forms differ from ccRNA 1 fast by the insertion of a single repeated sequence (Figure 4-6) and could be simply generated from the ccRNA 1 fast by processing and/or transcription mechanisms.

2) The ccRNA 1 slow isolates can differ in the size of their inserted sequence repeats (Figure 4-8, Table 4-1) suggesting separate origins for these ccRNA slow variants.

3) While most ccRNA fast isolates contain a sequence heterogeneity at residue 198, and consist of varying ratios of the 246 and 247 residue species, each of the nine sequenced ccRNA slow isolates consists of a single homogeneous population, either with or without a C residue at the position homologous to ccRNA 1 fast residue 198, and with only one size of repeated sequence

(Figure 4-9, Table 4-1).

These data are consistent with the generation of ccRNA slow forms from ccRNA fast by single, rare sequence duplication events occurring separately in each cadang-cadang infected palm. All ccRNA slow molecules would, therefore, originate from single parent molecules and may accumulate in preference to ccRNA fast species due to a competetive advantage in replication.

I. Origin of cadang-cadang disease

The ccRNAs share biological properties and sequence and structural homology with viroids so that application of the term coconut cadang-cadang viroid (CCCV) is fully justified. However, whereas other viroids consist of a single predominant infectious RNA species, CCCV consists of several variant RNA species. It is feasible that CCCV may have arisen from a pre-existing viroid and that mutation or infection of new hosts, such as the coconut palm and related host species (Randles et al., 1980), resulted in the production of the variant ccRNAs by abberant transcription and/or processing mechanisms which normally occur faithfully in the replicaiton of other viroids. The outbreak and subsequent apparent rapid spread of the cadang-cadang disease in the Philippines this century (Zelazny, 1979) is consistent with such an

origin of the ccRNAs.

As viroids do not appear to encode functional polypeptide products, it seems likely that these pathogens rely entirely on the interaction of the viroid RNA with host cell components for replication. If so, the homology between CCCV, which replicates in several species of the monocotyledonous plant family Palmaceae (Randles et al., 1980), and other viroids, which replicate in dicotyledonous plant hosts (Diener, 1979), may mirror similar homology between cellular components responsible for viroid replication in these different host plants. The exact nature and function of these possibly conserved host cell components is as yet unknown.

CHAPTER 5

VELVET TOBACCO MOTTLE VIRUS

AND

SOLANUM NODIFLORUM MOTTLE VIRUS

INTRODUCTION

As outlined in Chapter 1, a new unique group of plant viruses has been reported in Australasia (Randles *et al.*, 1981; Gould and Hatta, 1981; Tien Po *et al.*, 1981; Francki *et al.*, 1983). The viruses consist of 30 nm diameter polyhedral capsids containing two major single-strand RNA species; the RNA 1 species are linear molecules of about 4,500 residues (Mr 1.5×10^6) whereas the RNA 2 species are circular, covalently closed molecules of 300-400 residues (Mr 1.25×10^5) with a high degree of internal base-pairing and have been termed virusoids as they share physical properties similar to those of viroids.

So far, there are four members of this new group of plant viruses; velvet tobacco mottle virus (VTMoV), solanum nodiflorum mottle virus (SNMV), lucerne transient streak virus (LTSV) and subterranean clover mottle virus (SCMoV). The best characterized of these are VTMoV and SNMV which possess a bipartite genome since both RNA 1 and the virusoid RNA 2 are necessary for infection. The genetic functions provided by the two viral RNAs have not been determined except for that of coding for the coat protein (Gould *et al.*, 1981). VTMoV and SNMV are serologically related while cDNA-RNA hybridization analysis gave estimates of sequence homology for the viral RNAs 1 of between 20% and 50%

depending on the stringency of the assay conditons (Gould and Hatta, 1981). On the other hand, hybridization analysis indicated that the complete sequence of VTMoV RNA 2 (Mr 1.2 x 10^5) is contained in SNMV RNA 2 (Mr 1.3x 10^5) (Gould and Hatta, 1981). Despite the close sequence similarities between the RNAs 2 of VTMoV and SNMV, neither RNA will support the replication of the heterologous RNA 1 (Gould et al., 1981) which indicates a highly specific relationship between the RNA 1 and RNA 2 of each virus.

Although viroids and virusoids appear to share similar physical characteristics, viroids are not encapsidated and replicate autonomously (Diener, 1979; Gross and Riesner, 1980). In order to further investigate the intriguing relationships between the RNAs of VTMoV and SNMV, we have sequenced the RNA 2 species of each virus and compared their structures with those of viroids.

MATERIALS and METHODS

A. Viruses and RNA

VTMoV and SNMV were kindly provided by Drs. R.I.B. Francki, J.W. Randles and A.R. Gould. Viruses were purified from infected Nicotiana clevelandii and viral RNAs isolated and purified essentially as described by Randles et al. (1981).

B. RNase Fingerprinting

The RNase A and RNase T_1 fingerprints of VTMoV and SNMV RNA 2 were determined as described for CCCV in Chapter 4.

C. RNA sequence determination

1) Partial enzymic digestion

Specific linear RNA fragments were obtained from circular RNA 2 molecules by partial RNase digestion under non-denaturing conditions as described in Chapter 2, except that 150 units/ml of RNase T_1 and 0.25 units/ml RNase U_2 were required for VTMoV RNA 2 and 300 units/ml of RNase T_1 and 0.25 units/ml RNase U_2 for SNMV RNA 2. The resultant RNA fragments were 5'-^{32}P-labelled *in vitro*, fractionated by polyacrylamide gel electrophoresis and sequenced by the partial enzymic digestion technique as described previously.

2) Dideoxynucleotide chain termination

As described in Chapter 2, RNA fragments produced by RNase T_1 digestion were also sequenced using the dideoxynucleotide chain termination technique. Specific purified 5'-^{23}P-labelled fragments were dephosphorylated with calf intestinal phosphatase and polyadenylated, using *E.coli* poly(A) polymerase. Sequencing reactions were carried out using $d(T_8C)$ as

the specific primer.

D. Synthesis and cloning of double-strand cDNA

Double-strand cDNA was synthesized from SNMV RNA 2 as described in Chapter 2, and digested with the restriction endonuclease Sau3A I. The DNA fragment corresponding to residues 131 to 216 of SNMV RNA 2 was purified and ligated into the BamH I site of the replicative form of phage M13 mp7 using T_4 DNA ligase as described in Chapter 2. Recombinant phage were screened by sequence determination using a specific M13 primer ($GTA_4CGACG_2C_2AGT$) and the dideoxynucleotide chain termination sequencing technique. Recombinant M13 replicative form was isolated (Birnboim and Doly, 1979), digested with Sau3A I and the cloned insert purified on a 6% polyacrylamide gel (Sanger and Coulson, 1978; Maxam and Gilbert, 1980) and used as a primer for the sequencing of RNA 2 of VTMoV and of SNMV by the dideoxynucleotide chain termination technique (Zimmern and Kaesberg, 1978; Symons, 1978; 1981).

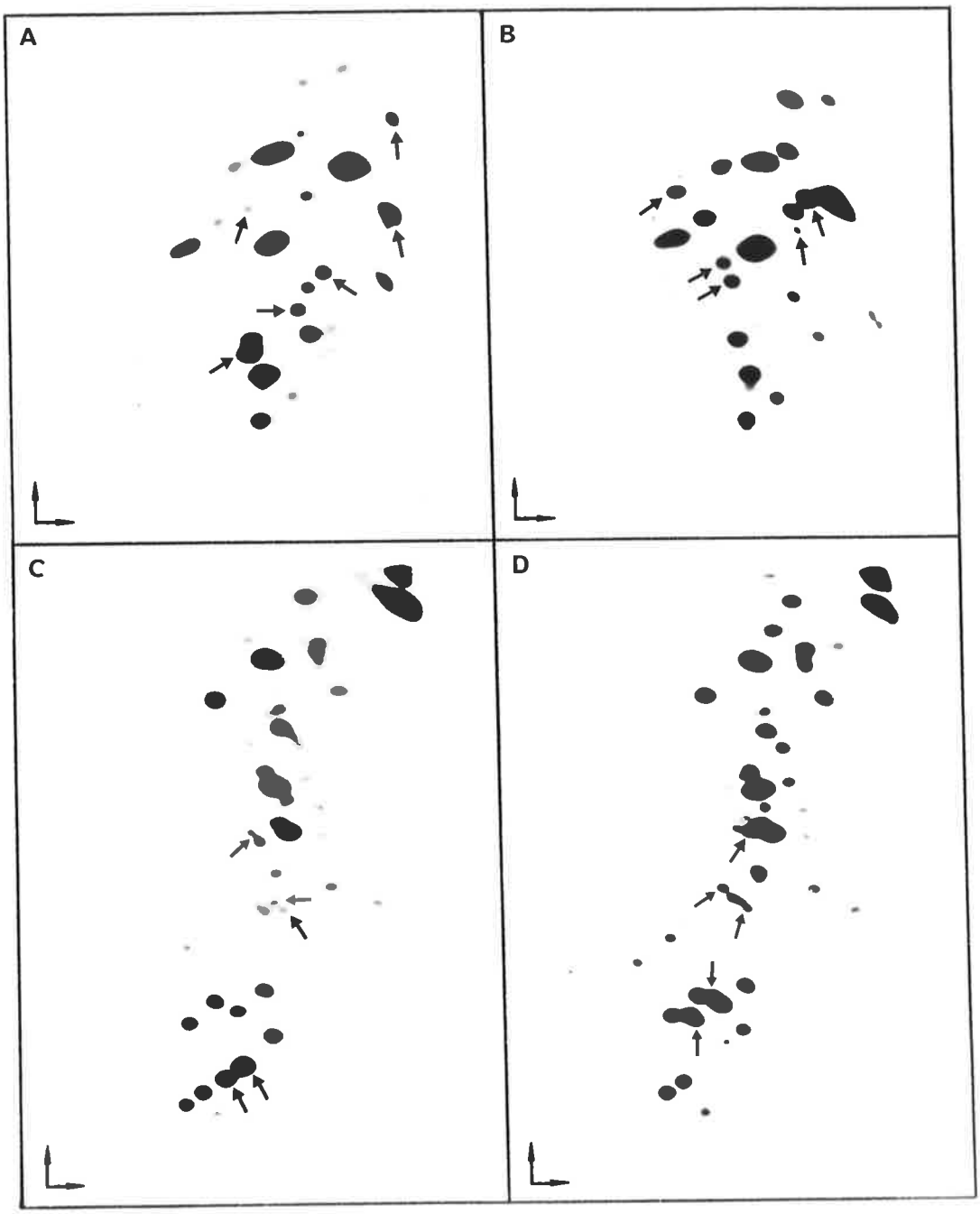
RESULTS

A. RNase fingerprints of VTMoV and SNMV RNA 2

Figure 5-1 shows the RNase A and RNase T_1 fingerprints of both the VTMoV and SNMV RNA 2 molecules. The fingerprints of the two RNAs share many spots in

Figure 5-1 RNase fingerprints of SNMV and VTMoV RNA 2.

Purified circular SNMV and VTMoV RNA 2 were digested with RNase A or RNase T_1, 5'-^{32}P-labelled and separated by two dimensional gel electrophoresis. The direction of first dimension electrophoresis is left to right, and the direction of second dimension electrophoresis is bottom to top. A, RNase A digested VTMoV RNA 2; B, RNase A digested SNMV RNA 2; C, RNase T_1 digested VTMoV RNA 2; D RNase T_1 digested SNMV RNA 2. Oligonucleotides unique to the VTMoV or SNMV fingerprints are indicated by arrows.


A
B
C
D

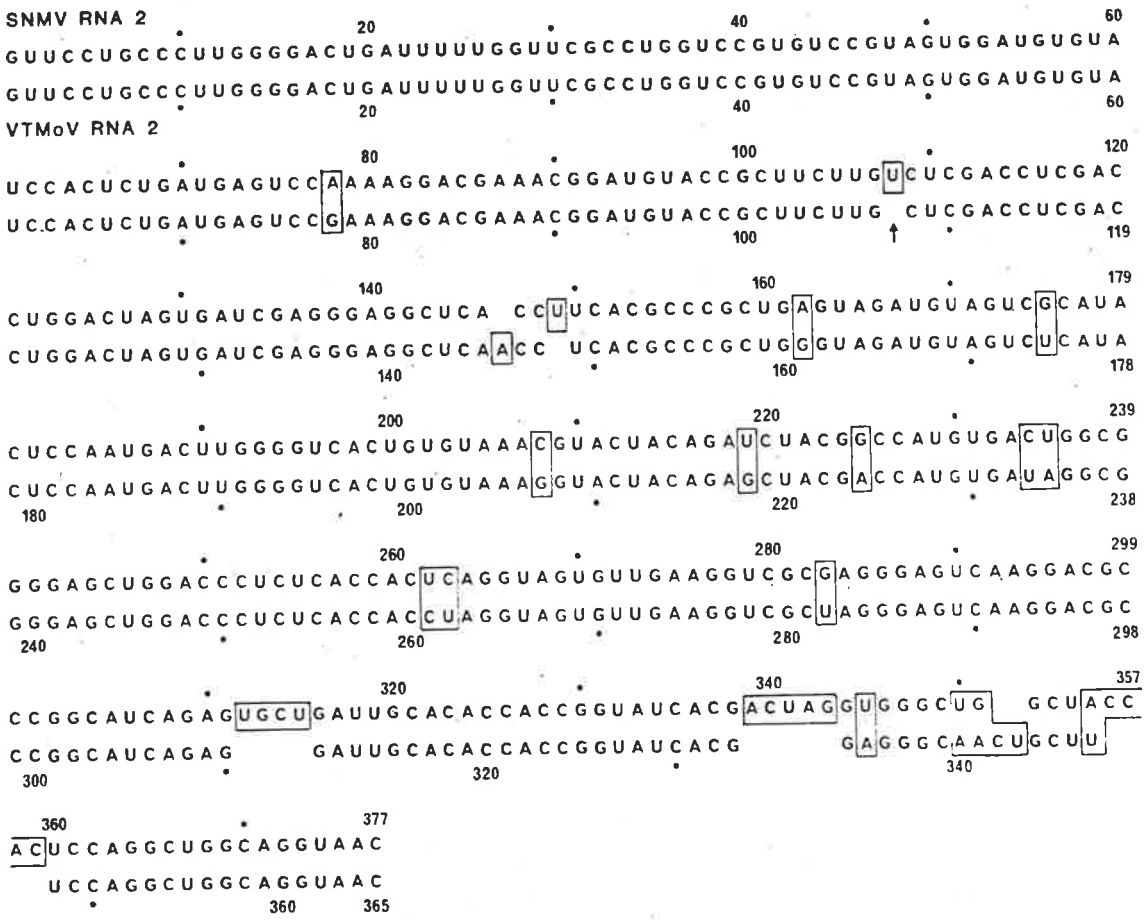
common, which confirms the existence of sequence homology between VTMoV and SNMV RNA 2 suggested by cDNA-RNA hybridization studies (Gould and Hatta, 1981). However, the fingerprints of each RNA species contain unique oligonucleotides showing that the smaller VTMoV RNA 2 is not wholly contained within SNMV RNA 2 and that the RNAs are related but distinct species.

