The EMBO Journal, Vol. 1 No. 2, pp.269–276, 1982

The tRNA-like structure of turnip yellow mosaic virus RNA: structural organization of the last 159 nucleotides from the 3′ OH terminus

C. Florentz, J.P. Briand, P. Romby, L. Hirth, J.P. Ebel, and R. Giegé*

Laboratoires de Biochimie et de Virologie, Institut de Biologie Moléculaire et Cellulaire du CNRS, 15, rue René Descartes, 67084 Strasbourg cedex, France

Communicated by J.P. Ebel
Received on 16 February 1982

The secondary structure of the isolated tRNA-like sequence (n = 159) present at the 3′ OH terminus of turnip yellow mosaic virus RNA has been established from partial nuclease digestion with S1 nuclease and T1, CL_3, and *Naja oxiana* RNases. The fragment folds into a 6-armed structure with two main domains. The first domain, of loose structure and nearest the 5′ OH terminus, is composed of one large arm which extends into the coat protein cistron. The second, more compact domain, is composed of the five other arms and most probably contains the structure recognized by valyl-tRNA synthetase. In this domain three successive arms strikingly resemble the Tψ, anticodon, and D arms found in tRNA. Near the amino-acid accepting terminus, however, there is a new stem and loop region not found in standard tRNA. This secondary structure is compatible with a L-shaped three-dimensional organization in which the corner of the L and the anticodon-containing limb are similar to, and the amino-acid accepting region different from, that in tRNA. Ethylnitrosourea accessibility studies have shown similar tertiary structure features in the Tψ loop of $tRNA^{Val}$ and in the homologous region of the viral RNA.

Key words: ethylnitrosourea/nuclease digestion/secondary structure/ tRNA-like/TYMV RNA

Introduction

That similar structures must fulfil similar functions is a key concept in molecular biology. This concept, however, seems to be invalidated with the tRNA-like structures present at the 3′ OH end of the RNA of many plant viruses which do not share apparent sequence analogies with standard tRNAs (reviewed by Haenni and Chapeville, 1979; Hall, 1979; Briand *et al.*, 1977; Silberklang *et al.*, 1977a; Guilley *et al.*, 1979; Ahlquist *et al.*, 1981). This is, in particular, the case for the 3′-terminal region of turnip yellow mosaic virus (TYMV) RNA (Briand *et al.*, 1977; Silberklang *et al.*, 1977a). This region of the viral RNA does not have the characteristic primary structural features of tRNAs, such as the presence of several strategic sequences and the modified bases. Moreover, it cannot be folded into a classical cloverleaf, despite the fact that its valylation, catalyzed by valyl-tRNA synthetase, has kinetic parameters similar to those of the $tRNA^{Val}$ (Giegé *et al.*, 1978; Florentz *et al.*, 1982). How do two structures as different as $tRNA^{Val}$ and the tRNA-like region of TYMV RNA behave in such a similar fashion in the presence of valyl-tRNA synthetase? This question might be answered, as proposed earlier (Giegé *et al.*, 1978), if one supposes that similar structural domains which are recognized by the synthetase exist in both $tRNA^{Val}$ and viral RNAs. We believe that such domains should exist at the level of the three-dimensional structure, but some similarities between the two types of molecules could already appear at the level of the secondary structures. The existing models of TYMV RNA secondary structure, however, do not reveal such similarities (Figure 1). These tentative models, although cloverleaf-like, contain ~105 nucleotides, a number markedly larger than the 77 nucleotides found in $tRNA^{Val}$. As a result, the size of the stem and loop regions is much larger than in tRNA. In fact the validity of these models of secondary structure can be questioned as they are based solely on the positioning of the putative CAC anticodon in a loop, without other experimental verification.

We report here the real secondary structure of the final 159 nucleotides present at the 3′ terminus of TYMV RNA; this structure might also be extended to the tRNA-like region of tobacco mosaic virus.

During the course of this work a chemical modification study, showing that the 3′ OH terminus of TYMV RNA adopts a highly compact structure, was published as a poster by Rietveld *et al.*, (1981) at the 5th International Congress of Virology held in Strasbourg in August 1981.

Results

Partial digestions with nucleases were performed on the pure ^{32}P end-labelled tRNA-like fragment comprising the last 159 3′-terminal nucleotides of TYMV RNA. To test all the nucleotides of the fragment and to discriminate between primary and secondary cleavage (Favorova *et al.*, 1981) both 5′- and 3′-labelled molecules were investigated. Prior to

