The EMBO Journal vol.10 no.3 pp.719-727, 1991

Formation of a thermodynamically metastable structure containing hairpin II is critical for infectivity of potato spindle tuber viroid RNA

P.Loss, M.Schmitz, G.Steger and D.Riesner

Heinrich-Heine-Universität Düsseldorf, Institut für Physikalische Biologie, Universitätsstrasse 1, 4000 Düsseldorf, FRG

Communicated by T.Jovin

The functional relevance of a hairpin II-containing structure of viroid RNA was studied by site-directed mutagenesis, thermodynamic calculations, experimental denaturation curves and infectivity tests. Hairpin II is formed during thermal denaturation of circular viroids or as part of a metastable structure during synthesis of viroid replication intermediates. In potato spindle tuber viroid (PSTVd), eight single-site mutations were generated in the segments which form hairpin II. From the mutated viroid cDNA clones, linear RNA transcripts of PSTVd unit length were synthesized. The relevance of hairpin II for the mechanism of denaturation was confirmed quantitatively by optical denaturation curves and temperature-gradient gel electrophoresis. Infectivity tests showed that the mutations in the core region of hairpin II reverted to the wild type sequence whereas the mutations in the peripheral regions of hairpin II remained genetically stable. These data are in accordance with the natural variance of hairpin II in other viroids of the PSTVd class. Thus, the integrity of the core of hairpin II is critical for infectivity. Hairpin II exhibits a strong similarity in sequence as well as in three-dimensional structure to certain DNA GC-clusters found in the 5'-upstream regions of some genes in man, animals, viruses and plants. A hypothesis about a function of hairpin II as a binding site for host cell transcription factors is proposed.

Key words: site-directed mutagenesis/stability calculations/temperature-gradient gel electrophoresis/transcription factor binding

Introduction

Viroids are plant pathogens distinguished from viruses by the absence of a protein coat and by their small size. They are circular single-stranded RNA molecules consisting of a few hundred nucleotides, the smallest having ~240 and the largest ~600 nucleotides. Obviously viroids only can possess a very limited coding capacity, and there is no experimental evidence for a viroid-encoded translation product; thus, one has to assume that viroid replication and pathogenesis completely depend on the enzyme systems of the host. Thus, their genetic information must be the RNA structure, the ability to undergo particular structural transitions, and the capability to interact with host cell factors (for reviews see Diener, 1979; Sänger, 1982; Riesner and Gross, 1985; Diener, 1987; Riesner and Steger, 1990).

Under native conditions viroids form a rod-like structure which may be described as a serial arrangement of short helices and small internal loops. The secondary structure of the potato spindle tuber viroid (PSTVd) is shown in Figure 1A. During thermal denaturation viroids undergo several structural transitions from the rod-like structure to the single-stranded circle without any intramolecular base pairing. The detailed mechanism of the structural transitions (summarized in Figure 1B) was described more than ten years ago (Henco *et al.*, 1979; Riesner *et al.*, 1979). In a highly cooperative main transition all basepairs of the native structure are disrupted and particularly stable hairpins are newly formed (HP I, II, III). This transition may be viewed as a switch from an extended to a branched structure with a marked loss of base pairing; thermodynamically, the extended structure is destabilized due to competition with the branched structure. At higher temperatures the stable hairpins dissociate independently from each other in the order of their individual thermal stabilities.

Hairpin I is formed in a region of the molecule which shows a strong sequence homology among all viroids of the PSTVd group and a structural similarity even to viroids of other groups; therefore, it is called the central conserved region (Sänger, 1982; Riesner and Gross, 1985). With slight sequence variations hairpin II is present in all viroids of the PSTVd group (Riesner and Gross, 1985). Viroids of this group show an overall sequence homology of 63-81% and can all use tomato plants as a host (Sänger, 1988). Hairpin III was only found in PSTVd and is therefore of marginal interest. Regarding position, length and GC content, hairpin I and hairpin II are more conserved between different viroids than the rest of the molecule. Therefore, it was argued that the ability to form these hairpins originates from biological selection and cannot be merely coincidental. During the last years evidence has accumulated that hairpin I may be involved in the processing of oligomeric replication intermediates to unit length circular viroids (Tabler and Sänger, 1984; Visvader *et al.*, 1985; Diener, 1986; Hecker *et al.*, 1988). For hairpin II a similar functional relevance regarding replication or pathogenesis could not be supported by experiments.

This work concentrates on hairpin II. Hairpin II—as well as hairpin I—is not part of the native structure but is stable only at higher temperature. Thus, any functional role of these hairpins has to be attributed to metastable structures which contain these hairpins even under native conditions. *In vivo*, a transition from the extended structure to the branched (i.e. hairpin containing) structure (cf. Figure 1B) may be facilitated by protein binding. At present, no experimental data are available supporting this possibility. One might also imagine that hairpins are generated as metastable conformations during the synthesis of replication intermediates. The mechanism of viroid replication (Branch *et al.*, 1988; cf. reviews by Branch and Robertson, 1984 and Diener, 1987) and physicochemical studies on replication intermediates lend some support to the second possibility (Hecker *et al.*, 1988). The circular [by definition (+)-strand]

viroid is transcribed into an oligomeric (−)-strand RNA. The (−)-strand acts as template for the synthesis of an oligomeric (+)-strand RNA. Both transcription steps are catalysed by a host enzyme, the DNA-dependent RNA polymerase II (Mühlbach and Sänger, 1979; Rackwitz *et al.*, 1981; Semancik and Harper, 1984; Schindler and Mühlbach, 1990). The (+)-strand oligomeric RNA is cleaved enzymatically to unit length molecules which are then ligated to the mature viroid circles. The process of structure formation during synthesis of linear oligomeric viroid RNA can be followed experimentally (Hecker *et al.*, 1988). A dimeric RNA (+)-strand was transcribed *in vitro* from a viroid cDNA-containing transcription vector, and the structures present immediately after synthesis were analysed in a series of temperature-gradient gel electrophoreses (TGGEs). The results revealed that first a metastable structure was formed which contained the three hairpins (HP I−III), and that in a slow process the transition to more stable structures occurred afterwards. A section of an oligomeric replication intermediate in a metastable structure is shown in Figure 1C. In summary, experimental data presently available suggest that the formation of hairpin II may occur during replication of PSTVd.

Most of the features outlined above hold true quite generally for viroids; there is, however, one exception: avocado sunblotch viroid, which behaves differently with regard to the hairpins and the mechanism of replication and processing (Keese and Symons, 1987).

The method of site-directed mutagenesis has been applied to viroids several times before but the target site was either the virulence modulating region (Owens, 1990; Visvader and Symons, 1986) or the hairpin I-forming region (Visvader *et al.*, 1985) or mutations have been distributed over the whole molecule (Owens *et al.*, 1986). In many cases viroids did not tolerate the mutations introduced, i.e. the majority of the mutations, even single-site exchanges, destroyed viroid infectivity. Mutants that remained infectious generally had the correct native structure as well as an unimpaired hairpin I, demonstrating the functional relevance of these structural elements.