B. Primary structures of VTMoV and SNMV RNA 2

The base sequences of VTMoV and SNMV RNA 2 were determined by using both the partial enzymic digestion and dideoxynucleotide chain termination techniques with linear RNA fragments derived from partial RNase cleavage of the native circular RNAs 2. The sequence determination of RNA fragments from both 5'-terminii, using the partial enzymic digestion technique, and 3'-terminii, using the dideoxynucleotide chain termination technique, allowed confirmation of sequences and the resolution of occasional band compressions (Kramer and Mills, 1978) which were seen on sequencing in one direction but not in the other. The complete sequences of the two RNAs were eventually obtained from the sequences of numerous overlapping RNA fragments.

The complete base sequences of the two RNAs are given in Figure 5-2. Although the RNAs are

Figure 5-2 The primary sequences of SNMV RNA 2 and the 365 residue form of VTMoV RNA 2 are shown in linear form and aligned for maximum sequence homology. The 366 residue form of VTMoV RNA 2 has an extra UMP residue at position 108 (arrowed). The sequence differences between VTMoV RNA 2 and SNMV RNA 2 are boxed. Residue 1 in each case corresponds to the left-hand end of the secondary structure model of Figure 5-3.


SNMV RNA 2
GUUCCUGCCCUUGGGGACUGAUUUUUGGUUCGCCUGGUCCGUGUCCGUAGUGGAUGUGUA
GUUCCUGCCCUUGGGGACUGAUUUUUGGUUCGCCUGGUCCGUGUCCGUAGUGGAUGUGUA
VTMoV RNA 2
UCCACUCUGAUGAGUCCAAAAGGACGAAACGGAUGUACCGCUUCUUGUCUCGACCUCGAC
UCCACUCUGAUGAGUCCGAAAGGACGAAACGGAUGUACCGCUUCUUG CUCGACCUCGAC
CUGGACUAGUGAUCGAGGGAGGCUCA CCUUCACGCCCGCUGAGUAGAUGUAGUCGCAUA
CUGGACUAGUGAUCGAGGGAGGCUCAACC UCACGCCCGCUGGGUAGAUGUAGUCUCAUA
CUCCAAUGACUUGGGGUCACUGUGUAAACGUACUACAGAUCUACGGCCAUGUGACUGGCG
CUCCAAUGACUUGGGGUCACUGUGUAAAGGUACUACAGAGCUACGACCAUGUGAUAGGCG
GGGAGCUGGACCCUCUCACCACUCAGGUAGUGUUGAAGGUCGCGAGGGAGUCAAGGACGC
GGGAGCUGGACCCUCUCACCACCUAGGUAGUGUUGAAGGUCGCUAGGGAGUCAAGGACGC
CCGGCAUCAGAGUGCUGAUUGCACACCACCGGUAUCACGACUAGGUGGGCUG GCUACC
CCGGCAUCAGAG GAUUGCACACCACCGGUAUCACG GAGGGC AACUGCUU
ACUCCAGGCUGGCAGGUAAC
UCCAGGCUGGCAGGUAAC

covalently closed circular molecules, the sequences are presented in linear form for convenience and ease of comparison. SNMV RNA 2 consists of 377 residues while VTMoV RNA 2 consists of two approximately equimolar species, one of 366 residues which, like SNMV RNA, has a U at residue 108, and another species of 365 residues where this residue is deleted. This sequence heterogeneity within RNA 2 of VTMoV was determined by sequence analysis of individual purified fragments which differed in size by one residue and which were derived from either the 366 or 365 residue species. Further confirmation of this sequence heterogeneity and an estimate of the relative proportions of the two species were obtained using a cloned DNA fragment (derived from residues 131 to 216 of SMNV RNA 2) as a primer on the mixture of the intact VTMoV RNA 2 species (see Figure 2-8). This 365 residue species is arbitrarily presented in Figure 5-2 and numbering of the VTMoV RNA 2 sequence will refer to this species.

The extensive sequence homology between VTMoV and SNMV RNA 2, originally reported on the basis of hybridization analysis with cDNA (Gould and Hatta, 1981) and show by RNase fingerprinting, is confirmed by the sequence data. As suggested by RNase fingerprinting data, each RNA 2 species contains unique sequences, thus 95% of VTMoV RNA 2 is homologous with SNMV RNA 2 and 92%

of SNMV RNA 2 is homologous with VTMoV RNA 2. The sequence differences are unevenly scattered throughout the two RNAs with a cluster of base differences around residues 339-359 of SNMV RNA 2 and residues 333-347 of VTMoV RNA 2, while there is almost complete sequence homology between residues 360-146 of SNMV RNA 2 and residues 348-145 of VTMoV RNA 2.

C. Secondary structures of VTMoV and SNMV RNA 2

Secondary structure models for the two RNAs were constructed as described by Tinoco et al. (1971) and are shown in Figure 5-3. Both RNAs form extensively base-paired rod-like structures which are similar to those described for viroids (Sanger et al., 1976; Gross et al., 1982; Haseloff et al., 1982). The structures are consistent with the known sites of high sensitivity to ribonuclease under the conditions of high salt concentration used to generate specific RNA fragments from the circular RNAs for sequencing. Thus, the terminal single-strand hairpin loops and the central single-strand regions of both RNAs (residues 70-100 and 285-305) were especially susceptible to RNase cleavage.

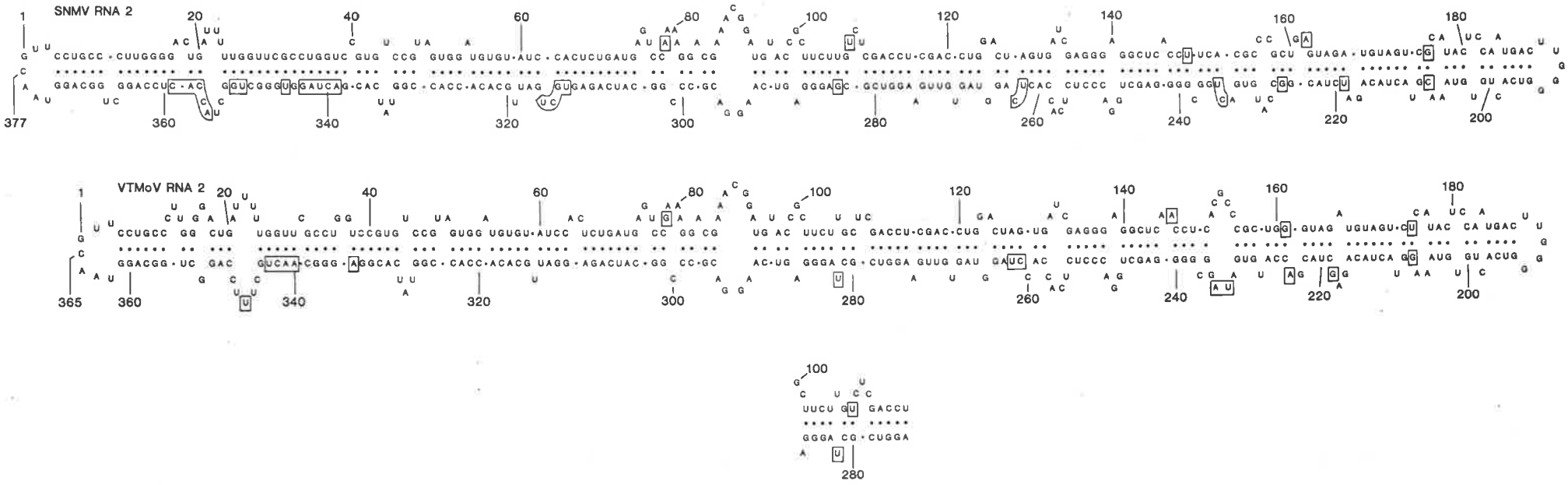
The properties of the proposed structures are summarized (Table 5-1) and are compared to those of the published structures of four viroids. The VTMoV and SNMV RNA 2 molecules possess proportions of G:C base

Table 5-1 Properties of proposed secondary structures for RNA 2 of VTMoV and SNMV compared with those of several viroids

RNA	No.of residues	No. of base pairs			G:C base pairs as % of total	Residues base paired %	ΔG* (KJ/mol at 25°C in 1M NaCl)
		A:U	G:C	G:U			
VTMoV RNA 2	365	38	72	13	59	67	-350
VTMoV RNA 2	366	38	71	14	58	67	-345
SNMV RNA 2	377	41	76	20	55	73	-455
ASBV	247	43	28	12	34	67	-280
PSTV	359	37	73	16	58	70	-610
CSV	356	44	64	16	52	70	-540
CEV	371	34	72	18	58	67	-590

*Parameters for calculation provided by Dr. D. Riesner (Steger, Gross, Randles, Sänger and Riesner, personal communication).

Figure 5-3 Predicted secondary structures of SNMV RNA 2 and the 365 residue form of VTMoV RNA 2 plus the segment of the 366 residue form of VTMoV RNA 2 containing the extra UMP residue (residue 108). The sequence differences between the RNAs are boxed.


SNMV RNA 2
1
20
40
60
80
100
120
140
160
180
377
360
340
320
300
280
260
240
220
200

VTMoV RNA 2
1
20
40
60
80
100
120
140
160
180
365
360
340
320
300
280
260
240
220
200

100
280

pairing which are lower than that of CCCV and higher than that of ASBV, but which are similar to those of the similarly sized PSTV, CSV and CEV. The thermodynamic stabilities of the proposed models were calculated using values kindly provided by Dr D. Riesner (Steger, Gross, Randles, Sanger and Riesner, unpublished data). The values of -455 KJ/mol for SNMV RNA 2 and of -345 and -455 KJ/mol for the two forms of VTMoV RNA 2 (Table 5-1) are consistent with their thermal denaturation properties; thus VTMoV RNA 2 gave a Tm of 57°C in 0.15M NaCl, 0.015M sodium citrate, pH 7, while SNMV RNA 2 gave a higher Tm of 64°C under the same conditions (Gould and Hatta, 1981; Gould, 1981). The predicted stabilities of the RNA 2 of VTMoV and of SNMV are lower than those of the similarly sized viroids PSTV, CSV and CEV but higher than those of the smaller CCCV and ASBV (Table 5-1).

D. Possible polypeptide translation products from RNA 2 species and their complements

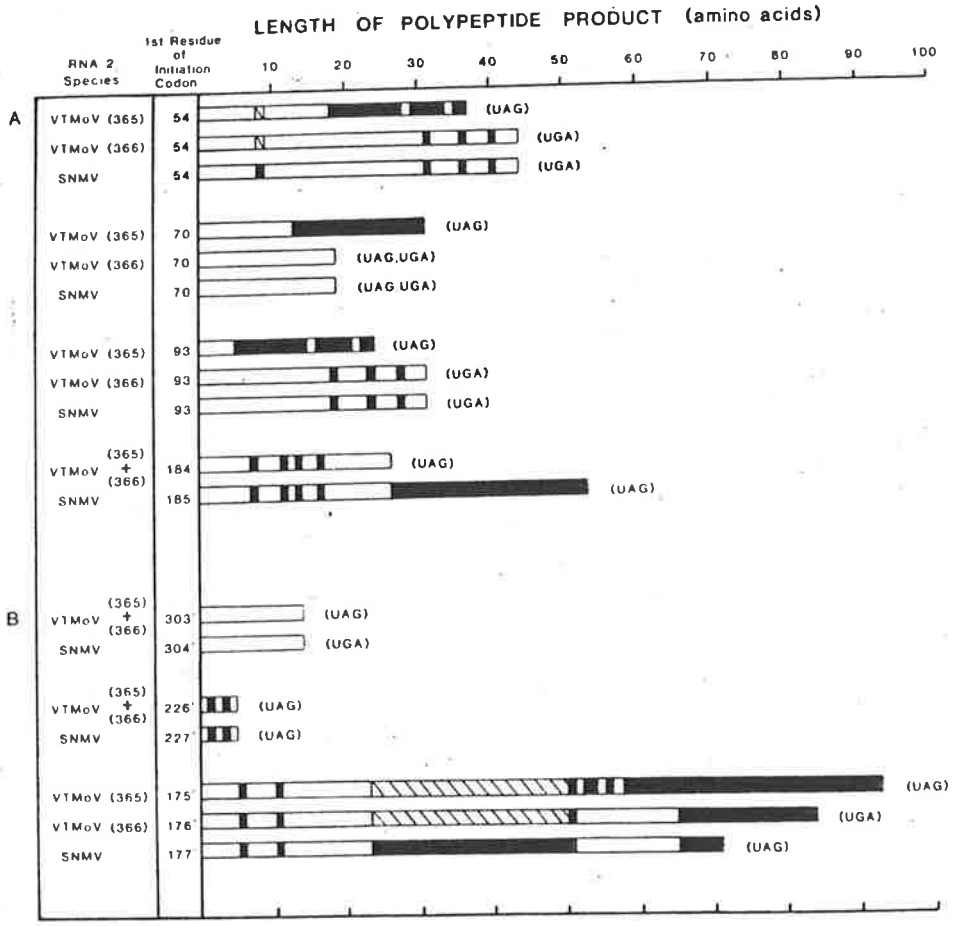
Since the RNA 2 species of VTMoV and SNMV are required with the homologous RNA 1 for viral infection (Gould et al., 1981), the RNA 2 molecules must code for some protein product(s) and/or contain structural information essential for viral replication. Evidence suggests that an AUG codon, not necessarily that nearest the 5' terminus of the mRNA, functions as the initiation

signal for eukaryotic mRNA translation (Baralle and Brownlee, 1978; Kozak, 1982; Lomedico and McAndrew, 1982) and that eukaryotic ribosomes do not interact with circular RNAs (Kozak, 1979; Konarska et al., 1981). Therefore, translation of the RNA 2 species would require the existence of specific linear RNA forms. Assuming that this condition is met, the extensive sequence homology between SNMV RNA 2 and the 366 residue form of VTMoV RNA 2, and their putative complementary RNA sequences, allows them to code for several similar polypeptide products (figure 5-4). However, only one small polypeptide product is shared between SNMV RNA 2 and the 365 residue form of VTMoV RNA 2 and their complements. All possible translation products are less than 100 amino acids in length and therefore the genes coding for the viral coat proteins (approximately 300 amino acids (Randles et al., 1981; Hollings et al., 1979)) must reside in the RNA 1 species.

