The Structure of Tobacco Rattle Virus Ribonucleic Acids: Common Nucleotide Sequences in the RNA Species

By G. DARBY AND A. C. MINSON

Department of Virology, Medical School,
University of Birmingham B15 2TJ, England

(Accepted 8 June 1973)

SUMMARY

It has been shown by competition hybridization experiments that common sequences of about 600 nucleotides are present in TRV long and short particle RNA species. Heteroduplex molecules have been formed by the hybridization of short particle RNA strands with long complementary strands derived from tissue infected with a ‘defective isolate’ of the virus. The heteroduplex molecules were shown to have a lower melting temperature than either the long–long or short–short homoduplexes, indicating some mismatching of base pairs. The results are discussed in the light of our present knowledge about the replication of viruses.

Theoretical equations have been derived to describe RNA–RNA competition hybridization and duplex formation.

INTRODUCTION

The CAM isolate of tobacco rattle virus (TRV) consists of two types of helical particles; a long rod (197 nm) and a short rod (52 nm) (Harrison & Woods, 1966). Both contain 5% RNA and both are required for a complete infection; the short particle contains the gene for the coat protein (Sänger, 1968, 1969).

Although information in both parts of the virus genome is required to achieve a successful infection, it is not known whether there is duplication of information. RNA–RNA competition hybridization experiments have provided evidence that partial sequence homology may exist between the two RNA species (Minson & Darby, 1973). However, the results of these experiments could not be interpreted unambiguously and to clarify the situation, further experiments have been done using alternative sources of material.

The long particles of TRV are infectious, but in the absence of short particles no coat protein is produced and infection results only in the replication of long RNA molecules (Frost, Harrison & Woods, 1967; Lister, 1966, 1968). The result is a ‘defective isolate’ of the virus which can only be recovered from infected plants by phenol extraction. Plants infected with this isolate provide a source of both TRV long RNA and its complement.

In this paper we describe experiments in which RNA from this source has been used to resolve the ambiguity of our previous work and establish that considerable sequence homology exists between long and short particle RNA species.
Biological materials. TRV isolate CAM was a gift from Dr B. D. Harrison. The preparation of high specific activity [\textsuperscript{3}H]-labelled TRV RNA species and also the preparation of double-stranded TRV RNA have been described previously (Minson & Darby, 1973). The preparation of high specific activity [\textsuperscript{32}P]-TRV RNA species was similar to that of [\textsuperscript{3}H]-RNA except that the virus was grown in leaf strips infiltrated with water and [\textsuperscript{32}P]-orthophosphate (The Radiochemical Centre, Amersham).

Preparation of TRV ‘defective isolate’. Long and short particles of purified TRV were separated by sucrose gradient sedimentation and the RNA was extracted from the long particles (Darby & Minson, 1972). The RNA was purified by sucrose gradient sedimentation and inoculated on leaves of Chenopodium amaranticolor at a concentration of 0.05 \mu g/ml. Leaves showing 1 or 2 local lesions were removed and the lesions isolated. Ten such lesions were pooled and ground in 1.5 ml 0.02 M-sodium phosphate, pH 7\textsuperscript{.}4, with 1.5 ml phenol and the aqueous phase inoculated on leaves of C. amaranticolor. The resulting local lesions were harvested after 8 days and divided into two equal groups. One group was extracted with phosphate buffer and the extract was left for several hours at room temperature and then tested for infectivity. The other group was extracted with buffer and phenol and the resulting aqueous phase was used as a stock inoculum of ‘defective TRV’.

Preparation of RNA from tissue infected with ‘defective TRV’. Leaves of Chenopodium amaranticolor were inoculated with ‘defective TRV’ to yield 50 to 100 lesions per leaf. After 10 days all leaves with more than 30 lesions were harvested. The total cell RNA was extracted and purified as previously described (Minson & Darby, 1973). The RNA was fractionated with 2 M-LiCl into a high mol. wt., single-stranded RNA precipitate (Def-ss-RNA) and a
Common sequences in TRV RNAs

A supernatant fraction containing low mol. wt., single-stranded RNA and double-stranded RNA (Def-ds-RNA).