In the present study the effect of mutations in the region of hairpin II was studied with thermodynamic as well as molecular biological and phytopathological methods. A detailed analysis of the linear RNA transcripts from mutant cDNA confirmed the significance of hairpin II in the mechanism of conformational transitions of viroids. In contrast to our expectation all mutants proved to be infectious, i.e. circular viroids could be isolated from tomato plants inoculated with mutagenised cDNA or RNA transcripts. Only some of the mutations, however, were genetically stable in the progeny, whereas the others reverted to the wild type sequence. The discussion of the sites of stable mutations and reversions, the comparison of hairpin II in different viroids, and the search for similar sequences and structures in eukaryotes led us to conclude that a core region of hairpin II is critical for viroid replication and that this structural element may possibly serve as a recognition signal for host transcription factors.

Results

Design and construction of mutants

Selection of the mutation sites. All mutations were introduced in the segments forming hairpin II (HP II; Figure 1A); only one mutation was located adjacent to these segments. The sites of mutations are shown in Figure 1A in the rod-like structure, in Figure 1B in the hairpin II-containing structure. All eight mutants were single-site mutations. It is obvious that these mutations all destabilize hairpin II, except the mutation at position 318 which forms an additional base pair adjacent to the stem. The site and type of nucleotide exchange were selected carefully in order to introduce as few perturbations in the native structure as possible (Figure 1A). Only mutation 318 transforms a mismatch between two helical regions into a base pair. Not only the (+)-strand sequences (Figure 1A,B) but also the (−)-strand sequences (not shown), which are synthesized during replication, were considered when selecting the mutations.

Synthesis of linear RNA-transcripts with single-site mutations. Using standard procedures for site-directed mutagenesis and cloning techniques, the respective mutants were introduced as monomeric and dimeric cDNA clones into the transcription vector pRH701, and from these templates monomeric and dimeric RNA transcripts were synthesized.

Thermodynamic analysis

As mentioned above the aim of the present work was two-fold, to confirm the denaturation mechanism (Figure 1B) and to determine the biological relevance of hairpin II. In view of the first question, the influence of the mutations on the structure and structural transitions of linear monomeric RNA transcripts was analysed. This was achieved by thermodynamic calculations, by optical melting curves and by temperature-gradient gel electrophoresis (TGGE). In the case of infectious mutants the circular progeny-viroids were studied as well.

Analysis by optical melting curves. For optical melting curves the monomeric RNA transcripts were purified by preparative polyacrylamide gel electrophoresis. In order to guarantee identical buffer conditions in the solution of each mutant, all samples were dialysed carefully. All melting curves exhibited one major transition. The curves in Figure 2 show that the T_m value and the shape of the curve of mutant 326 are not much, but significantly, different from that of the wild type. Such characteristics were evaluated for all mutants and will be discussed in connection with the results from TGGE and from thermodynamic calculations.

Analysis by TGGE. The analysis of the transcripts by TGGE had two advantages over the analysis by optical melting curves. Firstly, extracts from the *in vitro* synthesis could be applied to TGGE without further purification because the band of the full length transcript was clearly differentiated from those of fragments, vector DNA and enzymes. Secondly, and even more importantly, two different mutants could be applied to one and the same gel and thereby compared under exactly the same ionic conditions. Therefore, all mutants were analysed together with the wild type. The results depicted in Figure 3 are typical for all curves measured.

In Figure 3A the TGGE of a mixture of mutant 326 and wild type is shown. The main transition of the denaturation results in a drastic retardation of gel electrophoretic mobility. The curve of the wild type transcript is similar to that

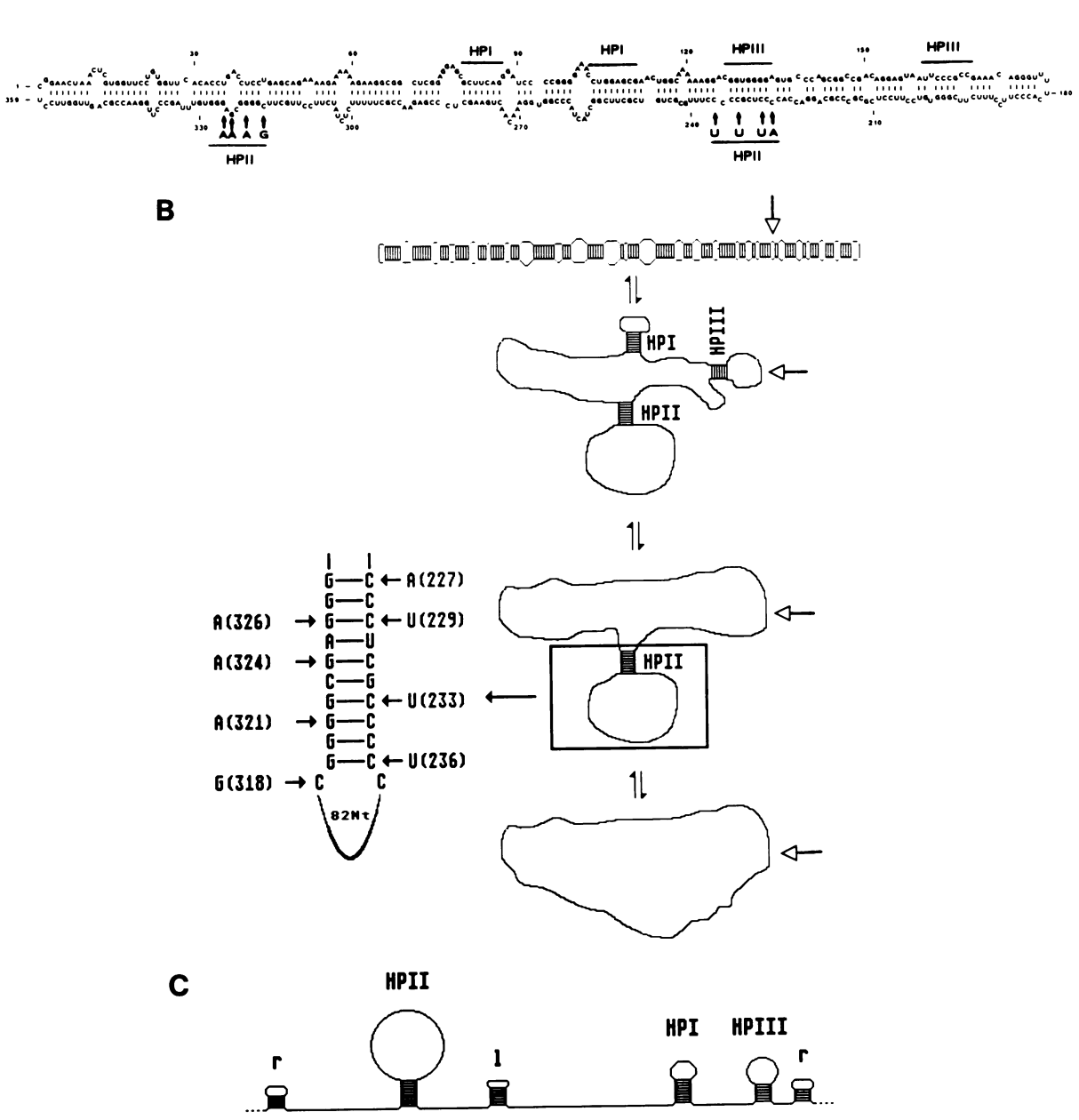

Fig. 1. Secondary structures of the potato spindle tuber viroid (PSTVd). (**A**) native, rod-like structure of circular PSTVd; (**B**) mechanism of denaturation of circular PSTVd; (**C**) section of an oligomeric (+)-strand replication intermediate in a metastable structure directly after synthesis. The hairpin structures I, II and III are designated by HPI, HPII and HPIII, respectively. The heavy arrows in the lower strand of the rod-like structure (A) and the hairpin II structure (B) indicate the nucleotide exchanges. The open arrows indicate the position of 5′ (nt 147)- and 3′ (nt 146)-ends in the linear transcript. The hairpins of the left and right terminal of the native structure are designated by l and r, respectively.

