Gene, 82 (1989) 43–52
Elsevier

GENE 03139

Sequences required for self-catalysed cleavage of the satellite RNA of tobacco ringspot virus *

(Recombinant DNA; plasmid vector; linker mutagenesis; transcription; RNA enzyme; self-cleavage; pathogenesis; replication; hammerhead domain; viroids)

Jim Haseloff and Wayne L. Gerlach

CSIRO Division of Plant Industry, Canberra, ACT 2601 (Australia)

Received by M. Belfort: 6 February 1989
Accepted: 12 April 1989

SUMMARY

The satellite RNA of tobacco ringspot virus (sTobRV) undergoes self-catalysed cleavage during replication. A plasmid for in vitro expression of sTobRV has been constructed and used to obtain a library of mutagenized sTobRV sequences. Screening of these mutants has allowed precise definition of the sequences required for (+) and (−) strand cleavage. The sequences and RNA structures associated with cleavage of each strand differ markedly. Cleavage of the (+) strand requires those sequences flanking the site for cleavage to form a 'hammerhead' domain, similar to those found in other satellite and viroid RNA. In contrast, cleavage of the (−) strand requires only a small region of 12 nucleotides (nt) at the site of cleavage, and a sequence of 55 nt positioned elsewhere in the molecule. Comparison with a closely related satellite suggests that a novel RNA structure may be involved in (−) strand cleavage.

INTRODUCTION

Satellite RNA of plant viruses are single-strand RNA molecules, which rely upon their helper virus for propagation through replication and encapsidation (see Francki, 1985, for review). They are often small, several hundred nt in length, and share little or no sequence similarity with either plant RNA or the RNA of their helper virus. The satellite RNA of the budblight strain of TobRV has been cloned and sequenced (Buzayan et al., 1986a). It consists of 359 nt and can be found as both linear and circular forms in infected tissues. It is thought that circular forms act as templates in a 'rolling-circle'-type mechanism for RNA transcription, giving rise to concatameric forms of (+) and (−) strand RNA intermediates during replication (Kiefer et al., 1982). The (+) and (−) strand concatamers undergo self-catalysed cleavage, both in vitro (Prody et al., 1986;

Correspondence to: Dr. W. Gerlach, CSIRO Division of Plant Industry, P.O. Box 1600, Canberra City, ACT 2601 (Australia) Tel. (062)465204; Fax (062)473785.

* Presented at the Albany Conference on 'RNA: Catalysis, Splicing, Evolution', Rensselaerville, NY (U.S.A.) 22–25 September, 1988.

Abbreviations: ArMV, arabis mosaic virus; bp, base pair(s); LMP, low melting point; nt, nucleotide(s); oligo, oligodeoxyribonucleotide; PEG, polyethylene glycol; sArMV, satellite RNA of ArMV; sTobRV, satellite RNA of TobRV; TobRV, tobacco ringspot virus.

Buzayan et al., 1986b) and in vivo (Gerlach et al., 1987), to produce unit-length 359-nt forms.

Self-catalysed cleavage of the (+) strand sTobRV requires only a neutral or higher pH and, in the presence of a suitable cation such as Mg^{2+}, occurs after the sequence GUC and results in the production of termini containing 5′ hydroxyl and 2′, 3′ cyclic phosphodiester groups (Prody et al., 1986). Similar self-cleavage reactions have been observed in a number of other plant-virus satellite RNAs (Forster and Symons, 1987a,b), a viroid (Hutchins et al., 1986) and a cellular RNA transcript in newt (Epstein and Gall, 1987). In each case a conserved-sequence and secondary-structure motif has been observed to be closely associated with the site for cleavage. This motif, which consists of three branched RNA helices bounding the susceptible phosphodiester bond and regions of highly conserved nt sequence, was originally termed the 'hammerhead' structure (Forster and Symons, 1987a), in view of its conserved base-pairing. However, it is likely that the conserved sequences around the site for cleavage adopt some complex tertiary structure, which allows reactive groups to be brought into close proximity to the susceptible bond, to catalyse cleavage.

In contrast, efficient cleavage of the (−) strand of sTobRV occurs after the sequence ACA, in the apparent absence of the conserved motif found in other RNA, and little is known of the sequences required for catalysis. In addition, the reverse of the reaction, spontaneous ligation of cleaved sTobRV (−) strand, has been readily observed in vitro (Buzayan et al., 1986b).

The presence of satellite RNA during TobRV infection leads to a marked amelioration of symptom development in host plants (Schneider, 1977; Gerlach et al., 1986; 1987). As the satellite RNA does not appear to encode functional polypeptide products, any functions or signals required for its propagation and effect on viral pathogenesis must reside solely within its small RNA genome. We have undertaken a reverse genetics approach to define regions of the sTobRV genome associated with its biological and molecular properties. In this paper we describe the in vitro mutagenesis of sTobRV and the use of these mutants to map sequences required for the self-cleavage of (+) and (−) strands of sTobRV.

MATERIALS AND METHODS

(a) Construction of a vector for in vitro expression of sTobRV

The 163-bp *Taq*I-*Spe*I fragment of sTobRV cDNA was isolated from pSP653 (Gerlach et al., 1986) and ligated to *Acc*I + *Xba*I-cut, phosphatase-treated pGEM4 to reconstitute the *Acc*I site. This plasmid was linearized with *Acc*I, phosphatase-treated and a 359-bp *Taq*I fragment of the sTobRV cDNA was inserted. The resulting clones were screened for the presence of a circularly permuted 522-bp sTobRV cDNA sequence containing the terminally redundant nt 277 to 81 (pTTS, see Fig. 1). The sTobRV sequence was flanked by promoters for T7 and SP6 RNA polymerases, and in vitro transcription gave rise to RNAs of (+) or (−) orientation, respectively, each of which contained two sites for self-cleavage (Buzayan et al., 1986a,b).

