# The initiation site for transcription of the TMV 30-kDa protein messenger RNA

Yuichiro Watanabe, Tetsuo Meshi and Yoshimi Okada

*Department of Biophysics and Biochemistry, Faculty of Science, University of Tokyo, Hongo, Tokyo 113, Japan*

Received 29 May 1984

The initiation site for transcripotion of the 30-kDa protein mRNA of tobacco mosaic virus was mapped uniquely at residue 1558 from the 3′-terminus on TMV RNA using the primer-extension and the S1-nuclease mapping method.

*Tobacco mosaic virus* *S1-nuclease mapping* *Primer extension* *30-kDa protein mRNA*

## 1. INTRODUCTION

The genome of tobacco mosaic virus (TMV) is a single-stranded RNA molecule about 6400 nucleotides long and codes for at least 4 proteins: the 130-, 180-, 30-kDa and coat proteins [1]. The synthesis of two of these proteins, the coat protein and the 30-kDa protein, is directed by respective subgenomic mRNAs (CP mRNA [2] and the 30-kDa protein mRNA [3]).

The synthesis of both subgenomic mRNAs starts at an early stage during infection. The 30-kDa protein mRNA is synthesized transiently along with the 30-kDa protein, while CP mRNA synthesis continues [4]. The mechanisms of synthesis of the subgenomic RNAs and of shut-off of 30-kDa protein mRNA synthesis are still unknown.

The CP mRNA was isolated from TMV-infected tobacco leaves and its sequence has been determined [5]. Compared with the sequence of the genomic RNA, the initiation site for transcription of the CP mRNA has been determined to be at residue 693 from the 3′-terminus of the genomic RNA.

Here, we have determined the initiation site for transcription of the 30-kDa protein mRNA by the primer-extension method and the S1-mapping method. It was located at the G base of residue 1558 from the 3′-terminus.

## 2. MATERIALS AND METHODS

Tobacco protoplasts were inoculated with TMV-RNA (common strain OM) by use of liposomes [6]. Total RNA was extracted from cells 6 h postinfection with lysis buffer as in [4] and subjected to electrophoresis on an 8 M urea–2.3% polyacrylamide gel. The 30-kDa protein mRNA was recovered from the gel as in [7]. Its messenger RNA activity was confirmed by in vitro translation in a rabbit reticulocyte lysate (Amersham) [8].

Two types of extension primers and S1 mapping probes were prepared as shown in fig.1. pOM5C6 [9] contains sequences of the 3′-portion of the genomic RNA of TMV OM. H-S1 (nucleotide 1940–1312 from the 3′-terminus) and B-S1 (nucleotide 1940–1425) are probes used for S1 mapping. H-Pr (nucleotide 1348–1312) and B-Pr (nucleotide 1447–1425) are primers used for primer-extension analysis. H-S1 and H-Pr were made from a *Hin*dIII fragment of pOM5C6, which was 5′-end phosphorylated and subjected to digestion by *Hae*III and *Hin*fI, respectively. B-S1 and B-Pr are made from a *Bgl*II fragment of pOM5C6 which was 5′-end phosphorylated and subjected to digestion by *Hae*III and *Taq*I, respectively.

The primer-extension method was performed as follows. 5′-$^{32}$P-labeled primer (0.01–0.05 pmol)

*Published by Elsevier Science Publishers B.V.*
00145793/84/$3.00 

and the 30-kDa protein mRNA (0.5 fmol) were annealed in a sealed capillary in 0.3 M KCl, 50 mM Tris–Cl (pH 7.9 at 42°C) by heating at 70°C for 10 min and cooling slowly to 30°C. The annealed mixture was diluted to 60 mM Tris–Cl, 10 mM $MgCl_2$, 10 mM DTT, 60 mM KCl, 1 mM of each dNTP. AMV reverse transcriptase (kindly provided by Dr A. Ishihama) was added to the mixture at 100 units/ml and incubated at 42°C for 45 min.

For S1-nuclease mapping the hybridization buffer and S1-digestion buffer were prepared as in [10]. Probes (0.05–0.2 pmol), the 30-kDa protein mRNA (1 fmol) and 20 $\mu$g carrier tRNA (brewer's yeast) were dissolved in the hybridization buffer, sealed in a capillary and heated at 90°C for 15 min. The mixture was kept at 50°C for 4 h, then poured into 30 vols precooled S1 digestion buffer. Hybrids were trimmed with S1 nuclease (Seikagaku Kogyo) at 1.5 units/ml at 45°C for 30 min.

After reverse transcription or S1-nuclease digestion the products were deproteinized with phenol/chloroform (1:1), precipitated with ethanol and resolved on the sequencing gels [7].

