The Dimerization Domain of Potato Spindle Tuber Viroid, a Possible Hallmark for Infectious RNA

Frank-Ulrich Gast,*,‡ Dirk Kempe,§ and Heinz L. Sänger§

Institut für Biochemie, Justus-Liebig-Universität Giessen, Heinrich-Buff-Ring 58, D-35392 Giessen, Germany, and Abteilung Viroidforschung, Max-Planck-Institut für Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, Germany

Received April 13, 1998; Revised Manuscript Received June 22, 1998

ABSTRACT: Covalently closed circular (+) RNA of the potato spindle tuber viroid (PSTVd) can efficiently dimerize noncovalently upon heating and slow cooling in the presence of monovalent cations or Mg^{2+}. In vitro transcription of subgenomic fragments reveals that the ability to dimerize resides in the "upper strand" of its self-complementary rod-like structure. Nuclease probing of these fragments, namely, molecules spanning either the upper or the lower strand of PSTVd, confirms the existence of the previously proposed hairpins I−III, of which hairpin I might contain noncanonical G·A and A·A base pairs. In addition, the upper and lower (+) strands contain large hairpin loops consisting of stretches rich in either adenosine or uridine. Dimerization of the upper (+) strand results in a nuclease-resistant core encompassing hairpin I and is inhibited by an antisense oligonucleotide spanning the entire hairpin; this palindromic domain thus represents the dimerization site. When upper and lower strands were heated and cooled together, no annealing to a viroid-like duplex of both molecules occurs, only dimerization of the upper strand. Therefore, the dimerization hairpin of viroid RNA represents a unique conformational signal that is homologous to similar regions in the human immunodeficiency virus and other retroviruses.

Viroids represent a distinct class of small plant pathogens in that they consist of only a small circular single RNA (+) strand without coding regions (reviewed in refs *1*−*3*). In contrast to plant satellite RNAs, which need helper viruses for replication, these RNAs replicate autonomously via the DNA-dependent RNA polymerase II of the host (*1*, *4*). The rolling circle replication of the viroid (+) RNA starts with the production of oligomeric (−) strands (*5*); these are then copied into oligomeric (+) strands which are subsequently processed into circular monomeric units, most likely by cleavage and ligation using a host endoribonuclease and an RNA ligase (*6*−*8*). Due to its self-complementarity, the circular viroid (+) RNA folds into a compact rod-like structure. This has been verified by analyzing its melting transitions (reviewed in ref *9*) and by site-directed mutagenesis, where it was found that, after deleting a 9 nt[1] stretch in the "upper strand", the rod-like structure was reestablished by a compensatory deletion in the "lower strand" acquired in planta (*10*). We have recently probed PSTVd-specific linear RNA transcripts that comigrate with the rod-like molecules of isolated circular PSTVd (*11*) and found only a few deviations from the theoretically proposed rod structure, namely, a small branch at the left terminus, noncanonical base pairs in the central or C region, and large internal A-rich loops (Figure 1).

Theoretical and experimental analysis has shown that upon thermal denaturation PSTVd RNA undergoes a transition into an array of three hairpins connected by single-stranded regions (*9*). Hairpin I (HP I, Figure 1) consists of an almost perfect 32 nt palindrome and is located in the upper C region which is conserved within each of the three different viroid groups (*3*). Because its apical loop contains a 10 nt palindrome, the entire HP I can base pair with a second HP I of another viroid unit of the oligomeric (+) strand intermediate, resulting in a so-called "trihelical region" (*12*). It was originally believed that the processing of an oligomeric (+) strand intermediate to a monomeric unit takes place within this structure (*1*, *6*, *12*). HP II with its GC-rich stem encompasses most of the lower strand in the rod model of Figure 1. HP III is positioned in a purine-rich region within the upper right part that varies in different viroids. All these thermodynamic domains have been proposed as being important for the viroid's life cycle (*9*), but their structure at ambient temperature, which is different from the one calculated for the temperature at the melting point, has not yet been determined.

It was shown by nondenaturing gel electrophoresis that isolated circular PSTVd RNA as well as linear PSTVd-specific transcripts are able to switch between different secondary or tertiary structures. A transition from a rod-like to a branched structure occurs upon thermal denaturation and subsequent renaturation in the presence of sufficient amounts of salt or Mg^{2+} (*11*, *13*, *14*). Our most recent unpublished data indicate that the branched conformer is a

* Address correspondence to this author at the Institut für Biochemie (FB 15), Justus-Liebig-Universität Giessen, Heinrich-Buff-Ring 58, D-35392 Giessen, Germany. Telephone: +49 (0) 641-9935403. Fax: +49 (0) 641-9935409. E-mail: Frank-Ulrich.Gast@chemie.bio.uni-giessen.de.

