The EMBO Journal vol.6 no.9 pp.2557–2563, 1987

Function of the 30 kd protein of tobacco mosaic virus: involvement in cell-to-cell movement and dispensability for replication

Tetsuo Meshi, Yuichiro Watanabe, Tetsuichiro Saito, Asako Sugimoto, Tatsuya Maeda and Yoshimi Okada

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

Communicated by A.L.Haenni

We have investigated the function of the 30 kd protein of tobacco mosaic virus (TMV) by a reverse genetics approach. First, a point mutation of TMV Ls1 (a temperature-sensitive mutant defective in cell-to-cell movement), that causes an amino acid substitution in the 30 kd protein, was introduced into the parent strain, TMV L. The generated mutant showed the same phenotype as TMV Ls1, and therefore the one-base substitution in the 30 kd protein gene adequately explains the defectiveness of TMV Ls1. Next, four kinds of frame-shift mutants were constructed, whose mutations are located at three different positions of the 30 kd protein gene. All the frame-shift mutants were replication-competent in protoplasts but none showed infectivity on tobacco plants. From these observations the 30 kd protein was confirmed to be involved in cell-to-cell movement. To clarify that the 30 kd protein is not necessary for replication, two kinds of deletion mutants were constructed; one lacking most of the 30 kd protein gene and the other lacking both the 30 kd and coat protein genes. Both mutants replicated in protoplasts and the former still produced the subgenomic mRNA for the coat protein. These results clearly showed that the 30 kd protein, as well as the coat protein, is dispensable for replication and that no *cis*-acting element for replication is located in their coding sequences. It is also suggested that the signal for coat protein mRNA synthesis may be located within about 100 nucleotides upstream of the initiation codon of the coat protein gene.

Key words: cell-to-cell movement/replication/reverse genetics/30 kd protein/tobacco mosaic virus

Introduction

Tobacco mosaic virus (TMV) has been well-characterized both biochemically and pathologically (Hirth and Richards, 1981; Van Regenmortel and Fraenkel-Conrat, 1986). The genome of TMV is a positive-sense, single-stranded RNA and encodes at least three non-structural proteins (130 kd, 180 kd and 30 kd proteins) and the coat protein (Goelet *et al.*, 1982; Ohno *et al.*, 1984). While the 130 kd and 180 kd proteins are translated directly from the genomic RNA, the 30 kd and coat proteins are translated from their respective 3′ co-terminal subgenomic mRNAs (30 kd protein mRNA and CP mRNA) (Hunter *et al.*, 1976; Guilley *et al.*, 1979; Watanabe *et al.*, 1984b).

The 30 kd protein was first discovered as a product of I_2-class RNAs using a cell-free translation system (Bruening *et al.*, 1976). Its *in vivo* expression has been demonstrated using antibodies raised against synthetic peptides corresponding to the predicted C-terminal amino acid sequence (Kiberstis *et al.*, 1983; Ooshika *et al.*, 1984). The 30 kd protein is known to be synthesized transiently at an early stage of infection (Joshi *et al.*, 1983), being controlled at the level of 30 kd protein mRNA synthesis (Watanabe *et al.*, 1984a).

TMV Ls1 was isolated as a spontaneously-occurring temperature-sensitive mutant from a Japanese tomato strain, L (Nishiguchi *et al.*, 1978). At a restrictive temperature Ls1 can replicate and assemble normally in the initially infected cells of leaves as well as in protoplasts, but it cannot spread from cell to cell in the inoculated leaves (Nishiguchi *et al.*, 1978, 1980). This defectiveness of Ls1 at a high temperature can be complemented not only by other temperature-resistant strains of TMV but also by potato virus X, a virus belonging to another taxonomic group (Taliansky *et al.*, 1982c). It has been indicated that a virus-coded diffusible (*trans*-acting) factor functions in viral cell-to-cell movement and suggested that the virus-coded transport function is rather general among plant virus groups (Taliansky *et al.*, 1982b; Atabekov and Dorokhov, 1984).

Leonard and Zaitlin (1982) detected a slight difference in the 30 kd protein sequence between L and Ls1 by comparing the peptide maps of the 30 kd proteins synthesized *in vitro*, and suggested involvement of the 30 kd protein in cell-to-cell movement. This idea has been supported by a comparative nucleotide sequence study of TMV L and Ls1 RNAs that revealed an amino acid substitution from Pro to Ser in the 30 kd protein (Ohno *et al.*, 1983). However, since the peptide mapping could not detect all the mutations that the 130 kd/180 kd proteins might also have, and since the nucleotide sequence determined accounted for only one-fourth of the genome, the possibility remained that another mutation might be responsible for the *ts* transport phenotype of TMV Ls1. Similar observations have been reported for a nitrous acid-induced mutant, Ni2519, derived from a common strain, A19 (Taliansky *et al.*, 1982a; Zimmern and Hunter, 1983).

Accumulating data strongly suggest the involvement of the 30 kd protein in cell-to-cell movement but they are not conclusive. Moreover, it has not been demonstrated that the virus-coded, *trans*-acting factor is a protein rather than RNA *per se*. There are also no available data indicating whether or not the 30 kd protein plays a role in viral replication or in other viral functions.

Recently, *in vitro* expression systems that allow production of infectious TMV RNAs from cloned full-length cDNA copies have been established (Dawson *et al.*, 1986; Meshi *et al.*, 1986) and as a result a reverse genetics approach has become possible for TMV research (Ishikawa *et al.*, 1986). We have constructed several kinds of TMV mutants *in vitro*, whose 30 kd protein genes were modified or deleted, and analyzed their biological properties in both plants and protoplasts.

Results

Precise localization of the mutation in TMV Ls1 on the genomic sequence

The comparison of about 1600 nucleotides from the 3′ end between L and Ls1 RNAs has revealed three base substitutions, one of which, at residue 5365 from the 5′ end of the L genomic

RNA, causes an amino acid change from Pro (CCU in L RNA) to Ser (UCU in Ls1 RNA) (Ohno *et al.*, 1983). This point mutation was introduced into pLFW3 (Meshi *et al.*, 1986), that carries a full-length cDNA copy of the genomic RNA of TMV L, by replacing the *Hha*I/*Nco*I fragment (residues 4780–5462) of pLFW3 by the corresponding fragment of pLs1-1-33 (Ohno *et al.*, 1983), a sequenced cDNA clone of Ls1 RNA (Figure 1). The generated plasmid, pLFS1, has the same structure as pLFW3 except for the base at residue 5365 (Figures 1 and 2A).

pLFW3 and pLFS1 were linearized at the *Mlu*I site immediately downstream of the TMV cDNA insert (Figure 1) and used as templates for *in vitro* transcription by *Escherichia coli* RNA polymerase. The infectivity of the *in vitro* transcripts was examined by local lesion assay using *Nicotiana tabacum* L. cv. Xanthi nc as described previously (Meshi *et al.*, 1986). The transcript derived from pLFS1 produced smaller lesions than those produced by TMV-L and by the transcript derived from pLFW3, 2–3 days after inoculation; this was consistent with the reported lesion characteristic of TMV Ls1 (Nishiguchi *et al.*, 1978).

