STUDIES ON THE BACTERIOPHAGE MS2. V. THE 5′-TERMINAL TETRANUCLEOTIDE SEQUENCE OF THE VIRAL RNA CHAIN

R.De WACHTER, J.–P.VERHASSEL and W.FIERS
Laboratory of Physiological Chemistry and of Molecular Biology, University of Ghent, Belgium

Received 5 June 1968

1. Introduction

The initial nucleotide of the RNA of the phages R17 [1], f2 (Dahlberg, personal communication), Qβ, R23 [2], and MS2 [3] is pppGp... . Hydrolysis of MS2 RNA with pancreatic RNase should produce, in addition to a large number of oligonucleotides of the form $(Pup)_nPyp$, one 5′-terminal oligonucleotide of the form $pppGp(Pup)_nPyp$ (where $n \geqslant 0$). Pancreatic RNase hydrolysates of ^{32}P-labelled MS2 RNA were separated into oligonucleotide groups of equal chain length by salt gradient chromatography on DEAE-cellulose columns and on DEAE-cellulose paper. pppGp was detected in alkaline hydrolysates of the heptanucleotide fraction isolated by column chromatography and of the pentanucleotide fraction isolated by paper chromatography. The different behaviour of the pppGp-containing oligonucleotide in the two separation methods allowed us to isolate it by a combination of both procedures. It was found to be a tetranucleotide with the sequence pppGpGpGpUp.

This means that no initiation codon for polypeptide synthesis coincides with the 5′-end of the RNA molecule. Comparison with the 3′-terminal sequence of MS2 RNA [4] reveals a striking complementarity, which suggests that both ends of the molecule are joined in a base-paired region.

2. Materials and methods

The preparation of unlabelled and ^{32}P-labelled MS2 RNA, the hydrolysis with pancreatic RNase, and the separation of the oligonucleotides on a DEAE-cellulose column were essentially as described previously [5].

Descending chromatography on DEAE-cellulose paper with a salt concentration gradient was carried out by means of a special trough [6]. Oligonucleotides were separated on 15 × 57 cm strips of Whatman DE 81 paper with a 40 ml linear gradient from 0.05 M Tris chloride + 0.025 M Tris + 8 M formamide to 0.35 M Tris chloride + 0.175 Tris + 8 M formamide. Nucleotide material was recovered from the paper by cutting out the spots, washing in 0.02 M NH_4HCO_3, elution with 200 μl 2 M triethylammonium bicarbonate, and lyophilization.

Desalted oligonucleotides were hydrolyzed in 0.3 M KOH for 18 h at 37° or in 1 N KOH for 1 h at 80°. The mixture was neutralized with $HClO_4$ after addition of carrier mononucleotides and pppGp. The latter was a synthetically prepared product (Messens and Van Montagu, in preparation). Descending paper chromatography was carried out on Whatman 3 MM paper with isobutyric acid – 1 M NH_3 (10:6 by vol) as solvent. Whatman 52 paper was used for electrophoresis; the buffer was 5% acetic acid and 0.5% pyridine by vol (pH 3.5), and the voltage gradient 100 V/cm.

3. Results

A pancreatic RNase hydrolysate of ^{32}P-labelled MS2 RNA (40 mg; 3.9×10^8 dpm) was separated on a DEAE-cellulose column at pH 7.9 in the presence of 7 M urea. Under such conditions, pppGp mi-

Table 1
pppGp-content of penta-, hexa-, and heptanucleotide fractions.

Fraction	percent pppGp found		Possible pppGp
	column chromatography	paper chromatography	content (a)
Pentanucleotides	0	0.5	1.3
Hexanucleotides	0	0	2.9
Heptanucleotides	2.5	0	3.9

Aliquots of the desalted oligonucleotides ($1 - 9 \times 10^5$ dpm) were hydrolyzed with alkali. The hydrolysates were paper-chromatographed with an isobutyric acid solvent, which separates in order of increasing R_F value pppGp, Gp+Up, Cp, and Ap. The pppGp-content was determined by strip scanning for ^{32}P and integration of the peaks.

(a) Calculated from the distribution of oligonucleotides in a pancreatic RNase hydrolysate and assuming a chain length of 3300 nucleotides for MS2 RNA.

grates with the tetranucleotides [1,3]. The 5′-terminal oligonucleotide pppGp(Pup)$_n$Pyp was therefore expected to move with the fragments of chain length 5 or higher. On the other hand, oligonucleotides of chain length 8 and higher have been extensively investigated [7] and have failed to reveal the presence of a pppGp-containing sequence. The penta-, hexa-, and heptanucleotides were each subjected to alkaline hydrolysis and the pppGp-content determined by paper chromatography. pppGp was only present in the heptanucleotides (table 1, left column). This suggests a sequence containing three more negative charges than pppGp, i.e. pppGpPupPupPyp.