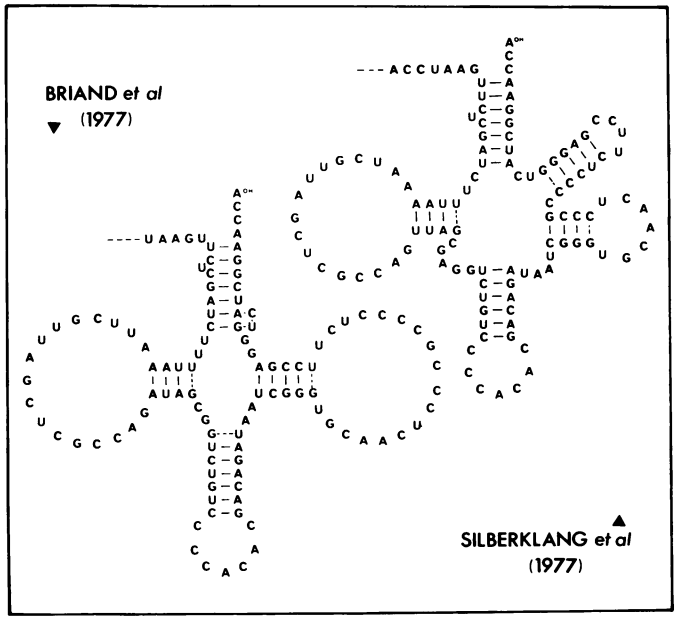
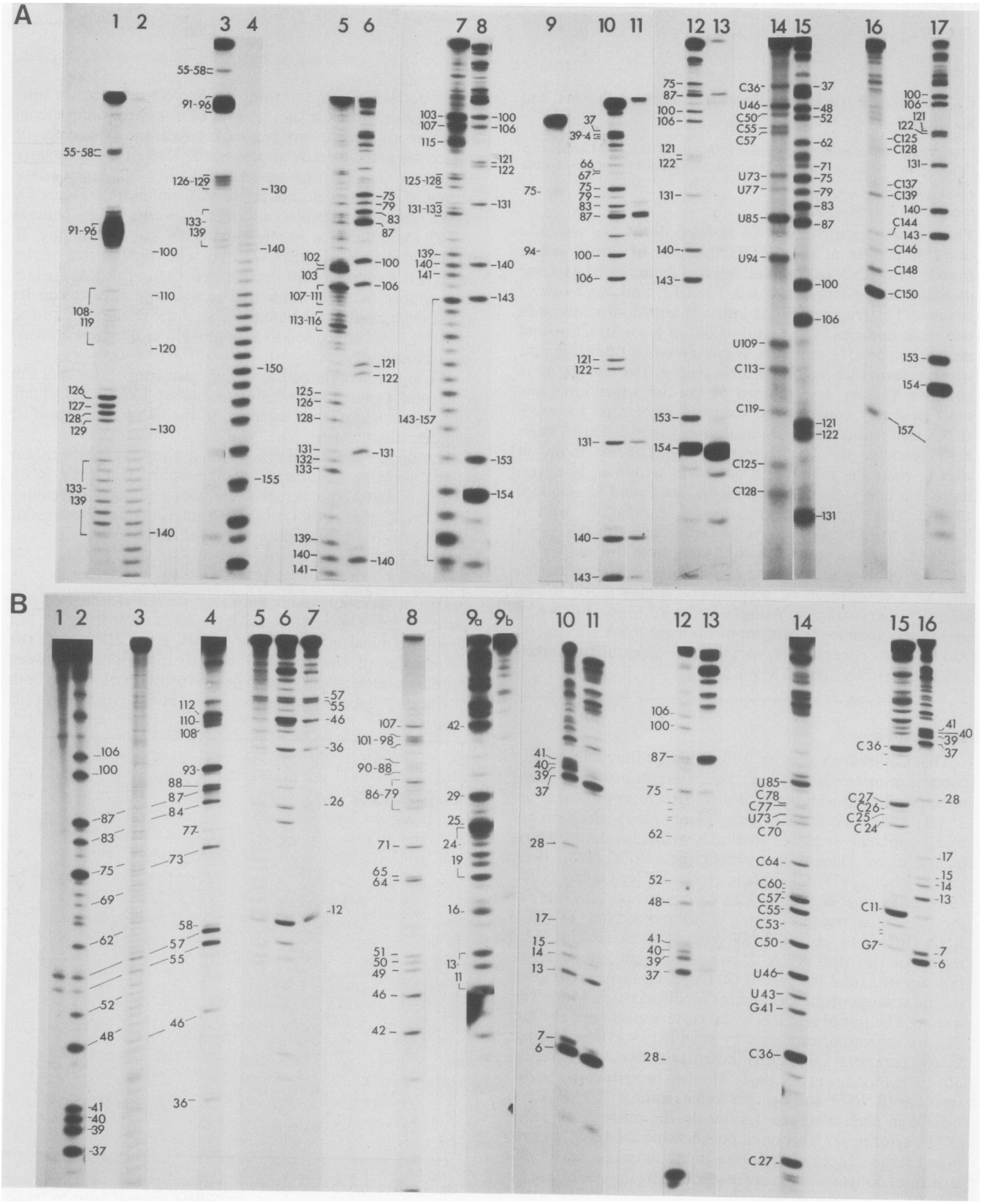


Fig. 1. Two hypothetical secondary structures of the tRNA-like fragment present at the 3′ OH terminus of TYMV RNA proposed earlier by Briand *et al.* (1977) and Silberklang *et al.* (1977a).

*To whom reprint requests should be sent.

 0261-4189/82/0102-0269$2.00/0


A
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
B
1 2 3 4 5 6 7 8 9a 9b 10 11 12 13 14 15 16

digestion the RNA samples were renatured for 5 min at 37°C in the presence of 10 mM $MgCl_2$, conditions which preserve, in full, the aminoacylation capacity of the molecules. Digests were analyzed on 10, 15, and 20% polyacrylamide gels and cuts were assigned by comparison with alkaline and RNase T1 ladders. Typical results, shown in Figure 2, are summarized in Figure 3, where only the strongest cuts are indicated.

The fragment is particularly sensitive to spontaneous degradation. As observed in numerous experiments, such degradation always occurs at the same positions in the nucleotide sequence, notably at positions 55, 57, 75, 85, and 94 (see Figure 3). If particular care is not taken during handling of the pure tRNA-like fragment, the extent of these spontaneous cuts can be important. Spontaneous degradation can, however, be greatly reduced, especially if the elution from the purification gels of the end-labelled RNA is carried out in the absence of phenol (see Materials and methods). The enzymatic accessibility studies described below were performed with molecules in which these degradations were minimal (see Figure 2A, lane 9 and Figure 2B, lanes 1 and 9b).

Accessibility to nuclease S1

Partial digestion of the 5' end-labelled fragment gives rise to two major series of cuts in the region 91–96 and 126–129; three other series of less frequent cuts arise at positions 55–58, 108–119, and 133–139 (Figure 2A, lanes 1 and 3). On the same gel lanes it is also apparent that five regions near the 3' end are accessible to nuclease S1 digestion. A partial S1 hydrolysis of the 3'-labelled tRNA-like fragment allows us to assign the exact position of these five regions. Figure 2B (lanes 4 and 7) shows that the cuts occur at position 12, 26, 36, 46, and 73, cut 26 being weaker than the others.

Accessibility to the RNase from cobra venom

Cobra venom RNase hydrolysis patterns for the 5' and 3' end-labelled fragments (Figure 2A, lanes 5 and 7, and Figure 2B, lanes 8 and 9) allow us to locate eight main regions of cleavage at positions 49–51, 64–65, 79–90, 98–101, 107–116, 125–128, 131–133, and 139–157, plus five single strong cuts at positions 25, 29, 42, 46, and 71. Most cuts, located in the 60–100 region, were found with both 5'-and 3'-labelled molecules.

Accessibility to RNase T1

Figure 2A (lanes 11 and 13) and Figure 2B (lanes 11 and 13) display a partial RNase T1 hydrolysis of the native tRNA-like fragment labelled at its 5' and 3' end, respectively. Among the 31 G residues present in the fragment, G154 is by far the most accessible. Frequent cleavages also occur at positions 6, 13, 37, 48, and 87. Seven G residues are clearly fully protected whereas 16 others react to some extent with RNase T1. In the middle of the sequence G75 and G83 are either weakly or not accessible in both 3'- or 5'-labelled RNA. This is indicative of secondary cuts.

Accessibility to RNase CL_3

Results of the partial hydrolysis of the tRNA-like fragment by RNase CL_3 are shown in Figure 2A (lanes 14 and 16) and Figure 2B (lanes 14 and 15) for the 5'- and 3'-labelled molecules, respectively. Many C residues present in the sequence are accessible to this enzyme; the strongest cuts are at positions 11, 27, 36, and 150. Interestingly, three strong cuts at positions 46, 85, and 94 and four weaker ones at positions 9, 73, 109, and 157 occur after U residues. Another cut also occurs after G41 (Figure 2B, lane 14). It is recalled that RNase CL_3 is not completely specific for C residues (Levy and Karpetzky, 1980). Cuts at positions 36, 46, 50, 55, 57, 73, and 85 were found with both 5'- and 3'-labelled molecules, the last two being U positions. Here also some cuts only appear with one type of labelling (i.e., positions 41, 43, 64, and 70) suggesting the appearance of secondary cuts.

