



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

## Subgenomic RNAs with Nucleotide Sequences Derived from RNAs 1 and 2 of Cucumber Mosaic Virus Can Act as Messenger RNAs *in Vitro*

KARL H. J. GORDON AND ROBERT H. SYMONS

*Adelaide University Centre for Gene Technology, Department of Biochemistry, University of Adelaide, Adelaide, South Australia 5000, Australia*

*Received August 13, 1984; accepted November 11, 1984*

Encapsidated RNAs of cucumber mosaic virus (CMV) were analyzed by hybridization to specific probes after gel electrophoresis. [<sup>32</sup>P]-complementary DNA (cDNA) probes were prepared by transcription of genomic RNA 1 and RNA 2 nucleotide sequences that had been cloned in a bacteriophage M13 vector. Probes that correspond to unique sequences near the 3' ends of RNA 1 and RNA 2 revealed over 20 smaller RNAs. The subgenomic RNAs derived from each genomic RNA were analyzed more definitively by hybrid selection from total encapsidated RNA, using minus DNA clones derived from sequences in either RNA 1 or RNA 2, and a cDNA probe for the 3' sequence conserved among all the genomic RNAs. Different patterns of over 20 minor RNA species, which were 3'-coterminal with RNAs 1 and 2, were detected, and they were reproducible irrespective of the host, cucumber or *Nicotiana clelandii*, from which the virus was isolated. The same RNA patterns were found in RNA extracted from the particulate fraction of CMV-infected cucumber or *N. clelandii*. In order to determine whether the subgenomic RNAs could function as messenger RNAs, hybrid-selected RNAs were tested by *in vitro* translation, using the rabbit reticulocyte lysate. The subgenomic RNAs from RNA 1 produced over 10 major polypeptides from  $M_r$  27,000 to  $M_r$  90,000 all of which could be translated from a few RNA species over about 2,300 nucleotides long. The 3'-coterminal subgenomic RNAs derived from RNA 2 gave less than 10 products from  $M_r$  17,000 to  $M_r$  85,000. The smallest product ( $M_r$  17,000) was produced by an RNA about 880 nucleotides long, whereas longer RNAs from 1400 to 2500 nucleotides were efficient mRNAs for polypeptides from  $M_r$  30,000 up to the largest translation products consistent with the size of the RNA. © 1985 Academic Press, Inc.

### INTRODUCTION

The genomic RNAs of plant RNA viruses are usually capable of directly expressing at least a portion of their genetic information through translation (Davies and Hull, 1982). However, they are subject to the usual constraints imposed upon the translation of eucaryotic mRNAs. In general, only the reading frame starting nearest the 5' end of the mRNA is available for translation (Kozak, 1981, 1984), so that open reading frames and initiation codons nearer the 3' end of a polycistronic mRNA cannot be expressed directly. A common strategy adopted by these viruses is to express internal and 3'-terminal genes by means of subgenomic RNAs which are derived from, and 3'-coterminal

with, the genomic RNAs. These subgenomic RNAs are used to express structural (coat protein) genes for many plant RNA viruses (reviewed by Davies and Hull, 1982; Bruening, 1981). Animal RNA viruses commonly employ subgenomic RNAs to express both structural and nonstructural genes, e.g., the 26 S RNA of the alphaviruses (Sawicki *et al.*, 1978; Riedel *et al.*, 1982), the mRNAs derived from RNAs 7 and 8 of influenza virus (Lamb and Choppin, 1983) and the mRNAs of the coronaviruses (Siddell *et al.*, 1982). At least one plant viral gene for a nonstructural protein, the  $M_r$  30,000 protein encoded by tobacco mosaic virus (TMV RNA, is also expressed by a subgenomic mRNA (Hirth and Richards, 1981). Hunter *et al.* (1983) have shown the existence of a fam-

ily of 3'-coterminal subgenomic RNAs which are capable *in vitro* of expressing the carboxyterminal regions of this TMV reading frame to give related proteins differing at their amino termini. It is still unknown whether this pattern also applies to other genes of TMV or other viruses, although other subgenomic RNAs have been detected for TMV (Goelet and Karn, 1982; Palukaitis *et al.*, 1983). It is also not known whether the multipartite plant RNA viruses use subgenomic RNAs to express genes other than for the coat protein, but Bol *et al.*, (1976) observed such possible subgenomic RNAs from alfalfa mosaic virus (AlMV).

The genome of cucumber mosaic virus (CMV) consists of three single-stranded, plus-sense RNAs designated RNAs 1 to 3. RNA 1 is 3387 nucleotides long (Rezaian *et al.*, 1985), RNA 2 is 3035 nucleotides long (Rezaian *et al.*, 1984), and RNA 3 2193 nucleotides long (Gould and Symons, 1982). All these RNAs are required for infectivity (Peden and Symons, 1973; Lot *et al.*, 1974) and are capable of acting as monocistronic mRNAs in cell-free translation systems (Schwinghamer and Symons, 1977; Gordon *et al.*, 1982). A single polypeptide of molecular weight 110,791, which must correspond to the *in vitro* translation product (estimated size  $M_r$  95,000), is encoded by 88% of RNA 1. RNA 2 encodes a polypeptide of molecular weight 94,333, which must correspond to the *in vitro* translation product (estimated  $M_r$  110,000) and is encoded by 83% of the RNA. RNA 3 is dicistronic; the 5'-terminal gene encodes the 3a protein of 333 amino acids (calculated  $M_r$  36,700) and the 3'-terminal gene encodes the coat protein of 236 amino acids (calculated  $M_r$  26,200). The coat protein is translated from a subgenomic RNA (RNA 4) of 1027 nucleotides (Gould and Symons, 1982) which is not required for infectivity, but is present as a major species in CMV encapsidated RNA.

Although the CMV genomic RNAs are efficient mRNAs *in vitro*, the translation products of RNAs 1, 2, and 3 have never been detected *in vivo* (Gonda and Symons, 1979; Gordon *et al.*, 1982; Symons *et al.*,

1982; and unpublished). This raises the interesting possibility that the genes on these RNAs may be expressed to some extent as partial polypeptides, possibly through the use of subgenomic mRNAs. The study of subgenomic RNAs derived from the CMV RNAs, other than the coat protein mRNA (RNA 4), is therefore of special interest.

In this paper we report the presence of subgenomic RNAs derived from both RNAs 1 and 2 of CMV. The RNAs were detected by hybridization after gel electrophoresis and transfer by blotting ("Northern" analysis), using [ $^{32}$ P]cDNA probes prepared from DNA templates representing segments of RNA 1 or RNA 2 cloned in a bacteriophage M13 vector. The subgenomic RNAs also were partially purified by hybrid selection using cloned minus-sense sequences and were translated *in vitro*.