(b) In vitro mutagenesis

The plasmid pTTS (50 mg) was linearized with *Bam*HI, treated with S1 nuclease and religated, to remove a unique *Bam*HI site. This construction, pTTS-B, was treated with 2×10^{-4} units DNase 1 in 20 mM Tris · HCl pH 7.0/15 mM $MnCl_2$ for 10 min at 37 °C. The linearized DNA was trimmed and end-filled using T4 DNA polymerase and purified by 0.7% LMT agarose gel electrophoresis and extraction. Kinased *Bam*HI linker sequences (pCGGATCCG) were ligated to the linearized plasmid overnight at room temperature in the presence of 5% PEG. The reactions were *Bam*HI-treated, and the linear plasmid DNA re-purified by 0.7% LMP agarose gel electrophoresis (this was necessary to remove last traces of circular plasmid, together with unligated linkers). Plasmids were recircularized using T4 DNA ligase and transformed into *Escherichia coli* DH-1. Colonies (more than 1000) were scraped from agar plates, grown in liquid culture to saturation and a mixed population of plasmid DNA was prepared. The mixed sTobRV cDNA inserts were excised by restriction enzyme digestion at flanking *Eco*RI and *Pst*I sites, purified by 1% LMP agarose gel electrophoresis, and subcloned into *Eco*RI + *Pst*I-cut, phosphatase-treated pGEM 4.

The resulting transformants were again pooled, grown in liquid culture and plasmid DNA was prepared. The plasmid DNA was treated with *Bam*HI, to cleave only those plasmids containing a *Bam*HI linker sequence, and the linear forms were again purified by two rounds of 0.7% LMP agarose gel electrophoresis, recircularized with T4 DNA ligase, and transformed into *E. coli* DH-1. Individual transformants were screened for the approximate position of the inserted *Bam*HI linker within the sTobRV sequence by restriction enzyme digestion, subcloned into M13mp19 and sequenced using the dideoxy chain-termination technique. Details of characterized mutants are shown in Table I.

(c) In vitro RNA transcription

For transcription of (+)strand sTobRV sequences, in vitro mutagenized plasmids were linearized with *Eco*RI and transcribed with T7 RNA polymerase. Correspondingly, (−) strand transcripts were produced using *Hin*dIII linearized plasmid and SP6 RNA polymerase. Transcriptions were for 2 h with 200 units/ml of T7 or SP6 RNA polymerase in 50 mM Tris · HCl pH 7.5/10 mM NaCl/6 mM $MgCl_2$/2 mM spermidine/1000 units per ml RNasin/500 mM ATP, CTP and GTP/100 mM [α-^{32}P]UTP. Radiolabelled RNA was analysed by 7 M urea/5% polyacrylamide gel electrophoresis. Efficient cleavage of sTobRV transcripts was observed under the conditions of in vitro transcription.

RESULTS AND DISCUSSION

(a) In vitro mutagenesis of sTobRV

To provide a substrate for the in vitro mutagenesis and expression of sTobRV sequences, a plasmid vector was constructed which contained a permuted, larger than unit-length copy of the satellite sequence (Fig. 1). The inserted sTobRV sequence contained a terminal sequence duplication of 163 bp, corresponding to a convenient *Taq*I-*Spe*I cDNA fragment. This duplicated sequence encompassed the known sites for self-cleavage of both (+) and (−) strand RNA (Buzayan et al., 1986b,c). Transcrip-

TABLE I

Details of mutant sTobRV sequences

Mutagenized sTobRV sequences were subcloned into M13mp19 and sequenced using the dideoxynucleotide chain-termination technique. All mutants contain a single inserted *Bam*HI linker sequence, and its position together with associated sequence deletion (D), insertion (I) or repetition (R) are indicated with their position numbered as for sTobRV RNA (Buzayan et al., 1986a). Repeated bases in mutants containing repetition (R) are shown in bold face.

Mutant	Position	Sequence alteration.
D-24	Δ #9-31	CGGAUG^{8} CGGAUCCG^{32} UGAGGA
D-69	Δ #17-49	CUUUCC^{16} CGGAUCCG^{50} GUCAGG
D-45	Δ #24-28	GUCUGA^{23} CGGAUCCG^{29} CCGUGA
D-30	Δ #25-36	UCUGAU^{24} CGGAUCCG^{37} ACGAAA
D-44	R #40	AGGACG^{39} **ACGGAUCCGA**^{41} AACAGG
D-41	Δ #50-53	AGGACU^{49} CGGAUCCG^{54} GGUGGC
B-15	Δ #52-67	GACUGU^{51} CCGAUCCG^{68} ACCACG
D-73	Δ #54-75	CUGUCA^{53} CCGAUCCG^{76} AACUAU
D-99	Δ #56-92	GUCAGG^{55} CCGAUCCG^{93} GCGUAG
D-17	R #72-79	CCACCA^{71} **CGUAAACU**CGGAUCCG**CGUAAACU**^{80} AGUGAA
D-52	R #74-79	ACCACG^{73} **UAAACU**CGGAUCCG**UAAACU**^{80} AGUGAA
D-70	R #103	AGCGUA^{102} GCGGAUCCGG^{104} GGGGUC
D-97	R #109-111	GGGGUC^{108} **UGC**CGGAUCCG**UGC**^{112} UACCUC
D-94	Δ #121-133	CUCGUU^{120} CGGAUCCG^{134} GUAGCC
D-43	Δ #148-164	CGUGUG^{147} CGGAUCCG^{165} AGUCAA
D-16	Δ #150-164	UGUGGG^{149} CGGAUCCG^{165} AGUCAA
D-101	Δ #161-186	CGGUGU^{160} CGGAUCCG^{187} UACCAC
D-28	R #173-183	UCAAGG^{172} **CGUACCAGGUA**CGGAUCCG**CGUACCAGGUA**^{184} AUAUAC
D-60	Δ #188-224	UAAUAU^{187} CGGAUCCG^{225} GUUGUG
D-86	R #191	UAUACC^{190} **ACGGAUCCGA**^{192} CAACGU
D-71	R #206-207	GUUUCU^{205} **CUCGGAUCCGCU**^{208} GGUUGA
D-75	Δ #227-246	GUUUCU^{226} CGGAUCCG^{247} UCGCAU
D-34	Δ #228-254	UUUGUU^{227} CGGAUCCG^{255} AGCGGC
D-22	Δ #242-299	GGUUCC^{241} CGGAUCCG^{300} GCGCCU
D-47	Δ #261-271	AGCGGC^{260} CGGAUCCG^{272} CUCAUU
D-25	Δ #274-280	UAUUCU^{273} CGGAUCCG^{281} CAUGGA
D-102	R #281-289	AUUCGA^{280} **CAUGGAAGU**CGGAUCCG**CAUGGAAGU**^{290} UUGAGA
D-88	Δ #281-298	AUUCGA^{280} CGGAUCCG^{299} CGCGCC
D-50	Δ #283-302	UCGACA^{282} CGGAUCCG^{303} CCUCUA
D-9	Δ #294-319	GUUUGA^{293} CGGAUCCG^{320} GCCGGG
D-76	Δ #303-307	ACCGCG^{302} CGGAUCCG^{308} ACACUA
D-66	Δ #304-307	CGCGGC^{303} CGGAUCCG^{308} ACACTA
D-27	I	CCGCGC^{303} **UCCGGAUCCGC**^{304} CUCUAC
D-79	Δ #311-344	UCUACA^{310} CGGAUCCG^{345} GCCCGA
D-13	Δ #334-340	CGAAUC^{333} CGGAUCCG^{341} UCUAGC
D-56	I	CCUGUC^{359} CGGAUCCG^{1} ACCGGA
D-51	R #7-9	ACCGGA^{6} **UGU**CGGAUCCG**UGU**^{10} GCUUUC
D-64	Δ #26-40	CUGAUG^{25} CGGAUCCG^{41} AACAGG
D-53	Δ #44-80	CGAAAC^{43} CGGAUCCG$^{(pGEM\ 4)}$ CGCCCG