DISCUSSION

In overall structure, VTMoV and SNMV RNA 2 resemble viroids in being small single-strand covalently closed circular RNA molecules which form rod-like native structures with extensive base-paired regions interspersed with single-strand regions. However, in contrast with viroids which replicate autonomously and

Figure 5-4 Possible polypeptide products of SNMV RNA 2 and the 365 and 366 residue forms of VTMoV RNA 2 (A), and their putative comlementary RNAs (B) are shown in schematic form. Each possible translation product is given with the residue number of the first base of the initiation codon plus the termination codon (s) in parentheses. For the complementary sequences, the same residue numbers are retained and therefore run in the 3' to 5' direction. The clear areas represent regions of amino acids sequence homology and the black areas regions of non-homology in the RNAs shown. The cross-hatched areas are regions of sequence homology between different products of the two VTMoV RNAs which correspond to regions of non-homology in SNMV RNA 2. Where an identical product is obtained from the two forms of VTMoV RNA 2, only one is shown.


LENGTH OF POLYPEPTIDE PRODUCT (amino acids)
RNA 2 Species
1st Residue of Initiation Codon
10
20
30
40
50
60
70
80
90
100
A
VTMoV (365)
VTMoV (366)
SNMV
54
54
54
(UAG)
(UGA)
(UGA)
VTMoV (365)
VTMoV (366)
SNMV
70
70
70
(UAG)
(UAG,UGA)
(UAG,UGA)
VTMoV (365)
VTMoV (366)
SNMV
93
93
93
(UAG)
(UGA)
(UGA)
VTMoV (365) + (366)
SNMV
184
185
(UAG)
(UAG)
B
V1MoV (365) + (366)
SNMV
303
304
(UAG)
(UGA)
VTMoV (365) + (366)
SNMV
226
227
(UAG)
(UAG)
VTMoV (365)
V1MoV (366)
SNMV
175
176
177
(UAG)
(UGA)
(UAG)

are not encapsidated (Diener, 1979; Gross and Riesner, 1980), the RNA 2 species are essential components of a bipartite genome and are encapsidated (Gould et al., 1981). While viroids do not appear to code for functional protein products (Conjero and Semancik, 1977; Davies et al., 1974; Semancik et al., 1977; Haseloff et al., 1982), there is no information available for the RNA 2 species.

The lack of conservation of possible translation products between SNMV RNA 2 and the 365 and 366 residue forms of VTMoV RNA 2, despite greater than 90% base sequence homology, suggests either that the 365 residue form of VTMoV RNA 2 may be non-functional or that RNA 2 coded translation products may have no function in viral replication. Although the involvement of RNA 2 coded translation products in viral replication cannot be excluded, it seems likely that the unique viroid-like structures of the RNA 2 molecules encode some function besides that of a template, and that this function is required for the replication of both RNA 1 and RNA 2 species.

Since neither VTMoV RNA 2 nor SNMV RNA 2 supports the replications of the heterologous RNA 1 species (Gould et al., 1981), the biological specificities of the RNA 2 species must be determined by differences in primary and/or secondary structures. The

only extensive region of sequence differences between VTMoV and SNMV RNA 2 lies around VTMoV RNA 2 residues 333-347 and SNMV RNA 2 resdues 339-359. Hence, this region may be involved in determining the specificity of the relationship between the RNA 1 and RNA 2 species, although the involvement of other structural differences cannot be excluded.

An unexpected complication during the sequence determination of the RNA 2 molecules was the occurence of sequence heterogeneity in VTMoV RNA 2 which consisted of two RNA species differing in the presence or absence of a U residue at position 108 (Figure 5-2, 5-3) and existing in approximately equimolar amounts. The two RNA species may have arisen either by a single mutation in a parent molecule followed by independent replication of the resultant two RNA species, or a mixture of the two species may be produced during each cycle of replication by transcriptional and/or processing events. In each case, it is possible that one of the two RNA species may be non-functional. Sequence heterogeneity within RNA populations has been reported for RNA phage Qβ (Domingo et al., 1978), vesicular stomatitis virus (Holland et al., 1979), satellite tobacco necrosis virus (Donis-Keller et al., 1981), citrus exocortis viroid (Gross et al., 1982) and coconut cadang-cadang viroid (Haseloff et al., 1982).

In the secondary structures of the viroids PSTV, CSV, CEV and CCCV there is a central region of the native rod-like structures which is highly conserved in both sequence and structure (Haseloff et al., 1982). ASBV (Symons, 1981) does not share this common structure except for the residues GAAACC (ASBV residues 45-50) which, as in the other viroids are present on a single strand loop in the central region of the native molecule. Interestingly, VTMoV and SNMV RNA 2 also contain the sequence GAAAC (residues 86-90 in both molecules) which is also present in a single-strand region in the centre of the proposed secondary structures. However, there is no extensive base sequence homology or complementarity between the viroids PSTV, CSV, CEV, ASBV and CCCV and the VTMoV and SNMV RNA 2 molecules.

It will be of considerable interest to determine the exact mechanisms by which VTMoV and SNMV support the replication of the homologous RNA 1 species as well as the molecular basis for the specificity between the RNA 1 and RNA 2 species, and to determine whether the common structural features of viroids and virusoids mirror some common function. Chapter 6 outlines work with subterranean clover mottle virus which allows some definition of the virusoid sequences involved in such relationships.

CHAPTER 6

SUBTERRANEAN CLOVER MOTTLE VIRUS

INTRODUCTION

Several isolates of subterranean clover mottle virus (SCMoV) from Western Australia have been described (Francki et al., 1983). As judged by electron microscopy, preparations of SCMoV consist of homogeneous populations of polyhedral virus particles about 30 nm in diameter, and serological tests using antiserum to purified SCMoV failed to reveal any antigenic differences between the isolates. However, the various SCMoV isolates appear to contain differing encapsidated RNA components. While all isolates contained a linear, single-stranded RNA species of approximately 4500 residues in size (RNA 1, Mr 1.5×10^6), each isolate contained either one or both of two viroid-like circular RNA species, RNA 2 (approximately 400 resides) and RNA 2' (approximately 300 residues).

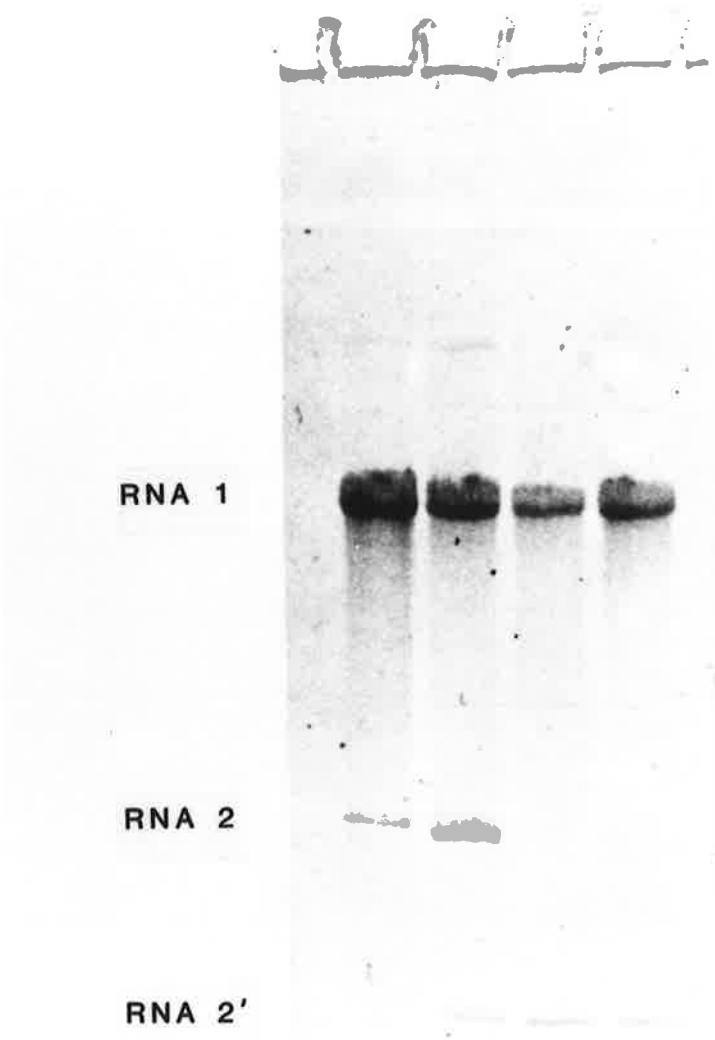
The SCMoV isolates used in this work were the isolates A, B, D and E of Francki et al. (1983). Some time after its isolation SCMoV-A was found to contain both RNA 2 and RNA 2' with the RNA 1 species. In contrast, SCMoV-E contains only RNA 1 and RNA 2 while SCMoV-B and SCMoV-D contain only RNA 1 and RNA 2'. The fractionated RNA components of these isolates are shown in Figure 6-1. Interestingly, while isolates A, D and E were obtained as field isolates, SCMoV-B was produced from SCMoV-A by passage through single lesions on pea leaves.

Figure 6-1 RNA components of SCMoV isolates.

RNAs extracted from SCMoV isolates E, A, B and D were fractionated by electrophoresis in a 4% polyacrylamide gel containing 7M urea. RNAs were detected by staining with toluidine blue. RNA 1 and the circular forms of RNAs 2 and 2' are shown. The linear forms of RNAs 2 and 2' migrated from the gel.

SCMoV isolates

E A B D



RNA 1

RNA 2

RNA 2'

4% TBE

+7M urea

The finding of SCMoV isolates which differed in their RNA 2 and/or RNA 2' components suggested either (1) the existence of two serologically indistinguishable viruses which each possessed different RNA components or (2) the existence of a single virus which could support the replication of either or both RNA 2 and RNA 2'.

The following work describes efforts to distinguish between these possibilities and to determine the structures of SCMoV RNAs 2 and 2'.

MATERIALS

Isolates of SCMoV were kindly provided by Dr. Richard Francki. Viral propagation, purification and extraction of viral RNAs were performed by Mr. Chris Davies as described (Francki *et al.*, 1983). SCMoV RNA species were purified by polyacrylamide gel electrophoresis (Gould, 1981), mostly by Mr. Davies.

METHODS

A. Synthesis and restriction endonuclease cleavage of ds cDNA

Random-primed first strand cDNA was transcribed from SCMoV RNA 1 essentially as described by Taylor *et al.* (1976). Purified SCMoV RNA 1 (2 µg) was resuspended in 25 µl containing 50mM Tris-HCl pH 8.3, 50mM KCl, 10mM $MgCl_2$, 10mM DTT, 100µM [α-^{32}P]dCTP

(50 µCi), 500µM dATP, dGTP and dTTP, 2 mg/ml DNase 1 treated salmon sperm DNA and 20 units avian myeloblastosis virus reverse transcriptase. After incubation at 37°C for 2 hours, the reaction was boiled and treated with RNase A before synthesis of second-strand cDNA as described in Chapter 2, Methods B-1. Synthesized double-strand cDNA was digested with various restriction endonucleases and fractionated by polyacrylamide gel electrophoresis as described also in Chapter 2, Methods B-1.

B. Fingerprinting of SCMoV RNAs

Purified circular SCMoV RNAs 2 and 2' were RNase A fingerprinted using techniques outlined in Chapter 4.

C. Sequence determination of SCMoV RNA 2 and RNA 2'

Purified SCMoV RNA 2 and RNA 2' were each subjected to partial digestion under non-denaturing conditions, as described in Chapter 2, using 150 units/ml RNase T_1, 0.1 µg/ml RNase A or 0.25 units/ml RNase U_2. The resulting linear RNA fragments were either 5'- or 3'-^{32}P radiolabelled, fractionated by polyacrylamide gel electrophoresis and sequenced using the partial enzymic digestion technique essentially as described in Chapter 2, except that sequencing gels

contained TBE buffer, 7M urea and 25% (v/v) deionized formamide. The sequences of overlapping RNA fragments were used to obtain the primary structures of the circular RNAs and the methods of Tinoco *et al.* (1971) used to predict the native secondary structures of the molecules.

