Complete nucleotide sequence of RNA 5 from cucumber mosaic virus (strain Y)

Soh Hidaka, Kiichi Ishikawa, Yoichi Takanami*, Susumu Kubo* and Kin-ichiro Miura

*Department of Biochemistry, Yamagata University School of Medicine, Zaoh-Iida, Yamagata 990-23, *Central Research Institute, the Japan Tobacco and Salt Public Corporation, Umegaoka, Yokohama 227* and +*Department of Industrial Chemistry, Faculty of Engineering, the Univerity of Tokyo, Hongo, Tokyo 113, Japan*

Received 19 June 1984

Abstract not received

Cucumber mosaic virus *Nucleotide sequence* *RNA 5*

1. INTRODUCTION

Cucumber mosaic virus, yellow strain (CMV-Y), contains 5 single-stranded RNA species including 3 genomic RNAs (RNA 1–3), a subgenomic RNA (RNA 4) and an RNA 5 [1]. RNA 5, the smallest segment, causes striking changes in symptoms on tomato and tobacco plants inoculated after being mixed with CMV free from RNA 5, and it has been proved to be a satellite RNA similar to CA RNA 5 [2,3]. Previously, we reported the sequence of 30 nucleotides from the 5′-terminus of the RNA 5 which differs in one nucleotide at position 24 (C⟶U) from the entire sequence of the CMV strain D (CMV-D) [4,5]. This RNA has simply a 10-nucleotide-long sequence as a noncoding region with cap structure and can stimulate slightly amino acid incorporation into polypeptides in an in vitro protein synthesizing system as compared with RNA 4, subgenomic RNA coding for the coat protein [6]. Small cucumovirus-associated RNAs were isolated from various strains [2,3,7]. Although these were reported to be similar in molecular mass, some were unable to induce a characteristic symptom in certain host plants [8].

In Japan, several species of RNA 5 were found among the cucumoviruses and its biological analyses are in progress (unpublished). We report here the complete sequence of RNA 5 of CMV-Y determined by rapid sequencing gel analysis and its comparison with that of CMV-Ds.

An AUG, closest to the cap structure, was located at position 11 and a UGA, termination codon, appeared at the 92nd nucleotide in phase with the readable codons. There were 8 AUGs and 7 termination codons downstream. The determined sequence contains 368 nucleotides and was compared with the sequence of RNA 5 of CMV-D [5]. The sequence was almost homologous in the 5′- and 3′-terminal regions. There was less homology in the middle of the RNA.

2. MATERIALS AND METHODS

CMV-Y was propagated in tobacco and its RNA was extracted as described [1,2]. RNA 5 was purified by electrophoresis through a 5% polyacrylamide slab gel containing 7 M urea in 50 mM Tris–borate (pH 8.3). After the RNA band, localized by ethidium bromide staining, was eluted from a gel piece, the 5′-terminal cap structure in the RNA was removed by periodate oxidation, β-elimination [9] and phosphatase treatment, then the naked 5′-terminus was labelled with polynucleotide kinase using [γ-^{32}P]ATP (3000 Ci/mmol, Amersham) [10]. For 3′-end labelling, the purified RNA 5 was incubated with RNA ligase and [5′-^{32}P]pCp, which was prepared from

+ To whom correspondence should be addressed

Published by Elsevier Science Publishers B.V.

[γ-^{32}P]ATP [11], at 4°C for 16 h [12]. The labelled RNA was purified again by electrophoresis in a 5% polyacrylamide gel as described above to remove degraded labelled RNA fragments during the incubation. To obtain the internal sequence of the RNA, it was partially digested with 0.5 units/ml of bacterial alkaline phosphatase (Worthington, BAPF) containing unknown nuclease [13] at 50°C for 30 min in 20 mM Tris–HCl (pH 8.2) and was purified with phenol treatment followed by precipitation with ethanol. The fragments, labelled at the 5′-end as described, were subjected to two-dimensional electrophoresis and the single spot detected by autoradiogram was recovered by essentially the same method as in [14]. For sequencing analysis, the labelled RNA and fragments were digested partially with ribonuclease T_1, which splits at guanylic acid (G), and with U_2 (A > G). Partial hydrolysis of RNA was carried out with alkali as in [4], with *Physarum* ribonuclease Phy M (A + U) and *Neurospora crassa* endonuclease (A + G + U) in the presence of 7 M urea [15,16]. Chicken liver ribonuclease CL3 (C > U) and *Bacillus cereus* ribonuclease (C + U) were also used to confirm the pyrimidine residue [17,18]. Partial chemical