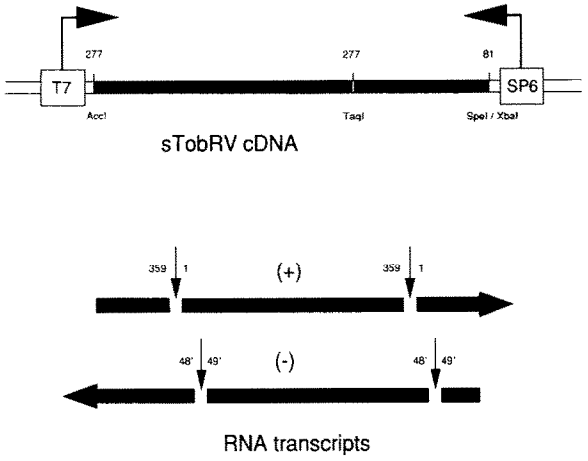

Fig. 1. Vector used for in vitro mutagenesis and expression of sTobRV. A cDNA insert containing a complete 359-nt sTobRV unit, flanked by regions corresponding to nt 277–359 and 1–81 was inserted into pGEM 4, as described in MATERIALS AND METHODS, section **a**. In vitro transcription with T7 or SP6 RNA polymerase yielded (+) or (−) strand RNA, respectively, as shown. Sites for self-catalysed cleavage of transcripts are arrowed. All sequence numbering refers to the corresponding nt in the (+) strand transcripts, with complementary sequences indicated by the prime symbols.

tion of sTobRV sequences by T7 or SP6 RNA polymerases gave (+) and (−) strand RNA, respectively, which each possessed two sites for self-catalysed cleavage. Subsequent self-catalysed cleavage of the RNA released unit-length, 359-nt sTobRV sequences. The (+) strand forms are indistinguishable from those found in vivo, and have been tested for biological activity (J.H. and W.L.G., unpublished results). In addition, different sized 5′- and 3′-flanking fragments are released as a result of cleavage, and these are diagnostic of cleavage at each of the two sites.

To physically map functional domains within the sTobRV genome, a method was chosen for the random introduction of mutations into the sTobRV sequence. Using a linker oligo mutagenesis protocol (Heffron et al., 1978), plasmid DNA was linearized with DNase 1, treated with T4 DNA polymerase to remove 3′ extensions and end-fill 5′ extensions, ligated to linker oligo and recircularized. The transformants resulted from a variety of linker insertion events, the majority of which involved insertion into non-essential regions of the plasmid other than the sTobRV sequence. To purify the recombinants of interest, a large number were pooled and their various sTobRV inserts were isolated and subcloned into wild-type pGEM4 plasmid. Among this pool of recombinants, only those plasmids carrying a linker sequence inserted into the sTobRV sequences contained a site for cleavage by *Bam*HI (Fig. 2). Thus, *Bam*HI digestion and purification of the resulting linearized plasmid allowed isolation of plasmids carrying mutant sTobRV sequences. The mutant sequences, summarized in Table I, were each characterized by restriction mapping and nt sequence determination. Each mutant contains an inserted 8-bp *Bam*HI linker sequence associated with a dele-

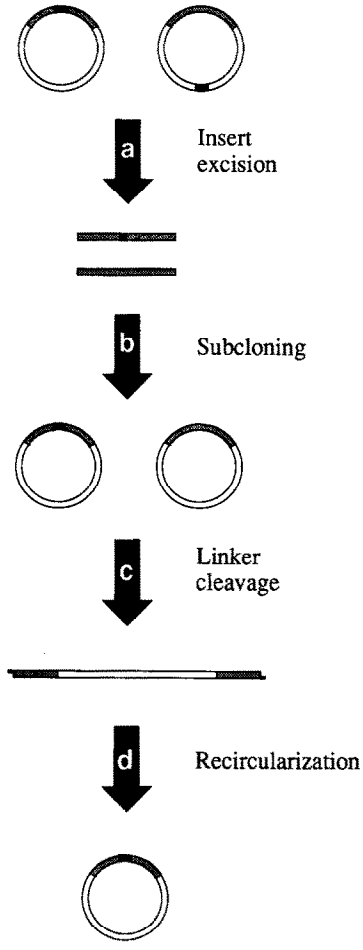

Fig. 2. Scheme for purification of in vitro mutants, as described in MATERIALS AND METHODS, section **b**. (**a**) *E. coli* harbouring various *Bam*HI linker-mutagenized plasmids were pooled, plasmid DNA isolated, and the cDNA inserts excised and purified. (**b**) The inserts were subcloned into unmodified pGEM 4. (**c**) Plasmids were treated with *Bam*HI and linear forms purified. (**d**) The linear species containing mutagenized sTobRV sequences were recircularized prior to transformation. Open segment, vector; shaded segment, insert; blackened segment, linker.

tion or duplication of sTobRV sequences of varying size. The associated alterations in sTobRV sequence can be attributed to the mutagenesis protocol, i.e., the imprecise nature of DNase I cleavage of double-strand DNA even in the presence of Mn^{2+}, and the subsequent use of T4 DNA polymerase to repair the termini. These alterations have proven useful in the mapping of sTobRV sequences required for self-catalysed cleavage.