The temperature sensitivity of TMV Ls1 in cell-to-cell movement can be easily visualized by its inability to cause halos around necrotic local lesions when, after necrotic lesions could be seen (usually 2 days after inoculation), plants are subjected to a non-permissive temperature for a few days and then returned to a permissive temperature (Nishiguchi *et al.*, 1978). The *ts* phenotype of the *in vitro* transcripts was tested for its ability to cause halos. As shown in Figure 3, halos were observed around necrotic lesions produced by the transcript derived from pLFW3 but not by the transcript derived from pLFS1. These results showed that the point mutation at residue 5365 in the 30 kd protein gene alone is sufficient to explain the *ts* phenotype of Ls1 and, in other words, that the 30 kd protein and/or its coding sequence itself are/is the *trans*-acting factor for cell-to-cell movement.

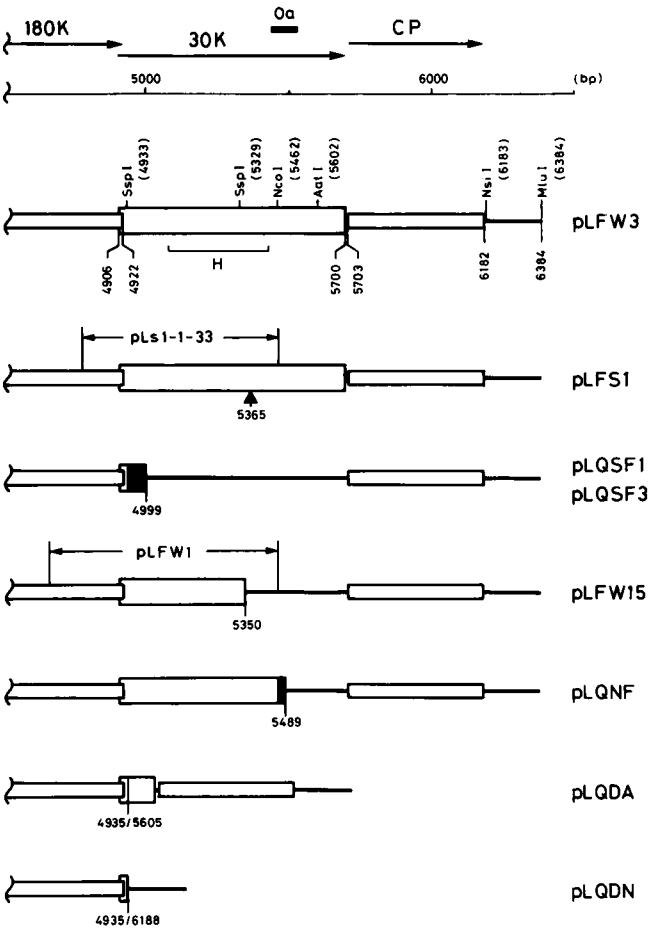

Fig. 1. Schematic representation of the constructed mutants. At the top of the figure, the regions coding for the C-terminal portion of the 180 kd protein (180K), the 30 kd protein (30K) and the coat protein (CP) are indicated. Oa denotes the assembly origin (Takamatsu *et al.*, 1983). Below the scale the coding and non-coding regions of each construct are shown by boxes and lines, respectively. The restriction sites with the first nucleotides of the recognition sequences in parentheses, the nucleotide residues at the first and last nucleotides of each gene and at the 3′ end of the genomic RNA are shown on the pLFW3. A bracket (H) under pLFW3 shows the region exhibiting relatively high homology among the 30 kd proteins of TMV strains and the 29 kd protein of tobacco rattle virus (Boccara *et al.*, 1986). The shaded boxes in pLQSF1, pLQSF3 and pLQNF show the amino acid sequences that are different from that of the authentic 30 kd protein. The point mutation in pLFS1 is shown by an arrow. The derivations of fragments in pLFS1 and pLFW15 are indicated above their constructs. The third nucleotides of the termination codons of the altered open reading frames in the frame-shift mutants and the nucleotides at the junctions of the deletion mutants are also indicated.

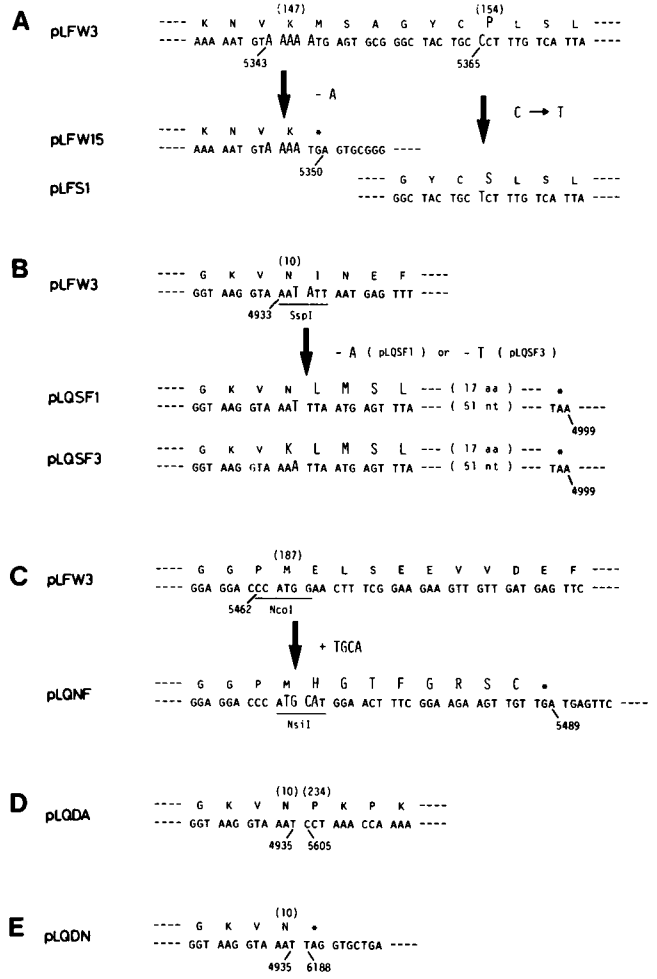

Fig. 2. The nucleotide and deduced amino acid sequences surrounding the mutations in pLFW15 and pLFS1 **(A)**, pLQSF1 and pLQSF3 **(B)**, pLQNF **(C)**, pLQDA **(D)** and pLQDN **(E)**. Nucleotide residues from the 5′ end of the genomic RNA and amino acid residues (in parentheses) from the N-terminal methionine of the 30 kd protein are indicated below and above the sequences, respectively. Deleted, inserted and substituted nucleotides are shown by large capitals. Amino acids that change with frame-shift mutations (in **A** and **B**) and the substituted amino acid found in TMV Ls1 (in **A**) are also shown by large capitals. Asterisks denote terminations.