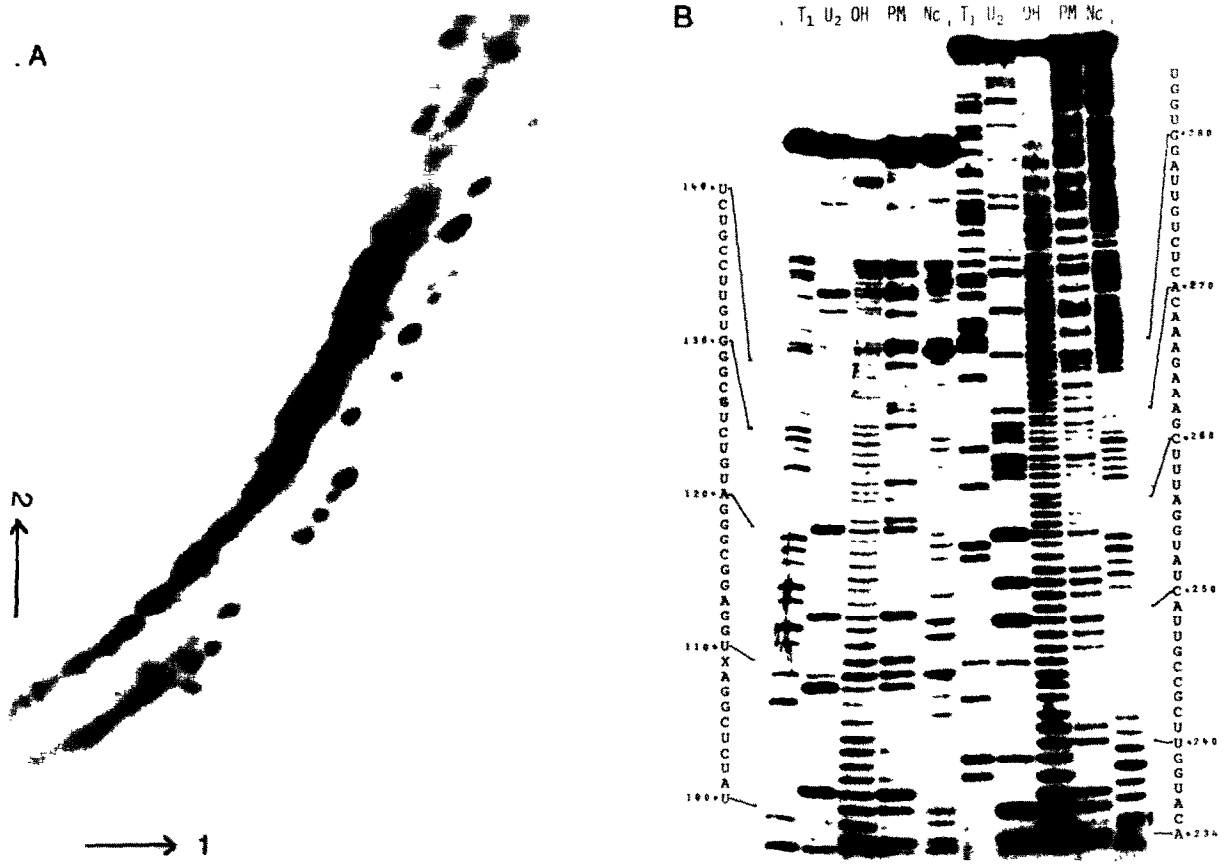


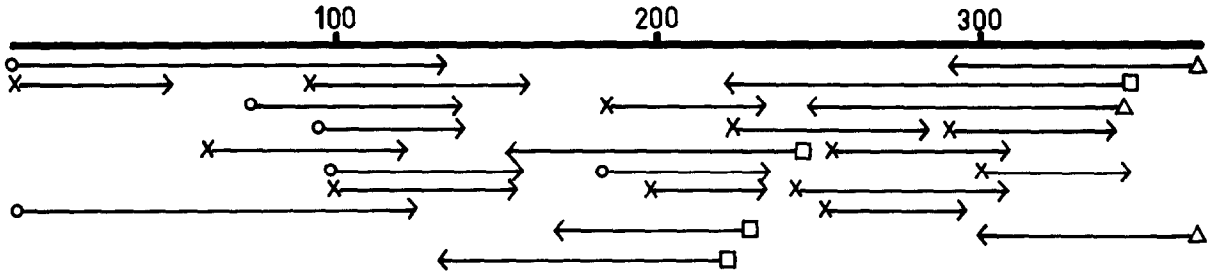
Fig.1. Sequence analysis of RNA 5. (A) Separation by two-dimensional electrophoresis of fragments produced by partial digestion of RNA 5 with unknown nuclease in bacterial alkaline phosphatase preparation, followed by labelling at the 5′-end with ^{32}P. First dimension was electrophoresis in 6% polyacrylamide gel (20 × 40 × 0.1 cm) at pH 3.5 and second dimension was in 12% gel (30 × 40 × 0.1 cm) at pH 8.3. (B) Autoradiogram of partial enzymatic digests of 5′-^{32}P-labelled fragments which were electrophoresed in 12% polyacrylamide slab gel (40 × 30 × 0.05 cm). Tracks: T_1, ribonuclease (RNase) T_1; U_2, RNase U_2 OH, alkali; PM, *Physarum* RNases Phy M; Nc, *Neurospora crassa* endonuclease. A band X (position 109) appeared in all tracks of this gel, but was confirmed as C in several repeated experiments illustrated in fig.2.

cleavage of the 3′-^{32}P-labelled RNA was done as in [19]. Electrophoresis of these partial digests were performed on 8 or 12% polyacrylamide gel.

3. RESULTS AND DISCUSSION

Fig.1A shows the separation of the 5′-labelled RNA fragments by two-dimensional gel electrophoresis. The efficiency of labelling of the 5′-end with ^{32}P seemed to be higher in the partial digest with nuclease in bacterial alkaline phosphatase preparation than with other enzymes such as ribonuclease T_1 under our conditions. To determine the sequence of the fragments, each of the major spots well-separated was analysed enzymatically and the results are shown in fig.1B.