(b) Mutant (+)strand sTobRV

Mutant sTobRV sequences were transcribed using T7 RNA polymerase to produce (+)strand RNAs, and the transcripts were assayed for their ability to undergo self-catalysed cleavage. The mutants fell into one of three classes depending on whether one or both of the sites for self-cleavage were active. In most cases, alterations to the sTobRV sequence had

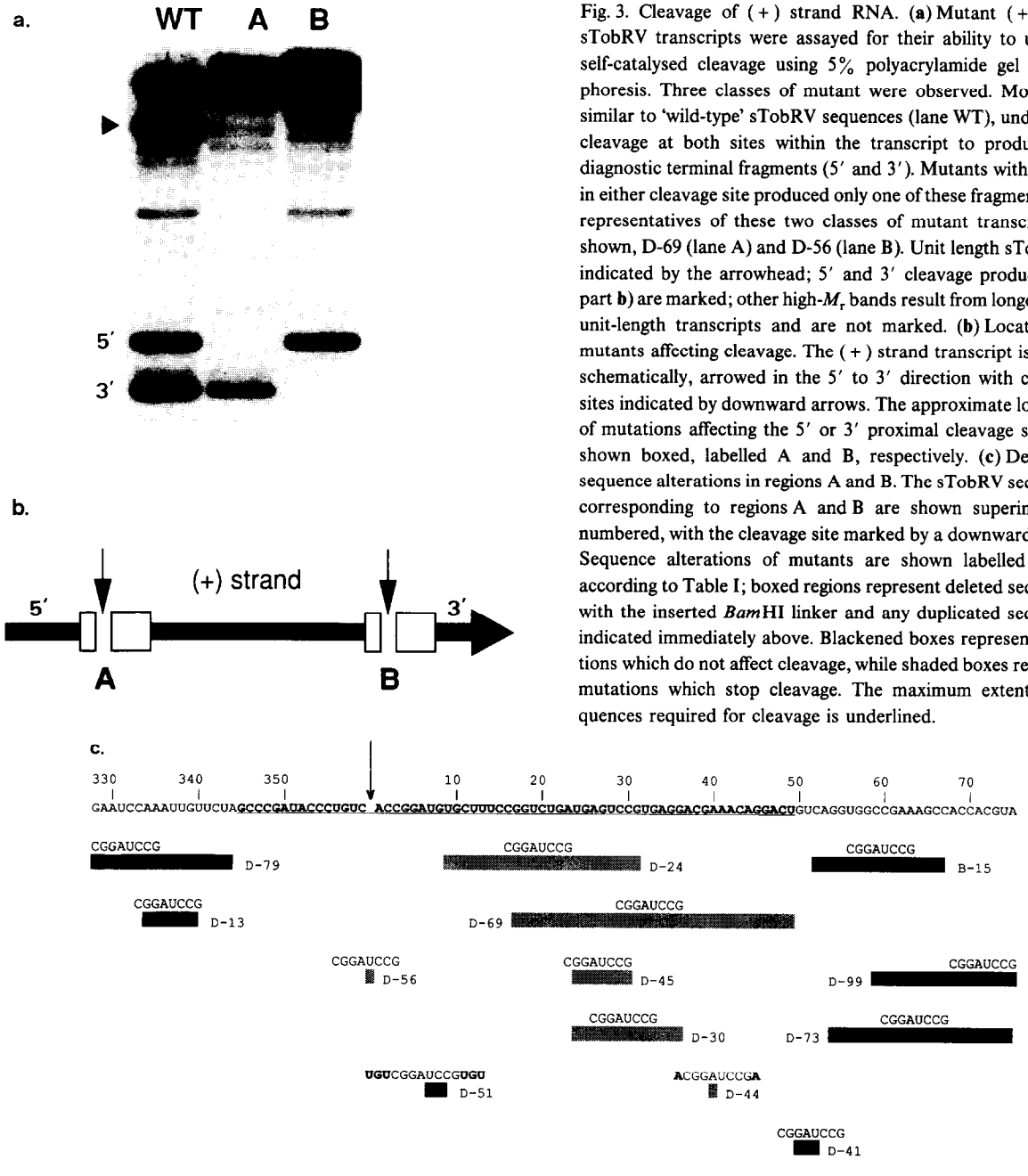

Fig. 3. Cleavage of (+) strand RNA. (**a**) Mutant (+)strand sTobRV transcripts were assayed for their ability to undergo self-catalysed cleavage using 5% polyacrylamide gel electrophoresis. Three classes of mutant were observed. Most were similar to 'wild-type' sTobRV sequences (lane WT), undergoing cleavage at both sites within the transcript to produce two diagnostic terminal fragments (5′ and 3′). Mutants with defects in either cleavage site produced only one of these fragments, and representatives of these two classes of mutant transcript are shown, D-69 (lane A) and D-56 (lane B). Unit length sTobRV is indicated by the arrowhead; 5′ and 3′ cleavage products (see part **b**) are marked; other high-M_r bands result from longer-than-unit-length transcripts and are not marked. (**b**) Locations of mutants affecting cleavage. The (+) strand transcript is shown schematically, arrowed in the 5′ to 3′ direction with cleavage sites indicated by downward arrows. The approximate locations of mutations affecting the 5′ or 3′ proximal cleavage sites are shown boxed, labelled A and B, respectively. (**c**) Details of sequence alterations in regions A and B. The sTobRV sequences corresponding to regions A and B are shown superimposed, numbered, with the cleavage site marked by a downward arrow. Sequence alterations of mutants are shown labelled below, according to Table I; boxed regions represent deleted sequences with the inserted *Bam*HI linker and any duplicated sequences indicated immediately above. Blackened boxes represent mutations which do not affect cleavage, while shaded boxes represent mutations which stop cleavage. The maximum extent of sequences required for cleavage is underlined.