‡ Justus-Liebig-Universität Giessen.

§ Max-Planck-Institut für Biochemie.

[1] Abbreviations: C region, central region; GYSVd, grapevine yellow speckle viroid; HIV-1, human immunodeficiency virus type 1; HP, hairpin; HS, high-salt; LS, low-salt; rNTP, ribonucleotide triphosphate; nt, nucleotide; PAGE, polyacrylamide gel electrophoresis; PSTVd, potato spindle tuber viroid; $PSTVd_{lh}$, lower half of PSTVd (nucleotides 174−359 and 1); $PSTVd_{uh}$, upper half of PSTVd (nucleotides 2−173); TBE, Tris-borate/EDTA buffer.

S0006-2960(98)00830-7 CCC: $15.00
Published on Web 09/12/1998

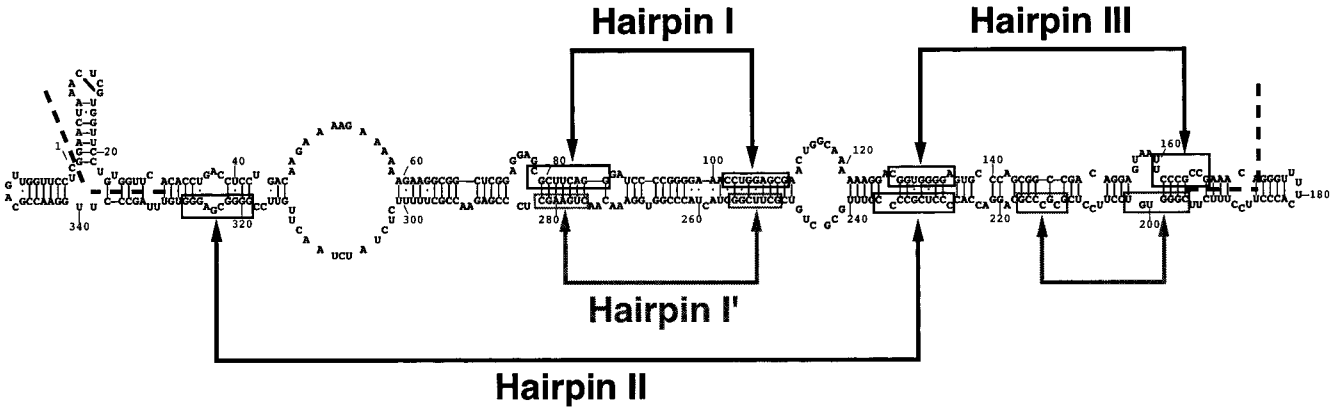

FIGURE 1: Proposed secondary structure of PSTVd based on structural probing and comparative sequence analysis (*11*). The model deviates slightly from earlier ones that were based on folding algorithms and temperature gradient gel electrophoresis (*9*). Potential hairpin regions, either taken from the literature (*9*, *23*) or obtained by theoretically folding subgenomic fragments at 37 °C using *mfold* (*17*), are defined by boxes and double arrows. Hairpin I′ is a conserved hairpin formed from the lower strand opposing hairpin I. Non-Watson−Crick base pairs are depicted by dots. The boundaries between the transcripts used in this work, $PSTVd_{uh}$ and $PSTVd_{lh}$, are illustrated by dashed lines.

cruciform in which HP I and a conserved hairpin on the opposite side of HP I, denoted HP I′ (Figure 1), stick out from the rod-like structure. While performing these experiments, we have observed a strong tendency of PSTVd RNA, even in the covalently closed circular form, to dimerize intermolecularly, especially after heating in the presence of salt or Mg^{2+}. By analyzing smaller fragments, we could find in the upper half of the viroid RNA molecule a single dimerization region. Here we describe the identification of this site and demonstrate its high homology to a similar domain in retroviral systems.