Alkylation of phosphates by ethylnitrosourea

Only part of the tRNA-like fragment has been investigated. Figure 4 shows an autoradiogram of a gel where the region between position 20 and 45 is probed. It can be seen that alkylation of phosphates 33 and 34 is hindered when the reaction is conducted at 20°C on native RNA, whereas at 80°C, under denaturing conditions, all phosphates are alkylated by ethylnitrosourea.

Discussion

Limited enzymatic digestions were used to probe the structure of the last 159 nucleotides at the 3' OH terminus of TYMV RNA. This approach has already been used to establish the secondary structures of several small RNAs (e.g., Branlant *et al.*, 1981; Toots *et al.*, 1981) and to check structural features in tRNAs and rRNA (e.g., Wrede *et al.*, 1979a, 1979b; Favorova *et al.*, 1981; Lockard and Kumar, 1981; Vassilenko *et al.*, 1981; Butorin *et al.*, 1982). The method depends on the differential accessibility of some residues or domains to nucleases. Primary cuts in the RNA obviously reflect the accessibility of structural features, nucleotide or domain, in the native molecule and can be easily interpreted in terms of secondary structure. However, the discrimination between the primary and the secondary cuts is not always clear (for a complete discussion, see Favorova *et al.*, 1981). To avoid this difficulty our digestions were conducted under very limiting conditions on both 5'- and 3'-labelled molecules. Technical difficulties restricted the search for secondary cuts to the central part of the 159 nucleotides of the tRNA-like fragment. In this region (results summarized in Figure 3), most cuts were primary. We are

Fig. 2. Autoradiograms of acrylamide gels of partial enzymatic digests of the purified 5'- or 3'-end labelled tRNA-like fragment of TYMV RNA. The figure is a composite of numerous experiments. For each type of hydrolysis and of gel, RNase T1 and alkaline ladders as well as controls without enzyme were run. Experiments were taken as significant only when the controls displayed very weak extra bands (see **lane 9** in **(A)**, **lanes 1** and **9b** in **(B)**). For the sake of clarity all controls and ladders are not shown in the figure; for the same reason only one or two types of gels are shown among the 10, 15, and 20% gels investigated. Numbering of residues is from the 3' towards the 5' end of the fragment. **(A)** 5'-end labelled fragment. Nuclease S1 digests: **lane 1** (15% gel), **lane 3** (20% gel). Cobra venom RNase digests: **lane 5** (15% gel), **lane 7** (20% gel). Partial RNase T1 digests: **lane 11** (15% gel), **lane 13** (20% gel). Partial RNase CL_3 digests: **lane 14** (10% gel), **lane 16** (20% gel). **Lanes 2** and **4** are alkaline ladders; **lanes 6, 8, 10, 12, 15,** and **17** are RNase T1 ladders, and **lane 9** represents a control migration on a 10% gel of the pure labelled fragment. **(B)** 3'-end labelled fragment. Nuclease S1 digests: **lane 4** (10% gel), **lane 7** (20% gel). Cobra venom RNase digests: **lane 8** (10% gel), **lane 9a** (20% gel), and the corresponding control **(lane 9b)**. Partial RNase T1 digest: **lane 11** (20% gel), **lane 13** (10% gel). Partial RNase CL_3 digests: **lane 14** (10% gel), **lane 15** (20% gel). **Lanes 3** and **5** are alkaline ladders; **lane 6** is a RNase P1 hydrolysis; **lanes 2, 10, 12** and **16** are RNase T1 ladders and **lanes 1** and **9b** are controls. When starting with the same number of counts deposited on the gels, autoradiography was about twice as long for hydrolysis patterns performed with the 3'-end labelled RNA than for those performed with the 5'-labelled fragment. For **lanes 9a** and **9b** autoradiography was even twice as long as for other experiments in the **(B)** series.

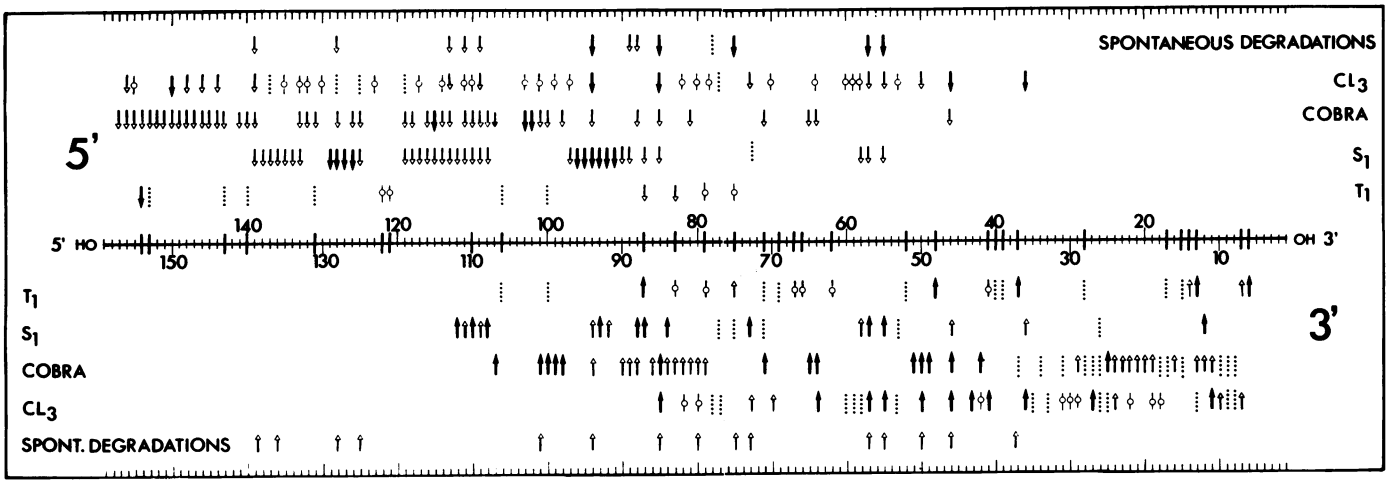


Fig. 3. Schematic representation of the accessibility studies against nucleases of the tRNA-like structure of TYMV RNA. The RNA sequence is symbolized by a line on which each nucleotide is represented by vertical light lines (A, U, or C residues) or heavy lines (G residues). The figure represents a summary of the results given in Figure 2 and of further experiments not shown; it also displays the location of the spontaneous degradations occurring in the RNA. ↑ strong cuts; ↟ less strong cuts; ⋮ weak cuts. Strength of cuts is estimated from the autoradiographs displayed in Figure 2; it is recalled, however, that exposition of the films was longer with 3'- than with 5'-end labelled RNA. For RNases T1 and CL_3 the symbol ⚲ represents the inaccessibility of the G and C residues against these nucleases.

