The EMBO Journal vol.12 no.7 pp.2969–2976, 1993

Evolution and replication of tobacco ringspot virus satellite RNA mutants

Christophe Robaglia, George Bruening[1], Jim Haseloff[2] and Wayne L.Gerlach[3]

Laboratoire de Biologie Cellulaire, INRA 78026 Versailles Cedex, France, [1]Department of Plant Pathology, University of California, Davis, CA 95616, USA, [2]Department of Molecular Biology, Massachussets General Hospital, Boston, MA 02114, USA and [3]Division of Plant Industry, CSIRO, PO Box 1600, 2601 Canberra, Australia

Communicated by H.L.Sänger

The replication properties of linker insertion–deletion mutants of tobacco ringspot virus satellite RNA have been studied by amplification in plants infected with the helper virus. Sequence analysis of the cDNAs corresponding to the replicated forms shows that only one of the original mutated molecules replicates unaltered, and in general new variants accumulate. Depending on the location of the original mutation three types of sequence modifications were observed: (i) deletion of the mutated region followed by sequence duplication, (ii) sequence duplication and deletion outside of the mutated region and (iii) limited rearrangements at the site of mutation. The mutant that replicates without sequence changes accumulates linear multimeric forms suggesting that self-cleavage is affected although the sequence alteration does not involve the hammerhead catalytic domain. Alternative RNA conformations are likely to play a role in the origin of this phenotype and in the formation of sequence duplications. These results demonstrate the great structural flexibility of this satellite RNA.

Key words: hammerhead ribozyme/molecular evolution/plant virus/RNA conformation/satellite RNA

Introduction

Satellite RNAs are small RNA molecules that depend on a helper virus for replication and propagation. They are not required by the virus and present no significant sequence homology with the viral genome. They can be viewed as molecular parasites of the virus infection process. Their presence can in some cases be associated with symptom reduction (for a review, see Francki, 1985). These properties have made them the basis of a strategy for virus resistance in plants (Gerlach *et al.*, 1987; Harrison *et al.*, 1987).

The satellite of tobacco ringspot virus (TobRV), a nepovirus, is an RNA molecule of 359 bases predominantly encapsidated as a linear monomer which is defined as (+) polarity (Kiefer *et al.*, 1982). Multimeric forms of (+) and (−) polarity have been observed in infected plants together with (+) and (−) circular monomers (Linthorst and Kaper, 1984). These observations have led to a model for replication where plus linear monomers are first circularized, then transcribed by a rolling circle mechanism into minus multimers which are in turn cleaved and circularized before being transcribed into plus multimers which are subsequently cleaved and encapsidated. This model has been further supported by the finding that both plus and minus monomers can be generated '*in vitro*' by autocatalytic RNA processing of multimeric substrates (Buzayan *et al.*, 1986; Prody *et al.*, 1986). The sequences necessary for plus and minus cleavage have been delimited (Feldstein *et al.*, 1989; Haseloff and Gerlach, 1989; Hampel *et al.*, 1990). The structure for the plus strand cleavage site has been shown to belong to the family of 'hammerhead' type ribozymes found in other plant virus satellites and viroids (for a review see Bruening, 1990), whereas the minus strand cleavage site appears to be a novel structural motif for RNA cleavage (Haseloff and Gerlach, 1989; Hampel *et al.*, 1990).

Despite its small size, the TobRV satellite RNA (sTobRV RNA) is therefore a complex RNA molecule carrying at least five structural features (+ and − strand replicase binding sites, + and − strand cleavage, encapsidation signal) in a highly compact form. Its high rate of replication and the availability of infectious cDNA clones make it a valuable model for RNA evolution studies. We report here that satellite RNA mutants introduced into TobRV-infected plants undergo sequence modifications leading to increasing fitness for replication.

Results

Mutant satellites are amplified upon virus inoculation

A linker scanning mutagenesis of an infectious clone of the TobRV (budblight strain) satellite RNA has been described (Haseloff and Gerlach, 1989). This mutant collection was used to map the domains necessary for self-cleavage on the plus and minus strands. It was shown previously (Gerlach *et al.*, 1987) that when the sTobRV satellite is expressed in both plus and minus orientations from the genome of the plant it can be efficiently amplified by tobacco ringspot virus inoculation. In an effort to characterize regions involved in other biological and biochemical functions of the molecule, selected mutants have been introduced into the genome of tobacco (var. Samsun) plants by *Agrobacterium* mediated transformation. It was anticipated that this would enable poorly replicating satellites to be amplified (see Figure 1B, for a summary of the mutants studied and their positions on the sTobRV molecule).

R1 plants from each transformed line were inoculated with the helper virus. RNA analysis done 4–8 weeks post-inoculation shows that the satellite RNA replicates. In most cases, a band with the same approximate mobility as sTobRV RNA was detected by ethidium bromide staining and its identity was further confirmed by Northern analysis using a sTobRV RNA probe (data not shown).

© Oxford University Press

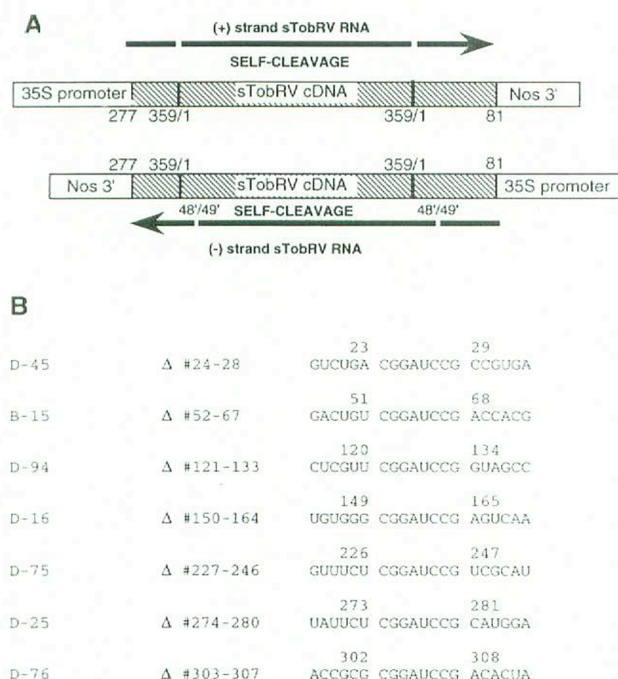

Fig. 1. (**A**) Genes expressing sTobRV RNA from plant genome. Transcription followed by autolytic cleavage produces (+) or (−) monomeric sTobRV RNAs. Numbers correspond to (+) strand, with complementary sequences indicated by prime symbols. (**B**) Details of sTobRV mutants used in this study. Δ indicates that the corresponding bases were deleted. In all cases the deleted bases were replaced by the inserted *Bam*HI linker sequence.