Analysis of frame-shift mutants

To clarify the involvement of the 30 kd protein in cell-to-cell movement, we constructed four kinds of frame-shift mutants, whose mutations are located at three different positions in the 30 kd protein gene (Figures 1 and 2). pLQSF1 and pLQSF3 have one-base deletions of residues 4936 and 4935, respectively, in the *Ssp*I recognition sequence encoding the 10−11th amino acid residues from the N terminus of the 30 kd protein (Figure 2B). pLFW15 also has a one-base deletion in five consecutive adenine nucleotides (residues 5343−5347), resulting in the immediate appearance of a termination codon and in a truncated 30 kd protein composed of 147 amino acids (Figure 2A), i.e. about half the size of the authentic 30 kd protein of TMV L which is composed of 264 amino acids. pLQNF has a 4-base insertion in the *Nco*I recognition sequence (residues 5462−5467) encoding the 186−188th amino acid residues of the 30 kd protein (Figure 2C). This mutation is located in the assembly origin (Figure 1).

To know whether or not the mutations introduced affected viral replication, the replicability of the frame-shift mutants was exam-

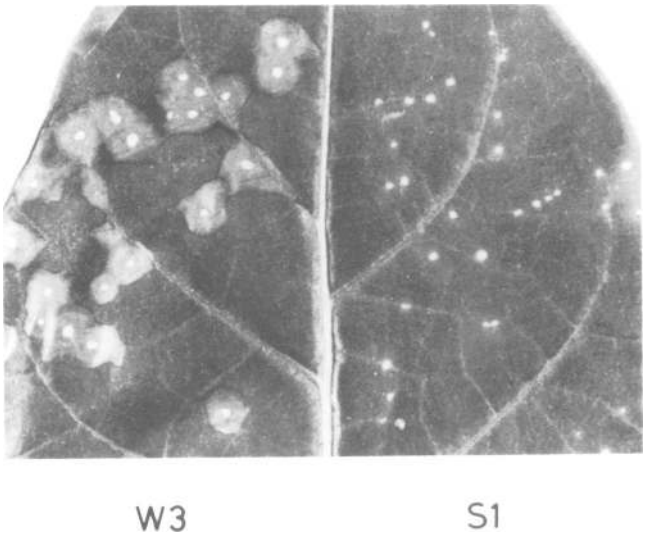

Fig. 3. Temperature-sensitive phenotype of the transcript from pLFS1. *In vitro* transcripts derived from pLFW3 and pLFS1 were reconstituted with the coat protein and inoculated on the left (W3) and right (S1) halves of a *Nicotiana tabacum* L. cv. Xanthi nc leaf, respectively. After inoculation, the plant was subjected to 25°C for 2 days, 32°C for 2 days and 25°C for 1 day, and then photographed.

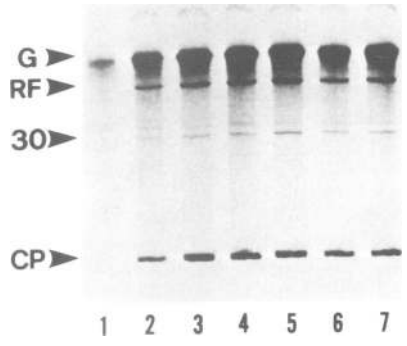

Fig. 4. Analysis of progeny viral RNAs of the frame-shift mutants. Protoplasts were mock-inoculated **(lane 1)** or inoculated with the L genomic RNA **(lane 2)**, and with the transcripts derived from pLQSF1 **(lane 3)**, pLQSF3 **(lane 4)**, pLFW15 **(lane 5)**, pLQNF **(lane 6)** and pLFW3 **(lane 7)**, respectively. The inoculated protoplasts were labeled with [^{3}H]uridine (25 μCi/ml) 6−8 h after inoculation. Labeled RNAs derived from 2×10^4 protoplasts were separated by 2.4% polyacrylamide-8M urea gel electrophoresis and detected by fluorography (the DMSO−PPO method). The positions of the genomic RNA (G), replicative form RNA (RF), the 30 kd protein mRNA (30) and the CP mRNA (CP) are indicated on the left. The RNA band just above the 30 kd protein mRNA is host-derived 18S rRNA.

