THE RNAs OF BROMOVIRUSES: 3′-TERMINAL SEQUENCES OF THE FOUR BROME MOSAIC VIRUS RNAs AND COMPARISON WITH COWPEA CHLOROTIC MOTTLE VIRUS RNA 4

Madeline R. GUNN and Robert H. SYMONS
Department of Biochemistry, University of Adelaide, Adelaide, SA 5001, Australia

Received 25 April 1980

1. Introduction

The ability of certain plant viral RNAs to be aminoacylated in the presence of aminoacyl tRNA synthetases implies a 3′-terminal transfer RNA-like structure [1]. In the case of the Bromoviruses and the Cucumoviruses, the four RNAs which constitute the tripartite genome of each virus all accept tyrosine [1]. Hence, it is of interest to compare the 3′-terminal sequence and structure of the RNAs of these two important groups of plant viruses.

The Cucumovirus cucumber mosaic virus (CMV) contains four single-stranded RNAs (RNAs 1–4 in order of decreasing molecular weight), the three largest of which are required for infectivity [2]. The sequence of 270 residues from the 3′-termini of these RNAs has been determined and showed regions of complete, partial or no homology between RNAs 1, 2 and 3 or 4 [3]; the terminal sequence of RNAs 3 and 4 was identical, as expected, since the sequence of RNA 4 is contained at the 3′-end of RNA 3 [4]. It was also possible to arrange the common 3′-terminal 138 residues of the four RNAs into a transfer RNA-like base-paired structure [3].

The best characterized of the Bromoviruses, brome mosaic virus (BMV), has many properties in common with CMV [5]. The sequence of the 3′-terminal 161 residues of BMV RNA 4 have been determined and evidence provided that the corresponding region of RNA 3 was identical while that of RNAs 1 and 2 showed only minor differences [6]. We have now essentially confirmed this sequence data and considerably extended it to longer sequences of the four BMV RNAs. We have also sequenced the 3′-terminal 166 residues of RNA 4 of another Bromovirus, cowpea chlorotic mottle virus (CCMV), to provide comparative data. The results have allowed important comparisons to be made between the 3′-terminal sequences and structures of the RNAs of BMV, CCMV and CMV.

Address correspondence to: R. H. S.

2. Materials and methods

Total BMV and CCMV RNAs were the very generous gift of Dr A. O. Jackson, Department of Plant Pathology, Purdue University, IN, and were isolated from viruses in his collection. BMV was originally obtained from the ATCC strain of Dr M. Brakke while CCMV was derived from an isolate of Dr J. Bancroft. The four RNAs of BMV and CCMV were purified [7] and polyadenylated at their 3′-termini with *Escherichia coli* ATP–RNA adenyltransferase (poly(A) polymerase) essentially as in [8]. Dr P. Palukaitis kindly provided some of these preparations.

The sequencing approach using the dideoxynucleotide chain termination technique [9] as applied to RNA was described in [3,10] and has been used here with modifications which have increased its reliability. For sequences from ~20 to >200 residues, reaction mixtures contained in 5 μl final vol.: 50 mM Tris–HCl (pH 8.0); 50 mM KCl; 8 mM $MgCl_2$; 10 mM DTT; 0.1–0.2 μg polyadenylated RNA; 0.05–0.1 μg $d(pT_8G)$; and 2.5 units of reverse transcriptase. Three of the four reaction mixtures also contained d[α-^{32}P]-GTP (5 μM, 10 μCi, 400 Ci/mmol [11]), 50 μM dATP, 50 μM dCTP, 5 μM dTTP, and either 8–75 μM

ddATP, 3–25 μM ddCTP or 0.4–4 μM ddTTP. The fourth reaction mixture contained d[α-^{32}P]CTP (5 μM, 10 μCi, 400 Ci/mmol), 50 μM dATP, 50 μM dGTP, 5 μM dTTP and 5–40 μM ddGTP. After incubation at 37°C for 30 min, labelled nucleotides were separated on the basis of chain length by electrophoresis on thin 77 cm long 6% polyacrylamide slab gels [3,10,12]. Autoradiography of gels was carried out at −70°C using preflashed Fuji RX medical X-ray film and Ilford fast-tungstate intensifying screens [13].