Hybridization techniques. These have been described previously (Minson & Darby, 1973). Briefly, the labelled RNA species (2 to 5000 cts/min) was mixed with double-stranded RNA in 2 × SSC (0.3 M NaCl, 0.03 M Na citrate), heated to 115 °C and then annealing was allowed to proceed at 70 °C for 20 h. The amount of acid-precipitable labelled RNA was measured after treatment with ribonuclease.

Thermal denaturation profiles. The method used was similar to that of Kingsbury (1966). Labelled RNA–RNA double-stranded hybrids were prepared by annealing labelled virus RNA with complementary strands at 70 °C for 20 h in 2 × SSC. They were cooled to 0 °C and diluted with water to 0.1 × SSC. Samples were heated at various temperatures for 5 min, cooled rapidly to 0 °C and then adjusted to 1 × SSC. Each sample was then digested with 50 units T1 ribonuclease ( Worthington) and 50 μg pancreatic ribonuclease (Sigma) for 20 min at 25 °C and the remaining TCA-precipitable labelled RNA was measured.

RESULTS

‘Defective TRV’

The preparation and testing of ‘defective TRV’ is summarized in Fig. 1. The purified long RNA, which was used initially in the preparation of the isolate, was inoculated on to leaves at 2.5 μg/ml to yield about 100 lesions/leaf. A buffer extract of 500 such lesions showed no ribonuclease resistant infectivity. Therefore at this high concentration of RNA no contaminating short RNA molecules can have complemented with long RNA to produce complete nucleoprotein particles. The long RNA was diluted to 0.05 μg/ml and was then used to prepare the ‘defective TRV’ isolate.

Competition hybridization studies

To show that Def-ss-RNA contained TRV long RNA sequences, it was used as competitor in hybridization experiments in which [3H]-labelled TRV long particle RNA was hybridized to complementary sequences in double-stranded RNA from TRV infected tissue. The results are shown in Fig. 2. Def-ss-RNA competed with [3H]-labelled long RNA and the data obtained fitted a theoretical homologous competition curve. A similar fraction of RNA from uninfected tissue did not compete. This experiment was repeated using unlabelled TRV long particle RNA as competitor and the data could be fitted to the same theoretical curve (Fig. 2). These data also allowed an estimate to be made of the TRV long RNA content of Def-ss-RNA. Fig. 2 shows that 1 μg Def-ss-RNA contains about 8 × 10⁻¹⁸ mol of TRV long RNA.

Having established that Def-ss-RNA contained TRV long RNA sequences, experiments were done to determine whether this RNA competed with [3H]-labelled short particle RNA in the formation of hybrids. Control experiments were done using short particle RNA as competitor. The results of these experiments are shown in Fig. 3 and it is clear that Def-ss-RNA competed in the formation of hybrids. For these competition curves \( f/c \approx 1/c \), where \( f \) is the fraction of label displaced by the competitor and \( c \) is the concentration of competitor (see appendix). The intercept on the \( 1/f \) axis when \( 1/c = 0 \) gives the fraction displaced when the competitor concentration is infinite. This is the fraction of the [3H]-labelled RNA whose sequences are present in the competitor.

Fig. 4a shows the reciprocal plot of the data in Fig. 3. When short particle RNA was used as competitor \( 1/f = 1 \) when \( 1/c = 0 \) and thus, as expected, all sequences present in [3H]-
short RNA were present in the homologous competitor. When Def-ss-RNA was used as competitor \( I/f \) was approximately 3.5 when \( I/c = 0 \). Therefore, \( f = 0.29 \) when \( c = \infty \), and hence about 29% of the \([\text{3H}]\)-labelled short RNA sequences were present in the heterologous competitor. Fig. 4b shows a similar treatment of the data taken from an experiment in which RNA from purified long particles was used as a competitor (Minson & Darby, 1973). When \( I/c = 0 \), \( I/f \) was approximately 3.7, indicating that about 27% of the \([\text{3H}]\)-labelled short RNA sequences were present in the competitor.

In these competition experiments almost identical results were obtained using entirely different sources of competing long RNA molecules.
Common sequences in TRV RNAs

Fig. 4. Reciprocal data from experiments in which $[^3H]$-short particle RNA was hybridized with denatured TRV double-stranded RNA in the presence of various competitors. (a) ○—○, short particle RNA; •—•, Def-ss-RNA. (b) ○—○, short particle RNA; •—•, long particle RNA.