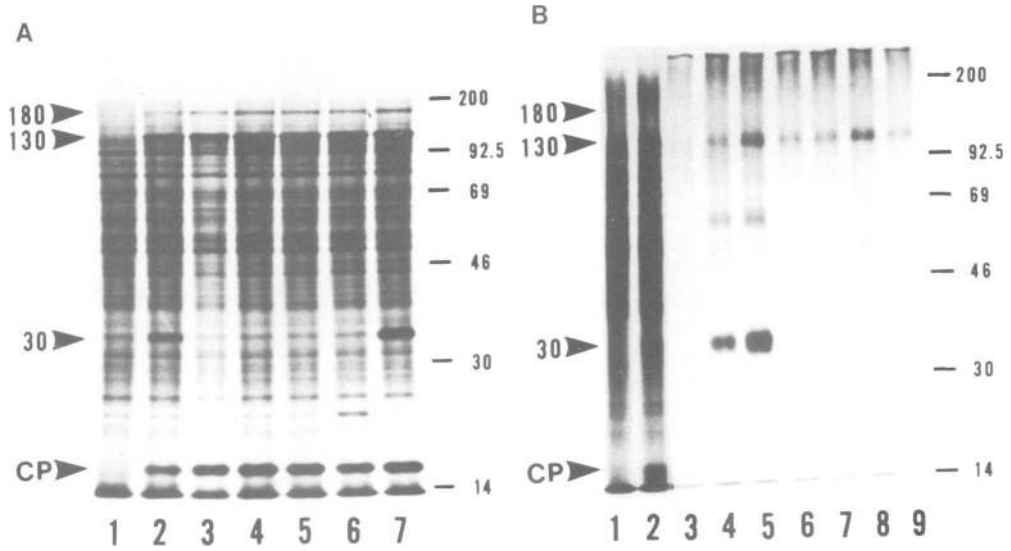

Fig. 5. Analysis of proteins synthesized in protoplasts inoculated with the frame-shift mutants. **(A)** Protoplasts inoculated were labeled with [^{35}S]methionine (2 μCi/ml) 6.5−7.5 h after inoculation and lysed. Proteins derived from 1×10^4 protoplasts were loaded on a 9.5% SDS−polyacrylamide gel. **Lane 1**, mock-inoculation; **lane 2**, inoculation with the L RNA; **lanes 3−6**, inoculation with transcripts derived from pLQSF1 **(lane 3)**, pLQSF3 **(lane 4)**, pLFW15 **(lane 5)**, pLQNF **(lane 6)** and pLFW3 **(lane 7)**. The positions of the 180 kd (180), 130 kd (130), 30 kd (30) and coat (CP) proteins are indicated at the left. At the right, the locations of mol. wt markers are shown. The 30 kd protein migrates to a different position from the 30 kd marker.
(B) Immunoprecipitation of the 30 kd protein. Protoplasts mock-inoculated **(lanes 1 and 3)** and inoculated with the L RNA **(lanes 2 and 4)**, and with the transcripts derived from pLFW3 **(lane 5)**, pLQSF1 **(lane 6)**, pLQSF3 **(lane 7)**, pLFW15 **(lane 8)** and pLQNF **(lane 9)** were labeled with [^{35}S]methionine (10 μCi/ml) 4−11 h after inoculation and lysed. Immunoprecipitation was performed using anti-30 kd protein antiserum that reacts with the C terminus of the 30 kd protein (Ooshika *et al.*, 1984) **(lanes 3−9)**, and therefore shortened 30 kd protein derivatives derived from the mutants could not be detected.

Table I. Summary of the biological characteristics of *in vitro* constructed TMV mutants[a]

Plasmid	Mutation	Production of local lesions[b]	Replication in protoplasts[c]	CP mRNA synthesis[c]
pLFW3	wild-type	+	+	+
pLFS1	substitution	+ (*ts*)	+	+
pLQSF1	frame shift	−	+	+
pLQSF3	frame shift	−	+	+
pLFW15	frame shift	−	+	+
pLQNF	frame shift	−	+	+
pLQDA	deletion	−	+	+
pLQDN	deletion	−	+	−

[a]The genomic organization of the constructed mutants and the sequences surrounding the mutation are shown in Figures 1 and 2. Each result was confirmed by using another independently isolated clone with the same structure, except for the cases of pLQSF1 and pLQSF3.
[b]Both reconstituted and naked *in vitro* transcripts (see Materials and methods) were used as inocula for pLFW3, pLFS1, pLQSF1, pLQSF3 and pLFW15. In the cases of pLQNF, pLQDA and pLQDN, only naked transcripts were used because of the low efficiency of reconstitution (pLQNF) or a lack of the assembly origin (pLQDA and pLQDN).
[c]Replication and CP mRNA synthesis in protoplasts were rated from Figures 4–8.

ined using a tobacco protoplast system. Protoplasts prepared from a tobacco suspension culture were inoculated with *in vitro* transcripts by electroporation, labeled with [^{35}S]methionine or [^{3}H]-uridine and sampled at intervals. Figures 4 and 5 show the *in vivo* synthesis of TMV-related RNAs and proteins, respectively. All the TMV-related molecules so far reported using our protoplast system (Watanabe *et al.*, 1984a, 1987a) except the 30 kd protein were labeled in protoplasts inoculated with the frame-shift mutants. The 30 kd protein could not be detected in protoplasts inoculated with the frame-shift mutants even by the use of anti-30 kd protein antibody (Figure 5B). These results indicate that all the frame-shift mutants constructed could replicate by themselves in single cells without the 30 kd protein (Table I). Although a slight difference could be seen in the band intensity for the genomic RNA and the CP mRNA between the frame-shift mutants and the parental pLFW3-derived virus (Figure 4, lanes 3–7), repeated experiments showed that this difference was mainly due to difference in the amounts of infectious transcripts in the inocula.

When the transcripts with a frame-shift mutation in the 30 kd protein gene were inoculated onto Xanthi nc tobacco leaves, no necrotic local lesions appeared (Table I). Considering that the mutation of Ls1 responsible for its defect in cell-to-cell movement was mapped in the 30 kd protein gene, it is likely that the apparent lack of infectivity (no production of lesions) was due to a defect in viral movement. However, the possibility remained that the mutants might spread without necrosis, because the determinant to induce necrotic lesions to tobacco plants with the *N* gene has not been elucidated. We therefore examined the propagation of the progeny viruses in the inoculated leaves of a systemic host, *N. tabacum* L. cv Samsun, 2 days or a week after inoculation as described in Materials and methods. For this purpose, two constructs, pLQSF3 and pLFW15, were chosen for the following reasons: no significant difference in replication was observed among the frame-shift mutants as described above, the mutations of pLQSF1 and pLQSF3 are located at the same position on the genomic sequence (Figure 2B), and the pLQNF-derived transcript was assembly-defective (data not shown).

Two days after inoculation, at the most about 10 ng of coat

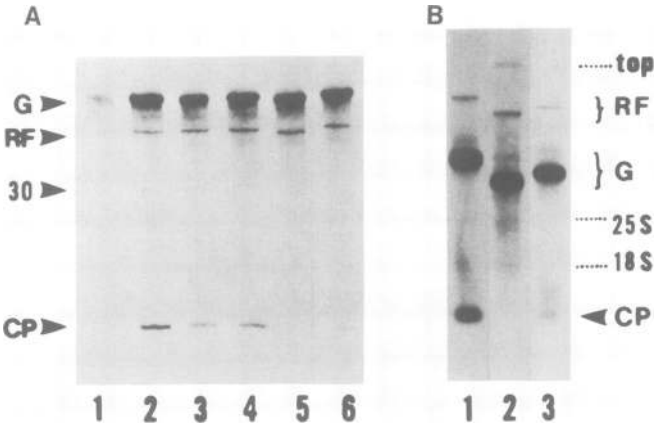

Fig. 6. Analysis of replication of the deletion mutants. **(A)** Progeny viral RNAs were analyzed as described in Figure 4, except that the labeling was performed 3–5 h after inoculation. **Lane 1**, RNA from mock-inoculated protoplasts; **lane 2**, RNA from TMV L-inoculated protoplasts; **lanes 3–6**, RNAs from protoplasts inoculated with the transcripts derived from pLFW3 **(lane 3)**, pLFS1 **(lane 4)**, pLQDN **(lane 5)** and pLQDA **(lane 6)**, respectively. Note that the difference in the mobility of the genomic RNA is not so obvious in this gel system. **(B)** 1% agarose gel electrophoresis of RNA from protoplasts inoculated with the transcripts derived from pLFW3 **(lane 1)**, pLQDN **(lane 2)** and pLQDA **(lane 3)**. Nucleic acids were fixed and detected by fluorography (the salicylic acid method). The intensity of bands does not correctly reflect the amount and the 30 kd protein mRNA cannot be seen because of the artificial loss of small sized RNAs when the agarose gel was dried up. The positions of 25S and 18S rRNAs are shown together with the TMV-related molecules.

protein could be detected in the concentrated virus fraction derived from 1 g of leaf tissue, which amounts to $10^{-2}-10^{-3}$ of the recovery of wild-type virus at the same time postinoculation (see Materials and methods). One week after inoculation, various amounts of coat protein (1 ng–5 μg per g leaf tissue) were detected, which amounts to about $10^{-3}-10^{-7}$ of the recovery of wild-type virus. These observations show that the frame-shift mutants infected tobacco plants and replicated at least in the initially infected cells of leaves. Taking account of the replicability of the frame-shift mutants in single cells, the reduced level of propagation in Samsun plants as well as the inability to produce necrotic lesions on Xanthi nc leaves would result from a defect (lost or decreased activity) in viral movement.