## MATERIALS AND METHODS

### Materials

Klenow fragment of *E. coli* DNA polymerase was obtained from Boehringer (Mannheim) or BRESA (Adelaide). Deoxy- and dideoxynucleoside triphosphates were from Sigma. M13-specific oligonucleotide (5' GTAAAACGACGCCAGT 3') and restriction endonucleases were obtained from New England Biolabs. [ $\alpha$ - $^{32}$ P]dATP and [ $\alpha$ - $^{32}$ P]dCTP (sp act of 2.4 Ci/mmol) were from BRESA. Low-melting-point agarose was obtained from BRL (Maryland). Nitrocellulose and Genescreen membrane filters were from Sartorius and New England Nuclear, respectively.

### Virus and RNA

Q-CMV (Francki *et al.*, 1966) was grown in either cucumber (*Cucumis sativus* L. cv. Supermarket) or *Nicotiana clevelandii* A. Gray. Viral RNA was prepared from purified virus as described by Peden and Symons (1973). RNAs 1, 2, and 3 were purified as described by Symons (1978). For some experiments, RNA was obtained from the particulate fraction of the CMV-induced RNA-dependent RNA polymerase

(Gill *et al.*, 1981) by extraction with phenol and ethanol precipitation.

*cDNA Clones from RNAs 1 and 2*

The clones (corresponding to segments of RNAs 1 or 2) used to synthesize [<sup>32</sup>P]cDNA probes and in hybrid-selection experiments for this study were kindly supplied by Ali Rezaian, Rhys Williams, and Allan Gould (Rezaian *et al.*, 1984, 1985). Either plus or minus clones in the bacteriophage vector M13mp7 were chosen from the library available. Their sizes and locations with respect to the RNA sequence are shown in Fig. 1.

*Northern Hybridization Analysis*

(i) [<sup>32</sup>P]cDNA probes. [<sup>32</sup>P]cDNA probes of high specific activity were synthesized from single-stranded M13 DNA containing inserts of plus polarity, using [ $\alpha$ -<sup>32</sup>P]dATP, [ $\alpha$ -<sup>32</sup>P]dCTP, and the Klenow fragment of

DNA polymerase, as described by Bruening *et al.* (1982). Restriction endonuclease *Eco*RI was used to excise the probes, which were purified by electrophoresis on 6% polyacrylamide, 7 M urea gels, in 90 mM Tris-borate, pH 8.3, 2 mM EDTA.

(ii) *Gel electrophoresis and blotting of RNA.* Nucleic acid samples were denatured with 1 M glyoxal in a total volume of 10  $\mu$ l of 10 mM sodium phosphate, pH 6.5, 0.1 mM EDTA, at 50° for 15 min (McMaster and Carmichael, 1977), before electrophoresis on 1.5% agarose gels in 10 mM Na phosphate, pH 6.5, 0.1 mM EDTA. The gel was both stained with ethidium bromide (5  $\mu$ g/ml) and destained in the electrophoresis buffer. The RNA was transferred to nitrocellulose or Genescreen filters by blotting, before hybridization with the [<sup>32</sup>P]cDNA probes as described by Thomas (1983). The filters were then washed in 1 $\times$  SSC (0.15 M NaCl, 0.015 M Na citrate, pH 7.0), 0.1% sodium dodecyl

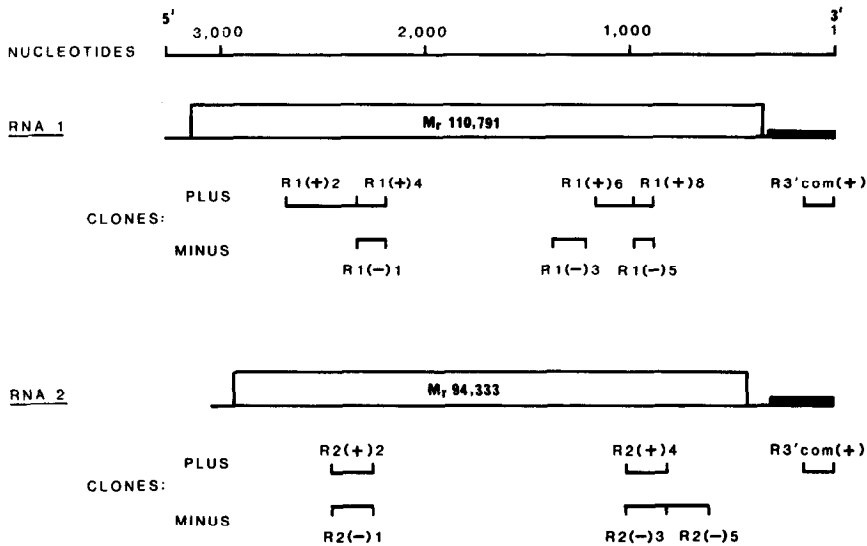


FIG. 1. Diagram showing the locations and lengths of clones used to analyze subgenomic RNAs from CMV RNAs 1 and 2. The scale in nucleotides is numbered from the 3' end. The RNA 1 sequence data are from Rezaian *et al.* (1985) and that of RNA 2 from Rezaian *et al.* (1984). Open reading frames are indicated by long open boxes and include the calculated sizes of each translation product. The 3'-conserved sequence is indicated by a heavy line. The positions of the clones used either for hybrid selection or to synthesize [<sup>32</sup>P]cDNA probes are shown. The clones are identified by the RNA from which they are derived, their orientation with respect to the RNA sequence and their relative order from the 5' end of the RNA, through the final digit given (even for (+) and odd for (-) clones).

sulfate (SDS) at room temperature, followed by 30 min in  $0.1\times$  SSC, 0.1% SDS at 50°. They were exposed to X-ray film at room temperature for autoradiography.

### Hybrid Selection of RNA

Appropriate bacteriophage M13 clones containing minus inserts (complementary to viral RNA) were grown in 500-ml cultures, and the DNA purified as described by Messing *et al.* (1981). From 0.3 to 0.5 mg of DNA was bound to twenty-four 3-mm diameter discs of nitrocellulose filters (Schleicher and Schuell, 0.2- $\mu$ m pore size) as described by Maniatis *et al.* (1982). Hybrid selection, using one such disc, was performed in 0.3 ml of a buffer containing 10 mM piperazine-*N,N'*-bis(2-ethanesulfonic acid) (PIPES)-KOH, pH 6.5, 0.4 M NaCl, 2 mM EDTA, 0.2% SDS, 65% (v/v) deionized formamide and up to 3 mg/ml CMV RNA in an eppendorf tube, for 3 hr at 50° (Maniatis *et al.*, 1982). Filters were washed at higher stringency than described by Maniatis *et al.* (1982): the washes comprised 6  $\times$  5 min in 10 mM Tris-HCl, pH 8.0, 0.15 M NaCl, 2 mM EDTA, 0.5% SDS at 65°, followed by 6  $\times$  5 min in 10 mM Tris-HCl, pH 8.0, 0.015 M NaCl, 2 mM EDTA, 0.5% SDS, at 50° (the last two minus SDS). The RNA was eluted by boiling in 2 mM EDTA, pH 8.0, for 2 min, recovered by ethanol precipitation, dried *in vacuo*, and resuspended in 0.1 mM EDTA. Aliquots were analyzed by RNA blotting (see above) or by *in vitro* translation (see below).

