# COMMON NUCLEOTIDE SEQUENCES ON LONG AND SHORT RNAs OF TURNIP YELLOW MOSAIC VIRUS

Kenneth RICHARDS, Jean-Paul BRIAND, Claudine KLEIN and Gérard JONARD
*Laboratoire de Virologie, Institut de Biologie Moléculaire et Cellulaire du CNRS, 15, Rue Descartes, 67084 Strasbourg Cédex, France*

Received 11 January 1977

## 1. Introduction

In a previous publication [1] we have shown that the RNA of turnip yellow mosaic virus (TYMV) can be separated into two discrete and functionally distinct species by moderate heat treatment in the presence of SDS or EDTA. The larger species of RNA (mol. wt $2 \times 10^6$) when translated in a cell-free protein synthesizing system, gave rise to polypeptides as large as 165 000 mol. wt but little if any coat protein whereas the short RNA (mol. wt $0.3 \times 10^6$) coded for the coat protein. Similar observations have recently been reported by Pleij et al. [2]. The short and long RNAs cannot be products of accidental in vitro degradation of a larger chain since both are 'capped' with $m^7G$ at their 5'-ends [1].

There is every reason to believe that the TYMV short RNA fulfills the same role in vivo as in the cell-free protein synthesizing system, i.e., that it is the functional messenger for coat protein in infected cells. This by no means excludes the possibility, however, that the coat protein cistron may likewise be present on the long RNA but remain inactive in translation. Indeed, a number of viruses are known in which the subgenomic coat protein messenger becomes encapsidated and can be copurified with the complete genome RNA which, itself, is nonfunctional in coat protein synthesis [3].

Pleij et al. [2] found that purified long RNA could induce local lesions on cabbage plants, suggesting that it contains the whole genome, including the coat protein cistron. Such infectivity tests may be misleading, however, in view of the low sensitivity of the test plant and because it is difficult to rule out trace levels of contamination of long RNA by short RNA molecules. Therefore, we have chosen to search for homologies of sequence between long and short RNAs by comparing their $T_1$ RNAase catalogues. In this paper we show that all the large $T_1$ RNAase oligonucleotides of short RNA, including sequences from the coat protein cistron and the amino acid accepting 3'-OH-extremity, are also to be found in purified long RNA.

## 2. Material and methods

$^{32}P$-Labeled TYMV RNA was prepared as described by Briand and collaborators [4]. The long and short RNA components were routinely separated by passage through a $100 \times 2.5$ cm column of Ultrogel AcA22 (LKB) after heating the unfractionated RNA to 60°C for 10 min [1]. For sequence analysis the short RNA was further purified on a second Ultrogel column. The purity of the two components was controlled on polyacrylamide–agarose (2.4%/0.5%) gels (fig.1).

Total $T_1$ RNAase digestion (1 unit $T_1$ RNAase/50 $\mu$g RNA) of purified long and short RNA was for 1 h at 37°C in 0.1 M Tris–HCl, 1 mM EDTA, pH 7.2. The digests were then analyzed by two-dimensional gel electrophoresis [5,6]. The first-dimension was a 10% polyacrylamide (bisacrylamide/acrylamide, 1:30, w/w) gel-slab ($40 \times 20 \times 0.4$ cm) in 6 M urea, 0.025 M citric acid, pH 3.5. The bromophenol blue marker was allowed to migrate 15 cm. The second-dimension was a 20% polyacrylamide slab-gel ($40 \times 30 \times 0.4$ cm) in 0.1 M Tris–borate, 0.0025 M EDTA, pH 8.3. The blue marker migrated about 22 cm. Oligonucleotides were located by autoradiography and

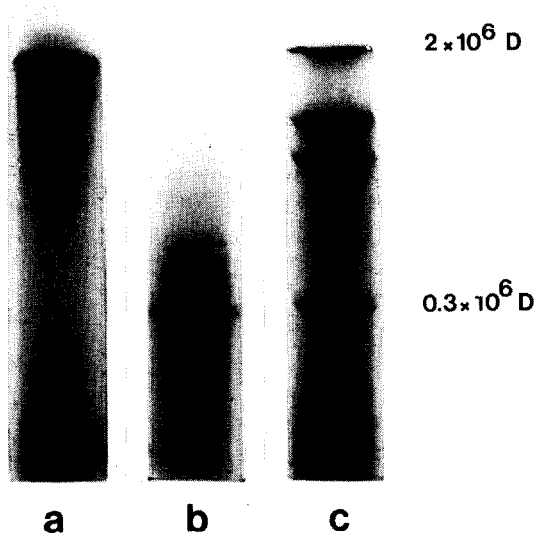


Fig.1. Analysis of Ultrogel column purified long and short TYMV RNA on polyacrylamide–agarose (2.4%/0.5%) gels. (a) Purified long RNA. (b) Purified short RNA. (c) Markers, TMV RNA (mol. wt $2 \times 10^6$) and BMV RNAs (mol. wt $1.1 \times 10^6$, $1 \times 10^6$, $0.7 \times 10^6$ and $0.3 \times 10^6$).

eluted from the gel electrophoretically onto disks of DEAE-paper. Further characterization of the oligonucleotides by pancreatic and $U_2$ RNAase digestion was performed under standard conditions [4].