The alignment of the fragments was based on the direct sequence from the terminus and on the overlapped sequence of themselves. The data obtained from the analyses of fragments were very helpful in eliminating the ambiguity which was often observed in the sequence derived from the terminal labelled RNA. Although almost all of the molecule was sequenced out by these methods, the autoradiogram around nucleotide residues 160–175 was unclear due to sequence compression.



1 10 20 30 40 50
m^{7}GpppGUUUUGUUUGAUGGAGAAUUGCGUAGAGGGGUUAUAUCUGCGUGAGGAUC

60 70 80 90 100
CAUCACUCGGCGGUGUGGGAUACCUCCCUGCUAAGGCGGGUUGAGAGUGU

110 120 130 140 150
AUCUCGGACUGGAGGCGGGAUGUCUGCGGGUGUUCCGUCUGCUGCCCACG

160 170 180 190 200
AUGGUGGGAGCCACCCCAGGGGCGACUUUUUCAGCUCUGCAUUUCUCAUU

210 220 230 240 250
UGAGCCCCCGCUCAGUUUGCUAGCAAAACCCGGACAUGGUUCGCCGUUAC

260 270 280 290 300
UAUGGAUUUCGAAAGAAACACUCUGUUAGGUGGUAUCGUGGAUGACGCAC

310 320 330 340 350
GCAGGGAGAAGCUAAGGCUUAUGCUAUGCUGAUCUCCGUGAAUGUCUAUA

360
CAUUCCUCUACAGGACCC

Fig.2. Derivation of complete nucleotide sequence of RNA 5 of CMV-Y. Illustration of the constitution of the complete sequence of RNA 5. Sequence analyses were made with 5′-^{32}P-labelled RNA (○) or 5′-^{32}P-labelled fragments (×) enzymatically, and with 3′-^{32}P-labelled RNA chemically (△) or enzymatically (□). In the case of ^{32}P-labelled RNA (○, □), labelling was located at the terminus of the RNA, the nucleotide chain between the terminus and the position indicated by (○ or □) has run a gel off in electrophoresis.