For sequence determination of the 10–30 residues adjacent to the 3′-poly(A) tail, it was necessary to replace the d[α-^{32}P]GTP or d[α-^{32}P]CTP by d[5′-^{32}P](pT$_8$G); reaction mixtures described above were modified to contain 50 μM dATP, dCTP and dGTP, 5 μM dTTP, d[5′-^{32}P](pT$_8$G) [3], and either 200 μM ddATP, 50 μM ddCTP, 75 μM ddGTP or 5 μM ddTTP. After incubation at 37°C for 30 min, reaction mixtures were fractionated on thin 34 cm long 20% acrylamide gels [3,10].

3. Results and discussion

3.1. *Sequencing procedure*

The specific aminoacylation of BMV RNAs by tyrosine in the presence of aminoacyl tRNA synthetases [1] indicates the presence of a 3′-CCA terminal sequence and this was confirmed by the sequence data in [6]. Hence, the dideoxynucleotide chain termination technique [9] was used to obtain extensive 3′-terminal sequences using the specific d(pT$_8$G) primer on the purified in vitro polyadenylated RNAs. The procedures used are outlined in section 2 and have been detailed in [3,10]. The sequencing gel patterns obtained were similar to those in [3,10] so examples are not given here. Ambiguities in sequencing gels due to minor band compression were resolved by replacing dGTP by dITP in the sequencing reaction mixtures [3,10].

3.2. *Sequences adjacent to the 3′-termini of the four BMV RNAs*

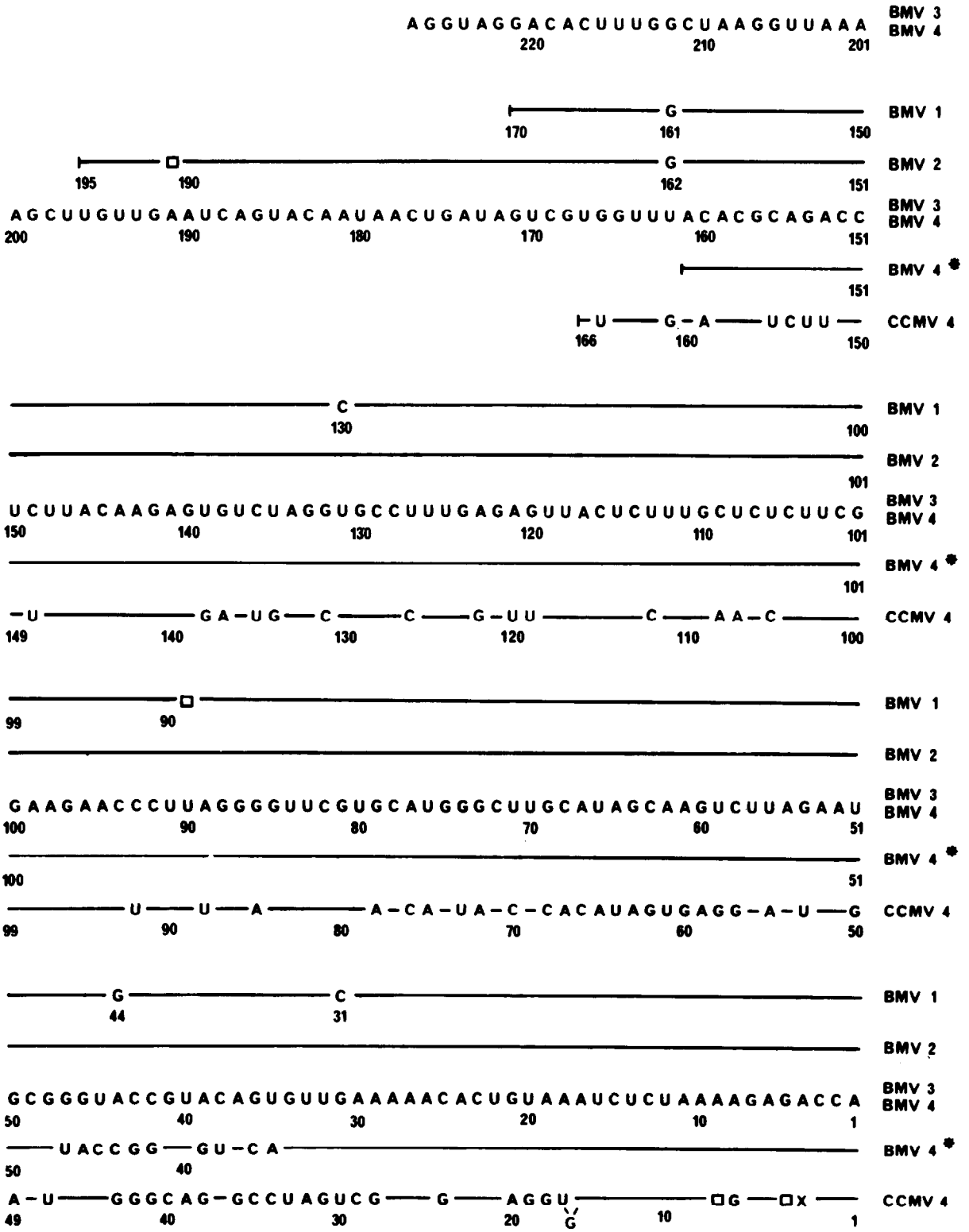
The sequence data for the 3′-terminal 170 residues of BMV RNA 1, 195 residues of BMV RNA 2 and for 227 residues of BMV RNAs 3 and 4 are given in fig.1. The sequence of RNA 4 was the same as that of RNA 3 which was expected since it is highly likely that the complete sequence of BMV RNA 4 is at the 3′-end of BMV RNA 3 [5,14]. The sequence of the 3′-terminal 161 residues of BMV RNA 4 reported [6] is identical to that obtained here except for a region between residues 35 and 47 (fig.1); this discrepancy appears to result from the incorrect ordering of these nucleotides in the sequence of [6] since the base composition of this region in the two RNAs is the same.