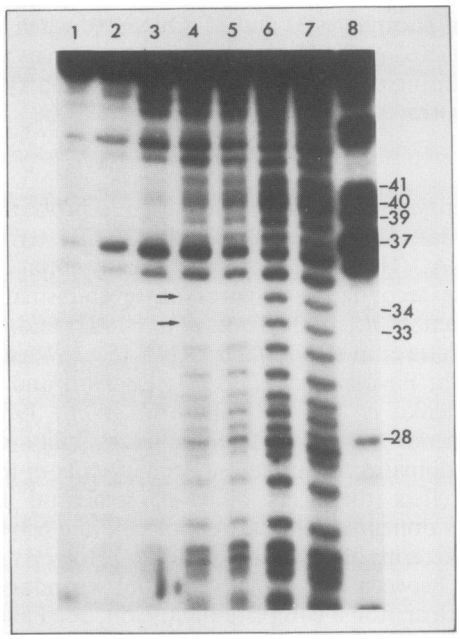


Fig. 4. Autoradiogram of a 15% acrylamide gel of a phosphate alkylation experiment with ethylnitrosourea of 3'-labelled tRNA-like fragment of TYMV RNA. **(1)** Control without treatment; **(2,3)** control incubations under conditions stabilizing **(2)**, or unfolding **(3)**, the structure of RNA; **(4,5)** alkylations (in duplicate) of native RNA; **(6,7)** alkylations (in duplicate) of unfolded RNA; **(8)** partial RNase T1 digest under denaturing conditions. The two arrows show the phosphates 33 and 34 inaccessible to alkylation by ethylnitrosourea.

confident, therefore, that the cuts at the two extremities of the molecule are also mostly primary.

To avoid any misinterpretation, we studied the accessibility of the tRNA-like fragment to four enzymes: the base-specific RNase T1 and chicken liver (CL_3) RNase (Levy and Karpetsky, 1980; Boguski *et al.*, 1980) and the structure-specific nucleases S1 (Ando, 1966) and cobra venom RNase (Vassilenko and Ryte, 1975). This allowed us to interpret the results obtained with one nuclease as a function of those obtained with the others. The experimental conditions used were adapted from earlier work on different types of RNAs (Wrede *et al.*, 1979a; Levy and Karpetsky, 1980; Favorova *et al.*, 1981; Branlant *et al.*, 1981) and optimized for the limited digestion of the native tRNA-like fragment. Since the pHs and the salt conditions vary with the enzyme, the possibility exists that the RNA has slightly different conformations in the different incubation conditions.

Secondary structure of the tRNA-like fragment

Figure 5 shows the secondary structure of the tRNA-like fragment of TYMV RNA deduced from the accessibility studies summarized in Figure 3. This model displays six main loop and stem regions and differs significantly from the earlier hypothetical proposals (see Figure 1).

The arms III, V, and VI are easiest to detect because nuclease S1 cuts at many sites in the loops. The cleavages of arm III by cobra venom RNase are consistent with a double-helical stem formed by residues 46 – 52 and 60 – 66. This stem could possess some structural flexibility since one primary cut occurs at C50 in the presence of RNase CL_3, which preferentially cleaves single strands (Levy and Karpetsky, 1980). The existence of loop III is supported by the frequent cleavage by nuclease S1 and RNase CL_3 at positions 55 and 57.

Six strong cuts produced by nuclease S1 at positions 91 – 96, as well as hydrolysis by RNase CL_3 at U94, reveal loop V. It is interesting to note the strong cleavage at U94 by RNase CL_3 even though this nuclease is 60 times less specific for U than for C residues (Levy and Karpetsky, 1980). This means that U94 is particularly exposed in the native structure. Splitting by cobra venom RNase (positions 98 – 101) and inaccessibility of G100 to RNase T1 and of C88 and C101 to RNase CL_3 support the existence of stem V.

The existence of loop VI is clearly shown by the frequent nuclease S1 cleavage at positions 126 – 129 as well as RNase CL_3 cuts at positions C125 and C128. The corresponding stem is more complex; it comprises a succession of double-stranded regions and internal loops referred to as loop VIa and VIb in Figure 5. The two double-stranded regions at the extremities of arm VI are well defined since only cobra venom RNase cleaves there. The region between them, positions 108 – 119 and 135 – 149, is less well defined since both