published earlier (Hecker *et al.*, 1988), although a transcript with different start and stop sites was studied previously. The shift of 1.1°C in the T_m value of the mutant 326 can be determined from TGGE with high accuracy (~0.1°C). Identification of the curve of the wild type and that of the mutant in TGGE was carried out in reference experiments where different concentrations of both RNAs were applied.

The transition of the mutant 318 (Figure 3B) is qualitatively different from that of the wild type. The retardation during the main transition is larger in 318 than in the wild type, and the final mobility of the linear molecule without internal basepairing is achieved only after performing a second transition (arrow). As expected for mutation 318, which elongates the stem of hairpin II by one base pair, the denaturation of the hairpin is stabilized and can be detected as an individual transition well resolved from the main transition. Moreover, an additional transition curve is observed for the mutant (318m). It represents a multi-branched conformer which characteristically undergoes a transition into the non-basepaired conformation within the temperature range of the second transition (cf. arrow). All other mutants exhibited transition curves similar to mutant 326 or wild type with slight variations in the steepness of the transition.

The cooperativity of the transition cannot be evaluated quantitatively from the steepness in TGGE, because the relationship between loss of basepairing and change in electrophoretic mobility depends on structural details and is not known. For comparing the thermodynamic behaviour of the different mutants and the wild type we therefore used

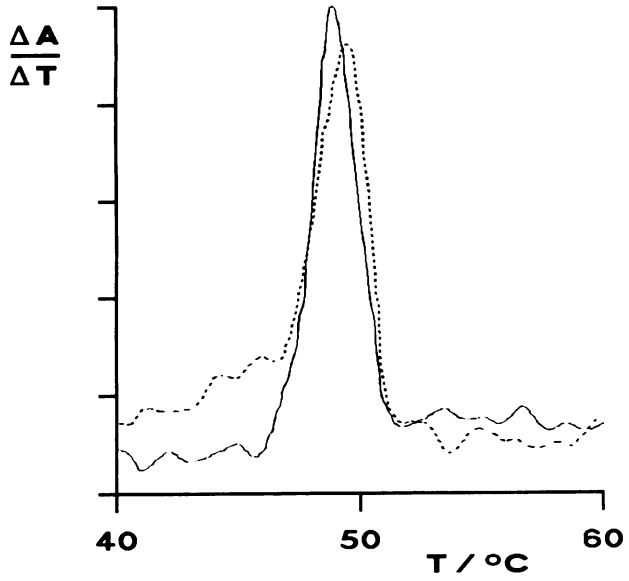

Fig. 2. Differentiated optical melting curves at 260 nm of wild type (——) and mutant 326 (---) linear transcripts. For better comparison curves from different experiments are depicted in the same graph. $\Delta A_{260}/\Delta T$ values are given in arbitrary units. Buffer conditions were 10 mM NaCl, 0.1 mM EDTA, 1 mM Na-cacodylate, pH 6.9; heating rate was 0.2°C/min. Differentiation was performed by linearization over 1°C.

TGGE analysis only to determine the T_m values, whereas the information about the cooperativity of the various transitions was derived from optical melting curves (Table I).

Comparison of experimental results with those from thermodynamic calculations. Optical denaturation curves of mutated and wild type linear transcripts were simulated by taking into account the absorption of the RNA secondary structures in dependence of temperature. In contrast to earlier calculations (Steger *et al.*, 1984) not only the secondary structure of lowest free energy was considered but the whole partition function was approximated by taking into account the 30 most favourable structures at every temperature (M.Schmitz and G.Steger, in preparation).

The calculated denaturation curve of the wild type transcript shows one cooperative main transition and a minor second transition at higher temperature. These transitions correspond to the transformation of the native structure to the branched structure and to the dissociation of the stable hairpins, respectively (see Figure 1B). As mentioned above, formation of hairpin II—and to a lesser extent of hairpin I—couples the denaturation of both halves of the native, rod-like structure which leads to the high cooperativity of the main transition. Furthermore, formation of hairpin II is competing with the native structure and thereby lowering the denaturation temperature of the native structure. Consequently, impairment of hairpin II by mutations leads to a decrease in the cooperativity and an increase in the stability of the main transition. This alteration could indeed be proved by comparing the experimental optical melting curves of wild type and the mutants (see Figure 2). The results are in accordance with the earlier interpretation of the denaturation of circular PSTVd (Riesner *et al.*, 1979).

In the case of the wild type transcript, the calculated midpoint temperature of the main transition is at 40% of total hypochromicity. Therefore, T_m values of all mutated transcripts are given as temperatures of 40% hypochromicity and listed as theoretical T_m values in Table I. The mechanism of denaturation is described in Table I by the feature of whether denaturation of both halves of the rod-like native structure is cooperative (i.e. coupled) or not (i.e. uncoupled).

Comparison of theoretical and experimental T_m values and cooperativity (Table I) shows good agreement. A better quantitative agreement cannot be expected because the experiments were performed at low ionic strengths whereas the calculations refer to 1 M salt. In most cases an increase in T_m is connected with less cooperativity. As expected from theory the situation is different with mutant 318, in which the native structure is stabilized by the mutation (Figure 1A) and the cooperativity is sustained due to the extension of hairpin II by one base pair (Figure 1B). The disagreement between theory and experiment in case of mutant 229 may be due to the fact that this mutation leads to a native structure of the linear transcript which deviates markedly from that of the wild type molecule near the 5′ and 3′ ends. This may result in preferred opening of base pairs from the sequence termini in low ionic strength.

Analysis of circular mutated PSTVd. As described in the next paragraph, infections with mutants 227, 236 and 318 produced stable circular progenies in the host plant. These were isolated and analysed by TGGE. The shifts in the T_m value were similar to those obtained with the linear transcripts. Figure 3C shows the TGGE of a mixture of mutant 318 and wild type PSTVd. The appearance of two different curves is a direct experimental proof of the fact that a mutant RNA is replicating in the infected plant.