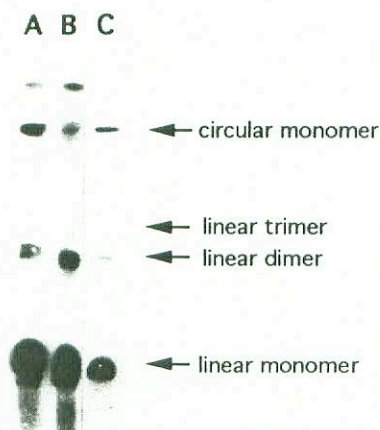

Fig. 2. Northern analysis of satellite RNAs replicating in plants expressing the mutant satellite D-94. (**A**) 2.5 M LiCl soluble RNA extracted from plant 80. (**B**) RNA from plant 82. (**C**) RNA from a plant expressing wild-type sTobRV RNA. RNAs were separated on a 6% acrylamide−50% urea gel. The probe is made of minus strand sTobRV sequences. The arrows point to the various forms of sTobRV observed during replication.

Mutants D-94 and D-16 have undergone deletions and sequence duplications

Northern analysis of the RNAs extracted from plants 80 and 82 expressing minus strand mutant D-94 reveals the presence of multiple bands of approximately linear monomeric sTobRV size (tracks A and B in Figure 2), suggesting the co-replication of several molecules. The satellite RNAs replicating in plant 80 were reverse transcribed using the primer 'Eco+', complementary to 19 bases at the 3′ end of plus strand sTobRV RNA. sTobRV cDNAs were amplified using the same primer and primer 'Pst+' complementary to 23 bases at the 3′ end of the minus strand. The PCR products were cloned and plasmids chosen randomly for sequence analysis. Four types of clones were observed (Figure 3): one (D80.6) contains a deletion of bases 111−132 encompassing the linker insertion. In the other three classes, the same deletion was accompanied by various extents of sequence duplication: bases 98−121 in clones D80.2, 93−113 in clone D80.3 and 109−134 in clone D80.4. A similar analysis performed on another plant also expressing the mutant D-94, allowed two new types of clones to be recovered, D83.1 which is similar to D80.2, and D83.2 which contains a remnant of the linker sequence in a 21 base rearrangement where bases 105−113 and 118−126 form a short direct repeat.

In contrast, in plant 82 only one type of sequence could be recovered in four independent clones in which the sequence UACCUCGUU**CGGAUCCG**GUA encompassing the insertion has been replaced by **GGAUC**GUA where the bases in bold type presumably represent remnants of the linker insertion. This variant may represent the major form of replicating satellite RNA in this plant.

Complementary DNAs were synthesized from RNAs extracted from plants expressing minus strand mutant D-16 and cDNAs were amplified and analysed as above (Figure 4). One class of clones was found to contain a 393 base cDNA insert with the *Bam*HI linker insertion maintained except for a G to U change at the eighth linker base, a deletion of bases 88−93 and the 84−132 region duplicated (Figure 4, variant E.8). Three other variant sequences deriving from mutant D-16 could be rescued from other plants. The variant E.59 displays the same features as E.8 except for the absence of the nt 84−132 duplication, and the variants J63.B1 and J63.B2 (both isolated from a single plant) have also conserved the original mutated sequence modified on the eighth linker base, J63.B1 harbouring a deletion of nucleotides 88−98 and J63.B2 being similar to the wild-type satellite in this region. The common feature of these four classes of variant satellites is the conservation of the *Bam*HI linker insertion modified in the eighth linker base.

Limited sequence rearrangements at the mutation site in mutants B-15, D-25 and D-76

In a plant expressing the plus strand mutant B-15, two variant molecules were found to be co-replicated, the variant B.5.14 retains the *in vitro* modification with the insertion of an A between the first (C) and second (G) bases of the *Bam*HI linker and with the sequence GUGC (nucleotides 88−91 of wild-type) replaced by CC. The other variant, B.5.15, has reverted to the wild-type satellite in the mutated region but has a tract of five instead of three A residues starting at position 75 (Figure 5A).

Three different molecules showing rearrangements at the site of mutation were isolated after virus inoculation of tobacco plants expressing minus strand mutant D-25 (Figure 5B). Furthermore, in cowpea plants co-inoculated with positive strand mutant D-25 from an *in vitro* transcription reaction and the helper virus, four other molecules were rescued, one of them (H.21) also bearing a duplication of 20 bases at position 65. Thus, a total of seven different sequence variants have been obtained from the mutant D-25,

WILD TYPE SEQUENCE

90 100 110 120 130 140

A C C G U G C U G C G U A G C G U A G G G G U C U G C U A C C U C G U U G G A G G U G G A G A U U G U A G C C U U C G
U G U G G G C G C G G C G G U G

IN VITRO GENERATED MUTANT

A C C G U G C U G C G U A G C G U A G G G G U C U G C U A C C U C G U U **C G G A U C C G** G U A G C C U U C G
U G U G G G C G C G G C G G U G

VARIANT D80.6

A C C G U G C U G C G U A G C G U A G G G G U C U G C C U U C G
U G U G G G C G C G G C G G U G

VARIANT D80.2 (= D83.1)