### In Vitro Translation of CMV RNAs

A cell-free lysate of rabbit reticulocytes was prepared and used for *in vitro* translation as described previously (Gordon *et al.*, 1982). The translation products were analyzed by discontinuous SDS polyacrylamide gel electrophoresis (Laemmli, 1970) and fluorography performed as described by Gill *et al.* (1981).

### Preparative Agarose Gel Electrophoresis of RNA

Total CMV RNA was fractionated on 1.5% low-melting-point agarose gels in 20

mM Tris-acetic acid, pH 7.5, 1 mM EDTA. The RNA bands were detected by staining briefly with 5  $\mu$ g/ml ethidium bromide, and excised. The gel slices were melted at 65° for 15 min and the RNA recovered by three phenol extractions as described by Kuhn *et al.* (1979).

## RESULTS

### Northern Analysis Discloses Subgenomic RNAs for CMV RNAs 1 and 2

Total, encapsidated CMV RNA was fractionated on agarose gels, blotted, and analyzed using [ $^{32}$ P]cDNA probes synthesized from clones of either RNA 1 or 2 sequences (Fig. 2). The locations of the clones chosen for this experiment are shown in Fig. 1. A cDNA probe from clone R1(+)<sub>8</sub>, which is located 902 to 963 nucleotides from the 3' end of RNA 1, disclosed about 20 RNA zones which migrated more rapidly than the RNA 1 zone (Fig. 2, track a). RNAs with sequences derived from RNA 2 were detected using a cDNA probe from clone R2(+)<sub>4</sub>, whose insert corresponds to the segment 813 to 1015 nucleotides from the 3' end of RNA 2. Over 20 RNAs were detected, all apparently smaller than RNA 2 and as small as 320 nucleotides (Fig. 2, track f).

Comparison of these smaller RNAs shown in tracks a and f of Fig. 2 with the main CMV RNA species detected by ethidium staining in the same tracks showed that in both cases some of the prominent RNAs detected by blotting and hybridization corresponded in position to the major stained RNAs. These apparent species (of 3000, 2200, and 1000 nucleotides in track 1, and of 2200, 1000, and 750 nucleotides in track f) could therefore have arisen from a shadow effect during electrophoresis. A similar phenomenon has been observed for TMV by Palukaitis *et al.* (1983).

In order to analyze the subgenomic RNAs derived from each genomic RNA more definitively in the absence of this electrophoretic artifact, it was necessary to eliminate the other genomic RNAs (and their related, shorter RNAs). This was achieved by hybrid selection from total



of the genomic RNAs, e.g., in the virion, rather than during their replication. There would then exist corresponding families of RNA fragments 5'-coterminally with the genomic RNAs. In order to characterize such putative RNAs, they were hybrid selected from total CMV RNA using minus clones R1(-)1 (RNA 1) or R2(-)1 (RNA 2), which are nearest the 5' ends of these RNAs, and analyzed using [<sup>32</sup>P]cDNA probes from clones R1(+)-2 (Fig. 2, track b) and R2(+)-2 (Fig. 2, track e), respectively. Both probes hybridized with the 3'-coterminally RNAs over about 2500 nucleotides in length, as expected, but showed different patterns for the smaller RNAs: about three RNAs of 1200 to 1500 nucleotides were present in the RNA 1 sample (track b), whereas there were about 10 from 800 to 1700 nucleotides in the RNA 2 sample (track e). Although these smaller RNAs do not appear to correspond to those expected if specific nicking is the cause of the 3'-coterminally subgenomic RNA patterns, it is possible that such nicking may occur more than once per molecule near the 5' end, resulting in no simple pair of related 5' and 3' segments. Hence, the situation with regard to specific 5'-terminal fragments was not clear and further work concentrated on the 3'-coterminally RNA fragments.

The patterns of the RNA species derived from either RNA 1 or RNA 2 were reproducible, irrespective of the host—cucumber or *N. clevelandii*—from which the virus was obtained. Furthermore, the same RNA patterns were detected in RNA extracted from the particulate fraction of CMV-infected cucumber or *N. clevelandii* plants, but not in extracts from uninfected plants (results not shown). These experiments have led to the conclusion that RNAs 1 and 2 each generate distinct families of about 20 or more specific, 3'-coterminally subgenomic RNAs, in addition to smaller numbers of subgenomic RNAs derived from the 5' ends of the genomic RNAs. All the discrete subgenomic RNAs detected in these experiments were distinct from a low background of randomly fragmented RNA.

### *In Vitro Translation of Hybrid-Selected RNA*

In order to determine whether the subgenomic RNAs could function as mRNAs, the hybrid-selected RNA was tested by *in vitro* translation, using the rabbit reticulocyte lysate. In control experiments, unfractionated viral RNA yielded four major translation products: the full-length translation products of each of the four major virus RNAs (Fig. 3, track a). Some minor bands were also visible, in addition to another major band of about  $M_r$  17,000. The presence of a number of minor translation products is in contrast to earlier translation studies of individual, purified genomic RNAs (Gordon *et al.*, 1982), where no significant early termination products were found. It seemed, therefore, that at least some of the minor bands could be translation products of the subgenomic RNAs. Note that a band, labeled E (Fig. 3, track a, of about  $M_r$  50,000), is an endogenous lysate product.