RESULTS

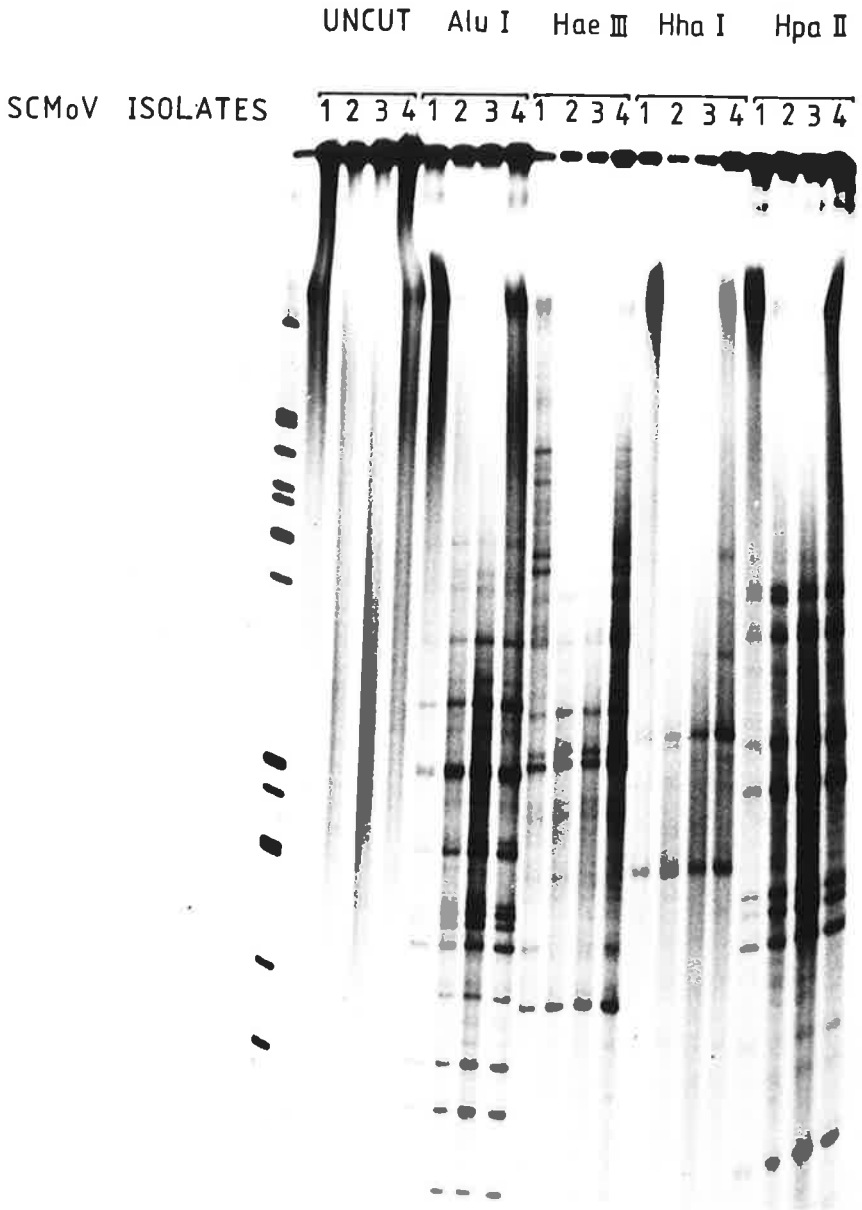
A. Analysis of SCMoV RNA 1 nucleotide sequences

In order to analyse the sequence relationships between the RNAs 1 of the SCMoV isolates, double-strand cDNAs were transcribed from purified RNAs 1 and then digested with the various sequence-specific restriction endonuclease, *Alu* I, *Hae* III, *Hha* I, *Hpa* II (see Figure 6-2) and *Acc* 1 (data not shown). Polyacrylamide gel electrophoresis of digested ds cDNAs results in gel patterns which reflect the arrangement of restriction enzyme recognition sequences on the original RNA 1 molecules. As shown in Figure 6-2, all four isolated RNA 1 species gave rise to indistinguishable patterns of ds cDNA restriction fragments for all enzymes used. While, (1) the synthesized ds cDNAs may not be wholly representative of the respective RNA 1 species, and/or (2) small nucleotide sequence differences between the RNAs 1 may exist which do not affect the size or number of observed ds cDNA

Figure 6-2 Restriction endonuclease digestion of SCMoV RNA 1 ds cDNA.

Double-strand cDNA was synthesized from RNA 1 purified from four different isolates of SCMoV. Isolates 1, 2, 3 and 4 correspond to SCMoV-E, SCMoV-A, SCMoV-B and SCMoV-D, respectively. The ds cDNAs were untreated or digested with Alu I, Hae III, Hha I or Hpa II and fractionated on a 6% polyacrylamide gel containing 2M urea. 3'-^{32}P-labelled Hpa II digested M13 mp7 RF was included as size markers, with sizes of 1596, 829, 818, 652, 545, 543, 472, 454, 357, 183, 176, 156, 129, 123, 79 and 60 base pairs.

RESTRICTION DIGESTS OF SCMoV RNA1 ds cDNA



restriction fragments, the results do suggest that the four isolated RNA 1 species are certainly closely related in nucleotide sequence, and may be identical.

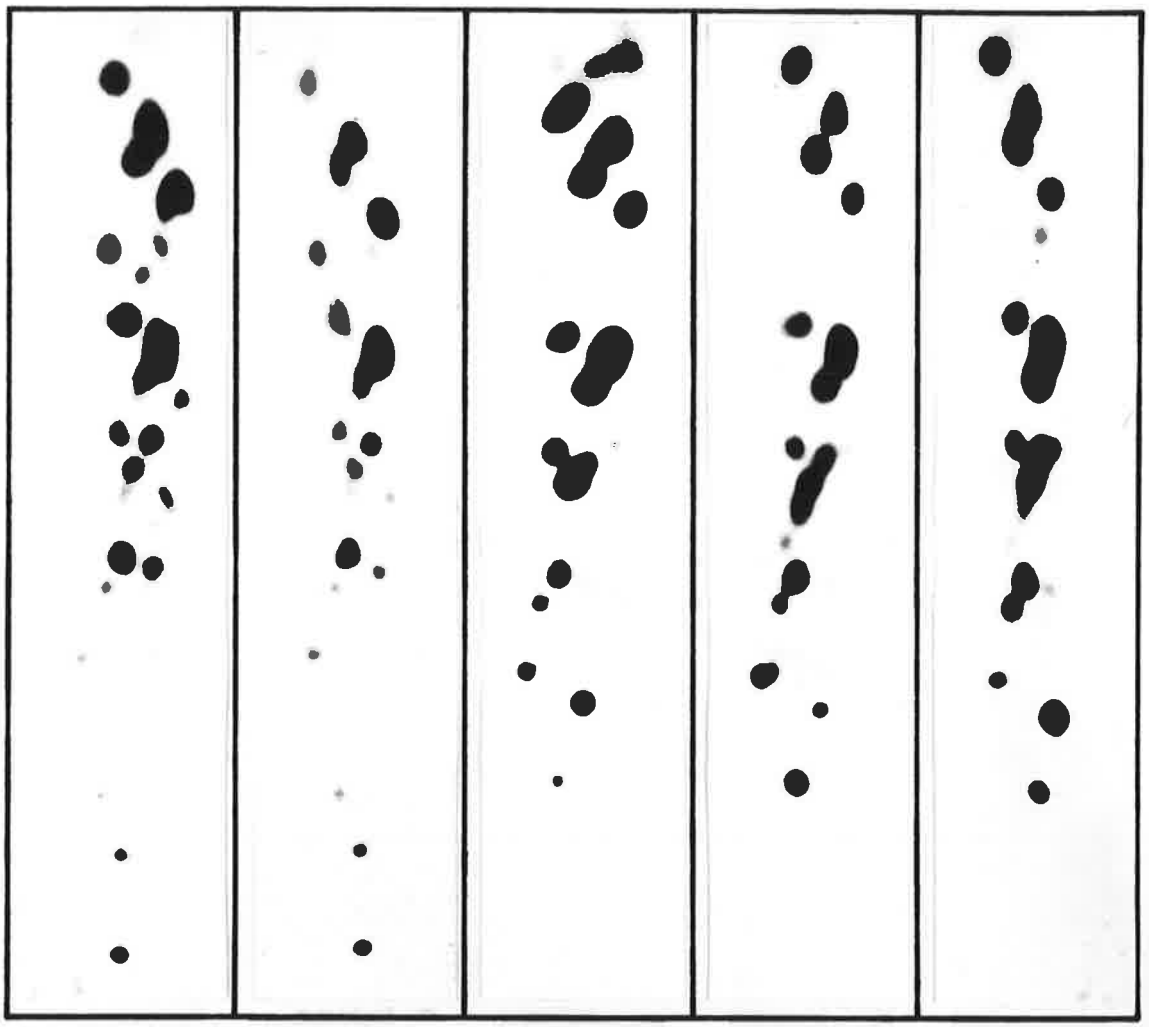
B. RNase fingerprinting of SCMoV RNAs 2 and RNAs 2'

In addition, the sequence relationships between the RNAs 2 of the SCMoV isolates were investigated using the RNase fingerprinting technique. Purified circular RNA 2 and RNA 2' species were digested to completion by RNase A, 5'-^{32}P-radiolabelled and fractionated by two-dimensional polyacrylamide gel electrophoresis as described in Chapter 4. The resulting oligonucleotide patterns are given in Figure 6-3 and show the following. (1) The oligonucleotide patterns obtained from the RNAs 2 of SCMoV-E and SCMoV-A are essentially identical. (2) The oligonucleotide patterns obtained from the RNAs 2' of SCMoV-A, SCMoV-B and SCMoV-D are also essentially identical. (3) Some oligonucleotides appear to be common to both the RNA 2 and RNA 2' species (see Figure 6-3).

Although RNase A (C and U specific) fingerprinting may not reveal minor sequence differences present in pyrimidine-rich regions of the molecules, the data suggest that, like the isolated SCMoV RNAs 1, each RNA 2 and RNA 2' species is either closely related or identical to similar species from different SCMoV

Figure 6-3 RNase A fingerprints of SCMoV RNAs 2 and 2'.

Circular RNAs 2 were purified from SCMoV-E and SCMoV-A, and circular RNAs 2' were purified from SCMoV-A, SCMoV-B and SCMoV-D. All RNAs were digested with RNase A, 5'-^{32}P-labelled and fractionated by 2-dimensional polyacrylamide gel electrophoresis. The resulting oligonucleotide fingerprints are shown with the directions of electrophoresis in the first dimension being left to right, and, in the second dimension, bottom to top.


SCMoV-E
RNA2
SCMoV-A
RNA2
SCMoV-A
RNA2'
SCMoV-B
RNA2'
SCMoV-D
RNA2'

isolates. The possible identity of the different RNA 2 or RNA 2' species is also supported by the identical fragments obtained after partial RNase T_1, U_2 or A cleavage under non-denaturing conditions (see Methods, this chapter) of RNA 2/2' from different isolates, and by preliminary sequence data (not shown).

In contrast, there are considerable differences between the oligonucleotide fingerprints, and thus primary sequences, of the RNA 2 and RNA 2'. So while some degree of sequence homology is indicated by the number of shared oligonucleotides, RNA 2 and RNA 2' each contain unique sequences and do not differ simply in the possession of repeated sequences.

From the evidence presented, it seems likely that the different isolates of SCMoV, which are serologically indistinguishable (Francki et al., 1983), contain essentially identical RNA 1 species and differ only in containing either or both RNA 2 and RNA 2' species. This is almost certainly the case for SCMoV-A, containing both RNA 2 and RNA 2' and SCMoV-B (which was derived directly from SCMoV-A) containing only RNA 2'. Furthermore, although RNA 2 and RNA 2' appear to differ significantly in nucleotide sequence, no sequence differences were observed between similar RNAs 2/2' obtained from different SCMoV isolates.

C. Sequence determination of SCMoV-A RNA 2 and RNA 2'

Linear RNA fragments were obtained from the RNA 2 and RNA 2' of the SCMoV-A isolate by partial ribonuclease digestion under non-denaturing conditions. These fragments were radiolabelled and sequenced using the partial enzymic cleavage method, and the sequences of overlapping fragments were assembled to give the complete primary structures of the circular molecules. The RNA 2 and RNA 2' species each consists of 388 and 327 residues respectively. In addition, the two RNA species share a single common region of about 220 residues of almost complete sequence homology. The primary structures of SCMoV-A RNA 2 and RNA 2' are presented in Figure 6-4 in a convenient linear form with the shared sequences shown boxed. Secondary structure models for these molecules were constructed using the methods of Tinoco et al. (1971) and are shown in Figure 6-5. Both RNAs may base-pair intramolecularly to form helical rod-like structures similar to those of VTMoV and SNMV RNA 2 and viroids. Strikingly, the nucleotide sequences conserved between SCMoV-A RNA 2 and RNA 2' are located so as to form the entire base-paired left-hand sides (as drawn) of the native molecules. Consequently, it is the differing lengths of the unconserved right-hand sides of the native molecules which account for their difference in size.

Figure 6-4 Primary structures of SCMoV RNA 2 and RNA 2'.

Nucleotide sequences common to the two RNAs are shown boxed.