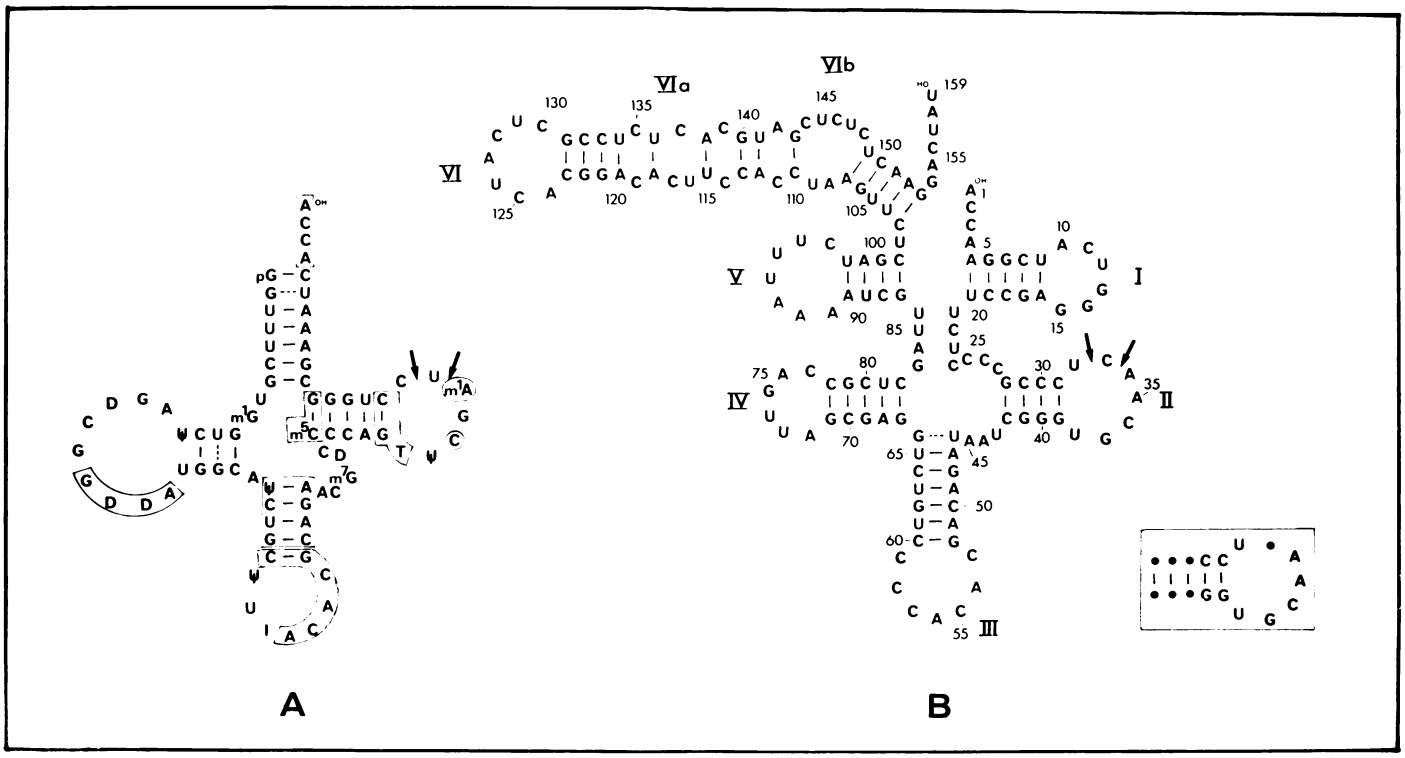


Fig. 5. Secondary structure of the tRNA-like fragment (n = 159) present at the 3′ OH terminus of TYMV RNA. **(A)** Secondary structure of yeast $tRNA^{Val}$ (Bonnet *et al.*, 1974) and **(B)** model of the viral RNA fragment. The nucleosides of $tRNA^{Val}$ common to the structure of the viral RNA are in boxes. The arrows indicate position of the phosphates which are inaccessible to alkylation by ethylnitrosourea; in the case of $tRNA^{Val}$ the ethylnitrosourea data are taken from Vlassov *et al.* (1981). The inset figure represents the homologous nucleosides present in arm II of the tRNA-like fragment and the Tψ arm of $tRNA^{Gly}$ from *S. epidermidis* (Roberts, 1974).

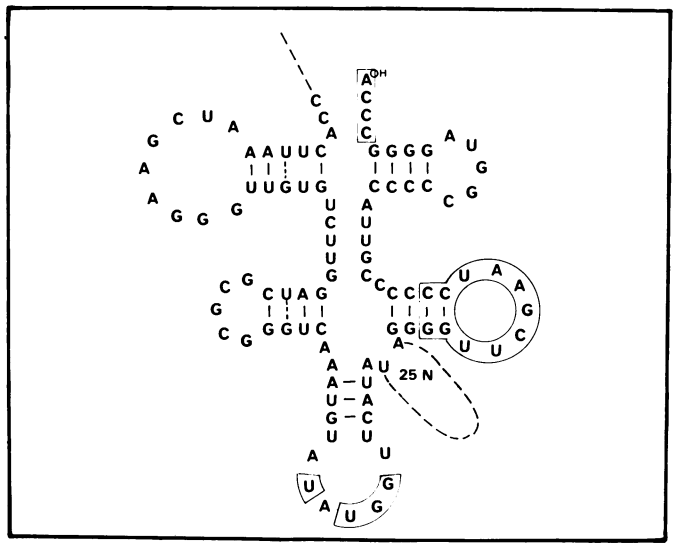


Fig. 6. Tentative model of the secondary structure of the tRNA-like fragment at the 3′ OH terminus of tobacco mosaic virus RNA. This RNA was sequenced by Guilley *et al.* (1979). The broken line corresponds to a 25-nucleotide long sequence which would fold as a long extra-loop. The nucleosides in boxes are those found at topologically similar positions in $tRNA^{His}$ from *Drosophila melanogaster* (Altweg and Kubli, 1980). It is recalled that $tRNA^{His}$ possesses a GUG anticodon whereas in the viral RNA another potential histidine anticodon, AUG, is found.

nuclease S1 and cobra venom RNase cleave at the same positions. This probably reflects some breathing of arm VI which can form different conformations in equilibrium. Since the most accessible G of the molecule is located at position 154, it is likely that the 5′-end sequence 154 – 159 is single-stranded despite the fact that cobra venom RNase cleaves, either secondarily or because of the stacking of the single-strand, in this region.

In contrast to arms V and VI, and to some extent to arm III, which are readily accessible to nucleases, suggesting rather loose structures, arms I, II, and IV are less accessible (see Figure 3). Indeed, the existence of loop II is only supported by cuts at positions 36 and 37 and that of loop IV by one cut at position 73. The corresponding stems are consonant with the cobra venom RNase sites and with the inaccessibility of several G and C residues to RNases T1 and CL_3.

The digestion experiments show that arm I, and more generally the region 1 to 27, is rather compact. Cuts by nuclease S1, and RNases CL_3 and T1 establish the existence of the six-nucleotide long loop I. The existence of the corresponding stem is supported by cobra venom RNase cuts and the inaccessibility of C8 and C19 to RNase CL_3 and of G17 to RNase T1 in the 5 – 9 and 16 – 20 sequences. Interestingly, cobra venom RNase splits in loop I and, in the stem, G6 and G7 are susceptible to RNase T1. Stem I could begin with an additional pair A4-U21; this possibility was not kept in the model since it engages the discriminator site A4 (Crothers *et al.*, 1972) in a double-stranded structure. Furthermore, the sequence 21 to 27, indicated as single-stranded in Figure 5, is cut by cobra venom RNase. One interpretation is that a double-stranded domain is formed between stretches 21 to 24 and 83 to 86. This is supported by cobra venom RNase cleavage at position 83 – 86 and by two possible base pairs (G-C and A-U) involving these distant parts of the molecule. All these results suggest a more complex organization of this region (see below).

The tRNA-like fragment undergoes "spontaneous" degradation of unknown origin when not handled with special care (Figure 3). Figure 5 shows that these breaks occur in single-stranded regions, particularly in loops III, IV, and V, and in the weakly structured arm VI. Preliminary experiments suggest that they are cation mediated.

The cloverleaf structure of yeast $tRNA^{Val}$ and the proposed secondary structure of the viral RNA fragment display striking similarities (Figure 5). The anticodon arm of $tRNA^{Val}$ resembles arm III in the tRNA-like fragment (Briand *et al.*, 1977; Giegé *et al.*, 1978). New analogies appear between the Tψ and D arms of $tRNA^{Val}$ and arms II and IV of the viral RNA, respectively. It is interesting to note the structural constants found in tRNA and conserved in the tRNA-like fragment, i.e., the CCA-end, the G-C base pair at the top of the Tψ stem, and the length of the Tψ and anticodon loops. Furthermore, the model shows a single-stranded region which is homologous to the extra-loop of a tRNA. The two types of molecules, however, present differences, especially in the amino-acid accepting region. Whereas the tRNA possesses a 7-bp stem, the viral RNA fragment has a much more complex structure composed of the two extra-arms I and V. Arm VI does not have any resemblance to tRNA.