The full-length translation product of RNA 1 ( $M_r$  110,791, Rezaian *et al.*, 1985) is one of the major polypeptides visible in track a of Fig. 3, where it is labeled as  $M_r$  95,000, the size estimated from SDS gels. The translation products of the RNAs selected by hybridization to minus clone R1(-)5 from total virus RNA, however, contained many polypeptides, ranging in size from the full-length product down to  $M_r$  27,000 (Fig. 3, track b). A similar polypeptide pattern was observed using RNA selected by minus clone R1(-)1 (Fig. 3, track c). All these polypeptides are thus most likely to be translated from the 3'-coterminally RNAs over 2300 nucleotides long, which are able to hybridize with both clones R1(-)1 and R1(-)5 (see Fig. 1). Some of the translation products, including the full-length polypeptide, appeared in variable amounts; the full-length polypeptide was never the most abundant translation product observed. The reduced translation of this full-length polypeptide from the RNAs hybrid selected by clone R1(-)1, compared to those hybrid selected by clone R1(-)5, may be due to competition by the small RNAs (of about 1200 to

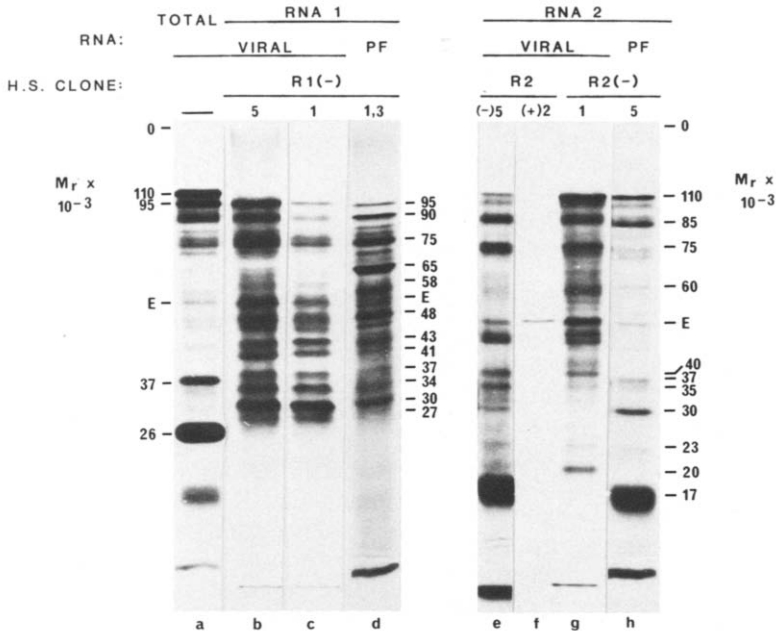


FIG. 3. *In vitro* translation products of hybrid-selected CMV RNAs. Track a shows the translation products obtained from total encapsidated RNA (150  $\mu\text{g}/\text{ml}$ ), using reticulocyte lysate. The other tracks show the translation products (RNA at 50  $\mu\text{g}/\text{ml}$ ) of RNA selected by hybridization to clones R1(-)5 (RNA 1, track b), R1(-)1 (RNA 1, track c), R2(-)5 (RNA 2, track e), R2(+2) (RNA 2, trace f), and R2(-)1 (RNA 2, track g). RNA from the particulate fraction (PF) of CMV-infected cucumbers was selected by hybridization to a mixture of clones R1(-)1 and R1(-)3 (RNA 1, track d) or to clone R2(-)5 (RNA 2, track h). Molecular weights of the translation products from CMV RNAs 1-4 are given to the left; those of the translation products in tracks b to d, next to track d; and those of the translation products in tracks e to h, on the right. The molecular weights of the RNA 1 and 2 full-length translation products are given as 95,000 and 110,000, respectively, since they migrate with these apparent sizes on SDS gels; the molecular weights of the RNA 3 and 4 translation products are the values calculated from the sequences. Translations were carried out in the reticulocyte lysate, as described under Methods, in the presence of [ $^{35}\text{S}$ ]methionine. E is an endogenous band due to the [ $^{35}\text{S}$ ]methionine.

1500 nucleotides—see track b, Fig. 2) which probably include the capped, 5' segment of RNA 1. In control experiments, the positive-insert clones R1(+4) and R1(+6) did not select any translatable RNA from total viral RNA (data not shown).

The RNAs hybrid selected from total RNA isolated from the particulate fraction of CMV-infected cucumber with clones R1(-)1 plus R1(-)3 also yielded many *in vitro* translation products. As shown in Fig. 3 (track d), these polypeptides corresponded in size to the many translation products already observed in tracks b and c but an additional major band was ob-

served ( $M_r$  65,000). Again, the full-length ( $M_r$  110,791) polypeptide was not the most abundant translation product. Identical translation product patterns were obtained with particulate fraction RNA obtained from CMV-infected cucumber or *N. cleavelandii* (data not given).

*In vitro* translation of the RNA selected from total viral RNA by hybridization to the clone R2(-)5 (i.e., RNA 2 and its subgenomic RNAs) yielded several products from  $M_r$  94,333 (the full-length translation product, Rezaian *et al.*, 1984, which is identified as  $M_r$  110,000, the size estimated from SDS gels) down to a major band at  $M_r$  17,000 (Fig. 3, track e). Similar



polypeptides were also produced by translation of RNA hybrid selected by clone R2(-)5 from the particulate fraction RNA (Fig. 3, track h), irrespective of the host plant involved (cucumber or *N. clevelandii*). The full-length product was always the longest polypeptide detected, but in variable amounts and never as the most abundant product. The most efficiently translated polypeptide was clearly that of  $M_r$  17,000. When total virus RNA was analyzed by hybrid selection with clone R2(-)1, which corresponds to a sequence located from 2261 to 2408 nucleotides from the 3' end of RNA 2, a different pattern of translation products was obtained (Fig. 3, track g). This included the full-length polypeptide as its major product, together with several other peptides over  $M_r$  40,000 in size. Polypeptides of  $M_r$  37,000 and  $M_r$  20,000 were very minor products. Clone R2(+2), with an insert complementary to that in clone R2(-)1, selected no translatable RNA at all (Fig. 3, track f). Note that the hybrid-selected RNAs were very efficiently depleted of the other CMV genomic RNAs, since their translation products were clearly absent (cf. track a in Fig. 3).

#### *Translation Products of Fractionated Subgenomic RNAs*

(i) *Translation products of subgenomic RNAs from RNA 1.* The above hybrid-selection data have shown that the subgenomic RNAs derived from CMV RNA 1 encode a number of *in vitro* translation products and suggest that the mRNAs for these proteins were the 3'-coterminal RNAs of over 2300 nucleotides in length, because they were able to hybridize to both clones R1(-)1 and R1(-)5. In order to test this, total virus RNA was first fractionated by gel electrophoresis. RNA in single fractions was then subjected to hybrid selection using clone R1(-)5, followed by *in vitro* translation (Fig. 4).

The RNA fraction with the greatest mRNA activity was that analyzed in track h of Fig. 4. This fraction was shown by Northern analysis, using a probe from clone R1(+8), to be enriched for RNAs

from 2100 to 2500 nucleotides in length and produced the same sized polypeptides, from  $M_r$  75,000 to  $M_r$  27,000, already detected in Fig. 3 (track b). The  $M_r$  90,000 polypeptide was not observed, presumably because its mRNA is over 2500 nucleotides long and therefore absent from this RNA fraction. The small polypeptides of  $M_r$  30,000,  $M_r$  43,000, and  $M_r$  48,000 shown in track f (Fig. 4) were produced by RNAs of 1600 to 1800 nucleotides. These RNAs were also present in smaller amounts among the RNAs producing the polypeptides shown in track h, probably due to aggregation of RNA on the nondenaturing gel. The small amounts of the  $M_r$  30,000 translation product in track e were translated from RNAs of 1400 to 1500 nucleotides in length; RNAs of these sizes also were present in the fraction analyzed in track h. These results therefore suggest that the smaller translation products (below  $M_r$  50,000 in size) can be translated either from the shorter subgenomic RNAs, or from the longer ones (of over 2100 nucleotides, depending on the polypeptide), e.g., through leaky initiation of translation of AUG codons occurring after that nearest the 5' end. The smallest RNAs of 800 to 1300 nucleotides, lacked any detectable mRNA activity (tracks a to d).