A C C G U G C U G C G U A G C G U A G G G G U C U G C C U U C G U G U G G G C G U A G G G G U C U G C C U U C G
U G U G G G C G C G G C G G U G

VARIANT D80.3

A C C G U G C U G C G U A G C G U A G G G G U C U G C C U G C G U A G C G U A G G G G U C U G C C U U C G
U G U G G G C G C G G C G G U G

VARIANT D80.4

A C C G U G C U G C G U A G C G U A G G G G U C U G C C U U C G U G U G G G C G C G G C A G U G U G C C U U C G
U G U G G G C G C G G C G G U G

VARIANT D83.2

A C C G U G C U G C G U A G C G U A G G G U U **C G G A U C C G** G U G U U **C G G A U C** G U A G C C U U C G
U G U G G G C G C G G C G G U G

VARIANT D82.1

A C C G U G C U G C G U A G C G U A G G G G U C U G **C** **G G A U C** G U A G C C U U C G
U G U G G G C G C G G C G G U G

Fig. 3. Nucleotide sequence of variant molecules derived from mutant D-94. Plus strand sequences are shown. Bold-typed bases represent *Bam*HI linker derived sequences. The black arrows represent sequence duplication. White arrows represent inverted repeats. Blank spaces define deletions.

WILD TYPE SEQUENCE

90 100 110 120 130 140 150

AACCGUGCUGCGUAGCGUAGGGGUCUGCUACCUCGUUGGAGGUGGAGAUUGUAGCCUUCCGUGUGGGCGCGGCGGUG

IN VITRO GENERATED MUTANT

AACCGUGCUGCGUAGCGUAGGGGUCUGCUACCUCGUUGGAGGUGGAGAUUGUAGCCUUCCGUGUGGGC CGGAUCCG AGUCAA

VARIANT E.8

AACC GCGUAGCGUAGGGGUCUGCUACCUCGUUGGAGGUGGAGAUUAACCGCGUAGCGUAGGGGUCUGCUACCUCGUUGG
AGGUGGAGAUUGUAGCCUUCCGUGUGGGC **CGGAUCCU** AGUCAA

VARIANT E.59

AACC GCGUAGCGUAGGGGUCUGCUACCUCGUUGGAGGUGGAGAUUGUAGCCUUCCGUGUGGGC **CGGAUCCU** AGUCAA

VARIANT J63.B1

AACC GCGUAGGGGUCUGCUACCUCGUUGGAGGUGGAGAUUGUAGCCUUCCGUGUGGGC **CGGAUCCU** AGUCAA

VARIANT J63.A1

AACCGUGCUGCGUAGCGUAGGGGUCUGCUACCUCGUUGGAGGUGGAGAUUGUAGCCUUCCGUGCGGGC **CGGAUCCU** AGUCAA

Fig. 4. Nucleotide sequences of variants derived from mutant D-16. The legends are the same as for Figure 3. The underlined base in the linker is changed from G to U in all sequenced molecules.

none of them leaving the original sequence unaltered. From plants expressing D-76 mutant satellites, two types of variant satellite were rescued both also bearing limited sequence rearrangements at the mutation site (Figure 5C). From plants expressing minus strand mutant D-45, wild-type RNA satellite was always recovered. In D-45, the location of the mutation is between the plus strand cleavage site and the minus strand cleavage site (at position 49 on plus strand coordinates, see Figure 1). This allows wild-type satellite minus strand to excise itself from the primary transcript.

A

WILD TYPE SEQUENCE

50 60 70 80 90 100

GACUGUGAGGUGGCCGAAAGCCACCACGUAAACUAGUGAACCGUGCUGCGUAGCG

IN VITRO GENERATED MUTANT B-15

GACUGU **CGGAUCCG** ACCACGUAAACUAGUGAACCGUGCUGCGUAGCG

VARIANT B.5.14

GACUGU **C**a**GGAUCCG** ACCACGUAAACUAGUGAACC cc UGCGUAGCG

VARIANT B.5.15

GACUGUGAGGUGGCCGAAAGCCAC ACGTAAAaa CUAGUGAACCGUGCUGCGUAG

B

WILD TYPE SEQUENCE

270 280

GGGUAUUCU CAUUCGA CAUGGAAC

IN VITRO GENERATED MUTANT D-25

GGGUAUUCU **CGGAUCCG** CAUGGAAC

VARIANT SEQUENCES RECOVERED FROM TRANSGENIC TOBACCOS

GGGUAUUCU cgaaucug CAUGGAAC

GGGUAUUCU **CGGAUC**UG CAUGGAAC

GGGUAUUCU ca**GAUC**g CAUGGAAC

VARIANT SEQUENCES RECOVERED FROM COWPEAS

GGGUAUUCU **CGGAUC** CAUGGAAC

GGGUAUUCU ca**GAU** CAUGGAAC

GGGUAUUCU **CGGAUCC** CAUGGAAC + deletion of nt 132-134

GGGUAUUCU cauucga CAUGGAAC + 20 bases repeat at nt 65 (H-21)

C

WILD TYPE SEQUENCE

300 310 320

CCGCGC CUCU ACACUAUGCGCGG

IN VITRO GENERATED MUTANT D-76

CCGCG **CGGAUCCG** ACACUAUGCGCGG

VARIANT SEQUENCES

CCGCG **CGGAUCC** ACUAUGCGCGG

CCGCG **CGGAU** ACUAUGCGCGG

CCGCG **CGGAUC** ACUAUGCGCGG

Fig. 5. Nucleotide sequences of variant satellite RNAs recovered from plants expressing mutants B-15 (**A**), D-25 (**B**) and D-76 (**C**). Lower-case letters show sequence modifications following replication. Bold-typed bases represent sequences derived from the *Bam*HI linker insertion.