EXPERIMENTAL PROCEDURES

PCR of Viroid-Specific cDNA. Templates for in vitro RNA transcription were PCR-amplified from a plasmid into which four cDNA units of a lethal 359 nt PSTVd isolate (KF 440-2) had been tandemly cloned (*15*). Different combinations of primers were used. Primers in the (+) orientation corresponded to position 2−17, 174−188, 197−310, and 265−282 of PSTVd and contained a T7 promoter immediately upstream of the PSTVd-specific sequence (*11*). The (−)-oriented primers were complementary to nucleotides 1/359−335, 330−310, 173−152, and 84−68. Different combinations of primers were used, and PCR was performed in a volume of 200 μL containing 4 ng of PSTVd-specific plasmid by applying 30 cycles of the following regimen: 94 °C for 1 min, 40 °C for 1 min, and 72 °C for 1 min. After the mixtures were checked by nondenaturing 6% PAGE, they were phenol/chloroform extracted and precipitated with 2-propanol. A cDNA fragment spanning nucleotides 2−146 was prepared from a larger one comprising nucleotides 2−173 by *Hae*III digestion.

Isolation of PSTVd RNA and in Vitro Production of Long and Short PSTVd-Specific Transcripts. Circular PSTVd RNA (isolate KF 440-2) was purified from infected tomato leaves by phenol/chloroform extraction and bidirectional gel electrophoresis (*11*). For in vitro production of PSTVd-specific RNA fragments, the DNA of a single PCR was transcribed with T7 RNA polymerase (purified from an overproducing strain) in 300−500 μL as previously described (*11*, *16*). After 90 min at 37 °C, 0.2 unit/μL (final concentration) RNase-free DNase I was added, and the DNA template was digested for 15 min. After addition of 50 mM EDTA (final concentration), the mixture was phenol extracted and 2-propanol precipitated. The centrifuged precipitate was desalted over NAP-5 columns (Pharmacia), and the desalted RNA was concentrated in a Speedvac. Transcripts were not treated further, but their purity was acceptable as judged from denaturing PAGE.

Transcription of short HP I molecules was done using gel-purified oligodeoxynucleotides as the template (*16*), using 1 μM template and 0.5 mM rNTPs (with and without [α-^{32}P]-UTP). After the reactions were quenched with EDTA, the mixture was separated on a 14% denaturing polyacrylamide gel. Bands corresponding to the full-length transcripts were excised and the gel pieces extracted with H_2O. The eluate was desalted, concentrated, and stored at −20 °C. The 5′-labeling of purified nonradioactive transcripts is described below.

Thermal Denaturation and Renaturation ("Annealing") of RNA Transcripts and Gel Analysis of Their Complexes. Denaturation was achieved by a 2−2.5 min incubation in buffer containing either low-salt (LS), high-salt, (HS), or Mg^{2+} buffer at 90−95 °C with subsequent slow cooling in a Styrofoam block over 10−15 min to room temperature. In these experiments, LS buffer consisted of 20 mM Tris-HCl and 0.2 mM EDTA (pH 8.0), HS buffer contained an additional 100 mM NaCl, and Mg^{2+} buffer consisted of LS buffer and 5 mM $MgCl_2$. RNA concentrations were usually 0.1−0.2 μg/μL (ca. 2−4 μM), but much less if radioactive RNA transcripts were used. Electrophoretic analysis (2% agarose gels or 4 to 8% PAGE) was performed at room temperature in 0.5× TBE [45 mM Tris-borate (pH 8.3) and 1.25 mM EDTA]. Nonradioactive gels were stained with ethidium bromide and photographed on a transilluminator; radioactive gels were silver stained to reveal the position of the unlabeled RNA in excess, dried, and autoradiographed at room temperature. To prove oligomerization, the same RNA quantity was annealed in different volumes over at least a 10-fold concentration range; bands disappearing upon dilution were thus identified as oligomers. Alternatively, a small amount of 5′-labeled transcript (see below) was annealed with increasing amounts of nonradioactive transcript. In competition experiments, 1−20 μM antisense oligodeoxynucleotides were present during melting. Their correct binding sites were confirmed after annealing in HS

buffer by digestion for 5−60 min with 0.5 unit/μL RNase H after addition of 10 mM $MgCl_2$ and 5 mM DTT; the length of the cleavage products was determined in 6% denaturing PAGE. The oligonucleotides were also used in the end-labeled form; the gels were then silver stained, dried, and autoradiographed to detect the position of bound oligonucleotides relative to the positions of the RNA transcripts.

