SEQUENCE HOMOLOGY AT THE 3'-ENDS OF ALFALFA MOSAIC VIRUS RNAs

L. PINCK and M. PINCK
Institut de Biologie Moléculaire et Cellulaire de CNRS, 15 rue Descartes, 67084 Strasbourg, France

Received 6 August 1979
Revised version received 7 September 1979

1. Introduction

We report that the four RNA molecules of alfalfa mosaic virus (AMV) have an extensive sequence homology at their 3'-ends. Our sequence data were obtained by the new direct chemical sequencing method [1].

AMV is a multipartite-genome plant virus. The virus particles contain four RNA molecules, conventionally numbered 1–4 in order of decreasing size. The first three constitute the true genome and contain copies of all the genes of the virus; for their replication either the coat protein or its messenger (RNA 4 in the virions) must be present [2]. Similarly bromegrass mosaic virus contains four RNA molecules but here genome expression is not known to be dependent on coat protein or on its messenger RNA. These four RNAs were shown to contain a common tRNA-like 3'-end sequence that is 161 nucleotides long [3]. In AMV RNA 4, the sequence of 91 nucleotides [4], corresponding to a 3'-end fragment that interacts strongly with the viral coat protein, has recently been reported. This site may be required for proper recognition of the RNA 4 3'-terminus by the viral replicase [5]. Our observations extend this possibility to the genomic AMV RNAs.

2. Materials and methods

2.1. *Purification of virus and RNA*

AMV (Strasbourg strain) was grown and extracted as in [6]. The phenol-extracted viral RNAs were separated on polyacrylamide–agarose (2.4–0.5%) gels and recovered from the gels as in [7].

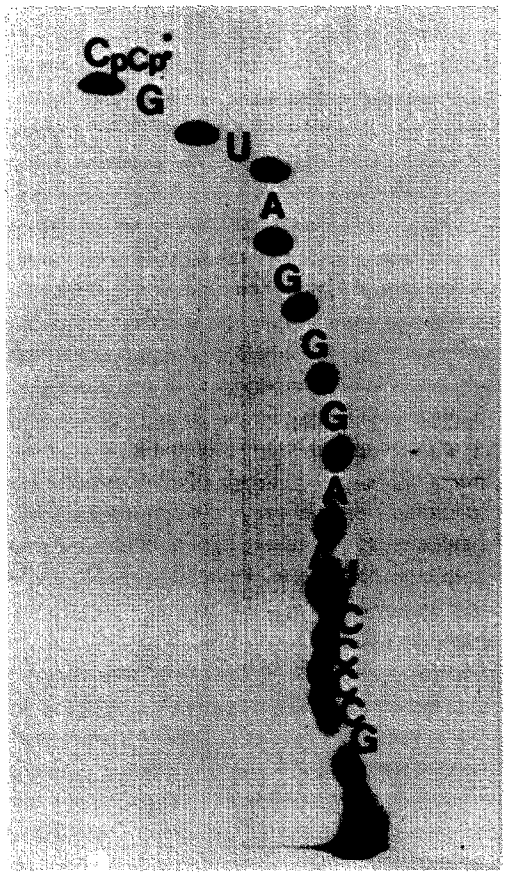
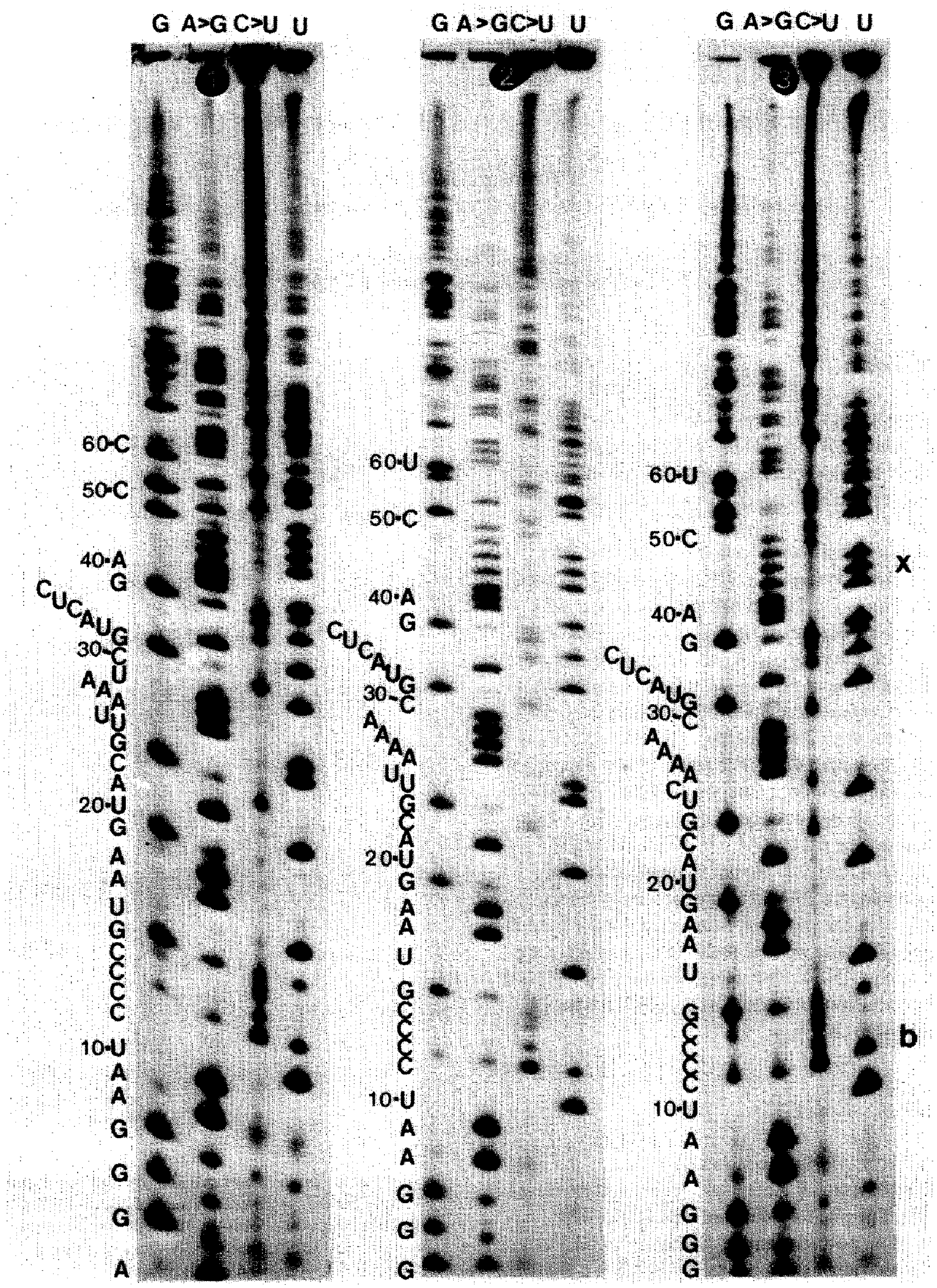