Infectivity and genetic stability

Infectivity of the mutants. Tomato seedlings were infected with dimeric transcripts of the wild type and of the eight mutants, and 4 weeks later the plants were tested for viroid infection by bidirectional gel electrophoresis. All mutants except 321 were infectious. Sequence analysis of the complete cDNA clones showed that in clone 321 a second mutation at position 203 (C→U) had been obtained unintentionally. Because this additional mutation was outside the scope of this work, clone 321 is not considered further.

Genetic stability of the mutants. The progenies of the infectious mutants were analysed by RNA sequencing with reverse transcriptase. Three of the infectious mutants maintained the mutation in the progeny while the other four mutants reverted to the wild type sequence of PSTVd during the first passage in the host plant. In Figure 4 the positions of the stable mutations and the reversions in hairpin II are shown. The reversions are located in the inner part of hairpin II, designated as 'core' in Figure 4.

In order to exclude the possibility that the reversions were due to errors during the *in vitro* RNA transcription prior to inoculation, infectivity of the reverting mutants was tested not only with RNA transcripts but also with cDNA which is synthesized in the bacterial cell with much higher fidelity. Using literature values for the error rates of RNA transcription (Blank *et al.*, 1986) and bacterial replication (Watson *et al.*, 1988), it was estimated that the RNA

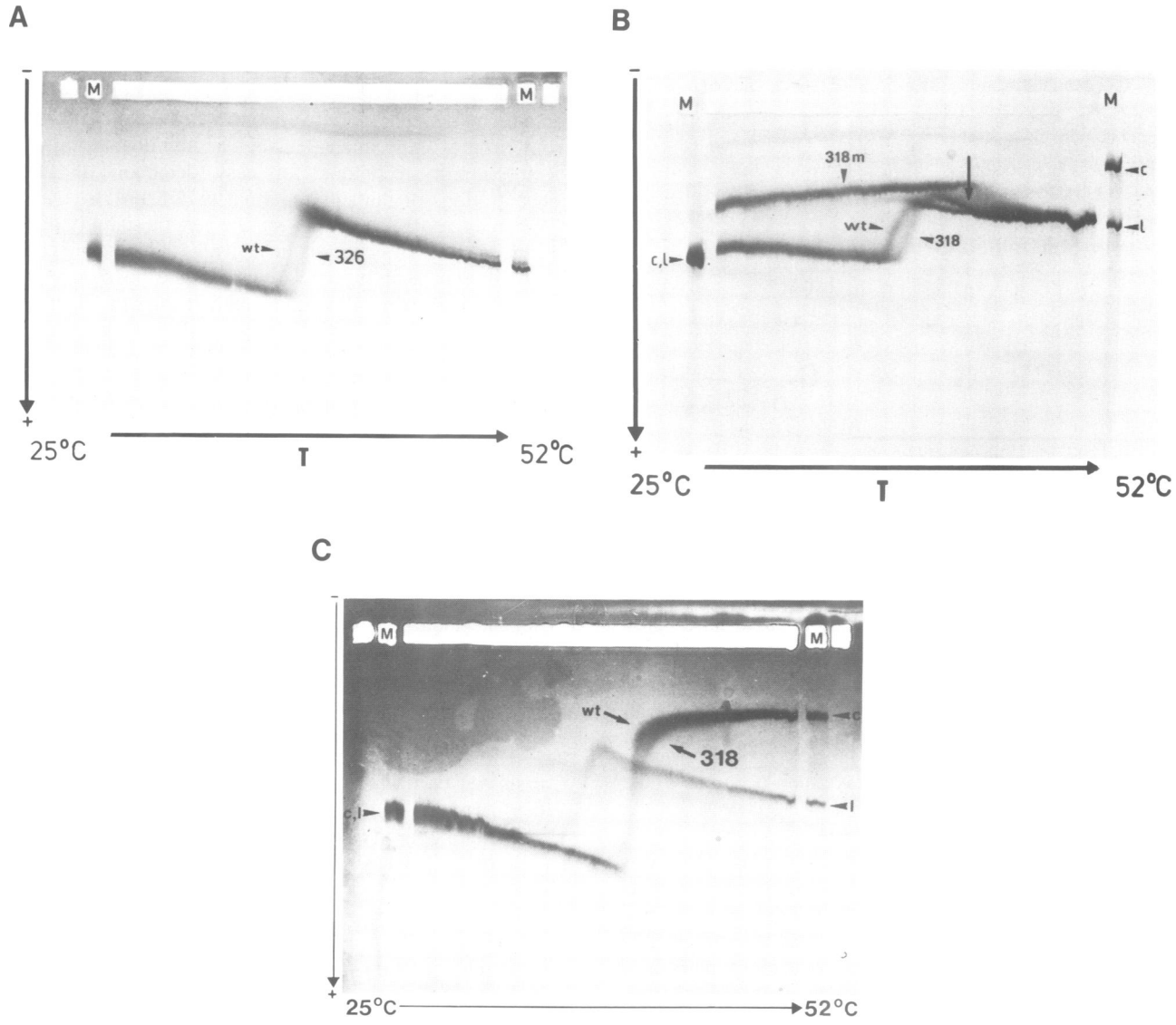

Fig. 3. Analysis of structural transitions of linear PSTVd-transcripts (**A,B**) and circular (**C**) PSTVd by temperature-gradient gel electrophoresis (TGGE) in 5% polyacrylamide. Buffer conditions were 17.8 mM Tris, 17.8 mM boric acid, 0.4 mM EDTA. The direction of the electrophoresis is from top to bottom. The lowest and highest temperatures of the linear temperature gradient are indicated at the bottom. Marker slots contain natural circular PSTVd (C) and natural linear PSTVd (L) (B and C), or linear transcripts (A). At temperatures below and above the main transition the electrophoretic mobility gradually increases with temperature (A and C, but not in B); this is most probably due to slight variations of the gel matrix and is not caused by differences of the viroid RNA. The designation of wild type sequence (wt) and mutations (326, 318) are as in the text. Curve 318m (B) corresponds to a multibranched conformer. The arrow indicates the high temperature transition of mutant 318 (see text). In C small amounts of linear molecules (right marker slot) are present in addition to the circular viroids, which are copurified during viroid preparation.

Table I. Thermodynamic analysis of the main transition of RNA transcripts from PSTVd mutants

Mutant	ΔT_m		Cooperativity (compared with wild type transcript)	
	Experimental	Theoretical	Experimental[a]	Theoretical[b]
227	0.0	+0.2	less cooperative	less cooperative
229	−1.3	0.0	same	less cooperative
233	0.0	+0.7	less cooperative	less cooperative
236	0.0	+1.9	less cooperative	less cooperative
318	+1.3	+0.6	same	same
324	+0.3	+0.8	less cooperative	less cooperative
326	+1.1	+2.7	less cooperative	less cooperative

T_m values were determined by TGGE and are listed as differences from that of the wild type (ΔT_m); cooperativity was evaluated from optical melting curves. For theoretical T_m values and cooperativity see text.
[a]Estimated from the half width of the main transition in optical denaturation curves.
[b]Cooperativity, i.e. coupling of the denaturation of the left and the right halves of the rod-like structure.