## 3. Results

Separation of oligonucleotides by two-dimensional gel electrophoresis is based upon a fractionation according to charge in the first-dimension and on the basis of size (with some influence from secondary structure) in the second-dimension. Figure 2 shows

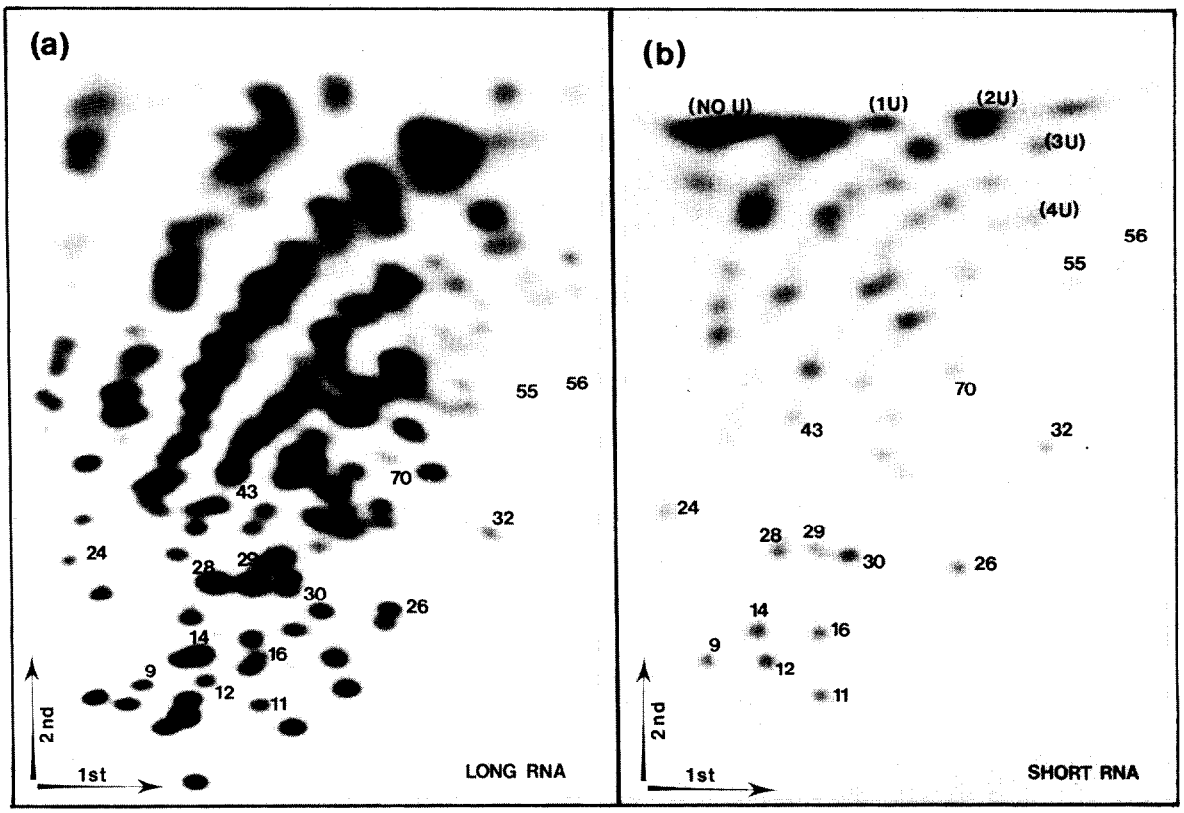


Fig.2. Two-dimensional polyacrylamide gel electrophoresis of total ribonuclease $T_1$ products of purified long (a) and short (b) TYMV RNA. Electrophoresis was from left to right in a 10% gel, at pH 3.5 and from bottom to top in a 20% gel-slab, at pH 8.3.

Table I
Partial characterization of some of the $T_1$ oligonucleotides common to long and short RNA