Fig.1. Autoradiogram of a two-dimensional electrophoretic and homochromatographic sequence analysis of a [5'-^{32}P]pCp-labeled 3'-end fragment of RNA 3. The labeled fragment, mixed with 2 μl tRNA, was hydrolysed in 10 μl water for 30 min at 100°C in a sealed capillary tube.


G A>G C>U U
G A>GC>U U
G A>GC>U U
1
2
3
60-C
50-C
40-A
30-C
20-U
10-U
60-U
50-C
40-A
30-C
20-U
10-U
60-U
50-C
40-A
30-C
20-U
10-U
X
b

2.2. *3′-end labeling with T_4 RNA ligase*

The purified AMV RNAs were labeled at their 3′-termini, using [5′-^{32}P]pCp (1000–3000 Ci/mmol, New England Nuclear) as a donor, with T_4 RNA ligase (Biolabs) as in [8].

2.3. *RNA sequencing method*

The direct RNA sequencing technique, by means of chemical base-specific reactions, was used on the 3′-end-labeled RNAs exactly as in [2]. Additional methods of analysis by electrophoresis on cellulose acetate and homochromatography were used to determine the nucleotide sequence adjacent to the 3′-end as in [9].

3. Results

Our first attempts to characterize the 3′-terminal nucleotides of the four AMV RNAs, by means of periodate oxidation, sodium boro[^{3}H]hydride reduction [10] and subsequent digestion with nucleases, gave the first indication of a common 3′-end GC_{OH}. This labeling technique is, however, difficult to use for extensive sequencing.