During the course of these experiments the wild-type molecule was frequently isolated from plants expressing the original mutant. Given the highly infectious nature of satellite RNAs and the fact that plants expressing wild-type sTobRV were present in our experiments, it is difficult to determine whether the appearance of non-mutated sequences results from reversion events or from contamination through transmission from plant to plant in the glasshouse. Nevertheless, we could not recover the satellite RNA in any form from two transgenic plants expressing the plus strand mutant D-45 within 3 weeks post-infection with TobRV. One plant was found eventually to contain a wild-type sTobRV 5 weeks post-infection. Therefore, it is very likely that the D-45 mutant was unable to generate a molecule capable of replication, as might be expected since this mutation completely suppresses *in vitro* cleavage activity (Haseloff and Gerlach, 1989).

Variant satellites replicate stably during further passage in cowpeas

In an attempt to study the stability during further passage of the rescued variant satellites, we designed a pair of primers which, after a PCR amplification step, will add at the ends of a monomeric satellite cDNA the sequences necessary for self-cleavage and the T7 promoter sequence for *in vitro* transcription (Figure 6). Single stranded cDNA clones of variants E.59, E.8, D80.2, D80.3 and D82.1 were used as templates and the resulting PCR products were cloned. Biologically active satellite RNAs were generated by *in vitro* transcription and co-inoculated on young cowpea plants with the helper virus. Northern analysis done 10 days post-inoculation shows that molecules with the expected mobility of the inoculated variant satellites are replicated (Figure 7). Subsequent cloning and sequence analysis reveals that they had not been further modified (not shown).

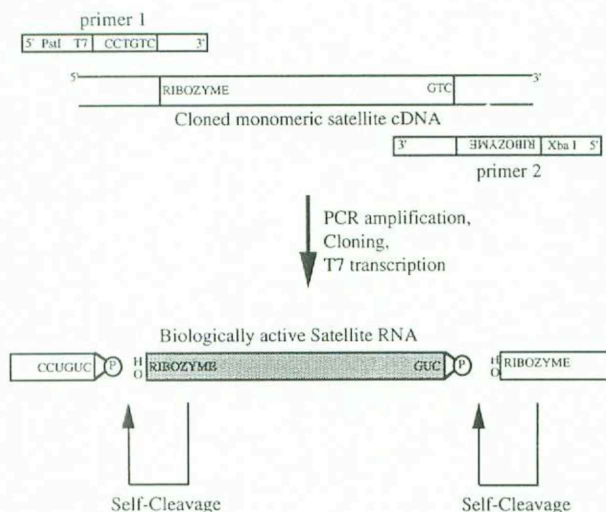

Fig. 6. PCR engineering of biologically active sTobRV RNAs from cloned monomeric cDNA copies. Primer 1 contains the T7 promoter for transcription *in vitro* and bases which, upon their incorporation into RNA and interaction with the adjacent satellite sequences, allow the formation of an active hammerhead ribozyme conformation. Primer 2 inserts into the transcribed sequences the bases 1–46 of sTobRV RNA resulting in self-cleavage at the 3′ end of the amplified molecule. P denotes the 2′–3′ cyclic phosphate generated during self-cleavage (Prody *et al.*, 1986).

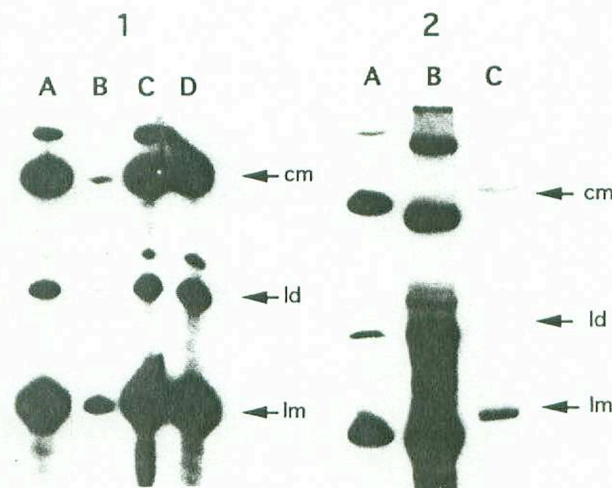

Fig. 7. Northern analysis (6% acrylamide–50% urea gel) of RNAs generated after co-inoculation of variant satellite RNAs and TobRV on cowpea plants. Biologically active satellite RNAs were generated as described in Figure 6 and co-inoculated on young plants with the helper TobRV. 2.5 M LiCl soluble RNA was extracted as in Materials and methods. The probe is the same as in Figure 2. Panel 1, track A, variant D82.1; track B, variant D80.3; track C, D80.2; track D, wild-type sTobRV RNA. Panel 2, track A, wild-type satellite RNA; track B, variant E.59; track C, variant E.8. lm, linear monomer; ld, linear dimer; cm, circular monomer.

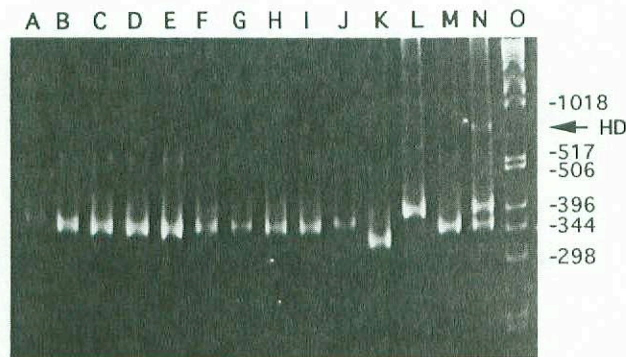

Fig. 8. Heteroduplex gel shift analysis of RT-PCR products of plants co-inoculated with different satellite RNAs and the helper virus. The following templates were used for amplification using primers Eco+ and Pst−, submitted to HGSA as described in Materials and methods and separated on 5% acrylamide non-denaturing gel. (**A**) Total RNA from a plant inoculated with wild-type sTobRV; (**B−D**) RNAs from three independent plants each co-inoculated with equal amounts of wild-type sTobRV (transcribed from the TTS-B plasmid, Haseloff and Gerlach, 1989) and variant D82.1 produced as described in Figure 6; (**E−G**) RNAs from three plants co-inoculated with wild-type sTobRV and variant D80.6; (**H−J**) RNAs from three plants co-inoculated with wt sTobRV and variant E.8; (**K**) 10 ng of cloned variant D80.6 cDNA; (**L**) 10 ng of cloned variant E.8 cDNA; (**M**) 10 ng of cloned variant D80.2 cDNA; (**N**) the amplification products of L and M were mixed in equal amounts and submitted to one cycle of denaturation/renaturation. The HD arrow points to the shifted heteroduplex band, and (**O**) BRL 1 kb ladder, sizes are in bp.