Back-inoculation with the concentrated virus fractions, of 14 plants assayed, onto Xanhi nc leaves revealed that considerable amounts of lesion-forming viruses were contained in all the fractions in which the coat protein detected was more than 0.5 μg per g tissue (four fractions: two each for the respective mutants). In the other ten fractions, in which the coat protein detected was less than 0.2 μg per g tissue, lesion-forming viruses could not be detected. Lesion-forming viruses recovered are thought to be revertants generated in the course of replication (not sequenced). They would be able to complement the transport defectiveness of the constructed frame-shift mutants pre-existing in Samsun plants, and might be a cause of the uneven recovery of progeny viruses at a week postinoculation. We do not know at present whether the frame-shift mutants were only restricted within the cells at the infection centres, or whether they spread slowly to a small number of cells in the immediate vicinity, for example due to very minute amounts of the 30 kd protein (or its derivative) that might be synthesized by frame-shift suppression.

Analyses of deletion mutants

The above finding that the frame-shift mutants can replicate in protoplasts also implies that the 30 kd protein is not necessary

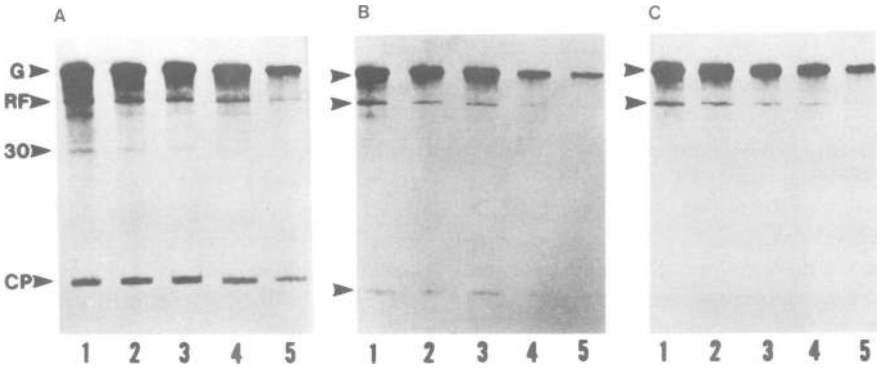

Fig. 7. Time course experiment on replication of the deletion mutants. Protoplasts were inoculated with the transcript derived from pLFW3 **(A)**, pLQDA **(B)** and pLQDN **(C)**, and labeled with [^{3}H]uridine for 2 h: 3–5 h **(lane 1)**, 6–8 h **(lane 2)**, 10–12 h **(lane 3)**, 16–18 h **(lane 4)** and 24–26 h **(lane 5)** after inoculation.

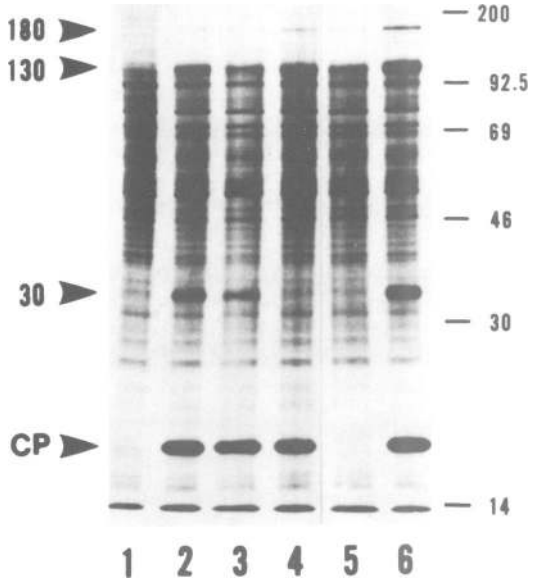

Fig. 8. Proteins synthesized in protoplasts infected with the deletion mutants. Proteins were analyzed as described in Figure 5A, except that the labeling was performed 3.5–4.5 h after inoculation. **Lane 1**, mock-inoculation; **lanes 2–5**, inoculation with the transcripts derived from pLFW3 **(lane 2)**, pLFS1 **(lane 3)**, pLQDA **(lane 4)** and pLQDN **(lane 5)**; **lane 6**, inoculation with the L RNA.

for replication. To demonstrate this more directly and to know whether any *cis*-acting element for replication is present in the 30 kd protein gene, we constructed two kinds of deletion mutants.

pLQDA has a deletion from the *Ssp*I site (residue 4936), 14 nucleotides downstream of the termination codon of the 180 kd protein gene, to the *Aat*I site (residue 5604), 96 nucleotides upstream of the termination codon of the 30 kd protein gene (Figures 1 and 2D). The other construct, pLQDN, has a deletion from the *Ssp*I site to the *Nsi*I site (residue 6187), just downstream of the termination codon of the coat protein gene (Figures 1 and 2E).

Neither transcript derived from pLQDA or pLQDN produced necrotic lesions on Xanthi nc tobacco leaves (Table I). The replicability of these deletion mutants was investigated in detail using the tobacco protoplast system. As shown in Figure 6, progeny viral RNAs with shortened sizes were labeled. Time course experiments did not show any significant difference of the synthetic patterns of the respective progeny viral RNAs from those observed on infection with the wild-type RNAs (Figure 7). The results clearly showed that both deletion mutants replicated by themselves in single cells. Therefore the 30 kd protein is not necessary for replication, and no *cis*-acting element for replication is located in the 30 kd protein coding sequence, except for the sequence encoding the N-terminal ten amino acids.

In protoplasts infected with the transcript derived from pLQDA, the CP mRNA and the resultant coat protein were synthesized (Figures 6A, 7 and 8). This observation implied that the sequence necessary for the CP mRNA synthesis is still retained in pLQDA.

Discussion

In this study, we have investigated the function of the 30 kd protein of TMV by a reverse genetics approach. From fine mapping of the mutation of TMV Ls1 responsible for the *ts* phenotype as to cell-to-cell movement and analyses of the *in vitro* constructed mutants with a frame-shift mutation in the 30 kd protein gene, the 30 kd protein has been confirmed to be involved in cell-to-cell movement. Moreover, the replication of both the frame-shift and deletion mutants in protoplasts has revealed that the 30 kd protein is not necessary for replication, and that there is no *cis*-acting element for replication in either the 30 kd protein coding sequence or the coat protein coding sequence (Takamatsu *et al.*, 1987).