Heteroduplex formation

As the long and short RNA species contained common sequences it was expected that Def-ds-RNA would contain complementary long RNA which would hybridize to both short particle RNA (heteroduplex) and to long RNA (homoduplex). This was tested by hybridizing $[^3H]$-labelled short RNA with increasing amounts of Def-ds-RNA. A control experiment was done in which $[^3H]$-labelled long RNA was hybridized with the Def-ds-RNA. The results (Fig. 5a) demonstrate that Def-ds-RNA contains sequences complementary to both long and short particle RNA. The same fraction from uninfected tissue contained no sequences complementary to either RNA.

The fraction of labelled RNA hybridized ($F$) is related to the concentration of double-stranded RNA ($d$) such that $1/d \propto 1/F$ (see appendix). The intercept on the $1/F$ axis when $1/d = 0$ gives the fraction of $[^3H]$-RNA hybridized at infinite double-stranded RNA concentration. This represents the fraction of the labelled RNA for which there are complementary sequences in the double-stranded RNA. Reciprocal plots of the data are shown in Fig. 5b. When $[^3H]$-labelled long RNA was hybridized to denatured Def-ds-RNA the reciprocal data fitted a linear plot which intercepted $1/F$ at approximately 1.2. This implied that about 80 to 85% of the sequences in the $[^3H]$-labelled long RNA were complementary.
to sequences in Def-ds-RNA. We would have predicted, of course, that the Def-ds-RNA would have contained sequences complementary to all sequences in the long particle RNA and that the $1/F$ intercept should have been 1. However, this theoretical value of 1.0 can only be obtained if the Def-ds-RNA contains equal amounts of long RNA and its complement, but not if it contains an excess of virus RNA. It is probable that LiCl precipitation which is used to remove the large excess of virus RNA in the preparations (Minson & Darby, 1973) is incomplete and that a small excess of virus RNA remains in the Def-ds-RNA fractions. This would then account for the failure of the data to fit a theoretical curve approaching 100% hybridization of the labelled RNA at infinite concentration of double-stranded RNA.

When denatured Def-ds-RNA was annealed with $^{3}H$-short particle RNA to form a heteroduplex, the reciprocal data (Fig. 5b) again fitted a linear plot but the intercept value on the $1/F$ axis was 4.8 implying that approximately 21% of the sequences in short particle...
RNA are complementary to sequences in Def-ds-RNA. However, since the reciprocal data for the homoduplex formation did not yield the theoretical $1/F$ intercept of $1.0$, the intercept of $1/F = 4.8$ for the heteroduplex may not be a true reflection of the extent of homology which may be somewhat greater than the calculated $21\%$.

**Thermal denaturation characteristics of hybrids**

In order to characterize further the common sequences present in TRV RNAs, a comparison was made of the thermal denaturation characteristics of homo- and heteroduplexes.

Three types of double-stranded hybrid were made:

(a) Hybrids formed by annealing $[^{32}P]$-long particle RNA to denatured Def-ds-RNA (long–long homoduplexes).

(b) Hybrids made by annealing $[^{32}P]$- or $[^{3}H]$-labelled short particle RNA to denatured double-stranded RNA from TRV infected tissue (short–short homoduplexes).

(c) Hybrids made by annealing $[^{3}H]$-short RNA with Def-ds-RNA (short–long heteroduplex).

$[^{3}H]$-labelled heteroduplexes were mixed either with $[^{32}P]$-labelled long–long homoduplexes or with $[^{32}P]$-labelled short–short homoduplexes. Thermal denaturation experiments were then done and the results are shown in Fig. 6. It is apparent from the data in Fig. 6 (a, b) that the short–long heteroduplex has a lower melting temperature than either homoduplex. The heteroduplex has a Tm of $69\, ^\circ\text{C}$ whereas that of both homoduplexes is $76\, ^\circ\text{C}$. The almost identical denaturation characteristics of the homoduplexes were expected as the GC contents of both are $42\%$ (Minson & Darby, 1973). Fig. 6c shows the results of a control experiment in which $[^{32}P]$-labelled short–short homoduplexes were mixed with $[^{3}H]$-labelled short–short homoduplexes and the denaturation profiles determined. The profiles were not significantly different indicating that the source of the labelled RNA used to prepare the hybrid did not alter the melting characteristics.
DISCUSSION

