STUDIES ON THE BACTERIOPHAGE MS2 NUCLEOTIDE SEQUENCE OF A 3′-TERMINAL FRAGMENT (n = 104)

R. CONTRERAS, A. VANDENBERGHE, W. MIN JOU,
R. DE WACHTER and W. FIERS
Laboratory of Molecular Biology and Laboratory of Physiological Chemistry
University of Ghent, Belgium

Received 29 July 1971

1. Introduction

Previous studies from this laboratory have established the 3 -terminal nucleotide sequence of Bacteriophage MS2 RNA up to position ω-16 [1, 2]. For the closely related bacteriophage R17, the terminal sequence up to position ω-51 has been reported [3]. Using another approach, viz. in vitro synthesis of a complementary strand, Weissmann and collaborators [4] were able to deduce the 3′-terminal sequence of the distantly related Bacteriophage Qβ up to position ω-52.

By hydrolysis of ^{32}P-labelled MS2 RNA with ribonuclease T_1 under milder conditions, we were able to obtain a set of larger fragments, all derived from the 3′-end. A partial sequence was recently presented [5], while we are now able to report the complete sequence of a 104 nucleotide long fragment.

2. Methods

^{32}P-labelled MS2 RNA was partially digested with ribonuclease T_1 at 0°, and the resulting products were separated on 12% or 6% polyacrylamide slab gels, as described previously [6]. Bands derived from the 3′-terminal end were further purified either by electrophoresis on acidic polyacrylamide gel in the presence of urea [7], or by two-dimensional electrophoresis on polyacrylamide gel, in which the first dimension is identical to the acidic system, while the second dimension is at neutral pH and with a higher polyacrylamide concentration (16% or 20%) (De Wachter and Fiers, in preparation).

Digests of the fragments with ribonuclease T_1 were separated either on DEAE-cellulose paper according to Sanger [8], or on polyethyleneimine plates based on the procedure described by Southern [9, 10]. Details on the latter method are described in the legend of fig. 1. Digests obtained by digestion with pancreatic ribonuclease A were separated by the electrophoresis-gradient chromatography bidimensional system [11].

3. Results

The 3′-terminal T_1 oligonucleotide is unique and results in a characteristic spot on two-dimensional fingerprints [1, 2].

Subsequently, we searched for longer fragments containing the aforementioned oligonucleotide, which was used as an identifying marker. This resulted in the identification of other unique oligonucleotides, which could in turn be used as markers in the screening for fragments derived from close to the 3′ end, but which missed the U–U–A–C–C–A–C–C–C–$A_{3'OH}$ sequence.

The fragments, so far obtained in pure form, are listed in table 1. The terminology used to designate the fragments has been described before [5]. Roman capital letters refer to bands from the 12% gel and Greek letters to bands from the 6% gel which