By analogy with the secondary structure of the 3′ OH terminal sequence of TYMV RNA, we have folded the tRNA-like structure of tobacco mosaic virus, which is chargeable with histidine (Oberg and Philipson, 1972). A strikingly similar five arm structure can be constructed provided that a large extra-loop composed of 25 nucleotides is included between arms II and III (Figure 6). Here again sequence analogies with $tRNA^{His}$ are found at the amino-acid accepting terminus, the Tψ arm, and the anticodon loop. Interestingly, an extra-arm I near the CCA end is also conserved in this model as well as in the tRNA-like structures found in the bromoviruses and cucumber mosaic virus RNAs (Ahlquist *et al.*, 1981).

Three-dimensional structure of the tRNA-like fragment

A three-dimensional structure can only be proposed for the last 100 nucleotides of the TYMV RNA fragment corresponding to arms I to V. Our data suggest that arm VI has a loose structure which probably, for functional reasons (see below), lacks a well-defined tertiary organization.

Loops II and IV are only weakly accessible to RNases T1 and CL_3 and nuclease S1, in contrast to loop III which is very accessible to these enzymes. This behaviour resembles that observed in several tRNAs where the D and Tψ loops are inaccessible and the anticodon loops are readily accessible to nuclease S1 (e.g., Wrede *et al.*, 1979a, 1979b; Lockard and Kumar, 1981). In the corresponding stem regions, the strongest cuts by cobra venom RNase are in stem III, similar to the tRNAs where the anticodon stem is the most easily split by this enzyme (e.g., Favorova *et al.*, 1981; Lockard and Kumar, 1981; Butorin *et al.*, 1982). These results suggest that the three-dimensional structure of arms II, III, and IV is similar to that of the corresponding parts in tRNA. To provide more direct evidence, an alkylation experiment of the phosphates in the 20 – 45 region was performed. As shown by Vlassov *et al.* (1981) two phosphates located at the 3′ side of the Tψ loop of all tRNAs investigated, including $tRNA^{Val}$, are inaccessible to alkylation by ethylnitrosourea. Interestingly there is a quasi-perfect identity between the phosphodiester reactivity and the steric accessibility of these phosphate groups as calculated from the X-ray crystal structure of $tRNA^{Phe}$ (Thiyagarayan and Ponnuswamy, 1979; Lavery *et al.*, 1980). In the tRNA-like structure of TYMV RNA two inaccessible phosphates are found in loop II at topologically similar positions (see Figure 4 and 5). These two phosphates are located in the hinge region of tRNA where the two helical domains form the characteristic L-shaped structure, more precisely in the cleft where the D loop overlaps the Tψ loop. The ethylnitrosourea experiment shows that a similar structure exists in the viral RNA: arms III and IV would form the anticodon limb of a L-shaped structure and arm II the beginning of the amino-acid accepting limb of the L. It is likely, however, that the overlap of loop IV on loop II is not as perfect as in $tRNA^{Phe}$ since positions 36 and 73 are weakly accessible to nucleases; the organization of these two loops resembles more that of the D and Tψ loops in yeast $tRNA^{Asp}$ (Moras *et al.*, 1980).

In the canonical tRNA, the helical domain containing the amino-acid acceptor end is constructed with the Tψ arm and the amino-acid stem. In the tRNA-like structure the amino-acid stem is replaced by a more complex structure (see Figure 5). Our data do not establish a definitive three-dimensional model for this region. We propose nevertheless two alternative structures: (i) the 3′ OH terminal sequence 1 – 27 including arm I has, by itself, through internal tertiary interactions, a compact amino-acid accepting structure, illustrated by cobra venom RNase cuts. This model is reminiscent of that of a three-quarter fragment of tRNA still aminoacylable (Renaud *et al.*, 1979a); and (ii) loops I and V would interact, stabilizing an embryonic double-helical structure built up with sequences 21 – 24 and 83 – 86. In this model, arms I and V would bulge out to the exterior of the L. Such an interaction between these loops would explain cleavage by cobra venom RNase at position 94 and, if this were the case, the strong nuclease S1 cuts found with 5′-labelled RNA (see Figure 3) at position 91 – 96 would essentially be of secondary origin except for cuts 92 – 94. In favour of this model, we note a certain base complementarity between loops I and V in TYMV and tobacco mosaic virus RNAs. The interaction of arm I with a distant complementary sequence was also proposed by Ahlquist *et al.*, (1981) for the tRNA-like structures of bromoviruses and cucumber mosaic virus RNAs.

Relationship between structural features and functions of the tRNA-like fragment

From sequence studies, the 5′ end of the tRNA-like fragment (region 107 to 159), which mostly forms arm VI, contains the 3′-terminal part of the coat protein cistron (Briand *et al.*, 1977). The first termination signal UAA is situated in loop VIb; the four other nonsense codons found at irregular intervals, but in the same phase (positions 92 – 94, 83 – 85, 71 – 73, and 44 – 46) along the RNA, are also located in loops or single-stranded regions.

The 3′-terminal fragment released by RNAse P contains a structure formed by arms I to V (see Figure 5) which can be esterified with valine (Prochiantz and Haenni, 1973). This fragment probably contains all the structural features which allow the efficient valylation of TYMV RNA (Giegé *et al.*, 1978). Studies on tRNAs, including yeast $tRNA^{Val}$, have shown that the domains recognized by the aminoacyl-tRNA synthetases are mainly formed by the Tψ, anticodon, and D arms, the amino-acid acceptor stem having a less important role in this process (e.g., Ebel *et al.*, 1979; Renaud *et al.*, 1979a, 1979b; Favorova *et al.*, 1981). The structural organization of the tRNA-like fragment of TYMV is com-

patible with the presence of a synthetase recognition site in a structure formed by arms II, III, and IV. Under certain conditions three-quarters of a tRNA can be efficiently aminoacylated (Renaud *et al.*, 1979a). In this process the flexibility of the molecule, and especially that of the amino-acid acceptor region, plays an important role in allowing the precise positioning of the CCA end in the catalytic site of the enzyme (Ebel *et al.*, 1973; Renaud *et al.*, 1979a). The last 85 nucleotides appear to form the minimum structure chargeable at an appreciable rate by valyl-tRNA synthetase. We cannot exclude, however, possible participation of arm V in an acceptor structure which would favour a better positioning of the CCA end in the catalytic site of the enzyme. This model is compatible with the interaction of valyl-tRNA synthetase with the inside of a L-shaped tRNA-like structure, similar to the interaction between tRNA and synthetase (Rich and Schimmel, 1977).

Aminoacylation of tRNAs is a prerequisite for the amino-acid donor function of these molecules during protein synthesis. The viral tRNA-like structure, however, does not seem to be involved in protein biosynthesis (Haenni and Chapeville, 1979). New features of the secondary structure of the tRNA-like fragment from TYMV support this statement. As seen in Figure 5, arm II, lacking modified bases, does not possess the characteristic features of the homologous Tψ loop of elongator tRNAs, but is strikingly similar to $tRNA^{Gly}$ from *Staphylococcus epidermidis*. Furthermore, arm IV does not contain the two adjacent G residues conserved in the D loop of elongator tRNAs, but, like the $tRNA^{Gly}$ from *S. epidermidis*, contains two adjacent U residues. This tRNA participates in cell wall peptidoglycan synthesis which does not use the ribosomal machinery (Stewart *et al.*, 1971).