Competition hybridization experiments have indicated that TRV long particle RNA contains nucleotide sequences which are homologous to approximately 28% of the sequences in the short particle RNA. On the other hand, heteroduplex formation showed that 21% of the sequences in short particle RNA would hybridize to sequences in complementary long RNA. These data suggest that 20 to 30% of the sequences in the short particle RNA are also present in the long. As the mol. wt. of TRV (isolate CAM) short RNA is $0.7 \times 10^6$ (Tollin & Wilson, 1971; Cooper & Mayo, 1972), it appears that sequences totalling 500 to 700 nucleotides are common to both RNA species. It was shown also that the heteroduplex molecules had a lower melting temperature than either the long-long or the short-short homoduplexes. Two simple interpretations of these results are possible: either the heteroduplex contains a proportion of mismatched bases and hence the common sequences in long and short RNA are similar but not identical, or the homologous regions are identical but are richer in guanosine and cytosine than the remainder of the RNA. A combination of these explanations is also a possibility.

Little is known about possible common nucleotide sequences in the RNA species of other viruses with fragmented genomes. Van Kammen (1971) concluded from competition experiments with the two RNA components of cowpea mosaic virus that there was no sequence homology between the RNAs. Content & Duesberg (1971) obtained evidence suggesting that there might be sequence homology between the RNA fragments of influenza virus and, more recently, a probable common sequence of 29 nucleotides has been detected by T1 mapping of the RNA fragments (Horst et al. 1972). It has been suggested that this short sequence might represent a binding site for a protein, but the possibility that more extensive sequence homology exists between the RNA fragments cannot be excluded by the T1 mapping data.

There are a number of possible explanations for the presence of homologous sequences in TRV RNA species and several are discussed below.

Common evolutionary origin. It is possible that both particles of TRV-CAM are derived from a common ancestor, divergent evolution resulting in some remaining sequence similarities.

Since all known strains of TRV are multicomponent it is reasonable to assume that they have evolved from a common multicomponent ancestor, and therefore divergence of the separate particles must have begun before divergence of different strains. It then follows that in the absence of any specific selection pressures (see below) the extent of divergence of the strains should not exceed that of the separate particles of a particular strain. In experiments we have done in collaboration with Dr D. J. Robinson, to compare sequence similarities between strains, we have found no detectable sequence homology between the CAM isolate and a North American strain (Oregon yellow) or a Scottish strain (PRN). These results suggest that the sequence homology we observe between TRV RNA species is unlikely to be due to the divergent evolution of the TRV RNA species from a common ancestor.

Common recognition sequences. Both RNA species must have recognition sites for binding coat protein, ribosomes and possibly a shared RNA polymerase. These untranslated regions may have similar sequences and account for the observed homology.

Little is known about the interaction of TRV RNA with its coat protein, but the observation of two-layered discs and stacked discs in TRV protein preparations (Semančík & Reynolds, 1969; Fritsch et al. 1972) suggests that the mechanism of TRV assembly may be similar to that of tobacco mosaic virus. TMV RNA is thought to contain a specific site at the 5'-end of the molecule at which assembly is initiated by the attachment of a two-layered
Common sequences in TRV RNAs

A two-layered disc of TRV protein would interact with about 200 nucleotides of TRV RNA (Tollin & Wilson, 1971) so it is conceivable that a fairly large recognition sequence exists for initiation of TRV assembly.

Information about ribosome and polymerase binding sites is available in bacteriophage systems. Ribosome binding sites appear to be less than 20 nucleotides in length and different sites in the same molecule often show little homology (Hindley & Staples, 1969; Steitz, 1969; Staples & Hindley, 1971). Polymerase binding sites may be much larger (Cory et al. 1972) but secondary structure is probably of greater importance than nucleotide sequence. Qβ RNA polymerase will efficiently replicate 6S ‘mini variant’ RNA (Prives & Silverman, 1972) but this RNA contains no sequence homology with Qβ RNA as measured by hybridization techniques (Spiegelman, 1971). It follows that long specific sequences are not required for polymerase binding.

Thus, while it is attractive to consider the sequence homology between long and short TRV RNA species to be due to the presence of common recognition sites, on the basis of currently available information it would seem unlikely that the specific sequences amounting to a total of 600 nucleotides could be accounted for in this way.