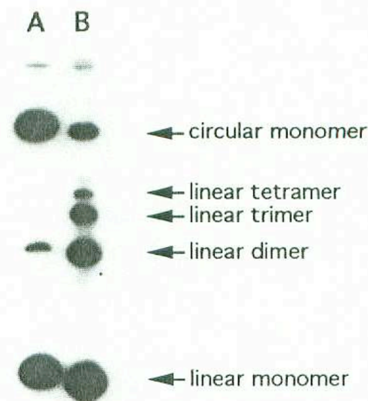

Fig. 9. Northern analysis (6% acrylamide−50% urea gel) of satellite RNAs replicating in plants expressing wild-type sTobRV (**A**) and mutant D-75 (**B**). The probe was the same as in Figure 3. Arrows show the different forms of replicating sTobRV RNA.

Competition experiments between wild-type and variant satellite RNAs

In order to determine the competitive value of some of the newly isolated sTobRV RNAs we devised a competition experiment between these variants and the wild-type sTobRV RNA. Cloned variants E.8, D80.2 and D80.6 were transcribed *in vitro* as described above and co-inoculated with an equal amount of wild-type satellite transcribed from the plasmid TTS-B and the helper virus on young cowpea plants. Three plants were used for each combination. RNA was extracted at 15 days post-inoculation. cDNAs were synthesized and amplified by PCR. The PCR products were analysed by heteroduplex gel shift assay (HGSA, Delwart and Mullins, 1991), the results are shown in Figure 8. This assay allowed us to identify those PCR products containing identical molecules or heterologous mixtures. It is particularly sensitive to the presence of insertions or deletions (E.Delwart, personal communication). In a control experiment, cloned satellites D80.6 and D80.2 were separately amplified, then mixed in equimolar amounts and submitted to the assay. This demonstrated the presence of a third heteroduplex band of slower mobility (Figure 8, lane N). Using RNAs from a co-inoculated plant, only a prominent band comigrating with wild-type satellite RNA was observed (Figure 8, lanes B−J). Thus, in all cases the wild-type satellite RNA was the predominant replicating molecule.

D-75 mutant replicates unaltered

Northern analysis of RNA from plants expressing the D-75 mutant satellite RNA using a plus strand specific probe shows an increased accumulation of multimeric forms and a decreased amount of circular satellite when compared to the replication pattern of the wild-type molecule (compare lanes A and B in Figure 9). The increased stability of uncleaved precursors strongly suggests that self-cleavage of the plus strand was less efficient in these molecules whereas the circularization of linear forms may be altered. Sequence analysis of PCR amplified cDNA clones reveals that in all cases the original D-75 mutant replicates unaltered (not shown).

Discussion

Sequences outside the catalytic domain can modify cleavage efficiency in vivo but not in vitro

In the mutant D-75 the nucleotides 227−246 of the sTobRV RNA were deleted and replaced by *Bam*HI linker sequences. The altered bases are not known to be involved in the self-cleaving domain, but can form three short RNA helices with the nucleotides 20−39 belonging to the active part of the plus strand hammerhead catalytic domain (Figure 10, WT). In the mutant molecule, the bases 20−28 which are an integral part of the plus strand hammerhead ribozyme structure (Forster and Symons, 1987a,b; Haseloff and Gerlach, 1988), can engage in a 9 bp structure with bases 221−229, the nucleotides 227−229 being provided by the *Bam*HI linker insertion (Figure 10, D-75). Such a strong interaction cannot be predicted from the most stable structure of the wild-type molecule (Figure 10, WT) and can induce an alternative RNA conformation destabilizing the catalytic domain. Similar observations concerning the role of alternative RNA secondary structures could be made in studies of the self-cleaving domain of lucerne transient streak virus satellite RNA (Forster and Symons, 1987b). Nevertheless, other types of destabilizing effects, like those involving tertiary structural alterations, or a direct action on the mechanism of catalysis, cannot be excluded. More information on the spatial structure of the hammerhead catalytic domain and on the atoms involved in catalysis would be necessary to answer these questions. Interestingly, previous studies did not show an altered *in vitro* cleavage for positive-stranded mutant D-75 (Haseloff and Gerlach, 1989). There is no clear explanation for the difference in behaviour between the *in vivo* and *in vitro* situations, but certain possibilities can be proposed. As the RNA used for

Fig. 10. Partial secondary structures of wild-type sTobRV and mutant D-75 around the positive strand cleavage site. In WT and D-75, self-cleaving sequences are represented in an active conformation whereas the insert shows the active hammerhead conformation. Bold-typed bases are those involved in the catalytic moiety of the hammerhead ribozyme. Thick lines materialize the possible base pairing in D-75 between nucleotides 20–28 and 221–229. Lower-case letters show bases derived from the *Bam*HI linker insertion. Boxed sequences in WT are those deleted and replaced by *Bam*HI linker sequences in D-75.