(ii) *Translation products of subgenomic RNAs from RNA 2.* The subgenomic RNAs from RNA 2 were fractionated as for RNA 1 (above), and the RNA in single fractions was hybrid selected using clone R2(-)5. *In vitro* translation of the RNAs (Fig. 5) yielded polypeptides similar in size to those shown in Fig. 3 (tracks e and h). The RNA fraction (Fig. 5a, track a) with the smallest RNAs present (shown by Northern analysis to be from 750 to 880 nucleotides long) contained a very efficient mRNA for a translation product of about  $M_r$  17,000. RNAs from 1000 to 1400 nucleotides gave small amounts of translation products of  $M_r$  23,000 and  $M_r$  30,000 (tracks b and c). The longer RNAs efficiently produced polypeptides of increasing size, from  $M_r$  30,000 to  $M_r$  85,000 and the full-length translation product (tracks d to h). Interestingly, these RNA fractions containing larger RNAs, from

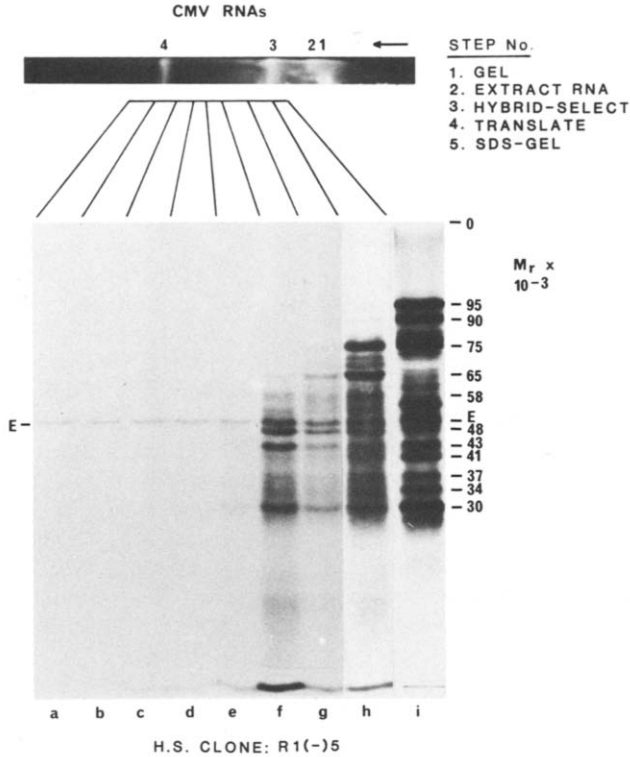


FIG. 4. Fractionation and translation of subgenomic RNAs from CMV RNA 1. Total encapsidated RNA (200  $\mu\text{g}$ ) was fractionated on a 1.5% low-melting-point agarose gel, which was then stained with ethidium bromide and photographed (one track shown at the top of the figure). The gel was sliced as shown, and the RNA was recovered by phenol extraction. Subgenomic RNAs derived from RNA 1 were selected from each fraction by hybridization to clone R1(-)5 (10  $\mu\text{g}$  DNA/fraction). Aliquots (30%) of the total hybrid-selected RNAs were translated *in vitro*, and the products were analyzed by 13% polyacrylamide-SDS gel electrophoresis. Tracks a to h refer to RNAs from the eight gel slices shown. Track i shows the translation products of RNA selected from total CMV RNA by hybridization to clone R1(-)5 (molecular weights shown on the right). E indicates the endogenous band due to [ $^{35}\text{S}$ ]methionine.

1400 to 2500 nucleotides, each included mRNAs for polypeptides from  $M_r$  30,000 up to the largest translation products consistent with the RNA size. This also may be due to leaky initiation of translation, as observed with the RNAs from RNA 1.

The possible mRNAs for the  $M_r$  17,000 translation product (Fig. 5a, track a) were shown by Northern analysis to be RNAs of 750, 800, and 880 nucleotides in length. The mRNA for the  $M_r$  17,000 polypeptide was selected by hybridization to DNA from clone R2(-)5 located 600 to 816 nucleotides from the 3' end of RNA 2 (Fig.

5, tracks a and j), but not from clone R2(-)3 (813 to 1015 nucleotides from the 3' end) (Fig. 5b, track i). The only open reading frame apparently available for the  $M_r$  17,000 protein corresponds to the carboxyterminal portion of the long open reading frame traversing RNA 2. This runs from an AUG codon 852 nucleotides from the 3' end of RNA 2 to the UGA codon at 426 nucleotides to produce a polypeptide of calculated size  $M_r$  16,008 (Rezaian *et al.*, 1984). It would thus be encoded by the 880-nucleotide RNA. We would expect this RNA to hybridize inefficiently to clone R2(-)3, since the overlap