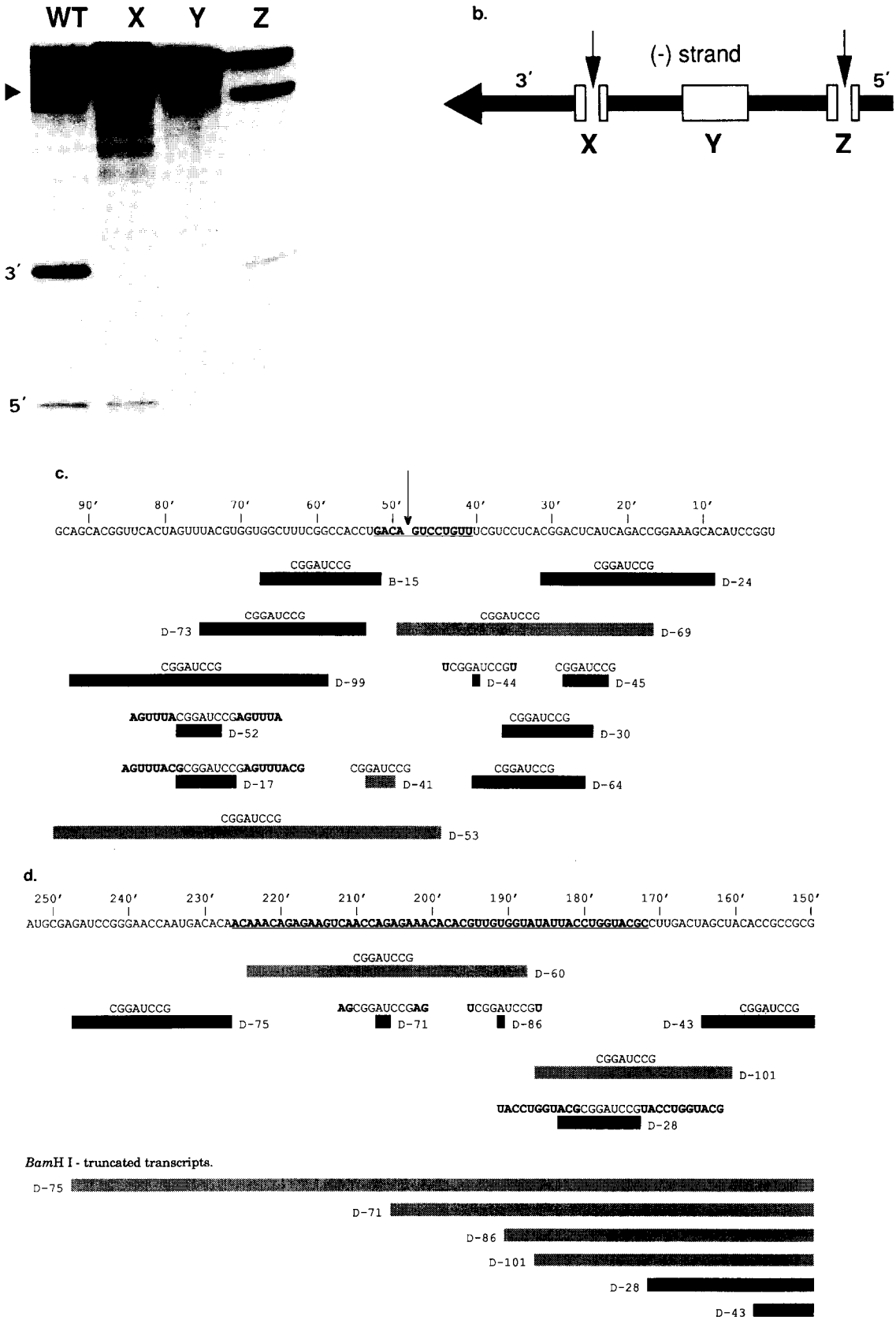
a.
WT
X
Y
Z
3′
5′
b.
(-) strand
3′
5′
X
Y
Z
c.
90′ 80′ 70′ 60′ 50′ 40′ 30′ 20′ 10′
GCAGCACGGUUCACUAGUUUACGUGGUGGCUUUCGGCCACCUGACA GUCCUGUUUCGUCCUCACGGACUCAUCAGACCGGAAAGCACAUCCGGU
CGGAUCCG B-15
CGGAUCCG D-24
CGGAUCCG D-73
CGGAUCCG D-69
CGGAUCCG D-99
UCGGAUCCGU D-44
CGGAUCCG D-45
AGUUUACGGAUCCGAGUUUA D-52
CGGAUCCG D-30
AGUUUACGCGGAUCCGAGUUUACG D-17
CGGAUCCG D-41
CGGAUCCG D-64
CGGAUCCG D-53
d.
250′ 240′ 230′ 220′ 210′ 200′ 190′ 180′ 170′ 160′ 150′
AUGCGAGAUCCGGGAACCAAUGACACAACAAACAGAGAAGUCAACCAGAGAAACACACGUUGUGGUAUAUUACCUGGUACGCCUUGACUAGCUACACCGCCGCG
CGGAUCCG D-60
CGGAUCCG D-75
AGCGGAUCCGAG D-71
UCGGAUCCGU D-86
CGGAUCCG D-43
CGGAUCCG D-101
UACCUGGUACGCGGAUCCGUACCUGGUACG D-28
BamH I - truncated transcripts.
D-75
D-71
D-86
D-101
D-28
D-43

Fig. 4.