End Labeling and Nuclease Probing of the Transcripts. One to two micrograms of the transcripts was 5′-labeled using alkaline phosphatase, polynucleotide kinase, and [γ-^{32}P]ATP or 3′-labeled using RNA ligase and [α-^{32}P]pCp, gel purified, and desalted. Before probing, transcripts were thermally renatured in LS buffer to obtain monomers and checked by 5% nondenaturing PAGE. Nuclease probing was performed essentially as described (*11*), i.e., with RNase T1 in 1× TE [10 mM Tris-HCl (pH 8) and 0.1 mM EDTA] or 5× TE, in the absence or presence of 5 mM $MgCl_2$ (25 °C), RNase T2 in 5× TE (25 °C), RNase A (Sigma) in 5× TE and 0−600 mM NaCl (37 °C), nuclease S1 either in 50 mM sodium acetate (pH 4.5), 3 mM $ZnCl_2$, and 300 mM NaCl or in 50 mM Tris-HCl (pH 7.4), 20 mM $MgCl_2$, and 300 mM NaCl (37 °C), and nuclease V1 in 25 mM Tris-HCl (pH 7.4), 10 mM $MgCl_2$, and 200 mM NaCl. A ladder was prepared by alkaline hydrolysis for 1−2 min of the RNAs in 50 mM $NaHCO_3$ at 95 °C. G-specific sequencing was performed by incubation with RNase T1 in 1× TE at 56 °C. Samples were run on standard 6 or 8% sequencing gels and analyzed by autoradiography. In the case of the short HP I transcripts, reactions were carried out at 0−25 °C, and the samples were analyzed on 20 cm × 20 cm × 0.05 cm, 20% denaturing gels.

The secondary structure of the RNAs was reconstructed for 37 °C using *mfold*, version 2.0 (*17*). Suboptimal structures within 10% of the minimal folding energy were generated with and without forcing certain bases to be single- or double-stranded, and those which fit the cleavage sites best were chosen. Structures were also generated in which certain bases were forced to be single- or double-stranded, according to the nuclease data. Only a few manual rearrangements (e.g., for noncanonical base pairs) were necessary.

Identification of Folding Domains and the Dimerization Site by Limited Nuclease Digestion and RNA Sequencing. For the identification of stable folding domains of the RNA transcripts as well as for the determination of their dimerization domain, 5 μg of transcript was melted in 50 μL of either LS or Mg^{2+} buffer to obtain monomers or dimers, respectively. To obtain similar ionic conditions, 5 mM $MgCl_2$ (final concentration) was added to the LS mixture. To identify thermodynamic folding domains, both mixtures were incubated for 100 min at room temperature in the same buffer with 2 units/μL RNase T1, after which they were loaded onto a 20 cm × 20 cm × 0.15 cm, 10% polyacrylamide/8 M urea gel. The major bands were cut out, extracted with H_2O, desalted and concentrated as described above, and 5′-labeled with [γ-^{32}P]ATP. Sequences were determined by enzymatic cleavage using an RNA sequencing kit (available at that time from U.S. Biochemical). To map the dimerization site, monomeric and dimeric PSTVd[nt 2−173] RNAs were treated with a larger amount (20 units/μL) of T1 in Mg^{2+} buffer for 16 h at ambient temperature. Cleavage products were separated on a 20% denaturing gel and the most prominent bands excised and sequenced as before.

RESULTS

Oligomerization of Monomeric PSTVd RNA. A band containing monomeric rod-like molecules (M in Figure 2A) is formed after thermal denaturation and slow cooling of full-length PSTVd-specific RNA transcripts in LS buffer (lanes 1 and 4 of Figure 2A; see Experimental Procedures for the buffer composition). However, when these RNAs are melted and renatured in HS or Mg^{2+} buffer, additional slowly migrating bands are found (lanes 2, 3, 5, and 6 of Figure 2A). Aside from the branched monomer (*11*−*13*; BM in Figure 2A) and a few minor bands of unknown origin (presumably dimers with alternative conformations), the most prominent band was identified as a dimer because of the concentration dependence of its formation (data not shown).

PSTVd RNA isolated from infected tomato leaves can also efficiently dimerize (lane 7 of Figure 2A). Because it contains about 70% circular and 30% nicked molecules, which comigrate in nondenaturing gels, one may ask whether the dimerization is limited to the linearized molecules. Therefore, we phosphorylated PSTVd RNA with T4 polynucleotide kinase; only the linear molecules can be labeled. The ratio of monomers:dimers after heating was identical for the entire silver-stained RNA population and for the labeled RNA (data not shown), indicating that all viroid molecules, including the circular ones, participate in dimerization.