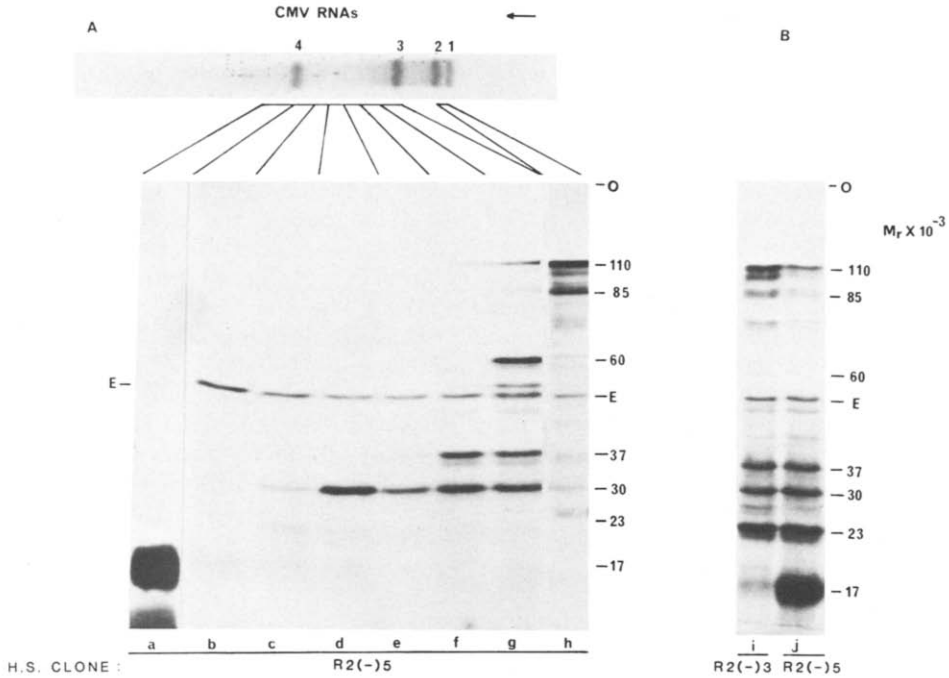


FIG. 5. Fractionation and translation of subgenomic RNAs from RNA 2. (A) Total CMV RNA (200  $\mu\text{g}$ ) was fractionated and recovered as described for Fig. 4. The preparative gel was sliced as indicated and subgenomic RNAs derived from RNA 2 were selected from the RNA fractions shown by hybrid selection with clone R2(-)5 (10  $\mu\text{g}$  DNA/fraction). Aliquots (30%) of the total RNAs selected from each of the fractions were translated and the products analyzed by 13% polyacrylamide SDS-gel electrophoresis (tracks a to h). Molecular weights of the major translation products are shown. E indicates the endogenous band due to [ $^{35}\text{S}$ ]methionine. (B) CMV RNA which cosedimented with RNA 4 on sucrose gradients was subjected to hybrid selection using clones R2(-)3 (track i) or R2(-)5 (track j) from RNA 2. The hybrid-selected RNAs were translated *in vitro* and the products analyzed by SDS-gel electrophoresis. Molecular weights of the major translation products are indicated.

would be only about 70 nucleotides. The cDNA probe from clone R2(+)+4 (which is complementary to clone R2(-)3) hybridized very poorly to this RNA (Fig. 2, track f). All these observations are consistent with the conclusion that the 880-nucleotide RNA is the mRNA for the  $M_r$  17,000 polypeptide. The next largest open reading frame on RNA 2 encodes a polypeptide of 100 amino acids (calculated  $M_r$  11,000), starting 627 nucleotides from the 3' end. This could be translated from the 750 nucleotide RNA (i.e., from the 3'-coterminous component if there is a mixture in this band) and may therefore be present below the  $M_r$  17,000 translation product (Fig. 5, track a). The variation in produc-

tion of some polypeptides (e.g., those of  $M_r$  23,000 to  $M_r$  60,000) may be due to competition in different subgenomic RNA fractions.

#### DISCUSSION

This work has shown that many subgenomic RNAs are derived from CMV RNAs 1 and 2. Both RNAs produced different patterns of at least 20 3'-coterminous RNAs. They also showed different patterns of subgenomic RNAs from their 5' ends. RNA 1 had about three such RNAs, from 1200 to 1500 nucleotides long, whereas RNA 2 showed about 10, from 800 to 1700 nucleotides in length.

Many of the subgenomic RNAs were found to have considerable *in vitro* messenger activity. Those derived from RNA 1 produced over 10 major polypeptides, all of which could be translated from a few RNA species over about 2300 nucleotides in size; some of the smaller polypeptides were also obtained from smaller RNAs over 1400 nucleotides long. Subgenomic RNAs derived from near the 5' or 3' ends of RNA 2 gave different patterns of translation products; the 3'-coterminal RNAs of increasing length gave products of increasing length (with some RNAs appearing to lack mRNA activity, as also noted for some of the RNA 1 subgenomic RNAs). These translation products were produced with great efficiency even in the presence of the appropriate genomic RNA or, in some cases, of total viral RNA.

RNA 1 contains a single long open read-

ing frame encoding a protein of  $M_r$  110,791 (Rezaian *et al.*, 1985). The next longest reading frame encodes a protein of only 69 amino acids. All the translation products of the subgenomic RNAs must therefore be derived from the  $M_r$  110,791 open reading frame, which contains a total of 31 AUG codons that could lead to many polypeptides upon either complete or prematurely terminated translation. The results in Fig. 4 showed that at least some of the translation products ( $M_r$  30,000,  $M_r$  43,000, and  $M_r$  48,000) of the 3'-coterminal subgenomic RNAs were produced by RNAs shorter than 1800 nucleotides. Correlation of the reading frame length required for these proteins with the RNA sizes (Table 1) showed that the proteins must be derived from the carboxyterminal region of the  $M_r$  110,791 reading frame traversing RNA 1. By the same analysis,

TABLE 1  
COMPARISON OF READING FRAMES WITH TRANSLATION PRODUCTS OF 3'-COTERMINAL SUBGENOMIC RNAs FROM CMV RNAs 1 AND 2

AUG codon (nucleotides from 3' end)	Number of amino acids	Molecular weight of translation product		Possible mRNA (min. size in nucleotides) (d)
		Calculated	Observed (b)	
RNA 1				
2662 (a)	781	87,316	90,000	2700
2323 (a)	668	75,528	75,000	2500
2077 (a)	586	65,011	65,000	2200
1976 (a)	552	61,020	55,000	2100
1522	401	44,294	48,000 (c)	1600
1450	377	41,386	43,000 (c)	1600
1336	339	37,075	37,000	1400
1117	266	29,521	30,000	1400
RNA 2				
2565	713	80,580	85,000	2600
2010	528	60,102	60,000	2050
1446	340	38,968	37,000	1450
1185	253	28,939	30,000	1200
1056	210	24,068	23,000	1200
852	142	16,008	17,000	880

<sup>a</sup> These initiation codons were chosen from several possible ones for each polypeptide because they best fitted the consensus sequence of Kozak (1984).

<sup>b</sup> Translation products observed in Figs. 4 (RNA 1) or 5 (RNA 2) are listed next to the calculated protein with whose size they best agree.

<sup>c</sup> The translation products of  $M_r$  48,000 and  $M_r$  43,000 (RNA 1) were allocated because they are present in track f of Fig. 4, while the  $M_r$  41,000 product is not.

<sup>d</sup> Based on comparison of RNAs present in the samples producing a particular translation product, with the distance of its AUG codon from the 3' end.

the larger proteins (over  $M_r$  50,000) would also share their carboxytermini with the  $M_r$  110,791 protein and possess staggered aminotermini. These proteins could be translated from the RNAs of 2100 to 3000 nucleotides in size, since most (22) of the 31 AUG codons in this reading frame are concentrated in a region covering 32% of the reading frame and starting 139 amino acids from the aminotermus. Some of the translation products, e.g., that of  $M_r$  41000 may, however, be due to prematurely terminated translation of the longer RNAs, since they were not translated from the RNAs shorter than 1800 nucleotides (Fig. 4, track f).