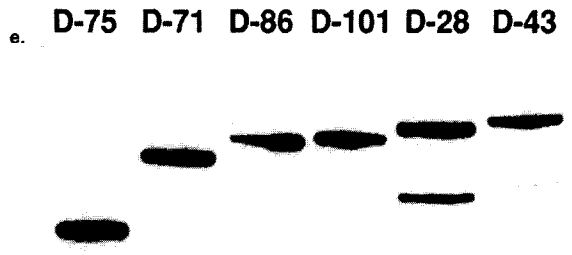

Fig. 4. Cleavage of (−) strand RNA. (**a**) As for Fig. 3, mutant (−) strand sTobRV transcripts were tested for their ability to undergo self-catalysed cleavage. Four classes of mutant were observed. Most were similar to 'wild-type' sTobRV sequences (lane WT), undergoing cleavage at both sites within the transcript to produce two diagnostic terminal fragments (5′ and 3′). Mutants with defects in either cleavage site produced only one of these fragments, and representatives of these two classes of mutant transcript are shown, D-41 (lane X) and D-53 (lane Z). An additional class of mutant, represented by D-101 (lane Y), prevented cleavage at both sites. Unit length (−) strand sTobRV is indicated by the arrowhead; 5′ and 3′ cleavage products (see part **b**, below) are marked; other high-M_r bands result from longer than unit length transcripts and are not marked. (**b**) Locations of mutations affecting cleavage. The (−) strand transcript is shown schematically, arrowed in the 5′ to 3′ direction with cleavage sites indicated by downward arrows. The approximate locations of mutations affecting the 5′ or 3′ proximal cleavage sites are shown boxed, labelled X and Z, respectively, while the location of mutants affecting both cleavage sites is shown labelled Y. (**c**) Details of sequence alterations in regions X and Z. The (−)strand sTobRV sequences corresponding to regions X and Z are shown superimposed, numbered according to the corresponding (+) strand sequence, with the cleavage site marked by the downward arrow. Sequence alterations are indicated as in Fig. 3c. The maximum extent of sequences required for cleavage is underlined. (**d**) Details of sequence alterations in region Y. The (−) strand sTobRV sequences corresponding to region Y is shown numbered according to the corresponding (+) strand sequence. Sequence alterations are indicated as in Fig. 3c, including 3′ truncations caused by transcription of *Bam*HI cleaved mutant plasmids. (**e**) The indicated plasmid DNA, which contained *Bam*HI linker sequences inserted near region Y, were linearized with BamHI, and transcribed using SP6 RNA polymerase to produce a series of 3′ truncated, (−) strand sTobRV RNA. These were assayed for their ability to undergo cleavage at the remaining 5′ site, using 5% polyacrylamide gel electrophoresis. Lanes are designated according to the mutant-bearing plasmid from which inserts were prepared. Only transcripts from D-28 and D-43 showed cleavage, as evidenced by two bands on the autoradiograph; transcripts from other plasmids did not, and only showed one band on the autoradiograph.

no effect on self-cleavage, giving rise to gel patterns similar to those seen with 'wild-type' transcripts, with the production of fragments corresponding to the 359-nt unit-length species and 5′ and 3′ termini (Fig. 3a). However, certain mutant RNAs with sequence alterations adjacent to either of the actual sites for cleavage (Fig. 3b, regions A and B) were defective in cleavage at that site. Thus a number of RNAs which contained altered sequences within the outlined region A did not produce the expected 5′-terminal RNA cleavage fragment. Similarly, RNA with alterations within region B did not produce the expected 3′ fragment. Gel fractionated 'wild-type' and representative mutant RNA transcripts are shown in Fig. 3a. Mutations within regions A and B only affect cleavage at the adjacent site, and each region appears functionally independent (Buzayan et al., 1986c). In contrast, self-cleavage of dimer avocado sunblotch viroid and newt RNA requires interaction of adjacent cleavage sites (Forster et al., 1988).

Regions A and B, which are part of the duplicated sTobRV sequences within the artificial expression vector, are identical in sequence and can be superimposed. The actual sequence of mutations around regions A and B are shown together (Fig. 3c). Those which adversely affect (+) strand cleavage are designated by cross-hatched boxes, and all encroach upon a 64-nt sequence whose extremities are delineated by the active mutants D-79 and D-41. Accordingly, this region is an estimate of the maximum extent of sequences required for (+)strand cleavage, and may include some nt not directly involved in cleavage. Only one mutant (D-51) within the defined regions A and B undergoes cleavage, although at somewhat lower rate than the unmodified sequence (Haseloff and Gerlach, 1988). The remainder of the mutations within these regions essentially eliminate any cleavage at the adjacent site.

Our results demonstrate that self-cleavage of sTobRV (+) strand RNA requires at most 64 nt, which include the susceptible phosphodiester bond. This sequence contains a 52-nt domain of conserved base-pairing and sequence which is similar to that found in other self-cleaving RNA (see Fig. 5a). One of the mutants, D-51, which contains an insertion of 11 nt within this domain, retains cleavage activity. The inserted nt are positioned within a single-strand loop region in the predicted secondary structure (Fig. 5a), and would not be expected to disrupt the

arrangement of conserved nt around the cleavage site. Thus, the observed activity of this mutant provides additional support for the secondary structure model for the self-cleaving domain. Further, we have used the *Bam*HI linker inserted into the D-51 mutant, as a means of splitting the self-cleaving domain into two RNA fragments, which separately possess substrate and catalytic activities (Haseloff and Gerlach, 1988). An RNA fragment containing sTobRV nt 354-9 underwent cleavage at the normal site when mixed in solution with an RNA containing the 'catalytic' sequence, including sTobRV nt 7-46. This observation was used as a basis for the design of RNA enzymes with new substrate specificities, whose activities were based on the highly conserved 'catalytic' sequences of the sTobRV self-cleaving domain.

(c) Mutant (−)strand sTobRV RNAs

Mutant sTobRV sequences were transcribed using SP6 RNA polymerase to produce (−) strand RNA and assayed for the ability to undergo self-cleavage. The mutant RNA fell into four classes. As for the (+) strands, most mutations had no effect on (−) strand cleavage and gave gel patterns similar to 'wild-type' (Fig. 4a), while mutations which altered sequences close to either site for self-cleavage adversely affected cleavage at the adjacent site (Fig. 4b, regions X and Z). In marked contrast, sequence alterations within another region (Fig. 4b, region Y) eliminated cleavage at both sites in the (−)strand RNA.

Regions X and Z contain the sites of (−)strand RNA cleavage, and are identical, duplicated sequences. The positions of mutant sequences around regions X and Z are shown superimposed, in detail (Fig. 4c). The only mutations which had a deleterious effect on RNA cleavage encroached on a 12-nt sequence (nt 52′–41′) spanning the cleavage site. These residues, flanked by the active mutants B-15 and D-64, must contain the sequences required to define a (−) strand cleavage site, and are indicated in Fig. 4c. Unlike the (+) strand RNAs, mutations distant from the actual sites of cleavage also affected the cleavage reaction. Two mutants, D-60 and D-101, which deleted sequences between nt 161′–224′, eliminated cleavage at both sites for (−) strand RNA cleavage (Fig. 4d). Interestingly, several mutants which result in the duplication and insertion of extra sequences in this region were found to have little effect on (−) strand cleavage. To further define these sequences involved in cleavage, plasmids containing mutant sequences were linearized with *Bam*HI before transcription with SP6 RNA polymerase and the resulting 3′-truncated transcripts were assayed for their ability to undergo cleavage at the remaining 5′ cleavage site (Fig. 4e). As a result of these experiments, a 55-nt region was defined which must contain sequences required for cleavage of the (−) strand. Again, the boundaries of this sequence were defined by the positions of the closest flanking mutations (D-75 and D-28), which retain cleavage activity.