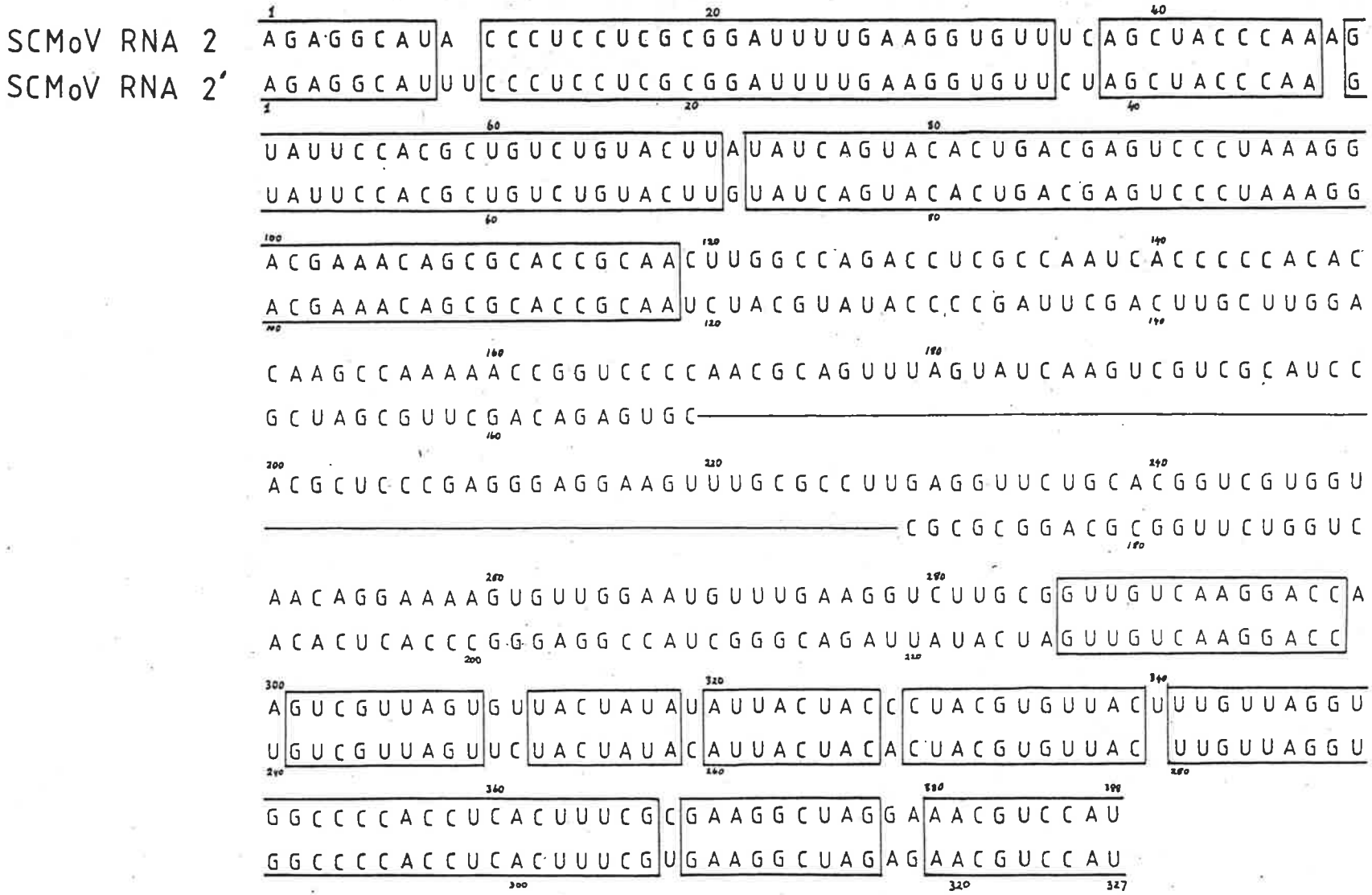
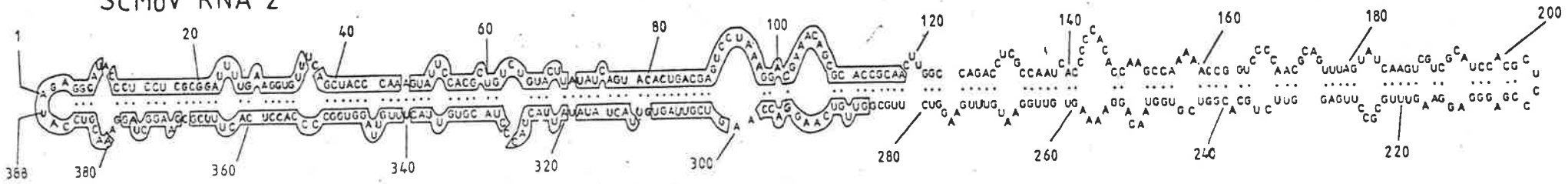

SCMoV RNA 2
SCMoV RNA 2'

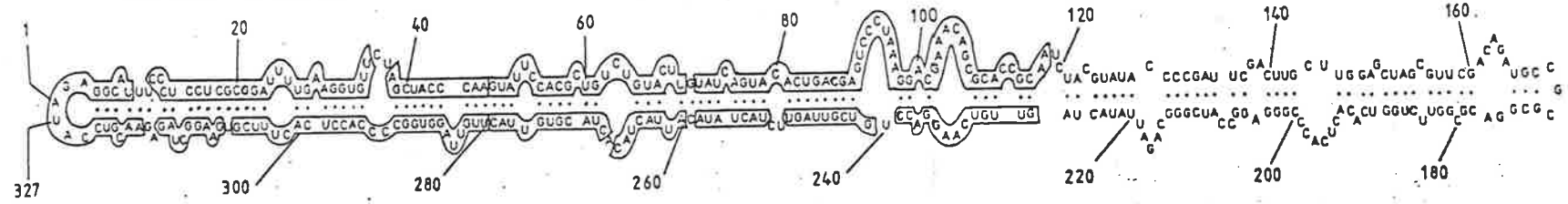
Figure 6-5 Secondary structures of SCMoV RNA 2 and RNA 2'.

Nucleotide sequences common to the two RNAs are shown boxed.

SCMoV RNA 2



SCMoV RNA 2'



DISCUSSION

A. Relationships between the various isolates of SCMoV

The four SCMoV isolates used in this work have now been shown to be indistinguishable serologically (Francki et al., 1983), to contain indistinguishable RNA 1 species and are likely to differ only in containing either one or both RNA 2 and RNA 2'. Given this, then irrespective of whether SCMoV RNAs 2/2' are required for viral infection, as appears to be the case with the RNAs 2 of VTMoV and SNMV (Gould et al., 1981), or are satellite RNAs, as appears to be the case for LTSV RNA 2 (Jones et al., 1983), it seems that SCMoV RNA 2 and RNA 2' must be functionally equivalent, that is, interchangeable.

B. Sequence homology between SCMoV RNA 2 and RNA 2'

The native structures of SCMoV RNA 2 and RNA 2' show remarkable conservation of the left-hand sides of the molecules (Figure 6-5). It therefore seems reasonable to propose that these common sequences and structures mirror the apparent interchangeable functions of the molecules. Thus, if RNA 2 and RNA 2' are satellite RNAs, the conserved left-hand sides of the molecules may contain recognition signals required for replication by viral and/or host components. If the RNAs are functionally similar to VTMoV and SNMV RNA 2,

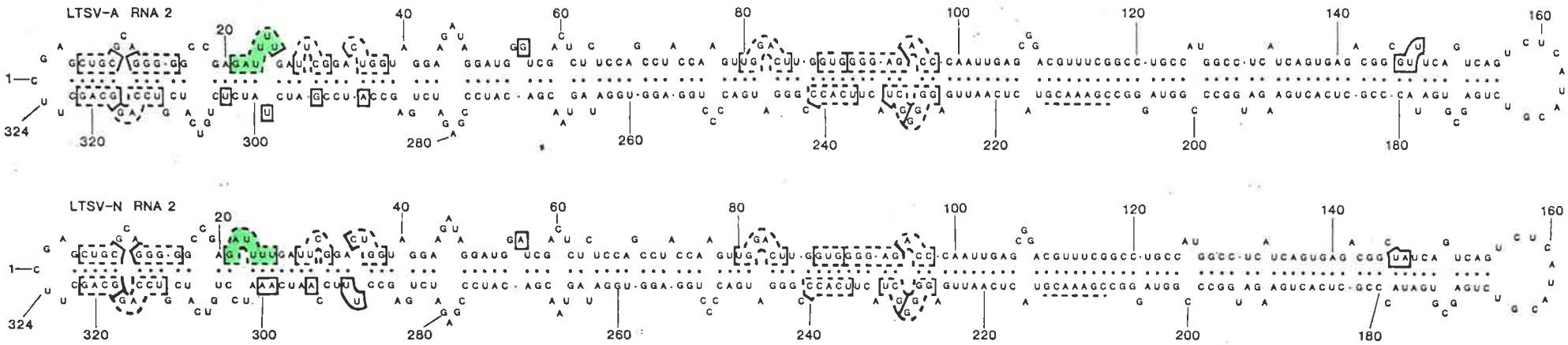
the left-hand sides of the molecules may also contain some undefined function required for viral replication. Interestingly, the conserved sequences of SCMoV RNA 2 and RNA 2' share homology with VTMoV and SNMV RNAs 2.

C. Sequence homology between SCMoV, VTMoV, SNMV and LTSV RNAs 2

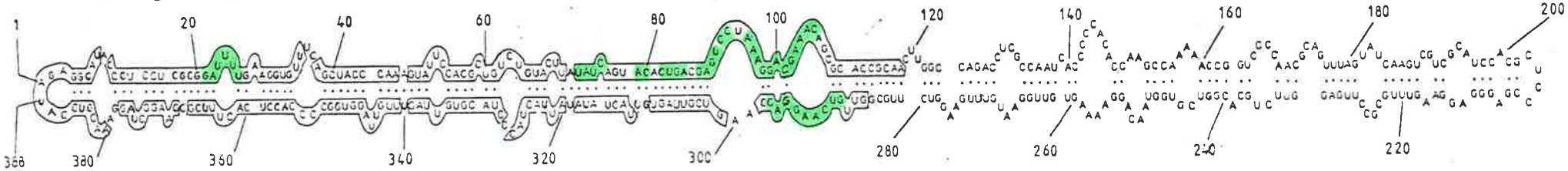
The determined sequences and predicted secondary structures of the viroid-like RNAs of SCMoV, VTMoV and SNMV and two isolates of LTSV (Keese and Symons, unpublished results) are shown in Figure 6-6, and common sequences indicated. First, the sequence GAUUUU is present on all RNAs in a similar position on the native structures (starting at approximately residue number 20 in all cases). The conservation of this sequence suggests that it may play some role which is common to the replication of all six RNAs, including the LTSV RNAs 2. Second, the virusoids of SCMoV, VTMoV, and SNMV all contain conserved sequences which are central to their native structures. These consist of two regions, one of 24 residues and the other of 9 residues, which are postioned on opposite sides of the rod-like molecules. Similarly, viroids also contain highly conserved sequences which are central to their rod-like native sturctures (Chapter 4), and the pentanucleotide sequence GAAAC is present, predominantly

Figure 6-6 Sequence and structural homology between SCMoV RNA 2 and RNA 2' and SNMV, VTMoV and LTSV RNAs 2.

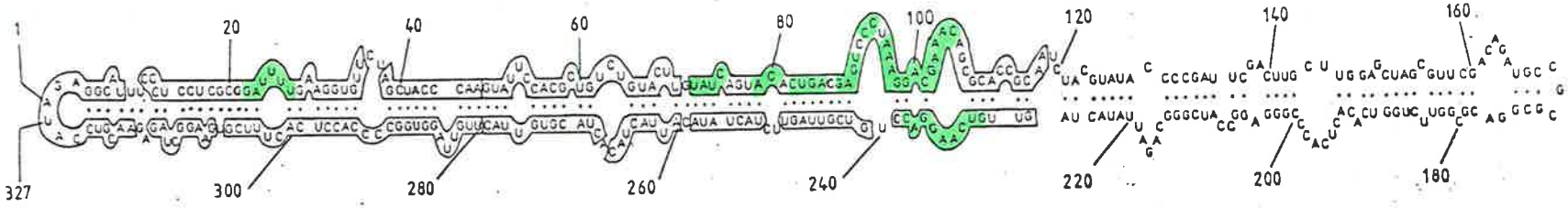
The proposed secondary structure models for the RNAs are shown with conserved sequences indicated in green.

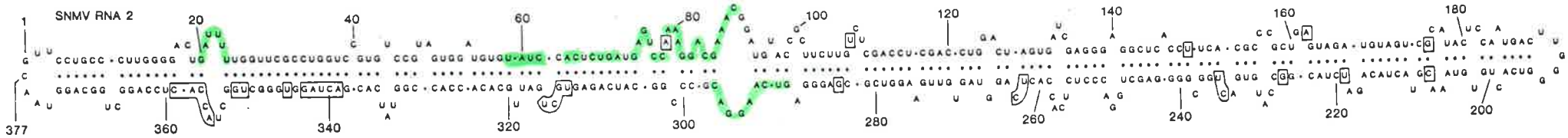


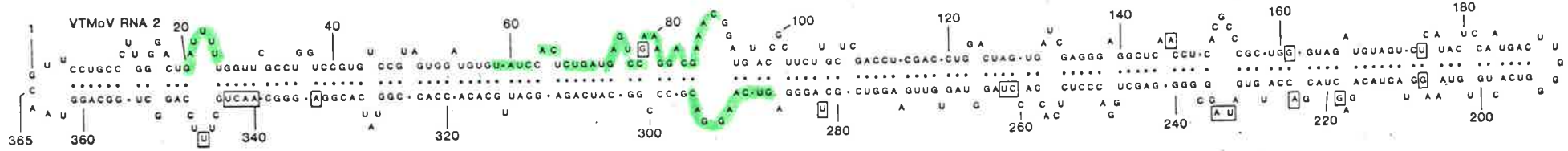
SCMoV RNA 2



SCMoV RNA 2'







single-stranded, in the central conserved sequences of SCMoV, VTMoV, SNMV virusoids and all sequenced viroids. However the virusoid molecules do not appear to contain the conserved stem-loop structure (Gross et al., 1982) or postulated U1a snRNA homology (Diener, 1981; Gross et al., 1982) which are present in all sequenced viroids except ASBV.