In [6], the 3′-terminal 161 residue fragment of BMV RNA 4 was obtained by partial ribonuclease T$_1$ digestion which indicates that residue 162 was a G. However, in our BMV RNA 3 and 4 sequences, residue 162 was a U, so a similar fragment would not be obtained with our RNAs. It is of interest that the corresponding residue in BMV RNAs 1 and 2 is a G (fig.1).

In comparing the sequences of BMV RNAs 1 and 2 to those of BMV RNAs 3 and 4, there are five differences in RNA 1 and two differences in RNA 2 (fig.1). Thus, in RNA 1 there are single base changes at residues 31, 44, 130 and 161, and a single base deletion after residue 90. In RNA 2, there is a single base change at residue 162 and a single base deletion after residue 190. Comparable single base deletions have already been reported in the 3′-terminal sequences of both CMV and AMV RNAs [3,10].

On the basis of T$_1$ and pancreatic ribonuclease maps of the 3′-terminal 161 residue fragments of BMV RNAs 1, 2 and 3, it was concluded [6] that the sequence of the RNA 3 fragment was identical to that of RNA 4, that there was one base change in RNA 2 and two base changes in RNA 3. The conclusion that the terminal sequences of RNAs 1 and 2 are

Fig.1. Sequences of the 3′-terminal regions of the four BMV RNAs and of CCMV RNA 4. BMV 4* is the sequence determined in [6]. Sequences of BMV RNAs 1 and 2 and of CCMV RNA 4 are arranged vertically with those of RNAs 3 and 4 to show maximum sequence homology. A horizontal line indicates the same sequence as in BMV RNAs 3 and 4. The absence of a residue is indicated by an open box. Residue 4 of CCMV could not be determined and is shown as X; residue 4 of the BMV RNAs was taken as A from the data in [6]. Subscript numbers are residues from the 3′-terminus.


AGGUAGGACACUUUGGCUAAGGUUAAA
220 210 201
BMV 3
BMV 4
G
170 161 150
BMV 1
G
195 190 162 151
BMV 2
AGCUUGUUGAAUCAGUACAAUAACUGAUAGUCGUGGUUUACACGCAGACC
200 190 180 170 160 151
BMV 3
BMV 4
BMV 4 *
151
U G A UCUU
166 160 150
CCMV 4
C
130 100
BMV 1
101
BMV 2
UCUUACAAGAGUGUCUAGGUGCCUUUGAGAGUUACUCUUUGCUCUCUUCG
150 140 130 120 110 101
BMV 3
BMV 4
101
BMV 4 *
U GA U G C C G U U C A A C
149 140 130 120 110 100
CCMV 4
99 90
BMV 1
BMV 2
GAAGAACCCUUAGGGGUUCGUGCAUGGGCUUGCAUAGCAAGUCUUAGAAU
100 90 80 70 60 51
BMV 3
BMV 4
100 51
BMV 4 *
U U A A C A U A C C A C A U A G U G A G G A U G
99 90 80 70 60 50
CCMV 4
G C
44 31
BMV 1
BMV 2
GCGGGUACCGUACAGUGUUGAAAAACACUGUAAAUCUCUAAAAGAGACCA
50 40 30 20 10 1
BMV 3
BMV 4
U ACCGG GU CA
50 40
BMV 4 *
A U GGGCAG GCCUAGUCG G AGGU G
49 40 30 20 10 1
CCMV 4