Several 30 kd-related proteins with the common C terminus to the authentic 30 kd protein are translated from I_2-class subgenomic RNAs *in vitro*, internal AUG codons being used as initiation codons (Hunter *et al.*, 1983; Ooshika *et al.*, 1984). In pLQSF1 and pLQSF3, frame shifts were introduced between the first and second AUG codons of the 30 kd protein gene, and therefore the whole coding sequences of the 30 kd-related proteins remain intact. However, we could not detect them in protoplasts inoculated with the transcript derived from either pLQSF1 or pLQSF3 even after immunoprecipitation, which was consistent with previous *in vivo* observations (Kiberstis *et al.*, 1983; Ooshika *et al.*, 1984). That these frame-shift mutants are non-infectious in tobacco plants indicates that the 30 kd-related proteins, if any, could not replace the authentic 30 kd protein.

It has been thought that the host range of a given virus is often determined by the functioning of the virus-coded transport protein (Taliansky *et al.*, 1982b; Atabekov and Dorokhov, 1984). When a given virus can replicate in the initially infected cells of a plant but cannot move to the adjacent cells, the plant will be judged as a non-host. Based on this idea, the 30 kd protein would be

at least one of the host-range determinants in the case of TMV. Until now, five 30 kd protein genes of tobamoviruses have been sequenced: common strains *vulgare* (Goelet *et al.*, 1982) and OM (Meshi *et al.*, 1982a), tomato strain L (Takamatsu *et al.*, 1983), cowpea strain Cc (Meshi *et al.*, 1982b) and cucumber green mottle mosaic virus (Meshi *et al.*, 1983, and unpublished results). The 30 kd proteins of these strains are less conserved as to both size and sequence than the other TMV-coded proteins (Ohno *et al.*, 1984). The sequence divergence of the 30 kd protein might relate to the difference in natural hosts of each strain. If so, the divergence might reflect the establishment of a strain with a distinct host range in the course of evolution. It would be advantageous for rapid evolution resulting in the acquisition of a new host that the 30 kd protein is dispensable for replication.

The mechanism of viral transport is still obscure. So far several cases of complementation in viral spread have been reported for combinations of viruses belonging to different taxonomic groups (Atabekov and Dorokhov, 1984). From these observations, the transport function was suggested to be rather general among viruses (Taliansky *et al.*, 1982b,c; Atabekov and Dorokhov, 1984). This idea is supported by recent findings showing homology among the 30 kd proteins of TMV, the 29 kd protein of tobacco rattle virus (TRV) (Boccara *et al.*, 1986; Cornelissen *et al.*, 1986) and the product of gene I of caulimovirus (Hull *et al.*, 1986). The homology between the TMV 30 kd protein and the TRV 29 kd protein extends over almost the whole protein sequences and, especially in the middle portion of the proteins (Figure 1), the homology is similar to that between the 30 kd proteins of TMV strains (Boccara *et al.*, 1986). Interestingly the responsible mutation in Ls1 is a change of a conserved amino acid residue in one of the most homologous regions (Figure 1) (Ohno *et al.*, 1984; Boccara *et al.*, 1986). These findings suggest that the mechanism of cell-to-cell movement of TRV is sufficiently similar to that of TMV, and probably the two proteins function in the same manner.

On the other hand, little homology was found between the TMV 30 kd protein and the products of tripartite viruses, possibly P3 of alfalfa mosaic virus (AlMV) and 3A of brome mosaic virus (BMV) (Ahlquist *et al.*, 1985). Nevertheless, complementation in viral spread has been reported for a combination of TMV and BMV (Hamilton and Dodds, 1970; Taliansky *et al.*, 1982b). The difference in the amino acid sequence is not necessarily attributed to their different host ranges, because TMV and AlMV can propagate in tobacco plants. One possible explanation is that there might be several mechanisms that make viral cell-to-cell movement possible. The functioning of each viral transport protein would result in free movement of wide-range genetic materials, which would be observed as complementation. Atabekov and colleagues have proposed two possible ways how the transport protein works (Taliansky *et al.*, 1982c; Atabekov and Dorokhov, 1984). One is based on the hypothesis that plants do not permit free movement of viruses by nature. The transport protein could modify the host structure directly or indirectly to enable a virus to move possibly through plasmodesmata. The other is based on the idea that the protein could repress a kind of host resistance response, postulated to be caused by viral infection and to result in the shutting off of intercellular communication, so that the infected virus can move from cell to cell. Neither proposal can be ruled out on the basis of the available data, including those related to protein localization; the 30 kd protein of TMV is found in the nuclear fraction of infected protoplasts (Watanabe *et al.*, 1986) and in plasmodesmata of infected leaves (Tomenius *et al.*, 1987), and AlMV P3 is in the cell wall fraction of infected tobacco leaves (Godefroy-Colburn *et al.*, 1986).

It remains unclear how the 30 kd protein interacts with a presumed host component(s), how the transient expression of the 30 kd protein is regulated, and so on. In addition to the 30 kd protein, the complicated process of viral movement must involve many other components, some of which may be virus-coded. Recently, the 130 kd/180 kd proteins have been suggested to be involved in the 30 kd protein mRNA synthesis and, as a result, in cell-to-cell movement (Watanabe *et al.*, 1987b). A host factor may be found in an investigation using resistant tomato lines with the *Tm-2* or *Tm-2²* gene, which are known to block TMV infection at the level of cell-to-cell movement (Motoyoshi and Oshima, 1977).

We have found that the CP mRNA is synthesized in protoplasts inoculated with the transcript derived from pLQDA that have a large deletion in the 30 kd protein gene. The frame-shift mutants lacking the 30 kd protein also produced the CP mRNA. Therefore, the CP mRNA synthesis and the resultant coat protein synthesis can be investigated separately from the 30 kd protein coding information. Takamatsu *et al.* (1987) recently detected a shortened CP mRNA in tobacco leaves infected with a coat protein-less mutant lacking most of the coat protein gene (a 451 nucleotide deletion downstream from residue 5710), and suggested the necessary signal for the CP mRNA synthesis may locate upstream of residue 5709. Together with their observations and the result of pLQDA, the signal would be contained within about 100 nucleotides corresponding to residues 5605–5709 of the genomic RNA. However, a part of the signal might be deleted in pLQDA, because the CP mRNA synthesis of the pLQDA-derived progeny would be slightly reduced, compared with that of the wild-type virus (Figures 6A and 7).

Materials and methods

Plasmid construction

All recombinant DNA techniques used were essentially according to Maniatis *et al.* (1982). Enzymes were purchased from Takara Suzo Co., Toyobo Co. and Nippon Gene Co.