| Spot | Pancreatic RNAase products (average of 3 determinations) | Some $U_2$ RNAase products | Approximate chain length (nucleotides) |
|---|---|---|---|
| 9 | 1 AAC, 2 AU, 1 AAC, 2 AC, 11–12 C, 2 U, 1 G | (4 C, 2 U)A, (C, U)A, CCG | 30 |
| 11 | 2 AAC, 3 AC, ~7 C, 8–9 U, G | (2–3 C, 3–4 U)A, (3 C, 2 U)A (2 C, U)A, CG | 29 |
| 12 | 2 AU, 1 AAC, 2 AC, ~12 C, 4 AU, G | | 28 |
| 14 | 1 AU, 1 AAC, 1–2 AC, ~7 C, 3 U, G | | 20 |
| 16 | 2–3 AC, 7–8 C, 4–5 U, G | (3 C, 3–4 U)A, (1 U, 4–5 C)A, (3–4 C, U)A, CCG | 20 |
| 24 | 1 AAC, 2–3 AC, ~9 C, 1 U, G | | 20 |
| 26 | 2 AU, 1 AC, 6–7 C, ~6–9 U, G | | 21 |
| 28 | 2 AAU, 5–7 AC, ≥14 C, 6–9 U, 1 G, 1 AAAG | | 2 × 22 |
| 29 | 1 AAU, 1 AU, 2 AAC, ~4 AC, ~12 C, ~9 U, 2 AG | | 2 × 22 |
| 30 | 2 AU, 2 AC, ~10 C, 6–7 U, G | (3 C, 1 U)A (2 C, U)A, (2 C, 3 U)G (C, U)A | 25 |
| 32 | 1 AC, 3–4 C, ~10 U, G | A, (4–5 C, 10 U)G | 17 |
| 43 | 3 AC, 3 C, 3 U, AAG | (2 U, 3 C)A, (2 C, U)A, | 15 |
| 55 | 3 C, 6 U, 1 G | | 10 |
| 56 | 1 AU, 1 C, 4 U, 1 G | | 8 |
| 70 | 1 AAAAU, 1 AU, 2 C, 3 U, 1 G | (4 U, C)A, (U, C)G | 13 |

the $T_1$ RNAase fingerprints of the purified short and long RNAs by two-dimensional gels. Note that most of the spots, particularly those which migrate rapidly in the second-dimension, fall into well-defined graticles depending upon the content of U. The extreme left-hand graticle is made up of oligonucleotides with no U, the second graticle contains products with one U, etc. (Table 1 plus further observations.)

It is evident that spots corresponding in mobility to all the large $T_1$ RNAase oligonucleotides in short RNA (fig.2(b)) are also to be found in long RNA (fig.2(a)). Further evidence for the identity of like-migrating products comes from a comparison of their pancreatic and $U_2$ RNAase catalogues. In every case corresponding oligonucleotides gave rise to the same $U_2$ and pancreatic RNAase end-products upon ionophoresis on DEAE-paper, at pH 1.9 (table 1). Among the well-characterized $T_1$ RNAase oligonucleotides we have found two oligonucleotides which have been previously sequenced, spots 43 and 70. Spot 43 has the composition expected for the oligonucleotide ACACUUCCACCUAAG, which is known to occur at the 3′-end of the coat protein cistron [4]. Thus the coat protein cistron must be present on the large as well as the small RNA. Preliminary observations suggest that oligonucleotides 9, 14 and 16 may also derive from the coat protein cistron. Oligonucleotide 70 produced AAAAU, AU, 2C, 3U and G upon pancreatic RNAase digestion and probably corresponds to the oligonucleotide AUCUUUAAAAUCG from the tRNA-like structure known to be present at the 3′-OH-end of TYMV RNA [4].

## 4. Discussion

The findings described in this paper leave little room for doubt that the sequence of the short RNA of TYMV is also present on long RNA. In particular, both RNAs contain the coat protein cistron. In view of this duplication of genetic information we would expect that the long RNA alone should suffice for pathogenicity. The absence of coat protein among the translation products of long RNA [1,2] suggests that secondary structure has rendered the binding site for ribosomes at the coat protein cistron inaccessible.

A subgenomic coat protein messenger has recently been discovered in tobacco cells infected with tobacco mosaic virus [7]. The messenger is a fragment of about 750 nucleotides which is produced from the 3′-OH-end of the full-length RNA in the course of infection and which does not become encapsulated. Full-length tobacco mosaic virus RNA is not active in coat protein translation. Thus the system of regulation of coat protein production in tobacco mosaic virus and TYMV can be seen to be rather similar, with the exception that the coat protein messenger is encapsulated in the case of TYMV. An important unanswered question is how the coat protein messenger in such viruses is produced from the full-length RNA.

## Acknowledgements

We wish to thank Professor L. Hirth for his continuous interest in this work. This research was in part supported by the Commissariat à l'Energie Atomique.

## References

[1] Klein, C., Fritsch, C., Briand, J. P., Richards, K. E., Jonard, G. and Hirth, L. (1976) Nucleic Acids Res. 3, 3043–3061.
[2] Pleij, C. W. A., Neeleman, A., Van Vloten-Doting, L. and Bosch (1977) L. Proc. Natl. Acad. Sci. USA in press.
[3] Jaspars, E. M. J. (1974) Advan. Virus Res. 19, 37–149.
[4] Briand, J. P., Jonard, G., Guilley, H., Richards, K. E. and Hirth (1977) L. Eur. J. Biochem. in press.
[5] De Wachter, R. and Fiers, W. (1972) Anal. Biochem. 49, 184–197.
[6] Frisby, D. P., Newton, C., Carey, N. H., Fellner, P., Newman, J. F. E., Harris, T. J. R. and Brown, F. (1976) Virology 71, 379–388.
[7] Hunter, T. R., Hunt, T., Knowland, J. and Zimmern, D. (1976) Nature 260, 759–764.