## 3. RESULTS AND DISCUSSION

The 30-kDa protein mRNA is synthesized transiently only at an early stage (2–9 h) in TMV infection [4]. Therefore, we isolated the 30-kDa protein mRNA from the cells 6 h postinfection and confirmed its messenger RNA activity for 30-kDa protein by in vitro translation in a reticulocyte lysate (not shown).

To determine the 5′-end of the 30-kDa protein mRNA precisely, we first carried out primer-extension experiments using two kinds of primers (fig.1). H-Pr or B-Pr was annealed with the 30-kDa protein mRNA or the genomic RNA. The reverse-transcribed products were then separated on a sequencing gel with an A + G ladder generated using H-S1 or B-S1 as in [7] (fig.2A). Only one band could be seen specific to the 30-kDa protein mRNA. This fact shows that the 30-kDa protein mRNA has no heterogeneity at the 5′-end. The cDNAs primed by H-Pr and B-Pr had identical termini (fig.2A). Considering that the sequencing ladder moves 1.5 bases faster than the reverse-transcribed DNA [11], the 5′-terminus of the

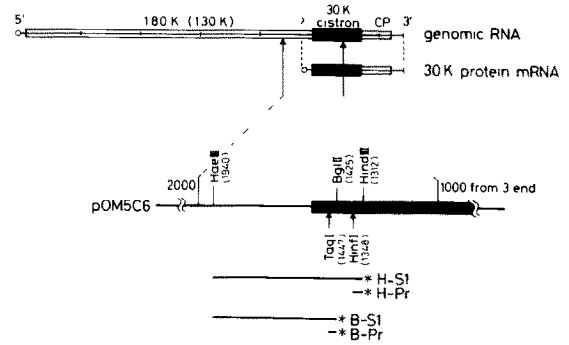


Fig.1. Primers for extension analysis and probes for S1 nuclease mapping. pOM5C6 [9] contains sequences of the 3′-portions of the genomic RNA of TMV OM. H-S1 and H-Pr were made from a *Hin*dIII fragment of pOM5C6. The fragment was 5′-$^{32}$P-labeled at the *Hin*dIII site with T4 polynucleotide kinase and subsequently digested with *Hae*III (H-S1) and *Hin*fI (H-Pr), respectively. B-S1 and B-Pr were made from a *Bgl*II fragment of pOM5C6. The fragment was 5′-$^{32}$P-labeled at the *Bgl*II site, and then digested with *Hae*III (B-S1) and *Taq*I (B-Pr), respectively. H-Pr and B-Pr were purified and the strands were separated on a 8.3 M urea–8% polyacrylamide gel [7]. H-S1 and B-S1 were purified on a 4% polyacrylamide gel and used without strand separation. Each fragment was 5′-end labeled at the terminus indicated by an asterisk (antisense strand). The nucleotide numbers in brackets indicate the position of the terminus of each fragment from the 3′-end of the genomic RNA.

30-kDa protein mRNA is deduced to be G at residue 1558 from the 3′-terminus.

In the S1-mapping method, only one band of S1 trimmed product was obtained specific to the 30-kDa protein mRNA (fig.2A). The band was, however, a few bases longer than the reverse transcript of the primer-extension method. The mobility of the band was not affected by a variety of S1 nuclease concentrations (0.3–3.0 units/ml). To resolve this discrepancy, we checked the initiation site for transcription of the CP mRNA by both the primer-extension and the S1-mapping method. By the primer-extension method, we could locate the initiation site of CP mRNA to be the same as reported in [12]. In the S1-mapping method, however, S1 trimmed products were a few bases longer than the reverse transcript (fig.2B). Therefore, on the basis of these results we conclude that transcription of the 30-kDa protein mRNA initiates at residue 1558 from the