The 3′-end [5′-^{32}P]pCp labeled fragments resulting from a mild T_1 RNase digestion of the purified RNAs (1 unit T_1/500 μg RNA in 10 mM Tris–acetate (pH 7.5), 1 mM Mg-acetate for 15 min at 0°C) were isolated from a polyacrylamide gel. Two-dimensional electrophoretic and homochromatographic analyses of these 3′-end fragments yielded the sequence of the final 15 nucleotides, identical for the four RNAs. Figure 1 shows the data obtained from the T_1 RNase fragment of RNA 3.

The direct chemical sequencing technique was applied on purified 3′-end-labeled RNA to determine a sequence of ~100 nucleotides. For each base-specific reaction, 0.6 μg labeled RNA and 10 μg carrier tRNA were used. Figure 2 shows the results obtained with RNAs 1, 2 and 3. The pattern obtained with RNA 4, not shown here, was identical to that

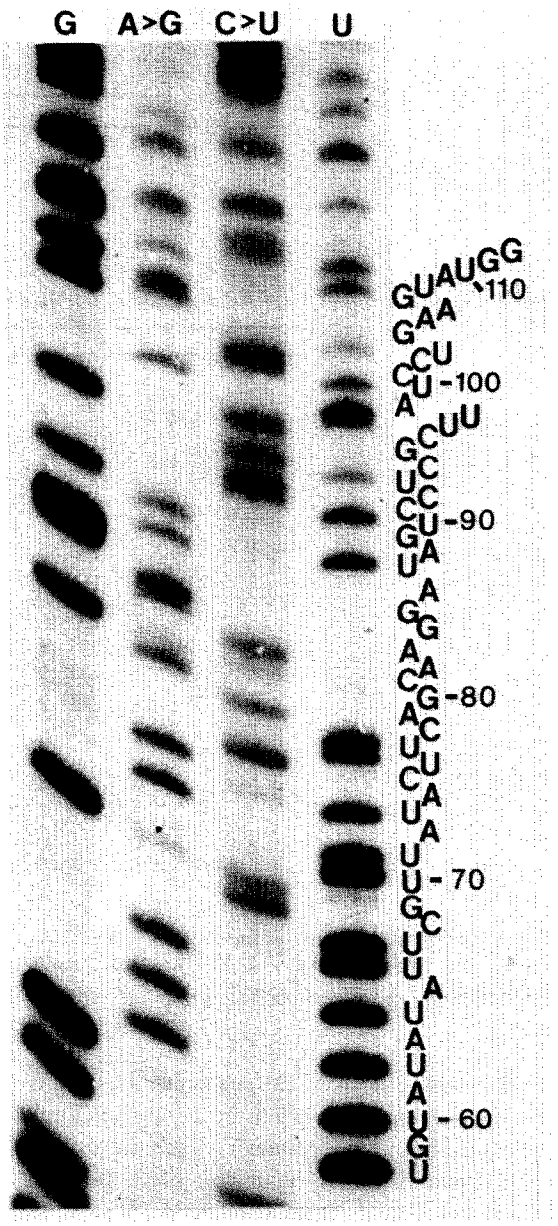


Fig.3. Autoradiogram of 3′-end labeled AMV RNA 3 analysed as indicated in legend of fig.2. The gel was run until the xylene cyanol marker had reached the bottom of the gel.

Fig.2. Autoradiogram of 3′-end labeled AMV RNA 1, 2 and 3, sequenced chemically on a thin (0.5 m/m) polyacrylamide gel (12% acrylamide, 0.6% bisacrylamide). The gel was run at 1.6 kV until bromophenol blue (b) and xylène cyanol marker (X) had reached the points indicated.