The 3′- and 5′-terminal regions of the 6600-nucleotide long genomic RNA of TYMV present complementary sequences (Briand *et al.*, 1978); this might help in the circularization of the viral RNA (Strazielle *et al.*, 1965). The eight-nucleotide long sequence at the 3′ end of the RNA, complementary with the 5′ end, is mostly located in loop I.

Finally, it appears that many plant viral RNAs which are able to be aminoacylated have a similar organization at their 3′ end. We propose a five arm tRNA-like structure for TYMV RNA; the 3′ terminus of tobacco mosaic virus RNA can be folded in the same way and bromoviruses and cucumber mosaic virus RNAs also present similar structural features in this region. In all cases the viral tRNA-like structures possess an extra-arm near the amino-acid accepting CCA which is never found in the canonical tRNAs. It is tempting to propose that this feature is not only involved in the aminoacylation of the viral RNAs but also acts as a signal for other functions, such as regulation of translation or initiation of replication. This regulatory role could be achieved through conformational changes of the viral RNAs governed by this extra-arm.

Materials and methods

Materials

TYMV and viral RNA were extracted by established procedures (Briand *et al.*, 1977). The (n = 159) 3′ OH terminal tRNA-like fragment was purified from a partial RNase T1 hydrolysis of the viral RNA by gel filtration on Ultrogel ACA_{54} columns followed by reversed phase chromatography (Florentz *et al.*, 1982). tRNA-nucleotidyl transferase from baker's yeast (EC 2.7.7.21 or 25) was prepared according to Rether *et al.* (1974). Ribonuclease T1 (EC 3.1.27.3) was from Sankyo (Tokyo); RNAse CL_3 was from Bethesda Research Laboratories (Neu Isenberg, GFR); nucleases S1 (EC 3.1.4), P1 (EC 3.1.4) and T4 polynucleotide kinase (EC 2.7.1.78) were from Boehringer (Mannheim). Cobra venom RNase was prepared according to Vassilenko and Ryte (1975) and had a specific activity of 4 units/μl (Vassilenko and Babkina, 1965). [α-^{32}P]ATP (410 Ci/mmol) and [γ-^{32}P]ATP (3200 Ci/mmol) were from the Radiochemical Centre, Amersham (UK). Urea RNase free, was from Schwarz Mann. Saccharose, RNase free, acrylamide and N N′-methylene-bisacrylamide were from BDH Chemicals. Other chemicals were of analytical grade. For autoradiography and fluorography Kodak X-omat-S films (30 x 40 cm) and M-R 400 X Ray intensifying screens from Agfa-Gevaert were used.

Labelling and handling of the tRNA-like fragment

The labelling of the tRNA-like fragment at its 5′ or 3′ end with radioactive ATP was done according to methods developed by Silberklang *et al.* (1977b) for tRNA labelling. Since the viral fragment originates from a RNase T1 hydrolysis it does not possess a phosphate residue at its 5′ end and thus can be directly labelled with [γ-^{32}P]ATP by T_4 polynucleotide kinase. At its 3′ end the TYMV RNA has a CC_{OH} terminus; thus it can be directly labelled with [α-^{32}P]ATP by tRNA-nucleotidyltransferase (Litvak *et al.*, 1970). After labelling, the fragments were purified by electrophoresis on polyacrylamide gels, eluted at +4°C according to Maxam and Gilbert (1977), and ethanol precipitated at −80°C. Labelling, as well as further experiments on the tRNA-like fragment, were carried out with sterilized vessel and pipettes, and quartz double-distilled water.

Despite these cautions some spontaneous degradations could occur in the fragment. We thus investigated, in a systematic study, different factors which could affect the integrity of the RNA (to be published); in particular we noticed that elution of the labelled material from the gel in the presence of freshly distilled phenol, which should protect from RNase degradations, on the contrary markedly intensified these degradations. We emphasize, however, that the experiments described in this paper were done in such a way that the level of degradation was minimized and kept much lower than that of the cuts due to the enzymes or chemicals used as structural probes. This was achieved by quick handling of the material eluted from the gels.

Limited digestions of labelled tRNA-like fragments in non-denaturing conditions

For each assay, 75 000 Cherenkov counts of ^{32}P-labelled material mixed to 2 μg of carrier tRNA-like fragment were used. Prior to use, the RNA fragments were partially denatured by heating at 50°C for 5 min and renatured by incubation at 37°C for a further 5 min in the enzymatic buffers which all contain 10 mM $MgCl_2$. Immediately after digestion equal parts of each reaction mixture were loaded on 10, 15 and 20% polyacrylamide slab gels (30 x 50 x 0.05 cm). Gels were run at 2000 volts and 20−40 mA.

Partial RNase T1 digestion was performed in 10 μl of 10 mM Tris-HCl pH 7.2, 10 mM $MgCl_2$ buffer. Incubations were for 2 min at 37°C with 0.02 unit and 0.01 unit of RNase for 3′- and 5′-labelled fragments, respectively.

Digestion with nuclease S1 was carried out at 37°C in 6 μl 50 mM KCl, 1 mM $ZnCl_2$, 25 mM acetate Na pH 4.5, 10 mM $MgCl_2$ for 2, 5, and 10 min with 1 unit of enzyme. It was stopped by addition of 4 μl 10 mM ATP.

Digestion with 0.1 unit of cobra venom RNAse was done in 10 μl 10 mM Tris-HCl pH 7.2; 10 mM $MgCl_2$ at 0°C for 2, 5, and 10 min and stopped by addition of 2 μl EDTA 100 mM.

Partial nuclease CL_3 digestion was done in 10 μl 10 mM potassium phosphate pH 6.5 for 10 min at 37°C with 0.5 unit of enzyme.

For the identification of the cuts, RNase T1 and alkaline ladders (Donis-Keller *et al.*, 1977) were run in parallel on the same gel. As small cobra venom RNase and nuclease S1 digestion fragments showed differences in electrophoretic mobilities when compared to T1 or alkaline ladders, they were identified by comigration with RNAse P1 partial hydrolysis products (Silberklang *et al.*, 1977b).

Alkylation of phosphates by ethylnitrosourea

Alkylation by ethylnitrosourea of accessible phosphodiester bonds and splitting of the tRNA-like fragment at these destablized bounds were carried out under native (20°C) and denaturing (80°C) conditions according to Vlassov *et al.* (1981). Analysis of the liberated oligonucleotides was performed as described above.

Acknowledgements

We are grateful to Drs. C. Branlant, A. Krol, and S. Vassilenko for providing us with cobra venom RNase and to Drs. B. Ehresmann and K. Dudley for stimulating discussion and critical reading of the manuscript. This work was supported by grants from the Centre National de la Recherche Scientifique and the Délégation Générale à la Recherche Scientifique et Technique.