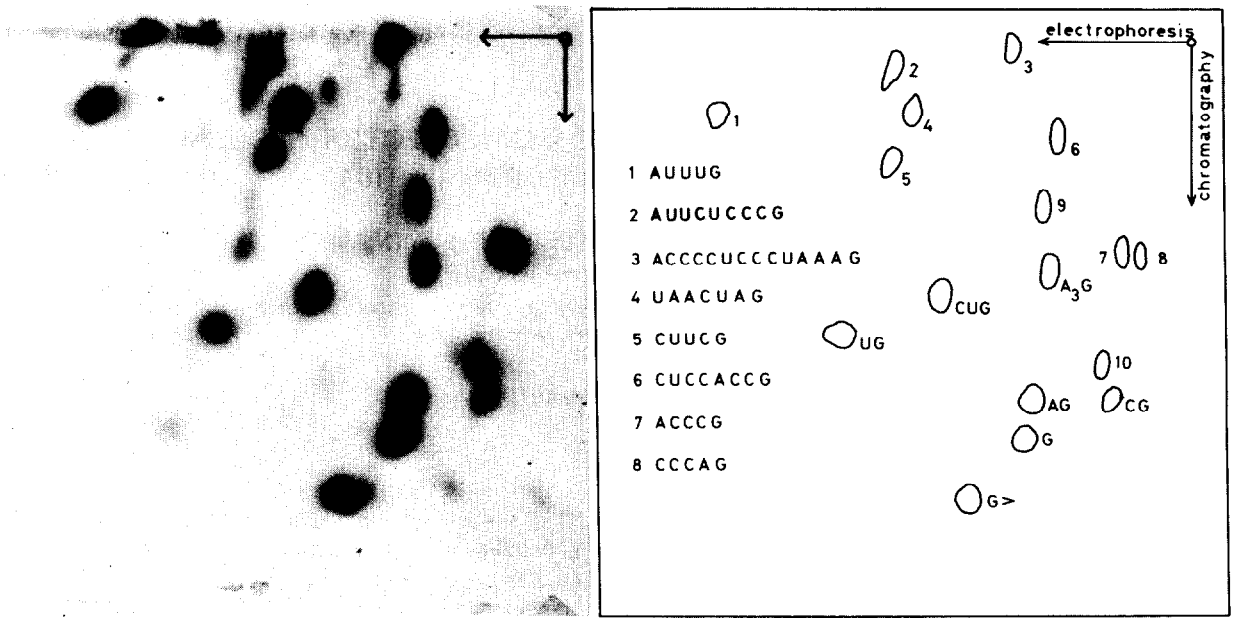

Fig. 1. T_1 – fingerprint of the 3′-terminal fragment on a thin layer polyethyleneimine plate. The fragment γ8 b6 was hydrolyzed with ribonuclease T_1 and the digest was separated by electrophoresis at pH 3.4 in the presence of urea [8]. The region of interest was then blotted over on a 20 cm × 20 cm PEI plate (Polygram Cel 300 PEI, Macherey and Nagel, Düren), which had been washed in 2 M formic acid–pyridine pH 2.2 and in water. The plate was then developed by ascending chromatography, using 2 M formic acid adjusted with pyridine to pH 3.4 as a solvent. Spots 9 and 10 are due to further degradation.

Table 1
Fragments derived from the 3′-terminal region.

Fragment	Splitting points (a)		Chain length *n*
D6zl; B12b21	0	1	15
D6z30 (b)	2	3	16
B12b20; γ4b17; γ8b21; δ1b13	0	2	19
Clz5; δ1b8	2	4	67
C3b3	3	6	68
γ8b7	2	5	76
B12b4α5; γ8b6	2	6	84
γ8b05	0	6	104

[a] Splits by ribonuclease T_1 are identified in fig. 2.
[b] Identical to spot no. 3 in fig. 1 of [2].

was used for analysis of milder digests. Most bands were purified by two-dimensional electrophoresis (indicated as b), while a few were obtained by acid electrophoresis alone (indicated as z). Thus a fragment, indicated as δ1b8 means spot no. 8 in the two-dimensional gel analysis of band no. 1 in the δ region of the 6% primary gel.

All fragments were analyzed by ribonuclease T_1 and by pancreatic ribonuclease fingerprinting. The sequence determination of the individual oligonucleotides will be reported elsewhere. It may be noted that two polypurine sequences, viz. G–A–A–A–G–G–U and A–A–A–G–A–G–A–G–G–A–C were characterized before [12, 13]. Furthermore, the latter oligonucleotide was known to be present in the 3′-two third part of the viral RNA chain [14]. Although the series of different fragments helped considerably in the reconstruction of the entire sequence, presented in fig. 2, more essential information was obtained by further partial digestion of some of the fragments (to be published).

The nature of the fragments, here described, are mainly determined by the secondary structure in this part of the molecule. Two hairpins, which are

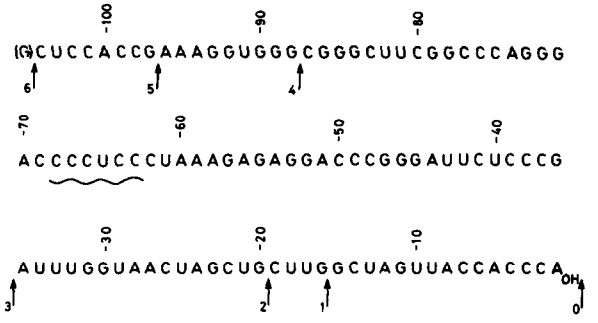

Fig. 2. Nucleotide sequence of the 3′-terminal fragment. The arrows indicate the ends of the fragments, listed in table 1. The sequence above the wavy line is still tentative.