very similar to those of RNAs 3 and 4 is consistent with the sequence data reported here, although the base changes listed are not.

3.3. *Sequence of 3'-terminal 166 residues of CCMV RNA 4*

It is of interest to compare the 3'-terminal sequences of RNA 4 from BMV and CCMV since both viruses are members of the Bromovirus group. The sequence for the 3'-terminal 166 residues of CCMV RNA 4 is shown in fig.1 and has been presented to show maximum sequence homology with BMV RNA 4. The alignment was identical except for two single base deletions after residues 4 and 7 in CCMV RNA 4 and a single base addition after residue 15. Although there are 68 base changes between CCMV RNA 4 and BMV RNA 4, as well as these deletions and additions, there are 16 regions where there are three or more homologous residues with the longest region being of 12 residues. The 3'-terminal sequences of CCMV RNAs 1, 2 and 3 have not been determined.

3.4. *Comparison of possible base-paired regions in 3'-terminal sequences of BMV RNA 4 and CCMV RNA 4*

Since the Bromovirus RNAs can be aminoacylated with tyrosine [1], possible base-paired regions were searched for with the aid of a computer programme [15] to see if tRNA-like structures existed. The clover-leaf arrangement for BMV RNA 4 (fig.2) is similar to that in [6] and 60 of the 3'-terminal 108 residues (56%) are base-paired. In spite of the considerable differences in the 3'-terminal sequences between BMV RNA 4 and CCMV RNA 4 (fig.1), it was possible to arrange the terminal 130 residues of CCMV RNA 4 in a very similar arrangement to