The GAUUUU and central conserved sequences which are shared by the SCMoV RNAs 2/2' and VTMoV and SNMV RNAs 2 are located within the left-hand sides of the native molecules which, in SCMoV RNAs 2/2', are highly conserved. In contrast to the SCMoV virusoids, VTMoV and SNMV RNA 2 will only replicate in conjunction with the RNA 1 species from the same virus, and the RNA 2 species are therefore not interchangeable (Gould et al., 1981) despite greater than 90% sequence homology. It is therefore interesting to note that the main sequence differences between VTMoV and SNMV RNA 2 lie clustered opposite the GAUUUU sequence and that the remainder of the left-hand side of the native structure is almost completely conserved (Chapter 5). Thus, these observations are consistent with both the functional similarity of the SCMoV, VTMoV and SNMV virusoids and the importance of the left-hand sides of these molecules in replication.

It is tempting to speculate that as the SCMoV

virusoids share common structures with VTMoV and SNMV RNA 2, they may also share common biological properties. That is, SCMoV RNA 2 and/or RNA 2' may be required for viral infection in a manner similar to that shown for the RNAs 2 of VTMoV and SNMV, and the central conserved sequences of these molecules may play some role in fulfilling this requirement. However, this proposition must be viewed with some scepticism; for while LTSV RNA 2 (which has been shown to behave as a satellite RNA (Jones et al., 1983) and is not required for viral infection) does not share the central conserved sequences of SCMoV, VTMoV and SNMV RNAs, another satellite RNA, that of tobacco ringspot virus (TobRV), does.

D. Satellite RNA of TobRV

TobRV belongs to the nepovirus group and consists of 28 nm isometric particles containing two single-strand RNAs of molecular weight 2.7×10^6 (RNA 1) and 1.3×10^6 (RNA 2) which comprise the entire genome of the virus. RNA 1 and RNA 2 each possess a 3'-polyA tract (Mayo et al., 1979a) and a 5'-covalently linked protein (Vpg) (Mayo et al., 1979b). In 1969, a novel RNA species was found in cultures of TobRV that had previously been apparently free of it (Schneider, 1969). This RNA species, which was dependent on TobRV for

replication (Schneider, 1971), shared almost no nucleotide squences in common with the supporting RNA (Schneider, 1977) and was classified as a satellite RNA. Since that time, at least 24 distinct isolates of TobRV have produced a satellite RNA of unknown origin during laboratory propagation (Kiefer et al., 1982). The satellite RNA consists of a single linear species (Diener et al., 1974; Schneider, 1977) of approximately 350 residues (Sogo et al., 1974), and does not share the 3'-polyadenylate and 5'-linked protein that are characteristic of the TobRV genomic RNAs (Kiefer et al., 1982), but instead bear 5'-hydroxyl and 3'-phosphate groups (Kiefer et al., 1982; G. Bruening, personal communication). Kiefer et al. (1982) have shown that TobRV also encapsidates a multimeric series of larger than unit length TobRV satellite RNA sequences and that double-strand RNA fractions isolated from infected plants can be denatured to produce similar multimeric series of both satellite and complementary RNA sequences. Schneider and Thompson (1977) have shown that the double-strand RNAs purified from infected tissue are infective only after denaturation and addition of TobRV, and Sogo and Schnieder (1982) demonstrated that while the double-strand RNA preparations contained predominantly linear molecules, circular and 'racket' shaped molecules were also

detected. Thus Kiefer et al.(1982) have proposed a rolling circle type mechanism for the replication of the satellite RNA of TobRV which would account for the production of circular and longer than unit length RNA sequences of the satellite RNA and its complement.

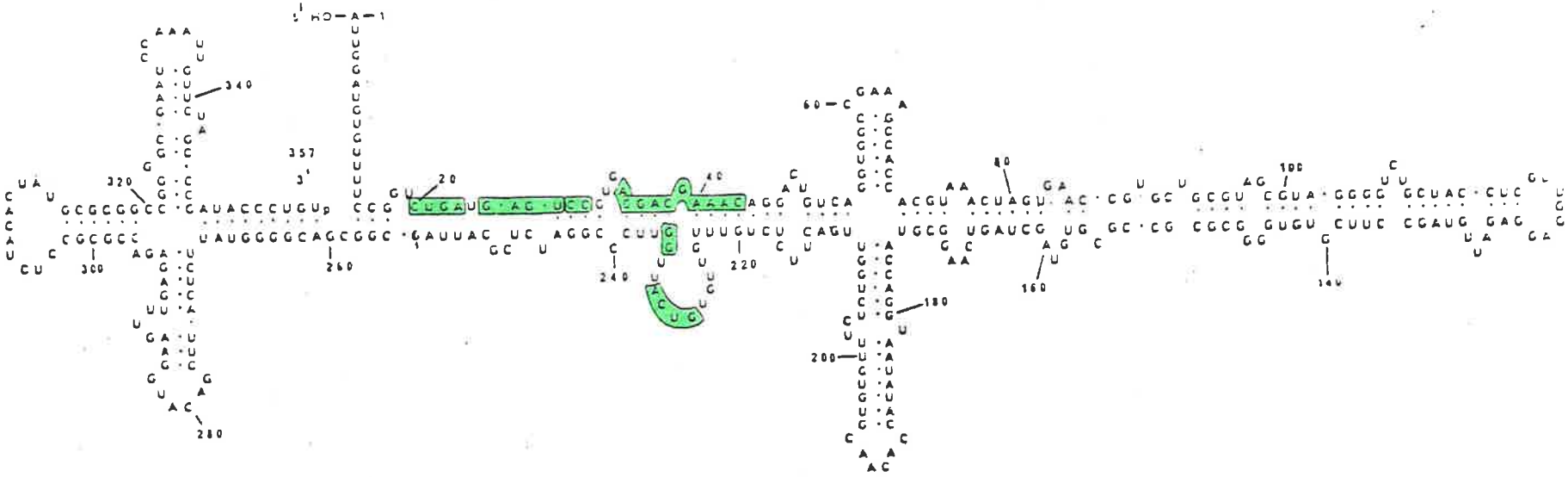
E. Sequence homology between TobRV satellite RNA and virusoids

The complete sequence of a satellite RNA associated with the budblight strain of TobRV has now been determined (Bruening, unpublished results) and was kindly provided by Dr. George Bruening. The sequence which consists of 357 residues is shown in Figure 6-7 with the predicted secondary structure of the molecule (Bruening, unpublished results). The predicted native molecule possesses an overall rod-like structure with four prominent stem-loop structures and a single-strand 5'-proximal region. Also indicated in Figure 6-7 is the remarkable extent of sequence homology between TobRV satellite RNA and the virusoid RNAs of VTMoV, SNMV and SCMoV. The homologous sequences correspond to the central conserved regions of the virusoids, and are also positioned in the rod-like centre of the TobRV molecule. Furthermore, the GAAAC sequence which is common to the central conserved regions of viroids and virusoids (except that of LTSV) is also present in the TobRV

Figure 6-7 Sequence and proposed secondary structure of TobRV satellite RNA.

This data was kindly provided by Dr. George Bruening (unpublished results). Sequences which are conserved between TobRV satellite, SCMoV RNAs 2 and 2', SNMV and VTMoV RNA 2 are shown indicated in green.

TobRV satellite RNA



satellite RNA, while tha GAUUUU sequence found in all virusoids, including that of LTSV, is absent.

The presence of sequences and structures which are shared by the TobRV satellite RNA and VTMoV, SNMV and SCMoV virusoids suggests the residence of common functions and/or signals within these RNAs. These RNAs replicate with viruses of quite different properties (nepovirus group versus VTMoV/sobemovirus group), and this may indicate that the possible conserved functions and/or signals shared by the RNAs are involved in interaction with host cell components rather than components of the different viruses. Kiefer et al. (1982) concluded from evidence outlined above that TobRV satellite RNA may replicate via circular and multimeric RNA intermediates through a rolling-circle type mechanism similar to that proposed for viroids (Branch et al., 1981; Owens and Diener, 1982; Kiefer et al., 1982; Bruening et al., 1982). The work presented in the final chapter is the result of preliminary attempts to determine whether virusoids also replicate via a rolling-circle type mechanism, with the attendant possibility of involvement of the conserved sequences.

CHAPTER 7

VIROIDS, VIRUSOIDS AND SATELLITES

INTRODUCTION

Larger than unit-length complementary (-) RNA intermediates have been detected in PSTV- and CEV-infected plant tissues (Branch *et al.*, 1981; Rohde and Sänger, 1981; Owens and Diener, 1982), and appear to exist mainly in extensively double-stranded RNA. Owens and Cress (1980) and Branch *et al.* (1981) have shown that RNase treatment of double-strand RNA intermediates from PSTV infected plants results in the production of complementary PSTV RNA of unit length or slightly larger while Owens and Diener (1982) demonstrated that denaturation of the double-strand RNAs released monomeric PSTV strands that had been complexed with multimeric complementary RNAs. In addition, multimeric series of both ASBV and its complement have been found in viroid-infected tissue (Bruening *et al.*, 1982), and dimeric RNAs have been shown for CCCV. Various workers have postulated rolling-circle type mechanisms for the replication of viroids (Branch *et al.*, 1981; Owens and Diener, 1982; Bruening *et al.*, 1982).

As outlined in the previous chapter, the replication of the linear satellite RNA of TobRV shares features common to viroid replication. Complementary RNA intermediates have been detected in high molecular weight double-strand RNAs, and RNase treatment of these duplex RNA intermediates reduced the molecules to a size

slightly but significantly larger than unit-length (Sogo and Schneider, 1982). Double-strand RNAs were shown to produce detectable satellite activity only if denatured (Schneider and Thompson, 1977), and to consist of multimeric series of apparently concatenate forms of RNAs of both polarities (Kiefer et al., 1982). Thus Kiefer et al. (1982) have also postulated a rolling-circle type mechanism for the replication of TobRV satellite RNA. It seemed feasible that virusoids may also replicate via a rolling-circle type mechanism and the remainder of this chapter is devoted to description of the experimental support for this notion and to its possible ramifications.

METHODS

A. Isolation of RNA

VTMoV and SNMV were kindly provided by Drs. R.I.B. Francki, J.W. Randles and A.R. Gould. RNAs were extracted from purified virus (Gould, 1981), and from virus infected Nicotiana clevelandii (Randles et al., 1981) using phenol-SDS extractions as previously described.

B. Blot hybridization

Nucleic acid samples were denatured by treatment with 1M glyoxal and 50% (v/v) dimethyl

sulphoxide (McMaster and Carmichael, 1977) and electrophoresed on 2.0% agarose slab gels (15x14x0.15 cm) in 10mM sodium phosphate pH 6.5 at 30 mA. Nucleic acids were transferred to nitrocellulose by blotting and baked *in vacuo* at 80°C (Thomas, 1980). Nitrocellulose sheets were prehybridized, hybridized and washed essentially as described by Thomas (1980). Complementary ^{32}P-DNA hybridization probe was prepared using recombinant M13 ss DNA, containing sequences corresponding to SNMV RNA 2 residues 131 to 216 (see Chapter 5), essentially as described by Bruening *et al*. (1982).

RESULTS

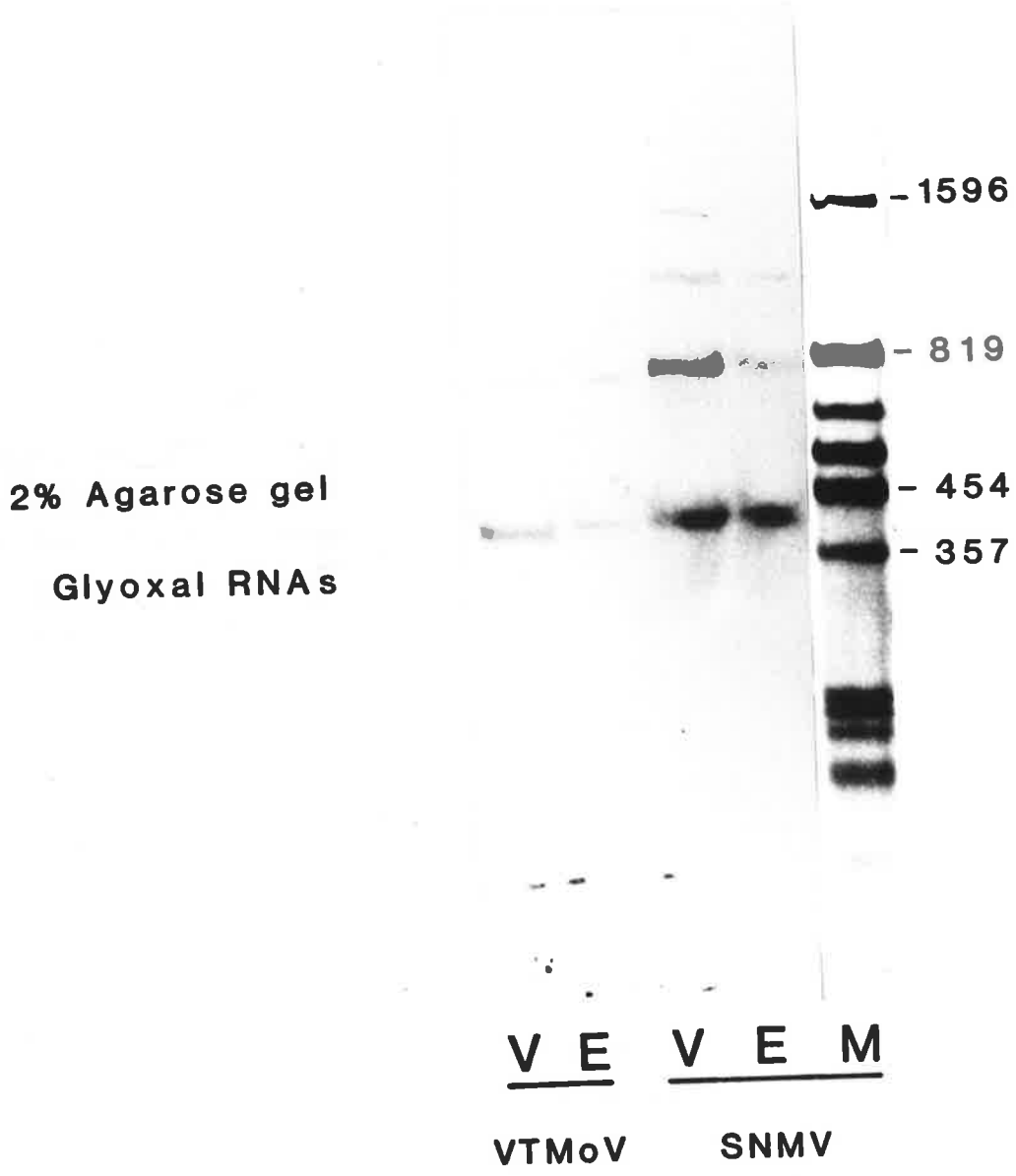
A. Analysis of VTMoV and SNMV RNA 2 sequences present in virus and infected tissues