For the (−) strand, the sequences required for cleavage consist of, at most, nt 52′–41′ flanking the site of cleavage, and nt 226′–171′. The mutagenized sTobRV contain a site for cleavage at each end of the transcript. Accordingly, there are two copies of the 12-nt domain per RNA, and only one copy of the 55-nt domain required for cleavage. This suggested that the single, larger domain might be acting in a catalytic fashion, possibly being able to participate in one or more cleavage reactions in *trans*. In an accompanying paper, Feldstein et al. (1989) demonstrate that this is so and that RNA fragments containing the 55-nt sequence can act as enzymes to catalyze cleavage of RNA containing the (−) strand cleavage site in *trans*. Our experiments confirm this result (J.H. and W.L.G., unpublished results).

Therefore, the sizes and activities of the two elements involved in (−)strand cleavage approximate those of the substrate and catalytic sequences, which have been artificially separated from the sTobRV (+)strand self-cleaving domain (Haseloff and Gerlach, 1988). One might predict that the two (−)strand elements would together form a conserved motif similar to those found in other self-cleaving RNA. However, little of the required base-pairing and conserved sequences are evident.

(d) A novel structural motif

The comparison of phylogenetically related species has proved useful in identifying secondary structures of probable functional importance for a number of cellular and viral RNA. Recently, the sequence of a 300-nt satellite RNA associated with

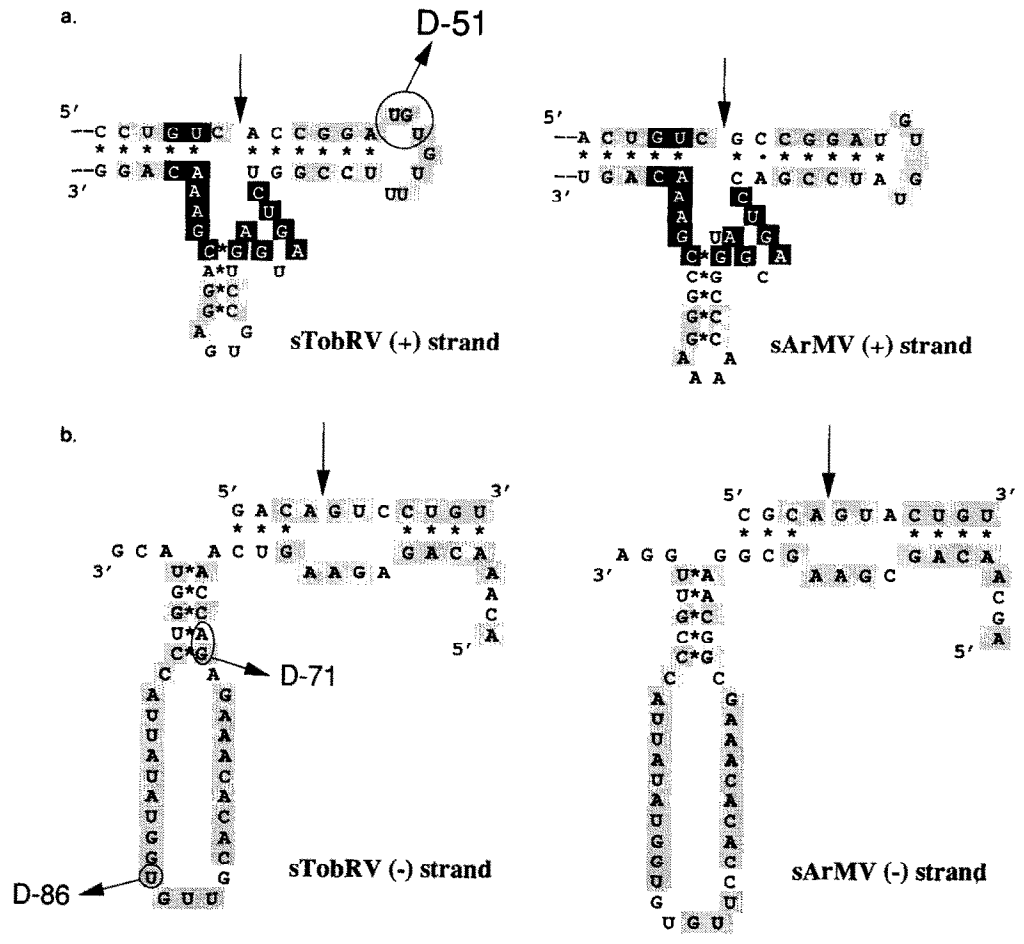

Fig. 5. Comparison of (+) and (−) strand cleavage. (**a**) Comparison of (+) strand cleavage sites from sTobRV and sArMV. Sequences common to sTobRV nt 355–46 and sArMV nt 295–48 are shaded, and those always found conserved in the 'hammerhead' domain (Forster and Symons, 1987a) are reverse-highlighted. Similar secondary structures can be formed in both sequences, and possible base-pairings are indicated by asterisks. In addition, a possible A–C interaction in sArMV is indicated by a heavy dot. The sites for cleavage, known for sTobRV and predicted for sArMV, are indicated by downward arrows, and the site of duplicated sequence and linker insertion for mutant D-51 in sTobRV is circled. (**b**) Comparison of (−) strand cleavage sites from sTobRV and sArMV. The two regions required for (−) strand cleavage of sTobRV (nt 52′–41′ and nt 226′–172′) are shown with the corresponding regions from sArMV (nt 53′–44′ and 159′–104′). Sequences common to sTobRV and sArMV are shaded, and conserved base-pairings are indicated by stars. The site for sTobRV (−) strand cleavage, and that predicted for sArMV are indicated by downward arrows. The sites of duplicated sequence and linker insertion for mutants D-71 and D-86 in sTobRV are circled.

ArMV was reported (Kaper et al., 1988). ArMV is a nepovirus related to TobRV, and its satellite RNA, sArMV, shares 50% sequence similarity with sTobRV. The presence of conserved sequences and potential base-pairing was used to identify the domain likely to be associated with sArMV (+)strand cleavage, which is shown in Fig. 5a. This structure is as shown by Kaper et al. (1988) with the addition of potential bp between nt 1–18 and 7–12, and a possible interaction between nt A and C (2–17). Possible A–C base interactions have been suggested for some tRNA species (Wetzel et al., 1979). Nucleotide sequences which are conserved between sTobRV and sArMV are indicated by shading, and those conserved among other satellite and viroid RNA are reverse-highlighted. Compensating base differences allow both sequences to adopt similar secondary structures and support a functional role for this structure.