A pancreatic RNase hydrolysate of MS2 RNA (0.75 mg; 5×10^7 dpm) was also separated by gradient chromatography on DEAE-cellulose paper at pH 8 in the presence of 8 M formamide (a similar separation is shown in fig. 1). The pppGp-containing sequence was now found in the pentanucleotide fraction (table 1, 2nd column). Under these conditions the separation seems to be based not only on the net charge but also on secondary binding forces, which depend on the number of bases and the purine-to-pyrimidine ratio. Thus pppUp moves at the tail of the dinucleotides while pppGp moves with the trinucleotides. Hence, the presence of the terminal oligonucleotide in the pentanucleotide region is compatible with a sequence containing three purine nucleotides more than pppUp.

Aliquots of the heptanucleotide fraction, prepared by column chromatography, were rechromatographed on DEAE-cellulose paper. As expected, a small fraction of the activity was separated from the main spot of heptanucleotides and travelled in the position of pentanucleotides (fig. 1). While the spot moving in the pentanucleotide position was hardly detectable on UV light, it appeared clearly on an autoradiogram. It contained 6.8% of the activity in the heptanucleotide spot. This result compares favorably with the 7.7% calculated assuming the heptanucleotide fraction to consist of 13 oligonucleotides [5] and assuming the formula pppGpPupPupPyp for the terminal sequence.

The spot corresponding to the putative terminal sequence was cut out and eluted. The material was hydrolyzed with alkali and the products separated by paper chromatography and electrophoresis in the presence of carrier pppGp, Gp, Up, Cp, and Ap. Fig. 2 shows the distribution of ^{32}P-activity on the chromatogram and on the electropherogram as determined by strip scanning. In another experiment the activity present in each carrier spot was determined by liquid scintillation counting. The results, listed in table 2, allow to conclude that the isolated oligonucleotide has the sequence pppGpGpGpUp.

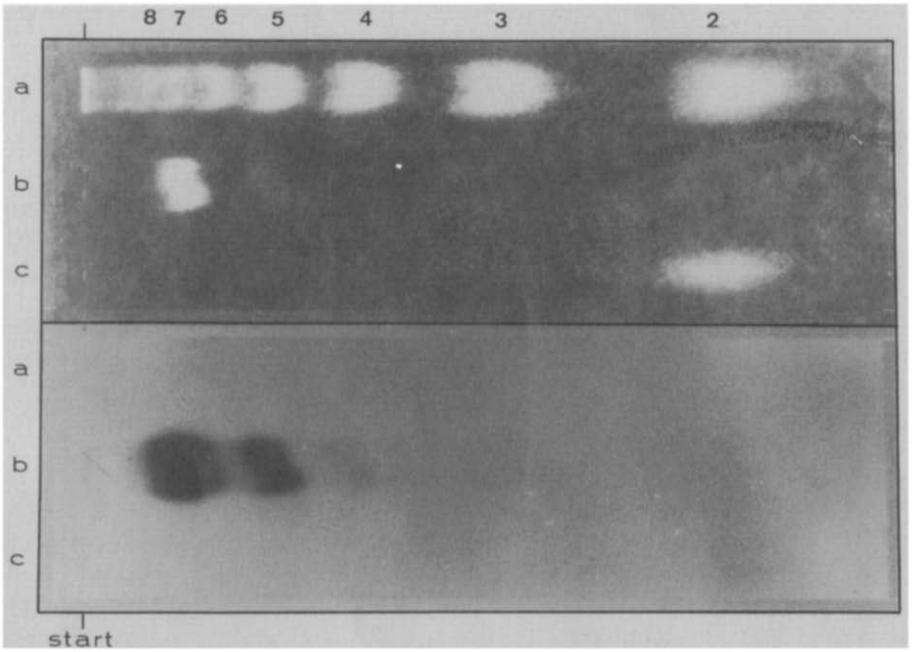

Fig. 1. *Separation on DEAE-cellulose paper of heptanucleotides obtained by column chromatography.* The upper half of the figure represents a UV contact print and the lower half an autoradiogram.
(a) Pancreatic RNase digest of yeast RNA (BDH; highly polymerized). The chain length of the components is indicated on top.
(b) Heptanucleotide fraction (0.029 mg; 3.14×10^5 dpm) obtained by column chromatography.
(c) Dinucleotides from a pancreatic RNase digest.