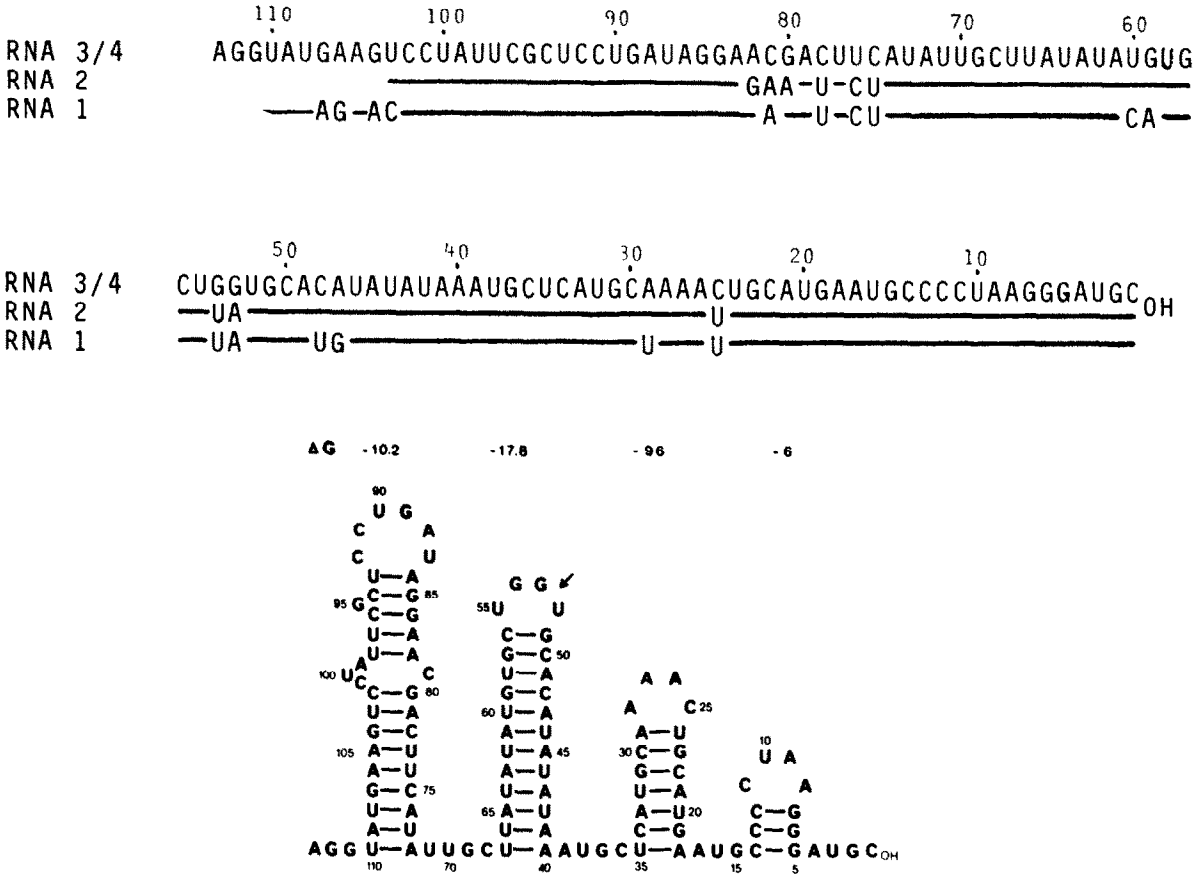


Fig.4. Nucleotide sequences of the 3′-end of AMV RNAs 1, 2 and 3 or 4 and possible secondary structure of RNA 3 with the values of ΔG calculated according to [12]. For RNAs 1 and 2, only nucleotides that differ from those in RNA 3 or 4 are indicated; lines indicate homologous sequences. The arrow indicates the main T_1 RNase cleavage point.

obtained with RNA 3. The sequences deduced from the sequencing gel shown in fig.2 and from duplicate gels run from a longer time, in which the xylene cyanol marker was at the bottom of the gel (fig.3 shows the sequencing gel from RNA 3), are given in fig.4. In this figure the common sequence of 113 nucleotides near the 3′-ends of RNAs 3 or 4 is indicated together with the nucleotide modifications that occur in RNAs 1 and 2. A possible secondary structure for the 3′-ends of RNAs 3 or 4 is also given, for which ΔG values were calculated according to [11]. The nucleotide difference at position 25, in the loop region, would not affect the stability of the second hairpin, but the difference in position 29 of RNA 1 would decrease its stability ($\Delta G = -7.6$). The modifications in positions 47, 48, 59 and 60 of RNA 1 leave the total number of GC and AU pairs constant and thus do not affect the stability of the third potential hairpin (ΔG unchanged).

The change of GG (53–54) to AU in RNAs 1 and 2 resulted in a significant difference in the T_1 RNase cleavage pattern, since G_{53} in RNAs 3 and 4 is the main T_1 cleavage point over a wide range of T_1 RNase concentrations.

Several nucleotide differences exist in the region 75–82 of RNAs 1 and 2, which greatly modify the structure of the fourth possible hairpin in RNA 1 and 2.

The nucleotide differences found with the chemical-sequencing method have been confirmed using the

method in [12] on fragments labeled at either 3′- or 5′-ends. With the enzymatic sequencing method, we observed band compression resulting from the persistence of secondary structure, especially in regions on the 5′-side of hairpin stems (12–14), (32–35), (57–60), (96–98) and also loop region (86–89). This phenomenon was much less pronounced with the chemical method. At the position of C_{11} in the nucleotide pattern (fig.2) there are spots in all four slots, a phenomenon peculiar to this sequence of four Cs (11–14) of which three are involved in GC pairs (fig.4) while the fourth (C_{11}) is presumably more chemically reactive. The sequence of four Cs is confirmed by the homochromatogram shown in fig.1.