Common gene product. A sequence of 600 nucleotides is sufficient to code for a protein of about 20000 mol. wt. The sequence homology observed might therefore be explained by the presence of a common essential gene in each RNA. This would mean that the virus carries apparently redundant information, but since little is known about the replication of the virus and the interaction between the RNA species, this possibility cannot be excluded.

None of these explanations for the common nucleotide sequences in TRV RNA species appear satisfactory and this presumably reflects our lack of knowledge about the origins and modes of replication of viruses with divided genomes.

APPENDIX

(a) General formula for competition hybridization. In competition hybridization experiments labelled RNA at a constant concentration \(L\) is hybridized to homologous complementary RNA (derived from double-stranded RNA added at a concentration \(d\)) in the presence of increasing concentrations of competitor \(c\).

If the fraction of the labelled species homologous with sequences in the competitor is \(H\), then for any concentration of competitor

\[
P = \frac{P}{P_{c=0}},
\]

where

\[
P = \frac{P_{c=0}}{P_{c=0}}.
\]

If the fraction of the label displaced by the competitor is \(f\), then,

\[P = 1 - f.
\]

Therefore,

\[1 - f = \frac{H \left( \frac{d}{d + L + c} \right) + (1 - H) \left( \frac{d}{d + L} \right)}{d + L}.
\]
$H(d((d+L)+c))$ is the fraction of the label which is hybridized to regions of the complementary strand that are complementary also to regions in the competitor. $(1-H)(d((d+L))$ is the fraction of the label hybridized to regions of the complementary strand that are not complementary to regions in the competitor.

Equation (1) can be re-arranged to yield a general equation for competition as follows:

$1 - f = \left(\frac{d+L}{d+L+c}\right) H + (1-H),

\therefore \frac{f d + f L + f c}{c H} = c H.$

Dividing by $f c H$

$\frac{1}{f} = \frac{1}{H} + \frac{d}{c H} + \frac{L}{c H}.$

Hence

$\frac{1}{f} = \frac{1}{H} + \left(\frac{d+L}{H}\right) \frac{1}{c}.$

As $(d+L)/H$ is a constant, $1/f$ plotted against $1/c$ is a straight line. Furthermore, at infinite competitor concentration (i.e. when $1/c = 0$), $1/f = 1/H$. Hence, the fraction of the labelled species homologous with the competitor ($H$) is the reciprocal of the intercept on the $1/f$ axis.

Simmons & Strauss (1972) derived an approximate equation to describe a special case of competition hybridization where heterologous competitor is contaminated with a very low level of homologous competitor. Their equation, $P = -(P \cdot c)/(d+L) + (1-H)$ (where $x$ is the degree of contamination), describes a situation where competition by the heterologous competitor is complete but where competition by the contaminating homologous species is incomplete. It can therefore only be applied for high levels of competitor concentration where, in the absence of contamination, the level of competition would be independent of the competitor concentration.

(b) General formula for duplex formation. In duplex formation a constant quantity $(L)$ of a labelled RNA species is hybridized to complementary strands (derived from double-stranded material added at a concentration $d$). If $H$ is the fraction of the labelled RNA homologous to the complementary RNA, the general formula for the fraction of the labelled material hybridized ($F$) is

\[(\text{Fraction hybridized}) \quad F = \frac{H d}{d+L}. \quad (2)\]

Equation (2) can be re-arranged as follows:

Inverting (2)

$\frac{1}{F} = \frac{d+L}{H d},

\therefore \frac{1}{F} = \frac{1}{H} + \left(\frac{L}{H}\right) \frac{1}{d}.$

If the quantity of double-stranded material ($d$) is varied and the quantity of labelled RNA ($L$) is kept constant, a graph of $1/F$ against $1/d$ would be a straight line. When the concentration of double-stranded RNA is infinite (i.e. when $1/d = 0$), $1/F = 1/H$. Hence, the fraction of the labelled species homologous with the double-stranded RNA ($H$) is the reciprocal of the intercept on the $1/F$ axis.
Common sequences in TRV RNAs

We wish to thank Miss Vicky Maxwell and Mr D. R. Smith for technical assistance, and Professor P. Wildy for helpful discussion and comment.

This work was supported by a grant from the Agricultural Research Council.

REFERENCES


(Received 17 April 1973)