4. Discussion

It is evident that the viral RNA chain does not start with an initiation codon for polypeptide synthesis. This means that the translation of the polynucleotide message starts at some distance from the 5′-terminus and stops before the 3′-terminus is reached. Indeed no chain-terminating codon was found in the 3′-terminal sequence either [4]. The length of the terminal RNA segments not involved in protein synthesis is at present unknown.

It is interesting to note that the bases 1 to 4 from the 5′-terminus are complementary with the bases 2 to 5 from the 3′-terminus. This makes it likely that the RNA molecule is folded in such a way that both ends form a double stranded region.

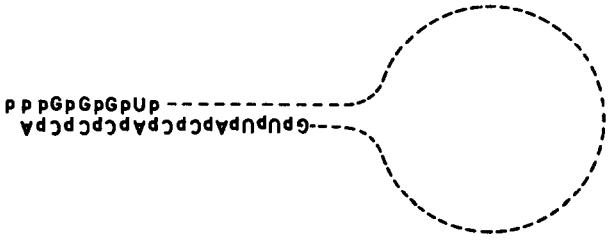

Table 2
Distribution of ^{32}P-activity in an alkaline hydrolysate of the isolated 5′-terminal sequence.

Spot	^{32}P-activity (dpm)	Ratio	Theoretical ratio (b)
pppGp	2182	3.19 (a)	4
Gp	1315	1.92	2
Up	685	1.00	1
Cp	27	0.04	0
Ap	67	0.09	0

(a) Some loss of pppGp seems to have occurred. The radioactivity moved definitely with carrier pppGp and not with the faster moving ppGp [3].
(b) Ratio expected for an alkaline hydrolysate of pppGpGpGpUp.

This terminal base complementarity may be important for the viral RNA replication. Indeed it would mean that the minus strand has the same terminal sequences as the plus strand (except perhaps

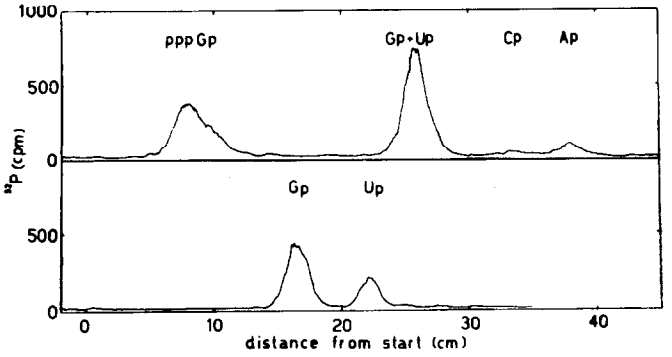

Fig. 2. *Separation of an alkaline hydrolysate of the isolated, 5′-terminal sequence.* An alkaline hydrolysate of the 5′-terminal sequence was separated by paper chromatography with an isobutyric acid solvent and scanned (upper curve). The spot containing Gp+Up was eluted, further resolved by electrophoresis, and scanned again (lower curve). The position of carrier substances detected in UV-light is indicated on each graph.

for the A residue). If then one or both of the extremities form a recognition site for a viral RNA polymerase, this site would be the same for the viral RNA and for its complement, so that both molecules may be recognized by a single enzyme.

Glitz [8] has recently claimed that the initial sequence of MS2 RNA is pppGpGpUp. However no evidence was given that this analysis pertained to a single, pure oligonucleotide.

Acknowledgement

This research was supported by grants from the U.S. National Institutes of Health (GM 11304-04) and from the Fonds voor Kollektief Fundamenteel Onderzoek.

References

[1] R.Roblin, J. Mol. Biol. 31 (1968) 51.
[2] M.Watanabe and J.T.August, Proc. Natl. Acad. Sci. U.S. 59 (1968) 513.
[3] R.De Wachter, J.–P.Verhassel and W.Fiers, Biochim. Biophys. Acta 157 (1968) 195.
[4] R.De Wachter and W.Fiers, J. Mol. Biol. 30 (1967) 507.
[5] W.Fiers, L.Lepoutre and L.Vandendriessche, J. Mol. Biol. 13 (1965) 432.
[6] R.De Wachter, J. Chromatog., in press.
[7] W.Min Jou, A.Raeymaekers and W.Fiers, Abstracts 5th Meeting of the Federation of European Biochemical Societies, Prague 1968, in press.
[8] D.G.Glitz, Biochemistry 7 (1968) 927.