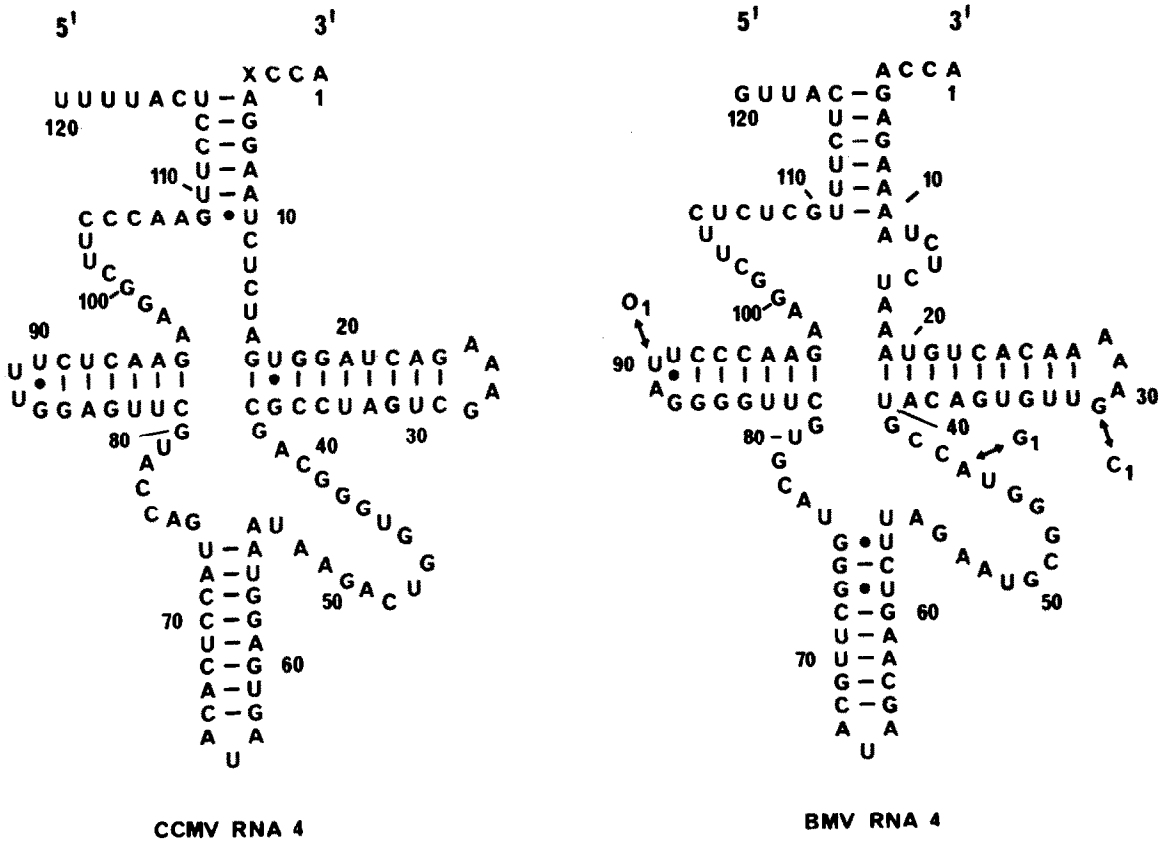


Fig.2 Possible base-paired structures of the 3'-terminal 120 residues of BMV RNA 4 and of CCMV RNA 4. Base substitutions in BMV RNA 1 at residues 31 and 44 are shown; the subscript indicates the RNA involved. The U–O substitution at residue 90 is a deletion.

that of BMV RNA 4 (fig.2) in which 60 residues are also base-paired. There are a number of important similarities between the two structures. The number of base pairs in each of the four base-paired regions is the same. The unpaired sequences AAAG and AUA are present at the ends of the first two arms of both structures; it may be significant that the latter sequence is an anticodon for tyrosine. The 26 residue sequence from residue 80–105 in BMV RNA 4 and 79–104 in CCMV RNA 4 is highly conserved; the only differences are the substitution within a hairpin loop of residue 89 (A) in BMV RNA 4 by a U at residue 88 in CCMV RNA 4 and the replacement of a G–C base pair at residues 86 and 93 in BMV RNA 4 by an A–U base pair at residues 85 and 92 in CCMV RNA 4.

In the proposed base-paired structure for BMV RNA 4 (fig.2), it is of interest that the two base substitutions at residues 31 and 44 and the base deletion at residue 90 that occur in RNA 1 are all in single-stranded regions.

3.5. *Comparison of 3′-terminal sequences of BMV RNA 4 with those of cucumber mosaic virus (CMV) RNA*

Although CMV is classified in a different family, the Cucumoviruses, to BMV and CCMV, a common feature is the ability of the CMV RNAs to be aminoacylated by tyrosine [1]. Hence, a comparison of 3′-terminal sequences and base-paired structures of the BMV (fig.1,2) and CMV RNAs [3] may indicate features of biological importance. Thus, the sequence 3′ GAAGGCUUCU 5′ at residues 96–105 from the 3′-termini of the four CMV RNAs is also found in the four BMV RNAs at residues 97–106, and the one residue shorter sequence 3′ GAAGGCUUC 5′ at residues 96–114 in CCMV RNA 4. This common sequence is also located at the same position in the proposed base-paired structures of BMV and CCMV RNAs (fig.2) and of CMV RNAs [3]. Further, the sequence 3′ GUGCUUCGGGA 5′ from residue 87–97 in CMV RNA 4 corresponds with one base change difference to the sequence from residue 79–89 of the BMV RNAs, although this sequence is not in the same position of the proposed base-paired structures [3] (fig.1). Another long sequence, 3′ AGAACAUUCU 5′ at residues 139–148 of CMV RNAs 3 and 4 only, is located at residues 141–150 of the BMV RNAs and with one base change at residues 140–149 of CCMV RNA 4.

When the 3′-terminal 51 residues of CMV RNA 4 were aligned with the 3′-terminal 55 residues of BMV RNA 4 to show maximum sequence homology (fig.3), 36 residues were identical (71%); for homologous sequences containing two or more adjacent residues, there were 32 common residues (63%). This is surprising since only 47% of residues in CCMV RNA 4 were homologous in the same region with those of BMV RNA 4. Extensive homology between BMV RNA 4 and CCMV RNA 4 only occurred from residue 79–166 of the CCMV RNA.