Fig. 11. Alternative secondary structures of Budblight and 62L isolates of sTobRV RNA. Boxed bases in Budblight satellite RNA sequence show those deleted by linker mutagenesis and replaced by *Bam*HI linker derived sequences CGGAUCCG.

the *in vitro* studies was only a partial multimer of sTobRV in which the altered sequence was present in only one copy (Figure 1A; Haseloff and Gerlach, 1989), the inhibiting effect observed *in vivo* may result from an interaction occurring either in a complete dimeric (or multimeric) precursor or between two separate molecules during replication. Intermolecular base pairing between structural elements of the self-cleaving hammerhead ribozymes has been shown to occur in avocado sunblotch viroid and newt catalytic RNAs (Forster *et al.*, 1988; Epstein and Pabon-Pena, 1991). Another possibility is that the mutation will modify a positive interaction occurring with a host or viral component. The decreased accumulation of circular forms in D-75 replicating satellite suggests that the ligation process is also affected. The activity responsible for the circularization of plus strand linear sToBRV has not been identified. Potential candidates are enzymes thought to be involved in the ligation of the exons during tRNA splicing (Kiberstis *et al.*, 1985). These enzymes have been characterized in wheat germ (Konarska *et al.*, 1981) and yeast (Phizicki *et al.*, 1986). Quantitative comparison of ligation efficiency for the wheat germ enzyme has not been performed and other plant enzymes like those from tobacco, in which our experiments were done, have not been studied. Some level of specificity may exist in plant cells and the satellite RNAs must have evolved to be correct substrates for these RNA ligating activities. The mutation introduced into the sToBRV D-75 molecule may thus modify the sequences and/or structures necessary for this circularization.

sTobRV can support sequence changes and still be efficiently replicated

These results reveal a flexible nature for the sTobRV molecule, since most of the mutants tested give rise to multiple variants competent for replication. The fact that the original mutant was never recovered, except in one case, shows that they have undergone a rapid sequence modification leading to the selection of a well adapted molecule. The frequent recovery of multiple variants of the same origin (four different sequences out of seven independent clones in plant D-80, seven variants in independent plants for mutant D-25) shows that multiple solutions to the replication constraints can be generated. The isolation of variant H-21 bearing a rearrangement at the site of mutation together with

a 20 base duplication at position 65 suggests also that some level of spontaneous rearrangement in the absence of selection pressure may occur. The same region of the satellite has also been shown to be rearranged in the 62L isolate of the satellite amplified by a different TobRV strain (Buzayan *et al.*, 1987). In the plants studied, the rescued and analysed molecules may represent only a small sample of the total number of different co-replicating satellite RNAs, as suggested by the large hybridizing zone corresponding to linear monomeric satellite in Figure 2 (tracks A and B). However, we were able to recover only one type of variant from plant 82 from the analysis of four clones. It is not clear whether at the time of harvesting (3−8 weeks post-inoculation with tobacco ringspot virus) the selection for replication is still acting or if only equally adapted molecules are surviving. Another possibility would be that variant molecules were continuously generated from mutant satellites expressed from the plant genome.

In competition experiments between wild-type (Budblight strain) satellite and variants E.8, D80.2 and D80.6 the wild-type molecule has always overcome the variant thus demonstrating that it is best adapted to replication in our experimental conditions.

Mechanism of sequence rearrangements and implications for the evolution of satellite RNA

In variants D80.2, D80.3 and D80.4 derived from mutant D-94 in plant 80, the region surrounding the linker insertion is replaced by two different sequence duplications bearing the same borders relative to the wild-type sequence (nucleotides 111−132, see Figure 3). The isolation from the same plant of variant D80.6 containing a deletion precisely encompassing the region of duplication in D80.2, D80.3 and D80.4 strongly suggests that a deletion event occurred first and was then followed by a duplication. Inspection of the sequences adjacent to deletion borders in D-94 shows that bases 110−115 and 128−133 can engage into a 6 bp stem structure (open arrows in Figure 3, the G128 being provided by the linker insertion). This structure is likely to be favoured by the disruption of the 11 bp discontinuous stem involving bases 106−117 and 122−132 in the wild-type molecule, as a result of the linker insertion (Figure 11, top). The replicating enzyme can then skip over the stem structure and resume synthesis in a downstream sequence. A similar model of deletion has been proposed in the case of the appearance of poliovirus defective interfering particles (Kuge *et al.*, 1986).

The deletion and duplication mechanism may involve either a recombination event between two molecules, where the replicase leaves one template together with the nascent RNA chain and reinitiates at another location on a second molecule, or a monomolecular event where the reinitiation occurs on the same template. Clearly, all these models are related in invoking the ability of the polymerase to dissociate from the replicated template and reinitiate RNA synthesis. This property seems to be a common feature of many viral RNA-dependent RNA polymerases (picornavirus, alphavirus, rhabdovirus, carmovirus) and may be induced by specific secondary structures of the RNA genome, as proposed by Romanova *et al.* (1986) and by Cascone *et al.* (1990). The latter authors proposed that replicase driven copy-choice mechanism was involved in the recombination events observed between two linear satellite RNAs replicated by turnip crinkle virus (TCV). These experiments also show that unequal crossing over had occurred at specific regions of sequence similarity between the two satellite RNAs and that, in some cases, non-templated nucleotides were present at the junctions. Sequence comparison between the right sides of the recombinant junctions, junction sites of TCV defective interfering RNAs, and the 5′ ends of TCV, revealed conserved sequence motifs which are proposed to be recognition sites for the viral polymerase. As the duplication events found in mutant sTobRV molecules are clustered to a single region (mutants D-94 and D-16, Figure 11), this suggests that it may contain specific replicase recognition sites.

In this light, it is of particular interest to note that the domain extending from nucleotides 71 to 175, where we observed the occurrence of duplications and rearrangements, can be folded into alternative secondary structures mutually excluding themselves (Figure 11). Moreover, this conformational switch is conserved in the two known isolates of sTobRV RNA, Budblight and 62L (Buzayan *et al.*, 1987). The mutations in D-16 and D-94 are disrupting paired structures favouring the formation of other conformers. Strong stops for replication can then be created, stimulating intra- or intermolecular recombination events. The state of a dynamic equilibrium between alternative secondary structures during RNA replication has already been pointed out by the early studies of Kramer *et al.* (1981). However, these models for alternative conformation do not take into account other types of stabilizing/destabilizing effects such as base stacking in unpaired regions, the formation of base triples between different structural domains or the occurrence of non Watson−Crick base pairs (apart from G-U). All of these structural features have been described in RNA molecules (for a review see Chastain and Tinoco, 1991) and can modify our understanding of dynamic interactions during RNA replication and catalysis.