A similar comparison can be made between the complementary strand of sArMV and the sequences which we have determined to be required for (−)strand cleavage of sTobRV. Common sequences are indicated by shading in Fig. 5b. Possible base-

pairing is only indicated where it is both conserved between the two satellite sequences and predicted, using an RNA folding algorithm (Freier et al., 1986). The locations of inserted sequences in active mutants D-71 and D-86 are also shown. Neither mutation results in the deletion of sTobRV sequences, and thus need not disrupt existing base-interactions, except by insertion. In each case, the insertion lies adjacent to sequences which differ between sTobRV and sArMV, and does not interfere with conserved base-pairing. As for the consensus (+)strand cleavage sites, the regions of sequence which are most highly conserved between sTobRV and sArMV (−)strand sites appear as unpaired bases in the secondary-structure model. This may be due to their involvement in tertiary interactions required for catalysis, which might impose sequence constraints.

(e) Conclusions

RNA catalysis results from RNA conformation bringing reactive groups into close proximity. A well-recognized secondary-structure motif underlies cleavage of sTobRV (+) strands, and it is likely that this highly conserved structure is directly involved in the catalysis. The RNA structure which we find associated with cleavage of sTobRV (−)strand may similarly play a role in catalysis and comprise a novel structural motif which will be found reiterated in other catalytic RNAs. Any model for the RNA structures involved in (−)strand cleavage must account for the activity of mutants D-71 and D-86, which contain sequences inserted within the 55-nt 'catalytic' region (Fig. 5b). Further manipulation of these mutants may provide additional information on the (−) strand catalysis, as did manipulation of the mutant D-51 for the (+) strand.

ACKNOWLEDGEMENTS

We thank George Bruening, Paul Feldstein and Jamal Buzayan for helpful discussion. J.H. was the recipient of a Queen Elizabeth II Research Fellowship.

REFERENCES

Buzayan, J.M., Gerlach, W.L., Bruening, G., Keese, P. and Gould, A.R.: Nucleotide sequence of satellite tobacco ringspot virus RNA and its relationship to multimeric forms. Virology 151 (1986a) 186–199.

Buzayan, J.M., Gerlach, W.L. and Bruening, G.: Non-enzymatic cleavage and ligation of RNAs complementary to a plant virus satellite RNA. Nature 323 (1986b) 349–353.

Buzayan, J.M., Gerlach, W.L. and Bruening, G.: A subset of the sTobRV sequence is sufficient for self processing. Proc. Natl. Acad. Sci. USA 83 (1986c) 8859–8862.

Epstein, L.M. and Gall, J.G.: Self cleaving transcripts of satellite DNA from the newt. Cell 48 (1987) 535–543.

Feldstein, P., Buzayan, J.M. and Bruening, G.: Two sequences participating in the autolytic processing of satellite tobacco ringspot virus complementary RNA. Gene 82 (1989) 53–61.

Forster, A.C. and Symons, R.H.: Self cleavage of plus and minus RNAs of a virusoid and a structural model for the active sites. Cell 49 (1987a) 211–220.

Forster, A.C. and Symons, R.H.: Self cleavage of virusoid RNA is performed by the proposed 55 nucleotide active site. Cell 50 (1987b) 9–16.

Forster, A.C., Davies, C., Sheldon, C.C., Jeffries, A.C. and Symons, R.H.: Self-cleaving viroid and newt RNAs may only be active as dimers. Nature 334 (1988) 265–267.

Francki, R.I.B.: Plant virus satellites. Annu. Rev. Microbiol. 39 (1985) 151–174.

Freier, S.M., Kierzek, R., Jaeger, J.A., Sugimoto, N., Caruthers, M.H., Neilson, T. and Turner, D.H.: Improved free energy parameters for predictions of RNA duplex stability. Proc. Natl. Acad. Sci. USA 83 (1986) 9373–9377.

Gerlach, W.L., Buzayan, J.M., Schneider, I.R. and Bruening, G.: Satellite tobacco ringspot virus RNA: biological activity of DNA clones and their in vitro transcripts. Virology 151 (1986) 172–185.

Gerlach, W.L., Llewellyn, D. and Haseloff, J.: Construction of a plant disease resistance gene from the satellite RNA of tobacco ringspot virus. Nature 328 (1987) 802–805.

Haseloff, J. and Gerlach, W.L.: Simple RNA enzymes with new and highly specific endoribonuclease activities. Nature 334 (1988) 585–591.

Heffron, F., So, M. and McCarthey, B.J.: In vitro mutagenesis of a circular DNA molecule by using synthetic restriction sites. Proc. Natl. Acad. Sci. USA 75 (1978) 6012–6016.

Hutchins, C.J., Rathjen, P.D., Forster, A.C. and Symons, R.H.: Self cleavage of plus and minus RNA transcripts of avocado sunblotch viroid. Nucleic Acids Res. 14 (1986) 3627–3640.

Kaper, J.M., Tousignant, M.E. and Steger, G.: Nucleotide sequence predicts circularity and self-cleavage of 300-ribonucleotide satellite of arabis mosaic virus. Biochem. Biophys. Res. Commun. 154 (1988) 318–325.

Kiefer, M.C., Daubert, S.D., Schneider, I.R. and Bruening, G.B.: Multimeric forms of satellite tobacco ringspot virus RNA. Virology 121 (1982) 262–273.

Prody, G.A., Bakos, J.T., Buzayan, J.M., Schneider, I.R. and Bruening, G.: Autolytic processing of dimeric plant virus satellite RNA. Science 231 (1986) 1577–1580.

Schneider, I.R.: Defective plant viruses. In Romberger, J.A. (Ed.), Beltsville Symposia on Agricultural Research — Virology in Agriculture. Allenheld, Osmun, NJ, 1977, pp. 201–219.

Wetzel, R., Kohli, J., Altruda, R. and Söll, D.: Identification and nucleotide sequence of the sup8-e UGA-suppressor leucine tRNA from *Schizosaccharomyces pombe*. Mol. Gen. Genet. 172 (1979) 221–228.