Such extensive 3′-terminal sequence homology between the CMV RNAs and those of the Bromoviruses suggests an evolutionary and functional relationship. The proposed base-paired structures for BMV and CCMV RNAs (fig.2) have much in common with the structure proposed for the CMV RNAs [3]. However, the complex arrangement of regions of homology, partial homology and non-homology found in the 3′-terminal 270 residues of the four CMV RNAs [3] was not evident in the BMV sequences determined so far and may represent a basic difference between the Cucumoviruses and the Bromoviruses. Further sequence data will need to be obtained to establish this.

In the case of the four CMV RNAs, regions of homology extended to 270 residues, and possibly to 300 residues, from the 3′-termini [3]. It will be of interest to determine how far the homology extends between the four BMV RNAs. It is unlikely to extend further than the termination codon of the BMV coat protein in RNA 4; on the basis of the size of

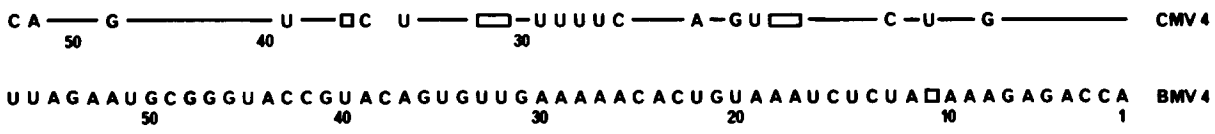


Fig.3. Comparison of the 3′-terminal sequence of BMV RNA 4 and of CMV RNA 4 [3]. The sequence of CMV RNA 4 has been arranged to show maximum sequence homology. The horizontal line indicates the same sequence and the open boxes deletions. Residue 4 of CMV RNA 4 is not known [3] but was taken as A for this comparison.

RNA 4 (M_r 3 × 10^5), the size of the coat protein (M_r 2.03 × 10^4) and the 9 residue untranslated region at the 5′-end [16], this would occur ~300 residues from the 3′-end.

Acknowledgements

The authors thank Dr R. I. B. Francki for the use of glasshouse facilities, Miss J. L. Rosey for assistance, Miss R. Lewis of the Computing Centre for extensive assistance with the computer programme, Dr J. C. Wallace for further help with its use, Dr A. O. Jackson for the gift of BMV and CCMV RNAs, and the Office of Program Resources and Logistics, National Cancer Institute, Bethesda, MD for the reverse transcriptase. This work was supported by the Australian Research Grants Committee.

References

[1] Hall, T. C. (1979) Int. Rev. Cytol. 60, 1–26.
[2] Peden, K. W. C. and Symons, R. H. (1973) Virology 53, 487–492.
[3] Symons, R. H. (1979) Nucleic Acids Res. 7, 825–837.
[4] Gould, A. R. and Symons, R. H. (1978) Eur. J. Biochem. 91, 269–278.
[5] Van Vloten-Doting, L. and Jaspars, E. M. J. (1977) Compr. Virol. 11, 1–53.
[6] Dasgupta, R. and Kaesberg, P. (1977) Proc. Natl. Acad. Sci. USA 74, 4900–4904.
[7] Symons, R. H. (1978) Aust. J. Biol. Sci. 31, 25–37.
[8] Gould, A. R., Palukaitis, P., Symons, R. H. and Mossop, D. W. (1978) Virology 84, 443–455.
[9] Sanger, F., Nicklen, S. and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[10] Gunn, M. R. and Symons, R. H. (1980) FEBS Lett. 109, 145–150.
[11] Symons, R. H. (1977) Nucleic Acids Res. 4, 4347–4355.
[12] Sanger, F. and Coulson, A. R. (1977) FEBS Lett. 87, 107–110.
[13] Laskey, R. A. and Mills, A. D. (1977) FEBS Lett. 82, 314–316.
[14] Ahlquist, P., Dasgupta, R., Shih, D. S., Zimmern, D. and Kaesberg, P. (1979) Nature 281, 277–282.
[15] Staden, R. (1977) Nucleic Acids Res. 4, 4037–4051.
[16] Lane, L. C. (1977) CMI/AAB Descriptions of Plant Viruses, no. 180.