Recombination and sequence rearrangement have been documented not only in viruses (see above), but also in viroids (Haseloff *et al.*, 1982; Rezaian *et al.*, 1990), linear satellite RNAs (Cascone *et al.*, 1990) and small RNAs capable of being replicated by the $Q\beta$ virus replicase (Munishkin *et al.*, 1988). The latter example demonstrates that the replicase can use host RNA as templates. The work described in this paper shows the rapid and variable nature of RNA variants which can be generated in a biological system if appropriate selection is applied.

Materials and methods

Plasmid construction

The plasmid TTS-B and linker insertion mutant derivatives containing a partial multimer of sTobRV (budblight strain) cDNA have been described (Haseloff and Gerlach, 1989). Selected mutant cDNAs (Figure 1B) were excised with *Hin*dIII and *Eco*RI, blunt-ended and cloned in both orientations in *Sma*I digested plant expression vector pJ35SCN (Walker *et al.*, 1987). The expression cassette was then excised using *Hin*dIII and cloned in the plant transformation vehicle pGA470 (An *et al.*, 1985). The constructions were introduced into the *Agrobacterium tumefaciens* strain LBA4404 (Hoekema *et al.*, 1983) by triparental mating. *Nicotiana tabacum* (cv Samsun) was transformed using a leaf disc transformation procedure (Horsch *et al.*, 1985).

RNA extraction and analysis

Leaf tissue (0.2 g) was ground in a microfuge tube in 0.3 ml of a mixture (1:2) of 2 M Tris, 0.2 M Na_2EDTA, 15% lithium dodecyl sulfate, 10% sodium deoxycholate, 10% Nonidet P-40, 2% β-mercaptoethanol (TE3D buffer) and phenol saturated with 20 mM Na_2EDTA pH 8, 0.4 ml of

chloroform−isoamyl alcohol and 0.25 ml of 3 M ammonium acetate−1 mM Na_2EDTA pH 8 were added. After vortexing and centrifugation, the supernatant (0.2 ml) was mixed with 0.8 ml of 3.6 M LiCl−10 mM Na_2EDTA pH 8 and RNA was allowed to precipitate on ice for 1−12 h, collected by centrifugation, washed in 70% ethanol, dried and resuspended in 20 μl of 0.1 mM Na_2EDTA pH 8.

For Northern analysis, 0.8 ml of the LiCl supernatant was mixed with 0.6 ml of isopropanol and nucleic acids precipitated at −20°C. This procedure was found to give a preparation highly enriched in all forms of sTobRV RNA which precipitate only partially in LiCl.

RNAs were separated on 6% acrylamide (38:2, acryl:bisacrylamide), 50% urea gels in 0.5 × TBE buffer, and electrotransferred in 0.5 × TBE buffer to Hybond N^+ membranes (Amersham). Hybridization to ^{32}P-labelled RNA probes was in 7% SDS, 50% formamide, 0.125 M Na_2HPO_4 pH 7.2, 0.05% Na-pyrophosphate, 100 μg/ml sheared denatured herring sperm DNA, 100 μg/ml heparin and 50 μg/ml tRNA at 60°C. Filters were washed twice in 2 × SSC, 1% SDS at room temperature and three times in 0.2 × SSC, 0.1% SDS at 60°C.

Polymerase chain reaction and cloning of satellite cDNAs

The cDNAs were synthesized from the LiCl precipitate of the extraction protocol described above. RNA (2−5 μl) mixed with 50 ng of primer 5'-CCGGAATTCCGACAGGGTATCGGGCTAGA-3' (Eco+ primer) in 10 μl of 50 mM Bicine−NaOH pH 8.3 at 37°C−50 mM NaCl were boiled for 10 s then annealed for 5 min at 50°C. Ten μl of buffer containing deoxynucleotides (2 mM each), DTT (10 mM), $MgCl_2$ (6 mM), RNasin (Promega, 10 units) and MMLV reverse transcriptase (BRL, 100 units) was added and the reaction placed for 30 min at 37°C. Three to five μl of this reaction was then submitted to 25−35 PCR cycles (45 s at 94°C, 30 s at 50°C and 2 min at 72°C) in 50 μl of 15 mM Tris−HCl, pH 8.8, 2 mM $MgCl_2$, 50 mM KCl, 0.2 mM d(ACGT)TP, 150 ng of each primer (Eco+ and Pst+: 5'-CCGCTGCAGACCGGATGTGCTTTCCGGTCT-GA-3') containing 2.5 units of Taq DNA polymerase (Cetus). After phenol−chloroform extraction and ethanol precipitation, the PCR products were digested by *Eco*RI and *Pst*I, separated on a 5% acrylamide gel, and fragments corresponding to monomer STobRV cDNA were eluted and cloned into pUC118 (Vieira and Messing, 1987). Single stranded DNA was generated by superinfection with the helper phage M13KO7 and sequenced with T7 DNA polymerase (USB or Pharmacia) and fluorescent primers using an Applied Biosystem 370A sequencer, at least two clones were analysed from each selected plant and the complete rescued satellite cDNA was sequenced in each case. In some cases, two independent PCR reactions were performed and their products cloned and analyzed separately to rule out any possibility of *in vitro* mutation.

For engineering biologically active satellites from clones obtained as above, single stranded DNAs or linearized plasmid containing cDNAs from variant satellites were amplified as above for 25 cycles with an annealing temperature of 60°C, using primers 5'-GCCTGCAGAATTAATACGACTCACTAT-AGGGAATTCCCTGTCACCGGATGTGCTTTCCGGTCTGA-3' and 5'-GACTCTAGACCTGTTTCGTCCTCACGGACTCATCAGACC-GGAAAGCACATCCGGTGACAGGGTATCGGCC-3' (primers 1 and 2 in Figure 6). The PCR products were digested with *Pst*I and *Xba*I, gel purified and cloned in pUC118 or 119. After *Xba*I digestion and *in vitro* transcription using T7 RNA polymerase, RNAs were inoculated to young cowpea plants together with a cucumber extract containing the helper tobacco ringspot virus (Gerlach *et al.*, 1986).