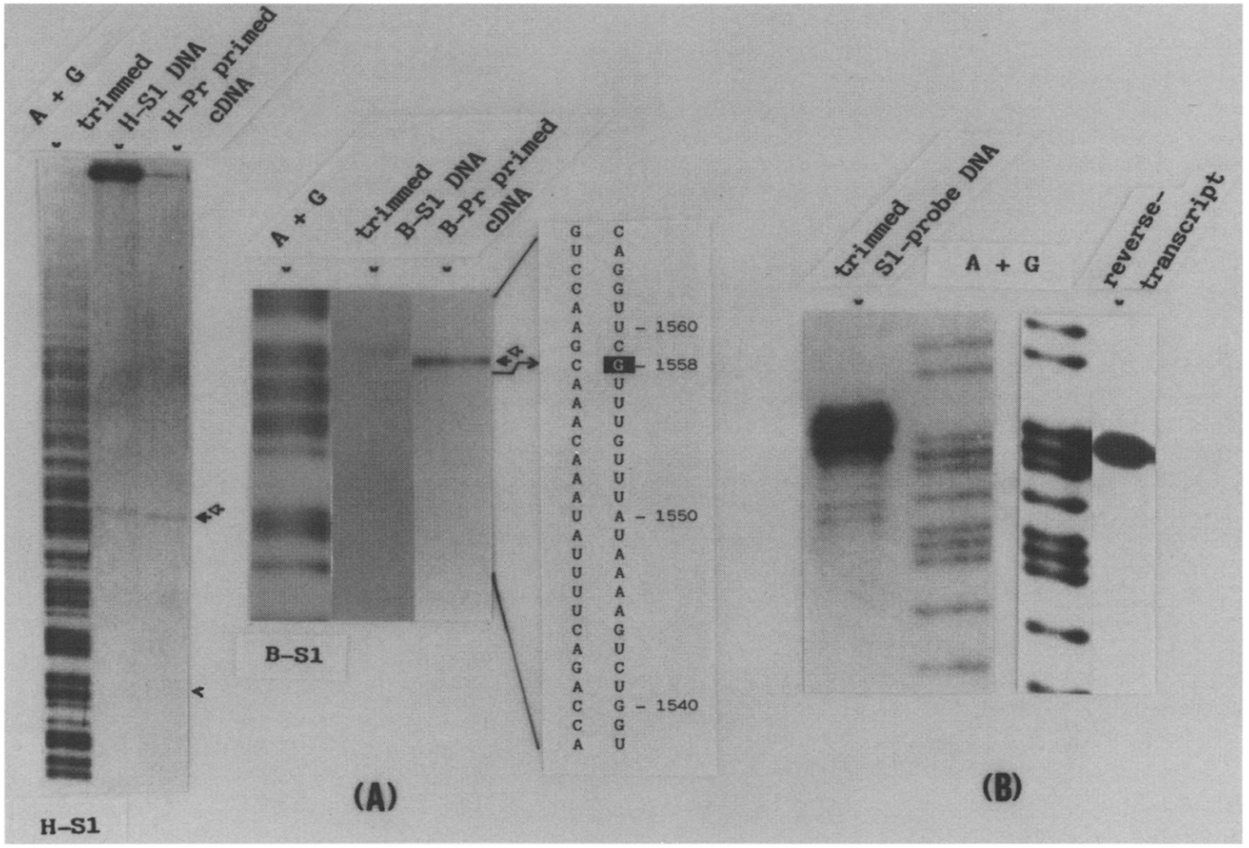


Fig.2. Analysis of the 5′-terminus of the 30-kDa protein mRNA (A) and CP mRNA (B) by primer-extension and S1 nuclease mapping analysis. A + G designate Maxam-Gilbert sequencing reactions [7] of each S1 probe (antisense strand). (A) The cDNA products of primer extension are shown by closed arrows. The S1 probes, which were hybridized with the 30-kDa protein mRNA and subsequently treated with S1-nuclease, are shown by open arrows. The cDNA bands migrated 1.5 bases behind the corresponding position of G (residue 1558 from the 3′-terminus of the genomic RNA). The arrow head indicates the banding position presumed in [1]. The sequences of the antisense strand are shown in the left array, those of sense strand are shown in the right array. The figures indicate the nucleotide number from the 3′-terminus of the genomic RNA. (B) The CP mRNA of the Cc strain was purified from rods as in [17] and analyzed by the S1-mapping and the primer-extension methods. The reverse transcript was unique, while the bands of S1 trimmed probe DNA were multiple probably due to transient breaking of base pairs at the end of the RNA:DNA hybrid during S1 nuclease digestion [18].

3′-terminus. Previously, authors in [13] reported that steric hindrance by the cap group caused the few deoxynucleotides corresponding to the region upstream of the 5′-terminus of the RNA to be resistant to S1 nuclease in the case of $\beta$-globin mRNA. The retardation of the mobility of the S1-trimmed product might suggest the existence of the cap structure in the 30-kDa protein mRNA. On the other hand, authors in [14] and [15] reported that the $I_2$ RNA appears to be uncapped. We tried to analyze chemically the 5′-terminal structure of the 30-kDa mRNA by labeling with vaccinia capping enzyme and [$\alpha$-$^{32}$P]GTP with or without prior $\beta$-elimination. However, the amount of the 30-kDa protein mRNA was so small that we could draw no conclusions on this point.