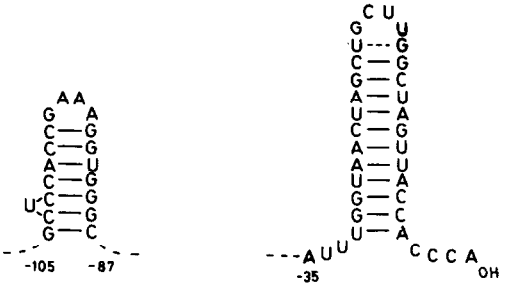

Fig. 3. Hairpin-loops present in the 3′-terminal fragment. These two hairpins have a tight secondary structure, as evidenced from partial enzymatic digestion. The bond 2, however, is very susceptible to enzymatic attack.

particularly strongly H-bonded and present in most digests, are shown in fig. 3. The bond at position 2 was usually broken, although the fragment 0-1 or 0-2 adhered to the rest of the molecule, until it was released in the pH 3.5-urea system.

4. Discussion

The sequence, presented in fig. 2, contains four potential nonsense triplets, two UAA's and two UAG's. Three of these occur in the last hairpin. The RNA-polymerase is the last cistron on the MS2 genome, and one would normally expect a termination signal for the latter polypeptide somewhere near the 3′-end. If neither of the above UAA triplets (the UAG's follow in phase) is functional as a stop, a very plausible hypothesis, it would mean that also at the 3′-end of the viral RNA chain an untranslated segment of over 100 nucleotides in length is present. This region may have a special function in replication.

The known 3′-terminal sequence of R17 RNA, 51 nucleotides long, is identical to the MS2 RNA counterpart [3]. This illustrates again the evolutionary stability of these untranslated regions, as noted before for the 5′-terminal fragments [6, 11].

The hairpins, shown in fig. 3, are only a part of a more intricate secondary and tertiary structure. A model for the latter, however, should preferably be based on further primary sequence results.

Acknowledgements

We thank Mrs. M. Borremans, Mr. R. De Baere, Mr. F. Deurinck and Mr. A. Raeymaekers, who participated in various aspects of this work. R.C. and W.M. hold fellowships from the Nationaal Fonds voor Wetenschappelijk Onderzoek. Financial support was provided by the Fonds voor Kollektief Fundamenteel Onderzoek and by NATO.

References

[1] R. De Wachter and W. Fiers, J. Mol. Biol. 30 (1967) 507.
[2] W. Min Jou, R. Contreras and W. Fiers, FEBS Letters 13 (1971) 105.
[3] S. Cory, P.F. Spahr and J.M. Adams, Cold Spring Harbor Symp. Quant. Biol. 35 (1970) 1.
[4] H.M. Goodman, M.A. Billeter, J. Hindley and C. Weissmann, Proc. Natl. Acad. Sci. U.S. 67 (1970) 921.
[5] W. Fiers, R. Contreras, R. De Wachter, G. Haegeman, J. Merregaert, W. Min Jou and A. Vandenberghe, Biochimie, in press.
[6] R. De Wachter, J. Merregaert, A. Vandenberghe, R. Contreras and W. Fiers, European J. Biochem., in press.
[7] R. De Wachter and W. Fiers, in: Methods in Enzymology, Vol. 21, Part D (1971) p. 167.
[8] F. Sanger and G.G. Brownlee, in: Methods in Enzymology, Vol. 12, Part A (1967) 361.
[9] K. Randerath and E. Randerath, in: Methods in Enzymology, Vol. 12, Part A (1967) 323.
[10] E.M. Southern and A.R. Mitchell, Biochem. J., in press.

[11] R. De Wachter, A. Vandenberghe, J. Merregaert, R. Contreras and W. Fiers, Proc. Natl. Acad. Sci. U.S. 68 (1971) 585.
[12] G. Haegeman, W. Min Jou and W. Fiers, J. Mol. Biol. 57 (1971) 597.
[13] W. Min Jou and W. Fiers, J. Mol. Biol. 40 (1969) 187.
[14] W. Min Jou, W. Fiers, H. Goodman and P. Spahr, J. Mol. Biol. 42 (1969) 143.