Oligomerization of Subgenomic RNA Transcripts of PSTVd. Efficient dimerization is realized with the upper half of the PSTVd molecule (nucleotides 2−173, termed $PSTVd_{uh}$; lanes 1−3 of Figure 2B). It occurs at high RNA concentrations in LS buffer (not shown) or at intermediate RNA concentrations in HS or Mg^{2+} buffer (Figure 2B,C). At 5 mM Mg^{2+} and RNA concentrations of 0.1 μg/μL (~2 μM), complete (100%) dimerization is observed (lane 3 of Figure 2B). The concentration dependence of dimer formation upon heating and slow cooling of $PSTVd_{uh}$ in Mg^{2+} buffer is shown in Figure 2C. Although the overall effect of monovalent cations (Na^+ and K^+ were interchangeable, not shown) resembles the effect of Mg^{2+}, the latter condition produces dimers more specifically, since only the lower of the two major dimer bands is visible (lane 3 of Figure 2B). However, heating in Mg^{2+} buffer and snap-cooling on ice produces both dimer bands, whereas heating and slow cooling of high concentrations of $PSTVd_{uh}$ in LS buffer leads preferentially to the upper of the two dimer bands (data not shown).

Dimerization is reversible. If a sample that was dimerized in Mg^{2+} is mixed with excess EDTA, heated, and slowly cooled, all dimers convert back into monomers (data not shown). Prolonged incubation in Mg^{2+} buffer at room temperature, without prior thermal denaturation, does not change preformed monomers into dimers. Heating at almost 60 °C is necessary to promote dimerization; ca. 50% of the monomers are converted into dimers after 30 min at 56 °C in Mg^{2+} buffer. In addition to dimer formation, the use of higher Mg^{2+} or RNA concentrations or heating in Mg^{2+} buffer with snap-cooling promotes the formation of slower migrating bands, presumably multimers, which are invisible in Figure 2B. Repeated freezing and thawing of $PSTVd_{uh}$