This was overcome by sequencing of the chemically modified RNA by sodium bisulfite which converts the C residue to U and destroys the hydrogen bond between G and the C residue in the secondary structure (not shown) [20]. The primary sequence of RNA 5 was constituted as shown in fig.2. Direct sequencing from each terminus was made with the terminal ^{32}P-labelled RNA by using 8 or 12% polyacrylamide gel. More than 150 nucleotides from the terminus can usually be read on an 8% gel. For the fragment analysis, 12% gel was mainly used for sequencing up to 120 nucleotides. Fig.2 gives the complete sequence of RNA 5. Previously, we reported the sequence of the 5′-terminal 30

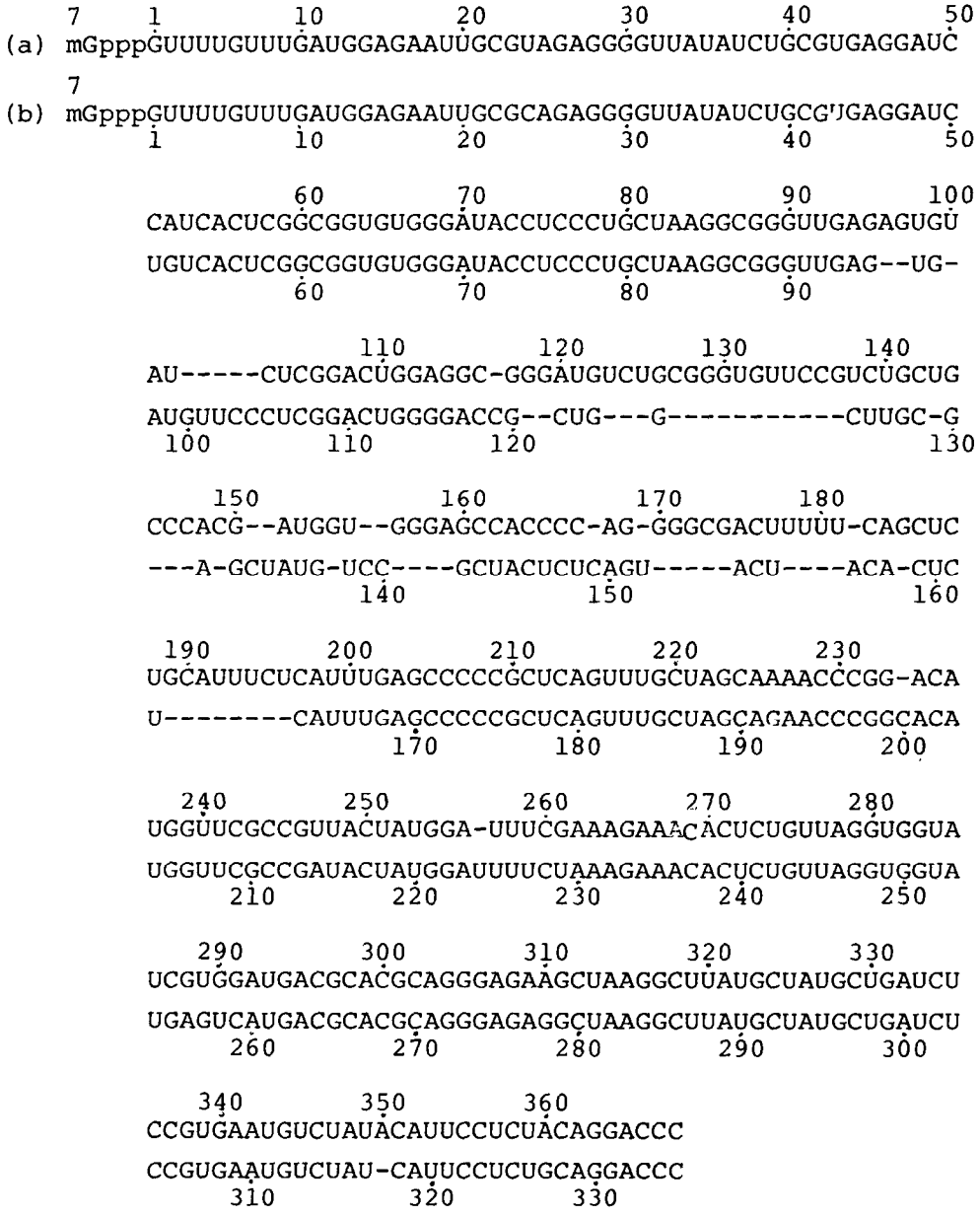


Fig.3. Comparison of primary structure of RNA 5s of CMV-Y (a) and CMV-D (b). Sequences are aligned for maximum homology.