References

Ahlquist,P., Dasgupta,R., and Kaesberg,P. (1981) *Cell*, **23**, 183-189.

Altweg,M., and Kubli,E. (1980) *Nucleic Acids Res.*, **8**, 3259-3262.
Ando,T. (1966) *Biochim. Biophys. Acta*, **114**, 158-168.
Boguski,M.S., Hieter,P.A., and Levy,C.C. (1980) *J. Biol. Chem.*, **255**, 2160-2163.
Bonnet,J., Ebel,J.P., Dirheimer,G., Shershueva,L.P., Krutilina,A.I., Venkstern,T.V., and Bayev,A.A. (1974) *Biochimie*, **56**, 1211-1213.
Branlant,C., Krol,A., Ebel,J.P., Gallinaro,H., Lazar,E., and Jacob,M. (1981) *Nucleic Acids Res.*, **9**, 841-857.
Briand,J.P., Jonard,G., Guilley,H., Richards,K.E., and Hirth,L. (1977) *Eur. J. Biochem.*, **72**, 453-463.
Briand,J.P., Keith,G., and Guilley,H. (1978) *Proc. Natl. Acad. Sci. USA*, **75**, 3168-3172.
Butorin,A.S., Remy,P., Ebel,J.P., and Vassilenko,S.K. (1982) *Eur. J. Biochem.*, **121**, 587-595.
Crothers,D.M., Seno,T., and Söll,D.G. (1972) *Proc. Natl. Acad. Sci. USA*, **69**, 3063-3067.
Donis-Keller,H., Maxam,A.M., and Gilbert,W. (1977) *Nucleic Acids Res.*, **4**, 2527-2538.
Ebel,J.P., Giegé,R., Bonnet,J., Kern,D., Befort,N., Bollack,C., Fasiolo,F., Gangloff,J., and Dirheimer,G. (1973) *Biochimie*, **55**, 547-557.
Ebel,J.P., Renaud,M., Dietrich,A, Fasiolo,F., Keith,G., Favorova,O.O., Vassilenko,S., Baltzinger,M., Ehrlich,R., Remy,P., Bonnet,J., and Giegé, R. (1979) in Söll,D., Abelson,J.N., and Schimmel,P.R. (eds.), *Transfer RNA:Structure, Properties and Recognition*, Cold Spring Harbor Monograph, Ser. 9A, Cold Spring Harbor Laboratory Press, NY, pp. 325-343.
Favorova,O.O., Fasiolo,F., Keith,G., Vassilenko,S.K., and Ebel,J.P. (1981) *Biochemistry (Wash.)*, **20**, 1006-1011.
Florentz,C., Mengual,R., Briand,J.P., and Giegé,R. (1982) *Eur. J. Biochem.*, in press.
Giegé,R., Briand,J.P., Mengual,R., Ebel,J.P., and Hirth,L. (1978) *Eur. J. Biochem.*, **84**, 251-256.
Guilley,H., Jonard,G., Kukla,B., and Richards,K.E. (1979) *Nucleic Acids Res.*, **6**, 1287-1308.
Haenni,A.L., and Chapeville,F. (1979) in Söll,D., Abelson,J.N., and Schimmel,P.R. (eds.), *Transfer RNA: Biological Aspects*, Cold Spring Harbor Monograph, Ser. 9B Cold Spring Harbor Laboratory Press, NY, pp. 539-556.
Hall,T.C. (1979) *Int. Rev. Cytol.*, **60**, 1-26.
Lavery,R., Pullman,A., and Pullman,B. (1980) *Theor. Chim. Acta*, **57**, 233-243.
Levy,C.C., and Karpetsky,P. (1980) *J. Biol. Chem.*, **255**, 2153-2159.
Litvak,S., Carre,D.S., and Chapeville,F. (1970) *FEBS Lett.*, **11**, 316-319.
Lockard,R.E., and Kumar,A. (1981) *Nucleic Acids Res.*, **9**, 5125-5140.
Maxam,A.M., and Gilbert,W. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 560-564.
Moras,D., Thierry,J.C., Comarmond,M.B., Fischer,J., Weiss,R., Ebel,J.P., and Giegé,R. (1980) *Nature*, **288**, 669-674.
Oberg,B., and Philipson,L. (1972) *Biochem. Biophys. Res. Commun.*, **48**, 927-932.
Prochiantz,A., and Haenni,A.L. (1973) *Nature*, **241**, 168-170.
Renaud,M., Ehrlich,R., Bonnet,J., and Remy,P. (1979a) *Eur. J. Biochem.*, **100**, 157-164.
Renaud,M., Dietrich,A., Giegé,R., Remy,P., and Ebel,J.P. (1979b) *Eur. J. Biochem.*, **101**, 475-483.
Rether,B., Bonnet,J., and Ebel,J.P. (1974) *Eur. J. Biochem.*, **50**, 281-288.
Rich,A. and Schimmel,P.R. (1977) *Nucleic Acids Res.*, **4**, 1649-1655.
Rietveld,K., Van Poelgeest,R., Pleij,C.W.A., and Bosch,L. (1981) Abstracts of the Fifth International Congress of Virology, P2/10.
Roberts,R.J. (1974) *J. Biol. Chem.*, **249**, 4787-4796.
Silberklang,M., Prochiantz,A., Haenni,A.L., and RajBhandary,U.L. (1977a) *Eur. J. Biochem.*, **72**, 464-478.
Silberklang,M., Gillum,A.M., and RajBhandary,U.L. (1977b) *Nucleic Acids Res.*, **4**, 4091-4108.
Stewart,T.S., Roberts,R.J., and Strominger,J.L., (1971) *Nature*, **230**, 36-38.
Strazielle,C., Benoit,H., and Hirth,L. (1965) *J. Mol. Biol.*, **13**, 735-748.
Thiyagarajan,P., and Ponnuswamy,P.K. (1979) *Biopolymers*, **18**, 2233-2247.
Toots,I., Metspalu,A., Villems,R., and Sarma,M. (1981) *Nucleic Acids Res.*, **9**, 5331-5343.
Vassilenko,S.K., and Babkina,G.T. (1965) *Biokhimiya*, **30**, 705-712.
Vassilenko,S.K., and Ryte,V.C. (1975) *Biokhimiya*, **40**, 578-582.
Vassilenko,S.K., Carbon,P., Ebel,J.P., and Ehresmann,C. (1981) *J. Mol. Biol.*, **152**, 699-721.
Vlassov,V.V., Giegé,R., and Ebel,J.P. (1981) *Eur. J. Biochem.*, **119**, 51-59.
Wrede,P., Wurst,R., Vournakis,J., and Rich,A. (1979a) *J. Biol. Chem.*, **254**, 9608-9616.
Wrede,P., Woo,N.H., and Rich,A. (1979b) *Proc. Natl. Acad. Sci. USA*, **76**, 3289-3293.