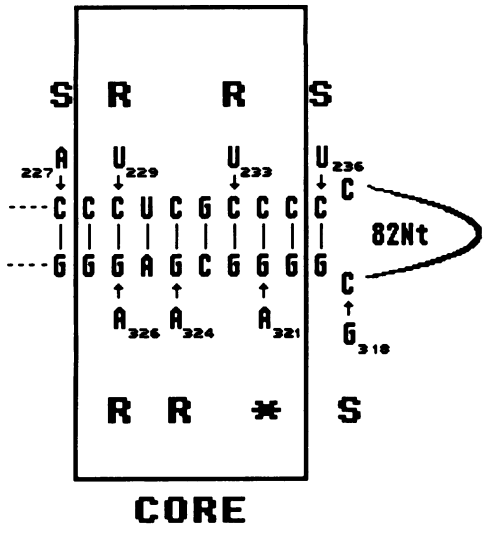

Fig. 4. Stable mutations (S) and reversions (R) in hairpin II of PSTVd. The core of hairpin II, depicted in a box, contains the reversions; its border extends to the positions of the stable mutants.

inoculum of 6×10^{11} RNA molecules (Materials and methods) contained $\sim 10^4$ revertants to the exact wild type sequence, whereas only 1–2 molecules with the reverted sequence were present in the cDNA inoculum of 3×10^{11} molecules. The cDNAs of the reverting mutants 229, 233, 324 and 326 were as infectious as the corresponding RNA transcripts which were analysed by gel electrophoresis. Thus, it could safely be concluded that reversion occurred in the plant after inoculation.

Pathogenicity of the stable mutants. Circular RNA was isolated from plants infected with the three stable mutants 227, 236 and 318 and used for another series of infections. The plants that were infected with the PSTVd mutants 227 and 236 showed the same symptoms as the plants inoculated with the wild type PSTVd (intermediate strain). In contrast, the plants inoculated with the mutant 318 clearly showed milder symptoms. These results reflect the positions of the mutations in the native structure. According to the literature (Schnölzer *et al.*, 1985) the extent of symptom expression is determined by the so called 'virulence modulating region' (Figure 1A, upper strand nt 42–60, lower strand nt 300–318). Only mutation 318 is located in or at the border of this region whereas mutants 227 and 236 are clearly distant and should not affect symptoms. Mutant 318 stabilizes the virulence modulating region, and thus the resulting milder symptoms are in accordance with the model of Schnölzer (Schnölzer *et al.*, 1985).

Discussion

Conserved core and variable periphery

Since the experimental tests for infectivity did not differentiate between cellular replication and intercellular transport, one has to consider both phenomena *a priori*. Obviously, the structural changes caused by stable mutations are tolerated during infection. Those from reversions, however, have to be part of a structural element in which changes are highly unfavourable for infection. A high selection pressure in favour of the wild type sequence has to be assumed so that the revertants do overgrow the mutants very quickly during infection. Many speculations could be presented regarding the mechanism of the high selection pressure observed, but this kind of discussion has to be delayed until more data about the exact concentrations of mutants and kinetics of reversions will become available from future experiments using a non-denaturing gel electrophoretic technique for easy detection of small concentrations of mutants (Zimmat *et al.*, 1989).

Whereas a correlation between the genetic stability of the mutation and its influence on the native viroid structure could not be drawn, such a correlation with the position of the mutations in hairpin II is quite convincing. In Figure 4, hairpin II is divided into a core region containing the reversions and two peripheral or variable segments where the stable mutations are located. Thus, it may be generalized from the present data, that the structural integrity of the core is highly required for infection, whereas mismatches or wobble pairs in the variable segments are tolerated. Since hairpin II can be formed in the (+)-strand sequence as well as in the (−)-strand sequence, and both polarities take part in replication, it cannot be differentiated whether the impairment of hairpin II in the (+)-strand or in the (−)-strand or in both is crucial.

Recent results of Owens and colleagues (Owens *et al.*, 1990) are in accordance with our emphasis on the relevance of hairpin II; they showed that constructs of viroid chimaeras of the PSTVd-group were infectious only if the hairpin II could be reestablished.

Conserved core and variable periphery in other viroids

When hairpin II of PSTVd is compared with that of other viroids belonging to the PSTVd-group, slight variations in length and sequences are obvious. The hairpins II may be aligned, however, so that it becomes obvious that the core (see Figure 4) indicated by the box in Figure 5 is identical in all of them. The only exception is an A:U pair in CSVd contrasting to a G:C pair in all other viroids. Thus, the concept we introduce here about a core region of hairpin II which is critical for infectivity, holds true for site-directed mutations as well as natural variations.

Homology of hairpin II with recognition sites of host genes

It is obvious to ask if structural elements similar to hairpin II do also exist in the host. In fact, GC-rich segments were found in the 5′ untranscribed regions of eukaryotic genes. Those have mostly been reported for human (Gidoni *et al.*, 1984; Levanon *et al.*, 1985), animal (Swick *et al.*, 1989) and viral systems (Gidoni *et al.*, 1984; McNight, 1984) whereas corresponding regions in plant genes have been reported in only two cases (Douglas *et al.*, 1987; Matsuoka and Minami, 1989). A GC-rich segment from the 5′ upstream region of the superoxide-dismutase gene from man (Levanon *et al.*, 1985) exhibits a striking similarity to hairpin II of PSTVd (Figure 6). Although depicted as a hairpin within one strand in Figure 6, the corresponding GC-rich double helices are formed in the eukaryote also as segments of the double-stranded genomic DNA without hairpin formation. Thus, a hairpin structure in the viroid could correspond to a segment of the genomic double-stranded DNA of the host. The other examples of GC-rich segments mentioned above show similar homology to hairpin II. Most of those are in so-called housekeeping genes (Watson *et al.*,

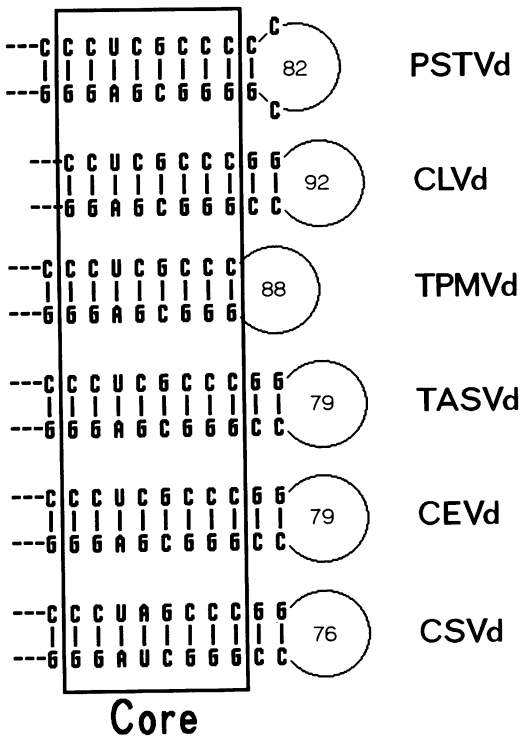

Fig. 5. Comparison of hairpin II in different viroids of the PSTVd-group. The core region (Figure 4) is identical in all hairpin II structures of this group with the exception of CSVd, in which one GC-pair is changed into an AU-pair. PSTVd, potato spindle tuber viroid; CLVd, *Columnea* latent viroid; TPMVd, tomato planta macho viroid; TASVd, tomato apical stunt viroid; CEVd, citrus exocortis viroid; CSVd, chrysanthemum stunt viroid.