4. Discussion

The results presented here indicate a near-perfect sequence homology of 74 nucleotides at the 3′-ends of the four AMV RNAs, where the few minor nucleotide differences probably would not significantly affect the high secondary structure of this region. In the regions that we have sequenced beyond nucleotide 75 (75–103 of RNA 2 and 75–111 of RNA 1), there are major similarities with the corresponding regions of RNAs 3 or 4. In addition, sequence data not reported here indicate that the non-coding region at the 3′-end of RNA 3 is nearly 200 nucleotides long, a length comparable to the 3′-end non-coding region of tobacco mosaic virus RNA [13].

There is an extensive homology at the 3′-end between the RNAs 3 or 4 of our strain and the RNA 4 of AMV strain 425 [4] in the region running from nucleotides 5–81. Differences appear only in the first hairpin and in the loop of the third hairpin.

Studies [14] on sequence homologies by means of DNA–RNA hybridization failed to reveal the homologies reported here. The method used in [14] could detect homologies $\gtrsim$50 nucleotides long. Since they found no homology between the four RNAs, one may deduce that a few nucleotide changes (1 in RNA 2, 4 in RNA 1) in the 50 nucleotides at the 3′-end suffice to prevent the formation of DNA–RNA hybrids.

The absence of a $-CCA_{OH}$ 3′-end may explain the failure of repeated attempts to aminoacylate these RNAs. We tried unsuccessfully to reconstitute the $-CCA_{OH}$ terminus using pure tRNA-nucleotidyltransferase from yeast (a kind gift of Dr Giegé) and labeled ATP and CTP, in the conditions of [15].

The homology between the four AMV RNAs at their 3′-ends is very interesting since there is a high affinity site for the coat protein at the 3′-end of the RNA 4 molecule. This site may be required for the viral replicase to recognise the RNA correctly [5]. Since high affinity sites for coat protein exist in all AMV RNAs [16] it is possible that such a binding is involved in initiating, or in inducing conformational changes necessary for the replication of the RNAs.

Acknowledgements

We thank Professor L. Hirth for his interest in this work and Dr A. Durham for improving the manuscript.

References

[1] Peattie, D. A. (1979) Proc. Natl. Acad. Sci. USA 76, 1760–1764.
[2] Bol, J. F., Van Vloten-Doting, L. and Jaspars, E. M. J. (1971) Virology 46, 78–85.
[3] Dasgupta, R. and Kaesberg, P. (1977) Proc. Natl. Acad. Sci. USA 74, 4900–4904.
[4] Koper-Zwarthoff, E. C. and Bol, J. F. (1979) Proc. Natl. Acad. Sci. USA 76, 1114–1117.
[5] Houwing, J. and Jaspars, E. M. J. (1978) Biochemistry 17, 2927–2933.
[6] Pinck, L. and Hirth, L. (1972) Virology 49, 413–425.
[7] Pinck, L. and Fauquet, C. (1975) Eur. J. Biochem. 57, 441–451.
[8] England, T. E. and Uhlenbeck, O. C. (1978) Nature 275, 560–561.
[9] Silberklang, M., Gillum, A. M. and RajBhandary, U. L. (1977) Nucleic Acids Res. 4, 4091–4108.
[10] Lockard, R. E. and RajBhandary, U. L. (1976) Cell 9, 747–760.
[11] Tinoco, I., Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M. and Gralla, J. (1973) Nature New Biol. 246, 40–41.
[12] Donis-Keller, H., Maxam, A. and Gilbert, W. (1977) Nucleic Acids Res. 4, 2527–2538.
[13] Guilley, H., Jonard, G., Kukla, B. and Richards, K. E. (1979) Nucleic Acids Res. 6, 1287–1308.
[14] Gould, A. R. and Symons, R. H. (1978) Eur. J. Biochem. 91, 269–278.
[15] Giegé, R., Briand, J. P., Mengual, R., Ebel, J. P. and Hirth, L. (1978) Eur. J. Biochem. 84, 251–256.
[16] Srinivasan, S. and Jaspars, E. M. J. (1978) Biochim. Biophys. Acta 520, 237–241.