Using recombinant M13 ss DNA containing cloned SNMV RNA 2 sequences, a ^{32}P-cDNA was synthesised and isolated as a probe specific for VTMoV and SNMV RNA 2. Figure 7-1 shows the pattern obtained when this ^{32}P-probe was used to detect VTMoV and SNMV RNA 2 sequences present in RNAs extracted from virions and infected plants. It can be seen that for both viruses, multimeric series of (+) RNAs are found both encapsidated and in tissue extracts. In addition, the dimeric forms of VTMoV and SNMV are detectable by

Figure 7-1 Multimeric RNAs containing VTMoV and SNMV RNA 2 sequences.

Nucleic acids were extracted from either purified virus (V) or virus infected plant tissue (E) for both VTMoV and SNMV. The nucleic acids were glyoxal treated, electrophoresed on a 2% agarose gel and subsequently transferred to nitrocellulose. VTMoV and SNMV RNA 2 sequences were detected using a ^{32}P-cDNA probe prepared from a cloned SNMV RNA 2 sequence as described in the text. 3'-^{32}P-labelled Hpa II cut M13 mp7 RF was contransferred to provide size markers.

PLUS OLIGOMERS OF VIRUSOIDS



toluidine blue staining of virion RNAs fractionated by denaturing polyacrylamide gel electrophoresis (results not shown).

DISCUSSION

A. Multimers of VTMoV and SNMV RNA 2

From blot hybridization experiments, such as that shown in Figure 7-1, it is apparent that multimeric series of RNAs containing VTMoV and SNMV RNA 2 sequences are found both packaged in virions and in infected plant tissues, as is the case for TobRV satellite RNA. On this evidence, it seems likely that the virusoids of VTMoV and SNMV replicate via a rolling-circle type mechanism similar to that proposed for TobRV satellite RNA (Kiefer et al., 1982) although the existence and properties of multimeric complementary RNAs and ds RNAs have yet to be determined.

Rolling-circle mechanisms require a circular template to allow transcription of multimeric RNA intermediates. Therefore, such a model would require that at some stage during replication, the linear TobRV satellite RNA be ligated to produce a circular template molecule. As one of the final steps in replication, unit-length linear virusoid or satellite RNA must be produced by either specific transcription or cleavage of multimeric RNAs. The 5'-hydroxyl and 3'-phosphate

groups present on TobRV satellite RNAs (Kiefer et al., 1982; G. Bruening, personal communication) suggest that these molecules are produced by specific cleavage rather than being primary transcripts. In the case of virusoids, the unit-length linear molecules must also be ligated to produce the final circular product. Thus, it is feasible that the TobRV satellite RNAs are simply defective in ligation and correspond to linear RNA intermediates in virusoid replication which, in contrast, are capable of circularization.

B. A possible site for RNA processing

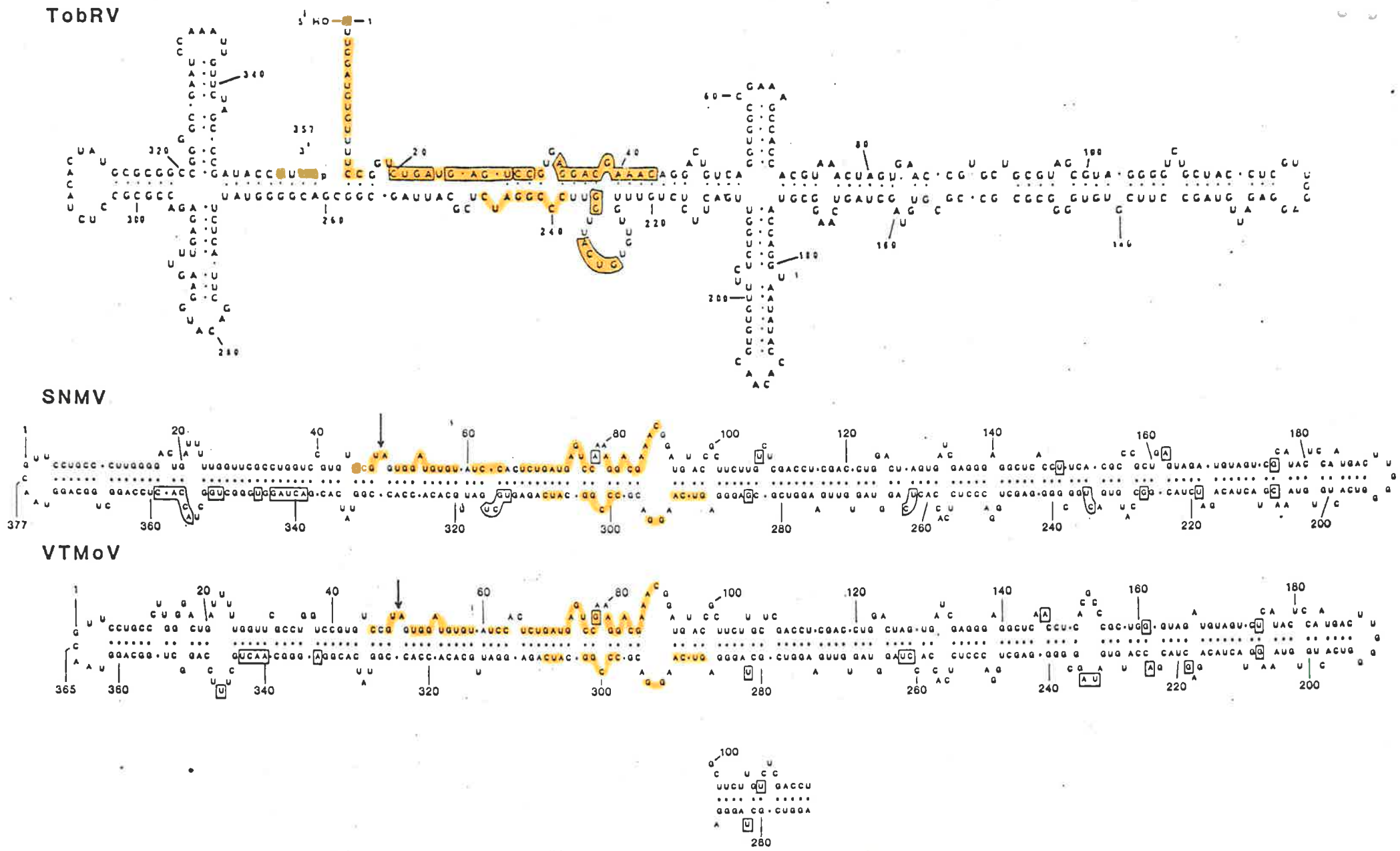
The 5' terminus of TobRV satellite RNA is adjacent in the molecule to the central conserved sequences shared with the virusoids of VTMoV, SNMV and SCMoV. When the sequences of these molecules are aligned as in Figure 7-2, extensive sequence homology between TobRV satellite and VTMoV and SNMV RNA becomes apparent. Homologous sequences extend from the central conserved regions to residues corresponding to the 5' terminus of TobRV satellite RNA (VTMoV and SNMV RNAs 2 residues 49) and include several residues corresponding to the 3' terminus. Therefore, it is proposed that the TobRV satellite and virusoid RNAs are produced by cleavage of multimeric RNA precursors at sites corresponding to between residues 357 and 1 for TobRV

Figure 7-2 A possible site for RNA processing.

A) The proposed structures of TobRV satellite RNA, SNMV and VTMoV RNAs 2 are shown with the nucleotide sequences conserved between these 3 RNAs indicated in colour. These sequences consist of those shown in Figures 6-6 and 6-7, and include additional sequences not shared by SCMoV RNAs 2 and 2'.

B) Comparison of the 3' and 5' proximal, and central conserved regions of TobRV with the corresponding regions in SNMV and VTMoV RNA 2 and SCMoV RNAs 2 and 2'. Regions of the RNAs are shown in linear form, and proposed sites for RNA processing of SNMV, VTMoV and SCMoV virusoids. Conserved residues are indicated in colour.

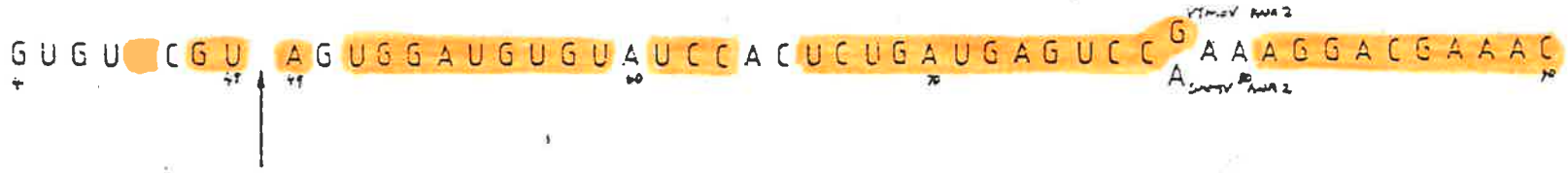
A



B

TobRV satellite

VTMoV and SNMV RNA 2



SCMoV RNAs 2 and 2'

satellite RNA and between residues 48 and 49 for VTMoV and SNMV RNAs 2. Unit length linear VTMoV and SNMV RNAs 2 would then be ligated to produce the mature circular forms of the RNAs. Implicit in this proposal is the assumption that the conserved sequences surrounding these putative sites for RNA processing are in some way functional, perhaps in determining the specific sites of cleavage. The conservation of these sequences in RNAs from different viruses may suggest their interaction with host, rather than viral, components (TobRV, VTMoV and SNMV share common host plants, such as Nicotiana clevelandii).

Interestingly, during sequence determination of VTMoV and SNMV RNAs 2, essentially complete termination of reverse transcription was observed at positions corresponding to residues 49 of the RNA templates. This was seen whether intact RNA 2 (see Figure 2-8) or purified linear RNA fragments were used as templates, and was presumably due to the presence of sequences and/or secondary structures capable of causing reverse transcriptase to chain terminate. For example, an 8-base-pair stem 3-base-pair loop structure can be formed in VTMoV and SNMV RNA 2 at residues 40 to 58, however the same structure cannot be formed at the corresponding sequences of TobRV satellite RNA (or SCMoV). It is unknown whether the precise coincidence

of sites for termination of reverse transcription with predicted processing sites in VTMoV and SNMV RNA 2 is a product of chance or, perhaps indirectly, of function.

In contrast, SCMoV RNA 2 and RNA 2' share little sequence homology with the other RNAs (Figure 7-2) outside the central conserved region, with the exception of several residues approximately corresponding in location to the 5' terminus of TobRV satellite RNA. Based solely on this limited sequence and structural homology, possible processing sites for SCMoV RNA 2 and RNA 2' are between residues 62 and 63 in each molecule. The lesser extent of sequence homology between the SCMoV virusoids and the other small viral RNAs (Figure 7-2) may be related to the limited and exclusive host range of SCMoV, which is not known to share host plant species with TobRV, VTMoV or SNMV (Francki et al., 1983). Thus functional nucleotide sequences might vary to accomodate the different requirements of host components in different species.

C. Viroid, virusoid and satellite RNAs

It now appears that there exists in plants a range of replicating RNA species which overall share many common features. Thus viroids, virusoids and TobRV satellite RNA all consist of ss DNA species of between 240 and 400 residues which, except that of TobRV

satellite, are rod-like base-paired circular molecules. These RNA species do not appear to code for functional polypeptide translation products, but replicate through multimeric RNA intermediates which are probably transcribed by a rolling-circle type mechanism. The unit-length progeny species must be produced by specific transcription or processing events and then, except in the case of TobRV satellite, ligated.

Given these similarities, the RNAs fall into one of two classes. The first contains viroids which are characteristically naked and capable of independent replication. The second consists of encapsidated RNAs, like those of VTMoV and SNMV which appear to contribute some function to a viral genome, or like those of TobRV and LTSV which are satellites. The members of each of these groups share at least some conserved sequences with others of the same group (see Chapter 4 and 6), and overall share remarkable conservation of regions central to their native structures (with the notable exception of LTSV RNA 2). Furthermore the pentanucleotide sequence GAAAC is present on all sequenced viroids, virusoids (except LTSV RNA 2) and TobRV satellite RNA, and is located within the central conserved regions of these molecules.