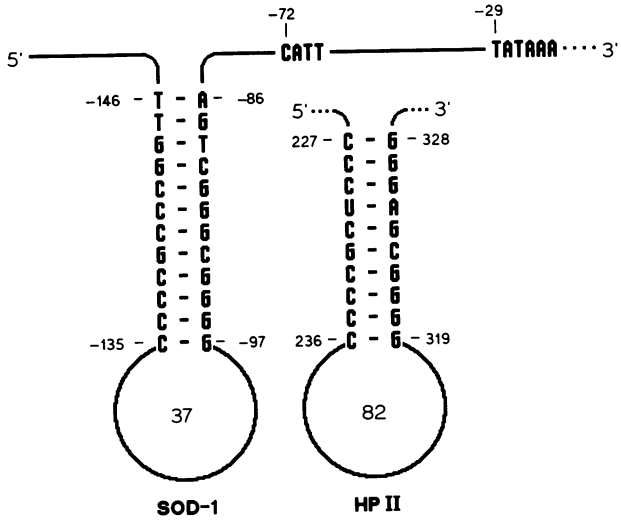

Fig. 6. Comparison of the hairpin II structure of PSTVd and a 5'-upstream GC-rich segment in the control region of the superoxide-dismutase I gene from man (Levanon *et al.*, 1985).

1988) and in some cases it was suggested that they act as binding sites for transcription factors (Gidoni *et al.*, 1984).

Both hairpin II RNA and the GC-rich DNA segment assume an A-form double helix

An RNA structural motif in viroids was compared with a DNA motif in host genes. Since the DNA segment as part of the 5'-untranscribed region does not lead to a corresponding RNA segment, one has to consider the relevance of this comparison. Recently the structures of several double-stranded DNA oligomers with high GC-content were studied by X-ray analysis (Heinemann *et al.*, 1987). It was found that these fragments of DNA assume an A-form double helix rather than B-form, which is otherwise typical for DNA. As far as it is known, double-stranded RNA always assumes the A-form. Thus the sequence homology between GC-rich segments of the host DNA and hairpin II of viroid RNA indeed corresponds to a structural similarity. From a close inspection of the structures as determined from X-ray analysis (Saenger, 1984) it is obvious that the size and shape of the large groove in particular are very similar in A-form RNA and in GC-rich DNA. The structural similarity suggests that a host protein factor which binds to the regulator GC-rich element may also bind to hairpin II of viroids.

Hairpin II as a hypothetical binding site of transcription factors

As mentioned above, the infectivity data of the present work leave open two possibilities for an involvement of hairpin II: it may be essential for the multiplication of viroids at the site of primary infection or it may enable viroids to be transported intercellularly to spread the infection. Whereas the knowledge about the molecular mechanism of viroid transport is nearly negligible and does not allow a more detailed discussion, more features of the different steps of viroid replication are known. Participation of hairpin II in that part of the infection process may indeed be considered based on our better knowledge of this process. The fact that some GC-rich segments of the host cell act as binding sites for transcription factors led us to suggest that hairpin II may fulfil a similar function as well. The replication enzyme itself, i.e. the DNA-dependent RNA polymerase II (see Introduction) does not seem to bind to hairpin II because it is attached specifically to one terminal hairpin of viroids in their rod-like structure as was seen on electron micrographs (Goodman *et al.*, 1984). Inside the cell, however, not only the circular viroids in the rod-like conformation but also the oligomeric (−)-strand intermediates in metastable, hairpin II-containing structures act as templates for polymerase II. Thus, the electron micrographs of complexes of polymerase and circular viroids do not exclude the possibility that polymerase could bind to hairpin II in replication intermediates.

Because of the analogy with the GC-rich elements in host DNA it seems nonetheless more probable that hairpin II is recognized by the transcription factors mentioned above. If a transcription factor did bind to hairpin II in circular viroids, one would have to assume that some denaturation of the native structure, probably induced by protein binding must occur. It is more obvious to assume that the transcription factor binds to hairpin II in the (−)-strand oligomeric viroid RNA which is synthesized in a hairpin II containing structure (Figure 3C). Since it is known from *in vitro* studies (Rackwitz *et al.*, 1981; Goodman *et al.*, 1984) that the polymerase II transcribes circular viroids without added transcription factors, it appears most probable that hairpin II is involved in optimization of the template function of the (−)-strand replication intermediate. Indeed, very recent studies (Schindler and Mühlbach, 1990) have shown that the (+)-strands are synthesized in higher concentrations in the cell as compared with (−)-strands, i.e. the (−)-strands have

to be the better templates. The high frequency of reversion as well as the low extent of replication of the hop stunt viroid, which lacks hairpin II, in tomato (Diener, 1979) additionally indicate that replication may occur on a low level without hairpin II formation, but with hairpin II the viroid is optimally adapted to tomato or potato as hosts.

Viroids are molecular parasites which form metastable structures and mimic cellular DNA

The core region of hairpin II as a critical structural element for viroid replication reveals two new aspects of viroid function, at least of those from the PSTVd group.

Hairpin II as part of a metastable structure is essential for cellular function. Metastable conformations of RNA had first been found in tRNA (Fresco *et al.*, 1966), but they were described as functionally 'wrong' structures. PSTVd, however, is an example of an RNA, for which not only the structure of lowest free energy but also metastable structures have to be considered under functional aspects. Possibly, the switch from the metastable, hairpin-containing structure to the low-energy, rod-like structure corresponds to functional changes during the life cycle of the viroid from a molecule active in replication to one which is organized in a stable and RNase-protected form inside the nucleolus (Harders *et al.*, 1989).

Hairpin II, although it is a double-stranded RNA segment, exhibits high homology in sequence and three-dimensional structure to GC-rich segments of genomic host DNA. Based on reports of the regulatory activity of GC-rich segments and in particular in their function as transcription factor binding sites, we suggest that hairpin II may act as a binding site for an as yet unidentified transcription factor of the host. This corresponds to the fact that all viroids capable of forming hairpin II share a common host range, i.e. potato and tomato. The concept that viroids function as molecular parasites by simulating host structures was mentioned earlier (Riesner *et al.*, 1979, 1983) and can now be confirmed by an additional phenomenon: the likeness to DNA of the native viroid RNA leads to recognition by the polymerase II; the likeness to RNA of the host GC-rich elements makes viroid replication even more effective by using host transcription factors.

Materials and methods

Bacterial strains

Escherichia coli strains BMH71-18, BMH71-18mutS and MK30-3 were used to propagate and mutagenize M13-clones. They were a gift from Drs L.Altschmid and W.Hillen, University of Erlangen. The *E.coli* strain HB101 was used in all steps with transcription vector pRH701. It was a gift from Dr R.Hecker (this laboratory).