Construction of pLFS1. pLFW3 carries a cDNA copy of TMV-L RNA just downstream of the P_M promoter, a modified P_R promoter (Ahlquist and Janda, 1984), from which an infectious RNA can be transcribed *in vitro* using *E. coli* RNA polymerase (Meshi *et al.*, 1986). pLs1-1-33 carries a cDNA insert derived from about 1600 nucleotides of the 3′ end of TMV-Ls1 RNA (Ohno *et al.*, 1983). The 0.86 kb *HhaI/NcoI* fragment (residues 4780–5462) of pLs1-1-33 was isolated and introduced into pLFW3 in place of the corresponding fragment to create pLFS1.

Construction of pLQNF. pLFW3 was linearized with *NcoI* (residue 5462), filled in with *E. coli* DNA polymerase I (large fragment), and then re-ligated to create pLQNF. The resulting four-base insertion was confirmed by generation of the *NsiI* recognition sequence (Figure 2C).

Construction of pLFW15. pLFW1 was first constructed in order to establish an *in vitro* expression system but the *in vitro* transcript derived from pLFW1 was non-infectious in tobacco plants (Meshi *et al.*, 1986). *In vitro* translation experiments showed that no 30 kd protein was produced with the transcript from pLFW1 (unpublished observation). Sequencing (Messing, 1983) of the *AccI/NcoI* fragment (residues 4661–5462) of pLFW1 revealed a one-base deletion in the 30 kd protein gene (one of the five consecutive adenine residues), which resulted in a frame-shift in the 30 kd protein gene (Figure 2A). pLFW15 was constructed by replacing the 0.8 kb *AccI/NcoI* fragment of pLFW3 by the corresponding fragment of pLFW1.

Constructions of pLQSF1 and pLQSF3. The digestion of DNA by *SspI* resulted accidentally in the inability of the DNA to be re-cut by *SspI* after ligation, probably because of a small amount of a contaminating exonuclease in the enzyme preparation. After *SspI* digestion of pLFW3, the DNA was purified and digested separately with *KpnI* (residue 4390) and with *FokI* (residue 5295). The generated 0.54 kb *KpnI/SspI* (residues 4395–4935) and 0.37 kb *SspI/FokI* (residues 4936–5308) fragments were ligated with the 0.88 kb *FokI/NsiI* (residues 5309–6187), 0.2 kb *NsiI/MluI* (residues 6188–6384) and 7.9 kb *MluI/KpnI* [including the

vector, promoter and TMV (residues 1–4394) sequences] fragments of pLFW3. After plasmids without the *Sspl* recognition site at residue 4933 had been selected, the 0.29 kb *Taql* fragments (residues 4687–4981) containing the lost *Sspl* site were subcloned into an M13 vector and then sequenced (Messing, 1983). pLQSF1 and pLQSF3 had one-base deletions (an adenine base at residue 4936 and a thymine base at residue 4935, respectively) in the *Sspl* recognition sequences (Figure 2B).

Constructions of pLQDA and pLQDN. pLQDA was constructed by ligating the 0.54 kb *KpnI/SspI*, 0.78 kb *AatI/MluI* (residues 5605–6384) and 7.9 kb *MluI/KpnI* fragments of pLFW3. pLQDN was constructed by using the 0.54 kb *KpnI/SspI*, 0.2 kb filled-in *NsiI/MluI* and 7.9 kb *MluI/KpnI* fragments of pLFW3. The junction of each deletion was confirmed by sequencing after sub-cloning (Messing, 1983).

In vitro transcription

In vitro transcription from linearized plasmids at the *MluI* site was carried out essentially according to Ahlquist *et al.* (1984) with the following modification. The reaction mixture contained 50 mM Tris–HCl (pH 8.3), 10 mM $MgCl_2$, 150 mM NaCl, 1 mM DTT, 0.2 mM each of ATP, UTP and CTP, 25 μM GTP, 0.5 mM m^7GpppG (New England Biolabs), 3.3% glycerol, 1200 units/ml of RNase inhibitor (Takara Shuzo Co.) and 10 nM *MluI*-cut template DNA. The amount of RNA polymerase, that was purified from *E. coli* A19 by the reported method (Burgess and Jendrisak, 1975), was experimentally adjusted so as to obtain the highest yield of full-sized transcripts. The reaction was chased with GTP (final 25 μM) at 20 min and 40 min, and with RNA polymerase at 50 min, and stopped by the addition of EDTA (final 15 mM) and subsequent phenol extraction at 70 min. When transcripts were used for protoplast inoculation, template DNA was digested by adding DNase I (DPRF, Worthington) directly to the reaction mixture (0.7 units per μg template) at 70 min, followed by a further 15 min incubation.

Infectivity assay

Reconstitution of capped *in vitro* transcripts and their inoculation on tobacco leaves was performed essentially as described previously (Meshi *et al.*, 1986; Ishikawa *et al.*, 1986). One ml of inoculum contained reconstituted and naked transcripts derived from 3–5 μg and 330 μg of template DNA, respectively. Progeny viruses were extracted from the inoculated leaves of *N. tabacum* L. cv. Samsun, and concentrated essentially as described previously (Fukuda *et al.*, 1981). The propagation of the progeny viruses was monitored as the amount of viral antigen (coat protein) in the concentrated virus fraction by the Western blotting method (Saito *et al.*, 1986) using anti-L antibody (generously supplied by F.Motoyoshi). The detection limit was about 1–3 ng per g leaf tissue. The amount of wild-type virus (and also progeny of the pLFW3-derived transcripts) recovered from 1 g of tissue was usually 1–10 μg and 1–10 mg at 2 days and a week postinoculations, respectively.

Protoplast inoculation

Approximately 2 μg of full-sized *in vitro* transcripts derived from 10 μg of template DNA, and 0.2 μg of the L genomic RNA were inoculated to 1×10^6 tobacco protoplasts prepared from a suspension culture (BY-2) (Watanabe *et al.*, 1982) by electroporation essentially as described by Okada *et al.* (1986). The procedure will be reported in detail elsewhere (Watanabe *et al.*, 1987a). In the case of pLFW3-derived transcript, the inoculum used usually contained the infectious RNA corresponding to 0.05–0.2 μg of the viral RNA, based on the results of the infectivity assay. However, in the cases of the mutants, the precise amounts of infectious transcripts in the inocula could not be determined because of the lack of a reliable bioassay system. Usually 30–75% of the cells were infected. Syntheses of proteins and RNAs in protoplasts were analyzed essentially as described previously (Watanabe *et al.*, 1984a). The immunoprecipitation of the 30 kd protein was performed as described (Saito *et al.*, 1986) using anti-30 kd protein antiserum (Ooshika *et al.*, 1984).

Acknowledgements

We are grateful to P.Ahlquist and Agrigenetics Research Associates for the use of the P_M promoter. We thank M.Ishikawa, N.Takamatsu and Y.Imai for their help and valuable comments, and M.Nishiguchi and F.Motoyoshi for comments and anti-L antibody. This work was supported in part by Grants-in-Aid from the Ministry of Education, Science and Culture and from the Ministry of Agriculture, Forestry and Fisheries, Japan.