A single question looms from this tangle of observatons. Do viroids share common functions, and

perhaps origins, with the second group of encapsidated RNAs? The possible involvement of conserved virusoid and satellite sequences in interaction with plant host components was inferred from data presented in Chapters 6 and 7, and viroids appear to rely entirely on host components for replication. It therefore seems reasonable to suggest that these biologically disparate RNA species may share some common mechanisms in replication which involve functionally similar, if not identical, host components. The two groups of RNA species may be derived from a common ancestral species or alternatively be products of convergent evolution.

These suggestions are of course based on inference rather than direct evidence as the approach taken in this work allows only a glimpse of the functions and origins of these molecules as reflected in their comparative structures. Confirmation or denial of these possibilities and, ultimately, answering of the three questions originally posed in Chapter 1 will rely upon studies of the host and viral components involved in the replication of these RNAs, rather than the RNAs themselves.

REFERENCES

Barrell, B. G. and Sanger, F. (1969) *FEBS Lett.* 3, 275-278.

Bigornia, A. E. (1977) *Philipp. J. Coconut Stud.* 2, 5-33.

Birnboim, H. C. and Doly, J. (1979) *Nuc. Acids Res.* 9, 6527-6537.

Boccardo, G., Beaver, R. G., Randles, J. W. and Imperial, J. S. (1981) *Phytopath.* 71, 1104-1107.

Boege, F., Rohde, W. and Sänger, H. L. (1982) *Biosci. Rep.* 2, 185-194.

Branch, A. D., Robertson, H. D. and Dickson, E. (1981) *Proc. Natl. Acad. Sci. USA* 78, 6381-6385.

Brederode, F. TH., Koper-Zwarthoff, E.C. and Bol, J. F. (1980) *Nuc. Acids Res.* 8, 2213-2223.

Brierly, P. and Smith, F. F. (1949) *Phytopath.* 39, 501.

Brierly, P. and Smith, F. F. (1951) *Plant Dis. Reptr.* 35, 524-526.

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

Camacho, H. A. and Sänger, H. (1982a) *Arch. Virol.* in press.

Camacho, H. A. and Sänger, H. (1982b) *Arch. Virol.* in press.

Conjero, V. and Semancik, J. S. (1977) *Virology* 77, 221-232.

Conjero, V., Picazo, I. and Segado, P. (1979) *Virology* 97, 454-456.

Davies, J. W., Kaesberg, P. and Diener, T. O. (1974) *Virology* 61, 282-286.

De Wachter, R. and Fiers, W. (1972) *Anal. Biochem.* 49, 184-197.

Diener, T. O. (1971a) *Virology* 43, 75-89.

Diener, T. O. (1971b) *Virology* 45, 411-428.

Diener, T. O. (1972) *Virology* 50, 606-609.

Diener, T. O. (1979) "Viroids and Viroid Diseases" Wiley Interscience, New York.

Diener, T. O. (1981) *Proc. Natl. Acad. Sci. USA* 78, 5014-5015.

Diener, T. O. and Lawson, R. H. (1973) *Virology* 51, 94-100.

Diener, T. O. and Raymer, W. B. (1967) *Science* 158, 378-381.

Diener, T. O. and Raymer, W. B. (1969) *Virology* 37, 351-366.

Diener, T. O. and Smith, D. R. (1975) *Virology* 63, 421-427.

Diener, T. O., Schneider, I. R. and Smith, D. R. (1974) *Virology* 57, 577-581.

Dimock, A. W. (1947) *N.Y. St. Flower Grower's Bull.* 26, 2.

Domdey, H., Jank, P., Sanger, H. L and Gross, H. J. (1978) *Nuc. Acids Res.* 5, 1221-1236.

Domingo, E., Sabo, D., Taniguchi, T and Weissman, C. (1978) *Cell* 13, 735-744.

Donis-Keller, H. (1980) *Nuc. Acids Res.* 8, 3133-3142.

Donis-Keller, H., Maxam, A. M. and Gilbert, W. (1977) *Nuc. Acids Res.* 4, 2527-2538.

Donis-Keller, H., Browning, K. S. and Clark, J. M. (1981) *Virology* 110, 43-54.

Efstratiadis, A., Vournakis, J. N., Donis-Keller., Chaconas, G., Dougall, D. K. and Kafatos, F. C. (1977) *Nuc. Acids Res.* 4, 4165-4174.

Flores, R., Chroboczek, J. and Semancik, J. S. (1978) *Physiol. Plant Pathol.* 13, 193-201.

Francki, R. I. B., Randles, J. W., Hatta, T., Davies, C., Chu, P. W. G. and Mclean, G. D. (1983) in press.

Frisby, D. (1977) *Nuc. Acid Res.* 4, 2975-2995.

Goodman, H. M. and MacDonald, R. J. (1979) *Methods Enzymol.* 68, 499-560.

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

Gould, A. R. (1981) *Virology* 108, 123-133

Gould, A. R. and Hatta, T. (1981) *Virology* 109, 137-147.

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

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

Gould, A. R., Francki, R. I. B. and Randles, J. W. (1981) *Virology* 110, 420-426.

Grill, L. K. and Semancik, J. S. (1978) *Proc. Natl. Acad. Sci. USA* 75, 896-900.

Gross, H. J. and Riesner, D. (1980) *Angew. Chem. Int. Ed. Engl.* 19, 231-243.

Gross, H. J., Domdey, H., Sänger, H. L. (1977) *Nuc. Acids Res.* 4, 2021-2018.

Gross, H. J., Domdey, H., Lossow, C., Jank, P., Raba, M., Alberty, H. and Sänger, H. L. (1978) *Nature* 273, 203-208.

Gross, H. J., Krupp, G., Domdey, H., Raba, M., Alberty, H., Lossow, C. H., Ramm, K. and Sänger H. L. (1982) *Eur. J. Biochem.* 121, 249-257.

Hadidi, A., Cress, D. E. and Diener, T. O. (1981) *Proc. Natl. Acad. Sci. USA* 78, 6932-6935.

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

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

Haseloff, J., Mohamed, N. A. and Symons, R. H. (1982) *Nature* 299, 316-322.

Henco, K., Riesner, D. and Sänger, H. L. (1977) *Nuc. Acids Res.* 4, 177-194.

Holland, J. J., Grabau, E. A., Jones, C. L. and Semler,

B. L. (1979) *Cell* 16, 495-504.

Hollings, M. and Stone, O. M. (1973) *Ann. Appl. Biol.* 74, 333-348.

Hollings, M., Stone, O. M., Barton, R. J. and Greber, R. S. (1979) *Rep. Glasshouse Crops Res. Inst. 1978*. 150-151.

Imperial, J. S., Rodriguez, J. B. and Randles, J. W. (1981) *J. Gen. Virol.* 56, 77-85.

Jones, A. T., Mayo, M. A. and Duncan, G. H. (1983) *J. Gen. Virol.* 64, in press.

Keller, J. R. (1953) *Cornell Univ. Agric. Exp. Stan. Memoir* 324, 40.

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

Klump, H., Riesner, D. and Sänger, H. L. (1978) *Nuc. Acids Res.* 5, 1581-1587.

Konarska, M., Filipowicz, W., Domdey, H. and Gross, H. J. (1981) *Eur. J. Biochem.* 114, 221-227.

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

Kozak, M. (1979) *Nature* 280, 82-85.

Kramer, F. R. and Mills, D. R. (1978) *Proc. Natl. Acad. Sci. USA* 75, 5334-5338.

Krupp, G. and Gross, H. J. (1979) *Nuc. Acids Res.* 6, 3481-3490.

Langowski, J., Henco, K., Riesner, D. and Sänger, H. L. (1978) *Nuc. Acids Res.* 5, 1589-1610.

Lockard, R. E., Alzner-Deweerd, B., Heckman, J. E.,

MacGee, J., Tabor, M. W. and Raj Bhandary, U. L. (1978) *Nuc. Acids Res.* 5, 37-55.

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

Maniatis, T. and Efstratiadis, A. (1980) *Methods Enzymol.* 65, 299-305

Maxam, A. M. and Gilbert, W. (1980) *Methods Enzymol.* 65, 499-560.

Mayo, M. A., Barker, H. and Harrison, B. D. (1979a) *J. Gen. Virol.* 43, 603-610.

Mayo, M. A., Barker, H. and Harrison, B. D. (1979b) *J. Gen. Virol.* 43, 735-740.

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

Miyazaki, M. (1974) *J. Biochem.* 75, 1407-1410.

Mohamed, N. A., Haseloff, J., Imperial, J. S. and Symons, R. H. (1982) *J. Gen. Virol.* 63, 181-188.

Mülbach, H. P. and Sänger, H. L. (1979) *Nature* 278, 185-188.

Niblett, C. L., Dickson, E., Horst, R. K. and Romaine, C. P. (1980) *Phytopath.* 70, 610-615.

Ocfemia, G. P. (1937) *Philipp. Agric.* 26, 338-340.

Olson, J. C. (1949) *Bull. Chrysanthemum Soc. Amer.* 17, 2-9.

Owens, R. A. and Diener, T. O. (1982) *Proc. Natl. Acad. Sci. USA* 79, 113-117.

Owens, R. A., Smith. D. R. and Diener, T. O. (1978) *Virology* 89, 388-394.

Palukaitis, P. and Symons, R. H. (1980) *J. Gen. Virol.* 46, 477-489.

Rackwitz, H. R., Rohde, W. and Sänger, H. L. (1981) *Nature* 291, 297-301.

Randles, J. W. (1975) *Phytopath.* 65, 163-167.

Randles, J. W. and Hatta, T. (1979) *Virology* 96 47-53.

Randles, J. W. and Palukaitis, P. (1979) *J. Gen. Virol.* 43, 649-662.

Randles, J. W., Rillo, E. P. and Diener, T. O. (1976) *Virology* 74, 128-139.

Randles, J. W., Boccardo, G. and Imperial, J. S. (1980) *Phytopath.* 70, 185-189.

Randles, J. W., Davies, C., Hatta, T., Gould, A. R. and Francki, R. I. B. (1981) *Virology* 108, 111-122.

Raymer, W. B. and Diener, T. O. (1969) *Virology* 37, 343-350.

Riesner, D., Henco, K., Rokohl, U., Klotz, G., Kleinschmidt, A. K., Gross, H. J., Domdey, H. and Sänger, H. L. (1979) *J. Mol. Biol.* 133, 85-115.

Rohde, W. and Sänger, H. L. (1981) *Biosci. Rep.* 1, 327-336.

Rubin, G. M. (1973) *J. Biol. Chem.* 248, 3860-3875.

Sanger, F., Nicklen, S. and Coulson, A. R. (1977) *Proc.*

Natl. Acad. Sci. USA 74, 5463-5467.

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

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

Sänger, H. L. (1972) *Adv. Biosci.* 8, 103-116.

Sänger, H. L., Klotz, G., Riesner, D., Gross, H. J. and Kleinschmidt, A. K. (1976) *Proc. Natl. Acad. Sci. USA* 73, 3852-3856.

Schneider, I. R. (1969) *Science* 166, 1627-1629.

Schneider, I. R. (1971) *Virology* 45, 108-122.

Schneider, I. R. (1977) Defective plant viruses. In "Beltsville Symposia in Agriculture" (J. A. Romberg, J. D. Anderson and R. L. Powell, eds.) 1, 201-219.

Schneider, I. R. and Thompson, S. M. (1977) *Virology* 78, 453-462.

Semancik, J. S. and Weathers, L. G. (1968) *Virology* 36, 326-328.

Semancik, J. S. and Weathers, L. G. (1970) *Phytopath.* 60, 732-736.

Semancik, J. S. and Weathers, L. G. (1972a) *Virology* 47, 456-466.

Semancik, J. S. and Weathers, L. G. (1972b) *Nature New Biology* 237, 242-244.

Semancik, J. S., Morris, T. J., Weathers, L. G., Rodorf,

B. F. and Kearns, D. R. (1975) *Virology* 63, 160-167.

Semancik, J. S., Conjero, V. and Gerhart, J. (1977) *Virology* 80, 218-221.

Sippel, A. E. (1973) *Eur. J. Biochem.* 37, 31-40.

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

Sogo, J. M., Schneider, I. R. and Koller, Th. (1974) *Virology* 57, 459-466.

Spohr, G., Mirault, M., Imaizumi, T. and Schrerrer, K. (1976) *Eur. J. Biochem.* 62, 313-322.

Symons, R. H. (1977) *Nuc. Acids Res.* 4, 4347-4355.

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

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

Symons, R. H. (1981) *Nuc. Acids Res.* 9, 6527-6537.

Takahashi, T. and Diener, T. O. (1975) *Virology* 64, 106-114.

Taylor, J. M., Illmensee, R. and Summers, J. (1976) *Biochem. Biophys. Acta.* 442, 324-330.

Tien-Po, Davies, C., Hatta, T. and Francki, R. I. B. (1981) *FEBS Lett.* 132, 353-356.

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

Tinoco, I. Jr., Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M. and Gralla, J. (1973) *Nature* 246, 40-41.

Visvader, J. E., Gould, A. R., Bruening, G. E. and

Symons, R. H. (1982) *FEBS Lett.* 137, 288-292.

Wild, U., Ramm, K., Sanger, H. L. and Riesner, D. (1980) *Eur. J. Biochem.* 103, 227-235.

Wurst, R. M., Vournakis, J. N. and Maxam, A. M. (1978) *Biochemistry* 17, 4493-4499

Zaitlin, M., Niblett, C. L., Dickson, E. and Goldberg, R. B. (1980) *Virology* 104, 1-9.

Zelazny, B. (1979) *Acta. Phytopath. Acad. Sci. Hung.* 14, 115-126.

Zelazny, B. and Niven, B. S. (1980) *Plant Disease* 64, 841-842.

Zelcer, A., Zaitlin, M., Robertson, H. D. and Dickson, E. (1982) *J. Gen. Virol.* 59., 139-148.

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