Vectors

The M13mp9/M13mp9rev-system (Kramer *et al.*, 1984) was used for site-directed mutagenesis. The plasmid pRH701 (Hecker *et al.*, 1988; Hecker *et al.*, 1989) which contains a T7-promoter was used to produce transcripts of the PSTVd-mutants generated.

Enzymes

Enzymes were obtained from Boehringer Mannheim, Pharmacia and New England Biolabs and were used according to the instructions of the suppliers.

Synthetic oligonucleotides

The oligonucleotides for mutagenesis and sequencing were synthesized with an Applied Biosystems 380A DNA synthesizer and were purified by preparative polyacrylamide gel electrophoresis. In total eight primers were used for site-directed mutagenesis and five primers for sequencing. Mutagenesis primers (5′→3′) were: (MP227) AGGGGGCGAGGTGT GGTCCTG; (MP229) AGGGGGCGAAGGGTGGTCCTG; (MP233) CGCAAAGGGGACGAGGGGTGG; (MP236) CAGCGCAAAGAGG-GCGAGGGG; (MP318) CCCTCGCCCCCAAGCAAGT; (MP321) AACACCCTCGCTCCGAAGCAA; (MP324) GGGCTAAACACCCT-TGCCCCGAAGC and (MP326) GGCTAAACACCTTCGCCCCGA. DNA sequencing primers (5′→3′) were as follows: (M13-primer) GTAAAACGACGGCCAGT and (PRI 322) GTATCACGAGGCCC TTTCG. RNA sequencing primers (5′→3′): (RGV1) CCTGAAGCGCTC-CTCCGAG; (RGV3) CCAACTGCGGTTCCAAGGGCTAAACACC; (RGV5) GGAAGGACACCCGAAGAAAGGAAGGGTGAAA.

Molecular cloning

The PSTVd (intermediate strain)-cDNA (Cress *et al.*, 1983) was cloned into the *Bam*HI site of an M13mp9amE vector (Kramer *et al.*, 1984) by using standard techniques (Maniatis *et al.*, 1982). From the vector the cDNA was fractionated as described by Lizardi (Lizardi *et al.*, 1984). After mutagenesis the altered cDNAs were self-ligated, monomers were cut out with *Hae*III and cloned into blunt ended T7 transcription vector pRH701 (Hecker *et al.*, 1988) (to give pPL7XYZ where XYZ represents the number of the mutated nucleotide). Head to tail dimers were obtained by inserting a mutated monomeric cDNA into the *Bam*HI site of the transcription vectors containing the monomer PSTVd cDNA described above (pPL7XYZD). For *in vitro* transcription the plasmid was cut at the unique *Eco*RI site of the vector; the resulting transcripts of the T7 RNA polymerase carried pppGG at the 5′-end and GGGAAUU at the 3′-end in addition to the inserted nucleotide sequence. Monomeric transcripts represent the nucleotide sequence 147−359/1−146 (see Figure 1A), dimeric transcripts the nucleotide sequence 147−359/1−359/1−146.

Site-directed mutagenesis

Site directed mutagenesis of the PSTVd cDNA was carried out as described by Kramer (Kramer *et al.*, 1984), as a slight modification a 2000-fold molar excess of the oligonucleotide over the phage was used for inducing the mutation, which is 10 times more than that described previously (Kramer *et al.*, 1984). The rate of mutation ranged from 10 to 50%, and the potential mutants were prescreened by dot blot analysis (Dalbadie-McFarland *et al.*, 1982) and verified by sequencing.

Infectivity assays

Tomato seedlings were inoculated with dimeric RNA transcripts, with excised dimeric cDNAs or, in three cases, with circular PSTVd. In the case of RNA transcripts, 125 ng RNA/tomato seedling were used, whereas in the case of dimeric cDNAs we used 100 ng DNA/tomato seedling, which was the same amount of nucleic acid as used for the inoculation with circular PSTVd. After 4 weeks 1 g of tomato leaves was analysed for the presence of circular PSTVd (Schumacher *et al.*, 1983).

Sequence analysis

After mutagenesis the mutations were analysed in the M13 vector. The complete cDNA of each mutant was sequenced after cloning it into the transcription vector. For sequencing the T7 sequencing kit (Pharmacia) and the appropriate primer (see above) were used. PSTVd isolates from infected plants were sequenced according to Zimmern and Kaesberg (Zimmern and Kaesberg, 1978) using a 10× higher concentration of primers (R.Gruner, personal communication).

DNA isolation

Small scale plasmid isolation was done according to the method of Holmes and Quigley (Holmes and Quigley, 1981). Large scale preparations of plasmids and M13 RF were carried out as described by Hillen (Hillen *et al.*, 1981). M13 single stranded DNA was isolated following the procedure of Messing (Messing, 1983).

RNA isolation

Small scale analysis of PSTVd infections in tomato plants was carried out by bidirectional gel electrophoresis (Schumacher *et al.*, 1983). For RNA sequencing a larger amount and greater purity of PSTVd was achieved by a preparation procedure including a HPLC purification step (Colpan *et al.*, 1983; Riesner *et al.*, 1987).

Thermodynamic analysis

Thermal denaturation of linear transcripts as well as of circular PSTVd was analysed by temperature gradient gel electrophoresis (TGGE) (Rosenbaum and Riesner, 1987); the transcripts were also analysed by conventional optical melting curves (Steger *et al.*, 1984).

Computer calculations

The secondary structure of the mutants in the linear as well as in the circular form were calculated by the Zuker−Nussinov algorithm in the modified version of Steger *et al.* (Steger *et al.*, 1984). All calculations were carried out with a μVAXstation II (Digital Equipment Corporation).

Acknowledgements

We thank Dr R.Owens (Beltsville) for providing the clone of the PSTVd intermediate, Mr R.Gruner for RNA sequencing primers and Professor Dr W.Hillen (Erlangen) for stimulating discussions. We are indebted to Ms Viola Thiel for her help in preparing the manuscript. The work was supported by grants from Deutsche Forschungsgemeinschaft and Fonds der Chemischen Industrie.

References

Blank,A., Gallant,J.A., Burgess,R.R. and Loeb,L.A. (1986) *Biochemistry*, **25**, 5920−5928.

Branch,A. and Robertson,H.D. (1984) *Science*, **223**, 450−455.

Branch,A.D., Benenfeld,B.J. and Robertson,H.D. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 9128−9132.

Colpan,M., Schumacher,J., Brüggemann,W., Sänger,H.L. and Riesner,D. (1983) *Anal. Biochem.*, **131**, 257−265.

Cress,D.E., Kiefer,M.C. and Owens,R. (1983) *Nucleic Acids Res.*, **11**, 6821−6835.

Dalbadie-McFarland,G., Cohen,L.W., Riggs,A.D., Morin,C., Itakura,K. and Richards,J.H. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 6409−6413.

Diener,T.O. (1979) *Viroids and Viroid Diseases*. Wiley and Sons, New York.