nucleotides including the AUG triplet at position 11; this AUG is closest to the 5′-terminus, and is presumably an initiation codon for protein synthesis. In phase with this codon, the termination codon, UGA, appears at position 92. There are 5 AUGs and 7 termination codons downstream. Comparison of the sequence of RNA 5 of CMV-Y with CMV-Ds is shown in fig.3. Both sequences are aligned for maximum homology. These homologous sequences are not evenly distributed along the molecule, but accumulated around the 5′-terminal region (positions 1–95) and 3′-terminal region (positions 197–368). The sequence was 97% homologous with RNA 5 of CMV-D in the former and 92% in the latter region. Less homology is observed in the middle of the RNA (positions 96–196). A number of the sequence changes may be a result of insertion or deletion of nucleotides during the replication of the RNA. It is of interest whether these observed homologies are functionally required for the satellite nature and/or are only the reflection of the phylogenetic relation. The satellite RNA has been reported to be present in several viruses such as tobacco ringspot virus [21], tomato black ring virus [22] and turnip crinkle virus [23]. However, not only its origin, but also its functional role are still obscure. These satellite RNAs including RNA 5 of CMV-Y in this report and that of CA RNA 5 [3] were found to lead to markedly different symptoms in certain host plants from those caused by the infection with the genomic RNA alone. Together with other RNA 5s incapable of inducing characteristic symptoms, it can be hoped that the sequence comparison of various RNA 5s will provide us some clues to solve these problems.

REFERENCES

[1] Takanami, Y., Kubo, S. and Imaizumi, S. (1977) Virology 80, 376–387.
[2] Takanami, Y. (1981) Virology 109, 120–126.
[3] Kaper, J.M. and Tousignant, M.E. (1977) Virology 80, 186–195.
[4] Hidaka, S., Shimotohno, K., Miura, K., Takanami, Y. and Kubo, S. (1979) FEBS Lett. 98, 115–118.
[5] Richard, K.E., Jonard, G., Jacquemond, M. and Lot, H. (1978) Virology 89, 395–408.
[6] Yamaguchi, K., Hidaka, S. and Miura, K. (1982) Proc. Natl. Acad. Sci. USA 79, 1012–1016.
[7] Gould, A.R., Palukaitis, P., Symons, R.H. and Mossop, D.W. (1978) Virology 84, 443–445.
[8] Kaper, J.M., Tousignant, M.E. and Thompson, S.M. (1981) Virology 114, 526–533.
[9] Miura, K., Watanabe, K. and Sugiura, M. (1974) J. Mol. Biol. 86, 31–48.
[10] Donis-Keller, H., Maxam, A.M. and Gilbert, W. (1977) Nucleic Acids Res. 4, 2527–2538.
[11] Cranton, J.W., Silber, R., Malathi, V.G. and Hurwitz, J. (1974) J. Biol. Chem. 249, 7447–7456.
[12] England, T.E. and Uhlenbeck, O.C. (1978) Nature 275, 560–561.
[13] Alberty, H., Raba, M. and Gross, H.J. (1978) Nucleic Acids Res. 5, 425–434.
[14] De Wachter, R. and Fiers, W. (1972) Anal. Biochem. 49, 184–197.
[15] Donis-Keller, H. (1979) Nucleic Acids Res. 7, 179–192.
[16] Krupp, G. and Gross, H.J. (1979) Nucleic Acids Res. 6, 3481–3490.
[17] Boguski, M.S., Hieter, P.A. and Levy, C.C. (1980) J. Biol. Chem. 255, 2160–2163.
[18] Lockard, R.E., Alzner-De Weered, B., Heckman, J., MacGee, J., Tabor, M.W. and RajBhandary, U.L. (1978) Nucleic Acids Res. 5, 37–56.
[19] Peattie, D.A. (1979) Proc. Natl. Acad. Sci. USA 76, 1760–1764.
[20] Okada, N., Sakamoto, K., Itoh, Y. and Ohshima, Y. (1982) J. Biochem. 91, 1281–1291.
[12] Schneider, I.R. (1971) Virology 45, 108–122.
[22] Murant, A.F., Mayo, M.A., Harrison, B.D. and Goold, R.A. (1973) J. Gen. Virol. 19, 275–278.
[23] Altenbach, S.B. and Howell, S.H. (1981) Virology 112, 25–33.