The initiation site for transcription of the 30-kDa protein mRNA is 65 nucleotides upstream from the AUG initiation codon of the 30-kDa protein cistron. Previously, we found a homology between the sequence around the capping site of the CP mRNA and the sequence upstream from the

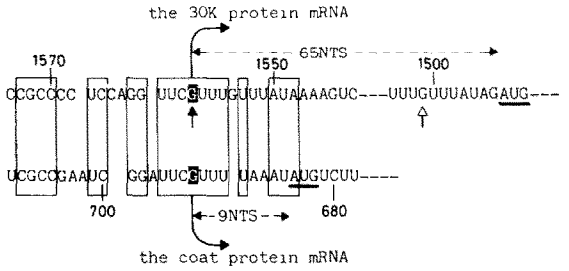


Fig.3. Comparison of the bordering regions of initiation sites for transcription of the two subgenomic RNAs on the genomic RNA. G bases in the black boxes are the initiation sites for transcription of each subgenomic RNA. Initiation codons are underlined. The closed arrow and the open arrow under the sequences indicate the positions we [14] and authors in [1] have predicted, respectively, as the initiation sites for transcription of the 30-kDa protein mRNA. The distances from the 3′-end are indicated in nucleotides.

30-kDa protein cistron and predicted that residue 1558 from the 3′-terminus might be the initiation site for transcription of the 30-kDa protein mRNA [16] (fig.3). The site of initiation of the 30-kDa protein mRNA was identical to our prediction. As shown in fig.3, both sequences share a common feature: A + U are abundant downstream of the initiation site and G + C are abundant upstream of the site. Although the mechanism of synthesis of the two subgenomic RNAs is still unknown, the sequences mentioned above possibly contain the recognition signal for the enzyme(s) which would participate in subgenomic RNA synthesis.

On the other hand, authors in [1] paid attention to a short U + A rich sequence, referred to as a 'Butler box', on the genomic RNA preceding the 5′-end of the CP mRNA. By analogy to the CP mRNA, they predicted that the G at nucleotide 1501 from the 3′-terminus could be the initiation site for transcription of the 30-kDa protein mRNA [1] (fig.3). Since we found no RNA species 1501 nucleotides long (fig.2A), the Butler box appears not to be significant in determining the initiation of transcription of the 30-kDa protein mRNA.

## ACKNOWLEDGEMENTS

We thank Dr A. Ishihama of Kyoto University for reverse transcriptase. This work was supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science and Culture, Japan.

## REFERENCES

[1] Goelet, P., Lomonossoff, G.P., Butler, P.J.G., Akam, M.E., Gait, M.J. and Karn, J. (1982) Proc. Natl. Acad. Sci. USA 79, 5818–5822.
[2] Hunter, T.R., Hunt, T., Knowland, J. and Zimmern, D. (1976) Nature 260, 759–764.
[3] Bruening, G., Beachy, R.N., Scalla, R. and Zaitlin, M. (1976) Virology 71, 498–517.
[4] Watanabe, Y., Emori, Y., Ooshika, I., Meshi, T., Ohno, T. and Okada, Y. (1984) Virology 133, 18–24.
[5] Guilley, H., Jonard, G., Kukla, B. and Richards, K.E. (1979) Nucleic Acids Res. 6, 1287–1308.
[6] Watanabe, Y., Ohno, T. and Okada, Y. (1982) Virology 120, 478–480.
[7] Maxam, A.M. and Gilbert, W. (1981) Methods Enzymol. 65, 499–560.
[8] Ooshika, I., Watanabe, Y., Meshi, T., Okada, Y., Igano, K., Inouye, K. and Yoshida, N. (1984) Virology 132, 71–78.
[9] Meshi, T., Takamatsu, N., Ohno, T. and Okada, Y. (1982) Virology 118, 64–75.
[10] Favaloro, J., Treisman, R. and Kamen, R. (1981) Methods Enzymol. 65, 718–749.
[11] Sollner-Webb, B. and Reeder, R.H. (1979) Cell 18, 485–499.
[12] Browning, K.S. and Clark, J.M. jr (1980) Biochemistry 19, 5922–5926.
[13] Weaver, R.F. and Weissmann, C. (1979) Nucleic Acids Res. 7, 1175–1193.
[14] Joshi, S., Pleij, C.W.A., Haenni, A.L., Chapeville, F. and Bosch, L. (1983) Virology 127, 100–111.
[15] Hunter, T., Jackson, R. and Zimmern, D. (1983) Nucleic Acids Res. 11, 801–821.
[16] Meshi, T., Ohno, T. and Okada, Y. (1982) J. Biochem. 91, 1441–1444.
[17] Fukuda, M., Okada, Y., Otsuki, Y. and Takebe, I. (1980) Virology 101, 493–502.
[18] King, T.C. and Schlessinger, D. (1983) J. Biol. Chem. 258, 12034–12042.