A single long open reading frame traverses RNA 2 (Rezaian *et al.*, 1984) and therefore also the subgenomic RNAs derived from this RNA. The next longest reading frame encodes a protein of only 100 amino acids. Since the 3'-coterminal subgenomic RNAs of increasing size analyzed in Fig. 5 were found to produce proteins of increasing size, their *in vitro* translation products must be encoded by the appropriately sized carboxyterminal portions of the gene for the  $M_r$  94,333 translation product. Six of the initiation codons in phase with the open reading frame of CMV RNA 2 (Rezaian *et al.*, 1984) would yield translation products whose calculated sizes correspond closely to those of major observed *in vitro* translation products (Table 1). Translation products of increasing size must therefore differ at their aminotermini. The alternative explanation, that the smaller translation products are produced by premature termination of translation, can be excluded because of the correlation between RNA length and translation product size.

All the AUG codons listed for RNA 2 in Table 1 match the consensus sequence for translation start sites (Kozak, 1984) reasonably well, but only some of those for RNA 1 do. However, not all plant viral mRNAs appear to fit this consensus sequence (Ravelonandro *et al.*, 1983). Some of the AUG codons in other reading frames have the adjacent nucleotides which may allow them to initiate protein

synthesis but would produce polypeptides too small (below 50 amino acids) to have been detected in these experiments.

The biological significance of the subgenomic RNAs remains unclear. The reproducible patterns of discrete species obtained by Northern analysis of encapsidated RNA argue that these RNAs are significant, for they show differences specific to the genomic RNA from which they were derived. Furthermore, at least some are active *in vitro* mRNAs, and they have also been detected in particulate fraction RNA extracted from plants. These observations suggest that at least some may be genuine subgenomic mRNAs, produced during the viral RNA replication cycle rather than by subsequent degradation. The critical tests of this proposition are to detect the translation products *in vivo* and to demonstrate their synthesis from the subgenomic RNAs of plant polyribosomes.

The gene products from RNAs 1 and 2 possibly consist of several functional domains, with the activities of catalytically important domains dependent upon which other domains are present in the polypeptide, as shown for aminoacyl:tRNA synthetases (Jasin *et al.*, 1984). If subgenomic RNAs indeed allow CMV to differentially express portions of these genes *in vivo*, they may also be instrumental in its evolution. First, they would afford the virus greater flexibility in responding to different hosts or conditions. Furthermore, they could allow the virus to profit from differing mutation rates along individual genes, by allowing expression of regions toward the carboxyterminal independently of the aminoterminal regions of a gene.

It is striking that regions of the genes on CMV RNAs 1 and 2 contain amino acid sequences homologous with corresponding RNAs of brome mosaic virus (BMV), AIMV and TMV (Rezaian *et al.*, 1984, 1985). Some of these homologous sequences are located in the aminoterminal part of the gene on RNA 1, but most are concentrated in the carboxyterminal 281 amino acids of this gene. In the case of RNA 2, the homologous region consists of about 400 amino acids from 545 to 146

residues from the carboxyterminus of the gene on CMV RNA 2. The homologous sequences in the carboxyterminal regions of the CMV RNA genes would therefore be expressed by the translation products of the 3'-coterminal subgenomic RNAs. An exception is translation of the  $M_r$  17,000 protein from RNA 2 since the carboxyterminal 145 amino acids from CMV RNA 2 show very little homology to the carboxyterminal stretch of the BMV gene and none to that of AlMV or TMV. The CMV  $M_r$  17,000 translation product therefore appears to be a polypeptide unique to CMV.

These homologous regions in corresponding genes of several plant RNA viruses may be responsible for fundamental functions, e.g., RNA replication, whereas the divergent regions may control specific host interactions. New viral genes, with the same fundamental function, but altered host specificities, could result from the linking of subgenomic RNAs from one virus with those from another, upon coinfection of a single host. It is therefore of great interest whether other RNA viruses generate similar families of subgenomic RNAs. Crude fractionation, followed by *in vitro* translation, of the encapsidated RNA from lucerne transient streak virus (Morris-Krsinich and Foster, 1983), solanum nodiflorum mottle virus (Kiberstis and Zimmern, 1984), tobacco etch virus (Otal and Hari, 1983), and turnip crinkle virus (Dougherty and Kaesberg, 1981) has disclosed products (other than the coat protein) whose mRNAs were smaller than the genomic RNA.

#### ACKNOWLEDGMENTS

We thank Dr. Richard Francki for glasshouse facilities, Jennifer Rosey and Sharon Freund for assistance, and Dr. Ali Rezaian, Rhys Williams, and Dr. Allan Gould for access to unpublished data and valuable discussions. This work was supported by the Australian Research Grants Scheme and a Commonwealth Government grant to the Adelaide University Centre for Gene Technology in the Department of Biochemistry.

#### REFERENCES

BOL, J. F., BAKHUIZEN, C. E. G. C., and RUTGERS, T. (1976). Composition and biosynthetic activity of

- polyribosomes associated with alfalfa mosaic virus infections. *Virology* **75**, 1-17.
- BRUENING, G. (1981). Biochemistry of plant viruses. In "The Biochemistry of Plants" (P. K. Stumpf and E. E. Conn, eds.), Vol. 6, pp. 571-631. Academic Press, New York.
- BRUENING, G., GOULD, A. R., MURPHY, P. J., and SYMONS, R. H. (1982). Oligomers of avocado sunblotch viroid are found in infected avocado leaves. *FEBS Lett.* **148**, 71-78.
- DAVIES, J. W., and HULL, R. (1982). Genome expression of plant positive-strand RNA viruses. *J. Gen. Virol.* **16**, 1-14.
- DOUGHERTY, W. G., and KAESBERG, P. (1981). Turnip crinkle virus RNA and its translation in rabbit reticulocyte and wheat embryo extracts. *Virology* **115**, 45-56.
- FRANCKI, R. I. B., RANDLES, J. W., CHAMBERS, T. C., and WILSON, J. B. (1966). Some properties of purified cucumber mosaic virus (Q strain). *Virology* **28**, 729-741.
- GILL, D. S., KUMARASAMY, R., and SYMONS, R. H. (1981). Cucumber mosaic virus-induced RNA replicase: Solubilization and partial purification of the particulate enzyme. *Virology* **113**, 1-8.
- GOELET, P., and KARN, J. (1982). Tobacco mosaic virus induces the synthesis of a family of 3'-coterminal messenger RNAs and their complements. *J. Mol. Biol.* **154**, 541-550.
- GONDA, T. J., and SYMONS, R. H. (1979). Cucumber mosaic virus replication in cowpea protoplasts: Time course of virus, coat protein and RNA synthesis. *J. Gen. Virol.* **45**, 723-736.
- GORDON, K. H. J., GILL, D. S., and SYMONS, R. H. (1982). Highly purified cucumber mosaic virus-induced RNA-dependent RNA polymerase does not contain any of the full length translation products of the genomic RNAs. *Virology* **123**, 284-295.
- GOULD, A. R., and SYMONS, R. H. (1982). Cucumber mosaic virus RNA 3: Determination of the nucleotide sequence provides the amino acid sequences of protein 3A and viral coat protein. *Eur. J. Biochem.* **126**, 217-226.
- HIRTH, L., and RICHARDS, K. E. (1981). Tobacco mosaic virus: Model for structure and function of a simple virus. *Adv. Virus Res.* **26**, 145-199.
- HUNTER, T., JACKSON, R., and ZIMMERN, D. (1983). Multiple proteins and subgenomic mRNAs may be derived from a single open reading frame on tobacco mosaic virus. *Nucl. Acids Res.* **11**, 801-821.
- JASIN, M., REGAN, L., and SCHIMMEL, P. (1984). Dispensable pieces of an aminoacyl tRNA synthetase which activate the catalytic site. *Cell* **36**, 1089-1095.
- KIBERSTIS, P. A., and ZIMMERN, D. (1984). Translational strategy of solanum nodiflorum mottle virus RNA: Synthesis of a coat protein precursor *in vitro* and *in vivo*. *Nucl. Acids Res.* **12**, 933-943.