Heteroduplex gel shift assay

PCR products from plants replicating sTobRV RNAs were extracted with chloroform, ethanol precipitated, resuspended in 10 μl of water and melted 1 min at 100°C (Delwart and Mullins, 1991). One microliter of 1 M NaCl was added and the samples were then allowed to reanneal for 1 h at 60°C before being separated on a 5% non-denaturing acrylamide gel and stained with ethidium bromide.

Computer analysis

The program FOLD of the Genetics Computer Group package (Devereux *et al.*, 1984) was used for secondary structure prediction, and the graphic representations were done using the program Loop Viewer (Gilbert, 1990).

Acknowledgements

We wish to thank Danny Llewellyn for help in plant transformation and Paul Keese, Mark Young and Peter Waterhouse for stimulating discussions and enthusiasm. We are grateful to Chris Hayes for running fluorescent sequencing gels and to Christian Meyer for help with the computer in Versailles. C.R. was supported by a NATO Fellowship.

References

An,G., Watson,B.D., Stachel,S., Gordon,M.P. and Nester,E.W. (1985) *EMBO J.*, **4**, 277−284.
Bruening,G. (1990) *Methods Enzymol.*, **180**, 546−558.
Buzayan,J.M., Gerlach,W.L. and Bruening,G. (1986) *Nature*, **323**, 349−353.
Buzayan,J.M., McNinch,J.S., Schneider,I.R. and Bruening,G. (1987) *Virology*, **160**, 95−99.
Cascone,P.J., Carpenter,C.D., Xiao Hua Li and Simon,A.E. (1990) *EMBO J.*, **9**, 1709−1715.
Chastain,M. and Tinoco,I. (1991) *Prog. Nucleic Acid Res. Mol. Biol.*, **41**, 131−177.
Delwart,E. and Mullins,J.I. (1991) *Proceedings of the Juan March Conference: Coevolution of Viruses, their Hosts and Vectors, Foundation Juan March*, 9−11 December.
Devereux,J., Haeberli,P. and Smithies,O. (1984) *Nucleic Acids Res.*, **12**, 387−395.
Epstein,L.M. and Pabon-Pena,L.M. (1991) *Nucleic Acids Res.*, **19**, 1699−1705.
Feldstein,P.A., Buzayan,J.M. and Bruening,G. (1989) *Gene*, **82**, 53−61.
Forster,A.C. and Symons,R.H. (1987a) *Cell*, **49**, 211−220.
Forster,A.C. and Symons,R.H. (1987b) *Cell*, **50**, 9−16.
Forster,A.C., Davies,C., Sheldon,C.C., Jeffries,A.C. and Symons,R.H. (1988) *Nature*, **334**, 265−267.
Francki,R.I.B. (1985) *Annu. Rev. Microbiol.*, **39**, 151−174.
Gerlach,W.L., Buzayan,J.M., Schneider,I.R. and Bruening,G. (1986) *Virology*, **151**, 172−185.
Gerlach,W.L., Llewellyn,D. and Haseloff,J. (1987) *Nature*, **328**, 802−805.
Gilbert,D.G. (1990) Program published electronically on the Internet, available via anonymous ftp to iubio.bio.indiana.edu.
Hampel,A., Tritz,R., Hicks,M. and Cruz,P. (1990) *Nucleic Acids Res.*, **18**, 299−304.
Harrison,B.D., Mayo,M.A. and Baulcombe,D.C. (1987) *Nature*, **328**, 799−802.
Haseloff,J. and Gerlach,W.L. (1988) *Nature*, **334**, 585−591.
Haseloff,J. and Gerlach,W.L. (1989) *Gene*, **82**, 43−52.
Haseloff,J., Mohamed,N.A. and Symons,R.H. (1982) *Nature*, **299**, 316−321.
Hoekema,A., Hirsch,P.R., Hooykaas,P.P.J. and Schilperoort,R.A. (1983) *Nature*, **303**, 179−180.
Horsch,R.B., Fry,J.E., Hoffman,N.L., Eichholtz,D., Rogers,S.G. and Fraley,R.T. (1985) *Science*, **227**, 1229−1231.
Kiberstis,P.A., Haseloff,J. and Zimmern,D. (1985) *EMBO J.*, **4**, 817−827.
Kiefer,M.C., Daubert,S.D., Schneider,I.R. and Bruening,G. (1982) *Virology*, **121**, 262−273.
Konarska,M., Filipowicz,W., Domdey,H. and Gross,H.J. (1981) *Nature*, **293**, 112−116.
Kramer,F.R. and Mills,D.R. (1981) *Nucleic Acids Res.*, **9**, 5109−5124.
Kuge,S., Saito,I. and Nomoto,A. (1986) *J. Mol. Biol.*, **192**, 473−487.
Linthorst,H.J.M. and Kaper,J.M. (1984) *Virology*, **137**, 206−210.
Munishkin,A.V., Voronin,L.A. and Chetrerin,A.B. (1988) *Nature*, **333**, 473−475.
Phizicki,E.M., Schwartz,R.C. and Abelson,J. (1986) *J. Biol. Chem.*, **261**, 2978−2986.
Prody,G.A., Bakos,J.T., Buzayan,J.M., Schneider,I.R. and Bruening,G. (1986) *Science*, **231**, 1577−1580.
Rezaian,M.A. (1990) *Nucleic Acids Res.*, **18**, 1813−1818.
Romanova,L.I., Blinov,V.M., Tolskaya,E.A., Viktorova,E.G., Kolesnikova,M.S., Guseva,E.A. and Agol,V.I. (1986) *Virology*, **155**, 202−213.
Vieira,J. and Messing,J. (1987) *Methods Enzymol.*, **153**, 3−11.
Walker,J.C., Howard,E.A., Dennis,E.S. and Peacock,W.J. (1987) *Proc. Natl Acad. Sci. USA*, **84**, 6624−6628.

Received on February 8, 1993; revised on March 18, 1993