References

Ahlquist,P. and Janda,M. (1984) *Mol. Cell. Biol.*, **4**, 2876–2882.
Ahlquist,P., French,R., Janda,M. and Loesch-Fries,L.S. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 7066–7070.
Ahlquist,P. Strauss,E.G., Rice,C.M., Strauss,J.H., Haseloff,J. and Zimmern,D. (1985) *J. Virol.*, **53**, 536–542.
Atabekov,J.G. and Dorokhov,Yu.L. (1984) *Adv. Virus Res.*, **29**, 313–364.
Boccara,M., Hamilton,W.D.O. and Baulcombe,D.C. (1986) *EMBO J.*, **5**, 223–229.
Bruening,G., Beachy,R.N., Scalla,R. and Zaitlin,M. (1976) *Virology*, **71**, 498–517.
Burgess,R.R. and Jendrisak,J.J. (1975) *Biochemistry*, **14**, 4634–4638.
Cornelissen,B.J.C., Linthorst,H.J.M., Brederode,F.Th. and Bol,J.F. (1986) *Nucleic Acids Res.*, **14**, 2157–2169.
Dawson,W.O., Beck,D.L., Knorr,D.A. and Grantham,G.L. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 1832–1836.
Fukuda,M., Meshi,T., Okada,Y., Otsuki,Y. and Takebe,I. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 4231–4235.
Godefroy-Colburn,T., Gagey,M.-J., Berna,A. and Stussi-Garaud,C. (1986) *J. Gen. Virol.*, **67**, 2233–2239.
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.
Guilley,H., Jonard,G., Kukla,B. and Richards,K.E. (1979) *Nucleic Acids Res.*, **6**, 1287–1308.
Hamilton,R.I. and Dodds,J.A. (1970) *Virology*, **42**, 266–268.
Hirth,L. and Richards,K.E. (1981) *Adv. Virus. Res.*, **26**, 145–199.
Hull,R., Sadler,J. and Longstaff,M. (1986) *EMBO J.*, **5**, 3083–3090.
Hunter,T.R., Hunt,T., Knowland,J. and Zimmern,D. (1976) *Nature*, **260**, 759–760.
Hunter,T., Jackson,R. and Zimmern,D. (1983) *Nucleic Acids Res.*, **11**, 801–821.
Ishikawa,M., Meshi,T., Motoyoshi,F., Takamatsu,N. and Okada,Y. (1986) *Nucleic Acids Res.*, **14**, 8291–8305.
Joshi,S., Pleij,C.W.A., Haenni,A.L., Chapeville,F. and Bosch,L. (1983) *Virology*, **127**, 100–111.
Kiberstis,P.A., Pessi,A., Atherton,E., Jackson,R., Hunter,T. and Zimmern,D. (1983) *FEBS Lett.*, **164**, 355–360.
Leonard,D.A. and Zaitlin,M. (1982) *Virology*, **117**, 416–424.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
Meshi,T., Ohno,T. and Okada,Y. (1982a) *J. Biochem.*, **91**, 1441–1444.
Meshi,T., Ohno,T. and Okada,Y. (1982b) *Nucleic Acids Res.*, **10**, 6111–6117.
Meshi,T., Kiyama,R., Ohno,T. and Okada,Y. (1983) *Virology*, **127**, 54–64.
Meshi,T., Ishikawa,M., Motoyoshi,F., Semba,K. and Okada,Y. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 5043–5047.
Messing,J. (1983) *Methods Enzymol.*, **101**, 20–78.
Motoyoshi,F. and Oshima,N. (1977) *J. Gen. Virol.*, **34**, 499–506.
Nishiguchi,M., Motoyoshi,F. and Oshima,N. (1978) *J. Gen. Virol.*, **39**, 53–61.
Nishiguchi,M., Motoyoshi,F. and Oshima,N. (1980) *J. Gen. Virol.*, **46**, 497–500.
Ohno,T., Takamatsu,N., Meshi,T., Okada,Y., Nishiguchi,M. and Kiho,Y. (1983) *Virology*, **131**, 255–258.
Ohno,T., Aoyagi,M., Yamanashi,Y., Saito,H., Ikawa,S., Meshi,T. and Okada,Y. (1984) *J. Biochem.*, **96**, 1915–1923.
Okada,K., Nagata,T. and Takebe,I. (1986) *Plant Cell Physiol.*, **27**, 619–626.
Ooshika,I., Watanabe,Y., Meshi,T., Okada,Y., Igano,K., Inouye,K. and Yoshida, N. (1984) *Virology*, **132**, 71–78.
Saito,T., Watanabe,Y., Meshi,T. and Okada,Y. (1986) *Mol. Gen. Genet.*, **205**, 82–89.
Takamatsu,N., Ohno,T., Meshi,T. and Okada,Y. (1983) *Nucleic Acids Res.*, **11**, 3767–3778.
Takamatsu,N., Ishikawa,M., Meshi,T. and Okada,Y. (1987) *EMBO J.*, **6**, 307–311.
Taliansky,M.E., Atabekova,T.I., Kaplan,I.B., Morozov,S.Yu., Malyshenko,S.I. and Atabekov,J.G. (1982a) *Virology*, **118**, 301–308.
Taliansky,M.E., Malyshenko,S.I., Pshennikova,E.S. and Atabekov,J.G. (1982b) *Virology*, **122**, 327–331.
Taliansky,M.E., Malyshenko,S.I., Pshennikova,E.S., Kaplan,I.B., Ulanova,E.F. and Atakbekov,J.G. (1982c) *Virology*, **122**, 318–326.
Tomenius,K., Clapham,D. and Meshi,T. (1987) *Virology*, in press.
Van Regenmortel,M.H.V. and Fraenkel-Conrat,H. (1986) *The Plant Viruses 2. The Rod-shaped Plant Viruses.* Plenum Press, New York.
Watanabe,Y., Ohno,T. and Okada,Y. (1982) *Virology*, **120**, 478–480.
Watanabe,Y., Emori,Y., Ooshika,I., Meshi,T., Ohno,T. and Okada,Y. (1984a) *Virology*, **133**, 18–24.
Watanabe,Y., Meshi,T. and Okada,Y. (1984b) *FEBS Lett.*, **173**, 247–250.
Watanabe,Y., Ooshika,I., Meshi,T. and Okada,Y. (1986) *Virology*, **152**, 414–420.
Watanabe,Y., Meshi,T. and Okada,Y. (1987a) *FEBS Lett.*, in press.
Watanabe,Y., Morita,N., Nishiguchi,M. and Okada,Y. (1987b) *J. Mol. Biol.*, **194**, 699–704.
Zimmern,D. and Hunter,T. (1983) *EMBO J.*, **2**, 1893–1900.

Received on May 4, 1987; revised on June 11, 1987