Diener,T.O. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 58−62.

Diener,T.O. (1987) *The Viroids*. Plenum Publishing Corporation, New York.

Douglas,C., Hoffmann,H., Schulz,W. and Hahlbrock,K. (1987) *EMBO J.*, **6**, 1189−1195.

Fresco,J.R., Adams,A., Ascione,R., Henley,D. and Lindahl,T. (1966) *Cold Spring Harbor Symp. Qaunt. Biol.*, **31**, 527−537.

Gidoni,D., Dynan,W.S. and Tijan,R.C. (1984) *Nature*, **312**, 409−413.

Goodman,T.C., Nagel,L., Rappold,W., Klotz,G. and Riesner,D. (1984) *Nucleic Acids Res.*, **12**, 6231−6246.

Harders,J., Lukacs,N., Robert-Nicoud,M., Jovin,T.M. and Riesner,D. (1989) *EMBO J.*, **8**, 3941−3949.

Hecker,R. (1989) *Thesis*. Heinrich-Heine-Universität, Düsseldorf, FRG.

Hecker,R., Wang,Z., Steger,G. and Riesner,D. (1988) *Gene*, **72**, 59−74.

Henco,K., Sänger,H.L. and Riesner,D. (1979) *Nucleic Acids Res.*, **6**, 3041−3059.

Heinemann,U., Lauble,H., Frank,R. and Blöcker,H. (1987) *Nucleic Acids Res.*, **15**, 9531−9550.

Hillen,W., Klein,R.D. and Wells,R.D. (1981) *Biochemistry*, **20**, 3748−3756.

Holmes,P.S. and Quigley,M. (1981) *Anal. Biochem.*, **114**, 193−197.

Keese,P. and Symons,R.H. (1987) In Semancik,J.S. (ed.), *Viroids and Viroid-like Pathogens*. CRC Press Inc., Boca Raton, pp. 2−44.

Kramer,W., Drutsa,V., Jansen,H.-W., Kramer,B., Pflugfelder,M. and Fritz,H.-J. (1984) *Nucleic Acids Res.*, **12**, 9441−9456.

Levanon,D., Lieman-Hurwitz,J., Dafni,N., Wigderson,M., Sherman,L., Bernstein,Y., Laver-Rudich,Z., Danciger,E., Stein,O. and Groner,Y. (1985) *EMBO J.*, **4**, 77−84.

Lizardi,P.M., Binder,R. and Short,S.A. (1984) *Gene Anal. Techn.*, **1**, 33−39.

Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Matsuoka,M. and Minami,E. (1989) *Eur. J. Biochem.*, **181**, 593−598.

McNight,S.L., Koinsbury,R.C., Spence,A. and Smith,M. (1984) *Cell*, **37**, 253−262.

Messing,J. (1983) *Methods Enzymol.*, **101**, 20−78.

Mühlbach,H.P. and Sänger,H.L. (1979) *Nature*, **278**, 185−187.

Owens,R. (1990) *Mol. Gen. Genet.*, in press.

Owens,R.A., Hammond,R.W., Gardner,R.C., Kiefer,M.C., Thompson,S.M. and Cress,D.E. (1986) *Plant Mol. Biol.*, **6**, 179−192.

Owens,R.A., Candresse,T. and Diener,T.O. (1990) *Virology*, **175**, 238−246.

Rackwitz,H.R., Rohde,W. and Sänger,H.L. (1981) *Nature*, **291**, 297−301.

Riesner,D. and Gross,H.J. (1985) *Annu. Rev. Biochem.*, **54**, 531−564.

Riesner,D. and Steger,G. (1990) In Saenger,W. (ed.), *New Series in Biophysics—Nucleic Acids*. Landolt-Börnstein,Vol. VII/1, Springer-Verlag, Berlin, pp. 194−243.

Riesner,D., Henco,K., Rohkohl,U., Klotz,G., Kleinschmidt,A.K., Gross,H.J., Domdey,H. and Sänger,H.L. (1979) *J. Mol. Biol.*, **133**, 85−115.

Riesner,D., Colpan,M., Goodman,T.C., Nagel,L., Schumacher,J., Steger,G. and Hofmann,H. (1983) *J. Biomol. Struct. Dyn.*, **1**, 669−688.

Riesner,D., Klaff,P., Steger,G. and Hecker,R. (1987) In Lee,J.H.J. and FrederickJ.F. (eds), *Endocytobiology III*. Annals of the New York Academy of Sciences, New York. Vol. 503, pp. 212−237.

Rosenbaum,V. and Riesner,D. (1987) *Biophys. Chem.*, **26**, 235−246.

Saenger,H.L. (1982) In Parthier,B. and Boulter,D. (eds), *Encyclopedia of Plant Physiology*. New Series, Vol. 14B, Springer-Verlag, Berlin, pp. 368−454.

Saenger,W. (1984) In Cantor,C.R. (ed.), *Springer Advanced Texts in Chemistry—Principles of Nucleic Acid Structure*. Springer-Verlag, Berlin.

Sänger,H.L. (1988) *Acta Horticulturae*, **234**, 79−87.

Schindler,J.-M. and Mühlbach,H.P. (1990) *Seventh International Congress on Virology*, Berlin, Aug. 26−31, 1990, p55-002.

Schnölzer,M., Haas,B., Ramm,K., Hofmann,H. and Sänger,H.L. (1985) *EMBO J.*, **4**, 2181−2190.

Schumacher,J., Sänger,H.L. and Riesner,D. (1983) *EMBO J.*, **2**, 1549−1555.

Semancik,J.S. and Harper,K.L. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 4429−4433.

Steger,G., Hofmann,H., Förtsch,F., Gross,H.J., Randles,J.W., Sänger,H.L. and Riesner,D. (1984) *J. Biomol. Struct. Dynam.*, **2**, 543−571.

Swick,A.G., Blake,M.C., Kahn,J.W. and Azizkhan,J.C. (1989) *Nucleic Acids Res.*, **17**, 9291−9304.

Tabler,M. and Sänger,H.L. (1984) *EMBO J.*, **3**, 3055−3062.

Visvader,J.E., Forster,A.C. and Symons,R.H. (1985) *Nucleic Acids Res.*, **13**, 5843−5856.

Visvader,J.E. and Symons,R.H. (1986) *EMBO J.*, **5**, 2051−2055.

Watson,J.D., Hopkins,N.H., Roberts,J.W., Argetsinger-Steitz,J. and Weiner,A.M. (1988) *Molecular Biology of the Gene*. The Benjamin Cummings Publishing Company, CA, USA. Fourth edition.

Zimmat,R., Gruner,R., Hecker,R., Steger,G. and Riesner,D. (1989) In Sarma,R.H. and Sarma,M.H. (eds), *Structure and Methods*. Vol. 3: *DNA and RNA*. Adenine Press, Guilderland, pp. 339−357.

Zimmern,D. and Kaesberg,P. (1978) *Proc. Natl. Acad. Sci. USA*, **75**, 4257−4261.

Received on October 1, 1990; revised on December 6, 1990