- KOZAK, M. (1981). Mechanism of mRNA recognition by eucaryotic ribosomes during initiation of protein synthesis. *Curr. Top. Microbiol. Immunol.* **93**, 81-123.
- KOZAK, M. (1984). Compilation and analysis of sequences upstream from the translational start site in eucaryotic mRNAs. *Nucl. Acids Res.* **12**, 857-872.
- KUHN, S., FRITZ, H. J., and STARLINGER, P. (1979). Close vicinity of IS1 integration sites in the leader sequence of the gal operon of *E. coli*. *Mol. Gen. Genet.* **167**, 235-241.
- LAEMMLI, U. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* **227**, 680-685.
- LAMB, R. A., and CHOPPIN, P. W. (1983). The gene structure and replication of influenza virus. *Ann. Rev. Biochem.* **52**, 467-506.
- LOT, H., MARCHOUX, G., MARRON, J., KAPER, J. M., WEST, C. K., VAN VLOTEN-DOTING, L., and HULL, R. (1974). Evidence for three functional RNA species in several strains of cucumber mosaic virus. *J. Gen. Virol.* **22**, 81-93.
- MANIATIS, T., FRITSCH, E. F., and SAMBROOK, J. (1982). "Molecular cloning." Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.
- MCMASTER, G. K., and CARMICHAEL, G. G. (1977). Analysis of single- and double-stranded nucleic acids on polyacrylamide and agarose gels by using glyoxal and acridine orange. *Proc. Natl. Acad. Sci. USA* **74**, 4835-4838.
- MESSING, J., CREA, R., and SEEBURG, P. H. (1981). A system for shotgun DNA sequencing. *Nucl. Acids Res.* **9**, 309-321.
- MORRIS-KRSINICH, B. A. M., and FORSTER, R. L. S. (1983). Lucerne transient streak virus RNA and its translation in rabbit reticulocyte lysate and wheat germ extract. *Virology* **128**, 176-185.
- OTAL, T., and HARI, V. (1983). Detection and cell-free translation of subgenomic RNAs of tobacco etch virus. *Virology* **125**, 118-126.
- PALUKAITIS, P., GARCIA-ARENAL, F., SULZINSKI, M. A., and ZAITLIN, M. (1983). Replication of tobacco mosaic virus. VII. Further characterization of single- and double-stranded virus-related RNAs from TMV-infected plants. *Virology* **131**, 533-545.
- PEDEN, K. W. C., and SYMONS, R. H. (1973). Cucumber mosaic virus contains a functionally divided genome. *Virology* **53**, 487-492.
- RAVELONANDRO, M., GODEFROY-COLBURN, T., and PINCK, L. (1983). Structure of the 5'-terminal untranslated region of the genomic RNAs from two strains of alfalfa mosaic virus. *Nucl. Acids Res.* **11**, 2815-2826.
- REZAIAN, M. A., WILLIAMS, R. H. V., GORDON, K. H. J., GOULD, A. R., and SYMONS, R. H. (1984). Nucleotide sequence of cucumber mosaic virus RNA 2 reveals a translation product significantly homologous to corresponding proteins of other viruses. *Eur. J. Biochem.* **143**, 277-284.
- REZAIAN, M. A., WILLIAMS, R. H. V., and SYMONS, R. H. (1985). Nucleotide sequence of cucumber mosaic virus RNA 1: Presence of a sequence complementary to the viral satellite RNA and homologies with other viral RNAs. *Eur. J. Biochem.*, submitted for publication.
- RIEDEL, J., LEHRACH, H., and GAROFF, H. (1982). Nucleotide sequence at the junction between the non-structural and the structural genes of the Semliki forest virus. *J. Virol.* **42**, 725-729.
- SAWICKI, D. L., KAARIANEN, L., LAMBEEK, C., and GOMATOS, P. J. (1978). Mechanism for control of synthesis of Semliki forest virus 26S and 42S RNA. *J. Virol.* **25**, 19-27.
- SCHWINGHAMER, M. W., and SYMONS, R. H. (1977). Translation of the four major RNA species of cucumber mosaic virus in plant and animal cell-free systems and in toad oocytes. *Virology* **79**, 88-108.
- SIDDELL, ST., WEGE, H., and TER MEULEN, V. (1982). The structure and replication of coronaviruses. *Curr. Top. Microbiol. Immunol.* **99**, 135-163.
- SYMONS, R. H. (1978). The two-step purification of ribosomal RNA and plant viral RNA by polyacrylamide slab gel electrophoresis. *Aust. J. Biol. Sci.* **31**, 25-37.
- SYMONS, R. H. (1979). Extensive sequence homology at the 3'-termini of the four RNAs of cucumber mosaic virus. *Nucl. Acids Res.* **7**, 825-837.
- SYMONS, R. H., GILL, D. S., GORDON, K. H. J., and GOULD, A. R. (1982). Gene content and expression of the four RNAs of cucumber mosaic virus. In "Manipulation and Expression of Genes in Eukaryotes" (P. Nagley, A. W. Linnane, W. J. Peacock, and J. A. Pateman, eds.), pp. 373-380. Academic Press, Sydney.
- THOMAS, P. S. (1983). Hybridization of denatured RNA transferred or dotted to nitrocellulose paper. In "Methods in Enzymology" (J. H. Law and H. C. Rilling, eds.), Vol. 110, pp. 255-266. Academic Press, New York.