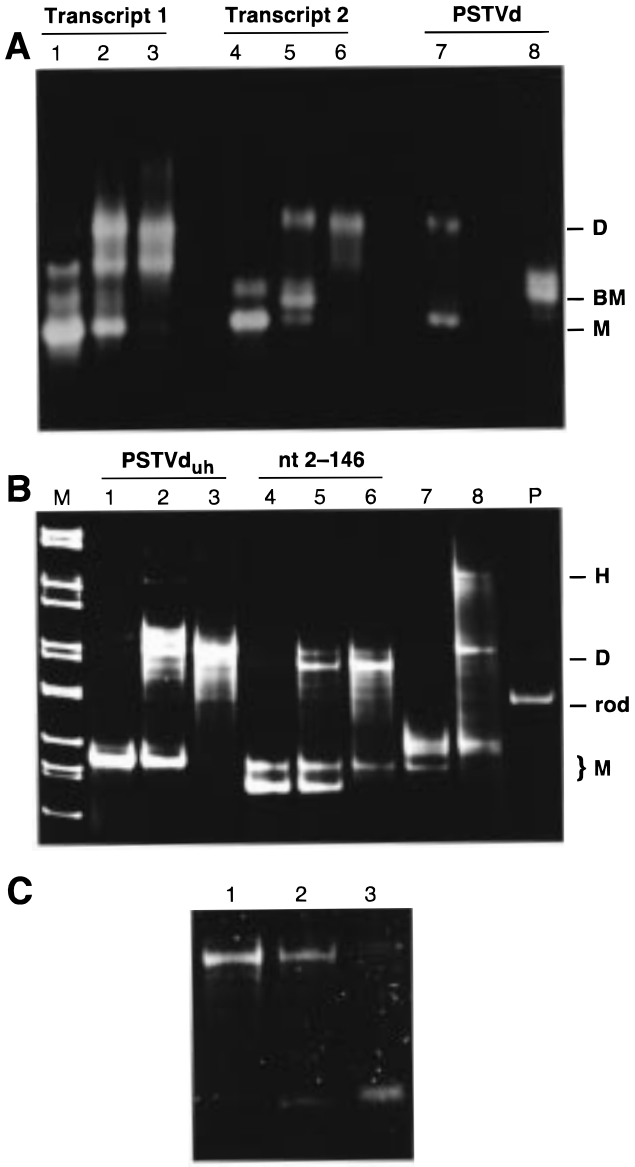

FIGURE 2: Dimerization of PSTVd-specific RNA transcripts (0.1−0.2 μg/μL) in a reaction volume of 10 (A) and 5 μL (B) and in a variable volume (C) through a heating−cooling cycle, visualized in nondenaturing gels. (A) A 2% agarose gel showing full-length PSTVd RNAs, namely, linear unit-length PSTVd transcripts commencing at nucleotides 2 and 1 (lanes 1−3) and commencing at nucleotides 174 and 173 (lanes 4−6). Heating and cooling of these molecules were performed in LS buffer (lanes 1 and 4), HS buffer (lanes 2 and 5), and Mg^{2+} buffer (lanes 3 and 6). In addition, the gel shows isolated PSTVd heated and slowly cooled in HS buffer in the absence (lane 7) and presence (lane 8) of 2.5 μM antisense oligodeoxynucleotide directed against hairpin I (~3-fold molar excess). Positions of rod-like monomer (M), branched monomer (BM, "cruciform"), and dimer (D) are marked. (B) A 5% PAGE showing subgenomic RNA fragments PSTVd$_{uh}$ (nucleotides 2−173) in LS (lane 1), HS (lane 2), and Mg^{2+} buffer (lane 3), a shortened form of PSTVd$_{uh}$ (nucleotides 2−146) in LS (lane 4), HS (lane 5), and Mg^{2+} buffer (lane 6), and PSTVd$_{lh}$ in HS buffer (lane 7). An equimolar mixture of PSTVd$_{uh}$ and PSTVd$_{lh}$ in Mg^{2+} buffer (lane 8) produces only low amounts of the PSTVd$_{uh}$−PSTVd$_{lh}$ duplex. P is PSTVd RNA isolated from infected tomato leaves; M is a DNA marker. Positions of monomer (M), dimer (D), (hetero)duplex (H), and the rod-like form of PSTVd (rod) are indicated. (C) Part of a 5% polyacrylamide gel showing the concentration dependence of the dimerization of PSTVd$_{uh}$ in Mg^{2+} buffer. A constant amount of 0.1 μg (1.8 pmol) of RNA was annealed in a volume of 2.5 (lane 1), 10 (lane 2), and 40 μL (lane 3). The slight mobility shift in the right lane is due to the high salt content of the 40 μL sample.

in H_2O also produces a ladder of oligomeric molecules (up to pentamers, data not shown).

When the cDNA template for PSTVd$_{uh}$ was digested with *Hae*III, transcription yielded a shortened RNA (nucleotides 2−146). This molecule could dimerize, albeit with lower efficiency (lanes 4−6 of Figure 2B), probably due to formation of alternative conformers. A "three-quarter" PSTVd molecule spanning the region of nucleotides 265−359 and 1−173 also showed dimerization (data not shown). No dimerization occurs with transcripts harboring only nucleotides 2−84, i.e., the upper left quarter of PSTVd, or encompassing the left half of PSTVd, i.e., starting at nucleotide 265 and ending at nucleotide 84 (cf. Figure 1; data not shown). In nondenaturing PAGE, the latter molecule ran as expected for an elongated stem−loop (data not shown) when compared to double-stranded RNA markers (described in ref *14*). Since molecules lacking HP I (nucleotides 2−84 and nucleotides 265−359/1−84) cannot dimerize whereas those molecules including it (nucleotides 265−359/1−173, nucleotides 2−173, and nucleotides 2−146) can, this stem−loop is likely to be responsible for dimerization. We will show below that this is indeed the case.

The lower half of PSTVd (nucleotides 174−359 and nucleotide 1, termed PSTVd$_{lh}$; see Figure 1) remains monomeric. In nondenaturing gels, this RNA forms a few fuzzy bands (lane 7 of Figure 2B) whose distribution was independent of RNA concentration, salts (NaCl, KCl, and $MgCl_2$), and temperature ramps (slow and snap-cooling). Similar results were obtained with a molecule encompassing only nucleotides 197−330 (i.e., HP II and 30 nucleotides at the 3′-terminus; data not shown).

Duplex RNAs Consisting of the Upper and Lower Half of PSTVd. Annealing the isolated half-molecules PSTVd$_{uh}$ and PSTVd$_{lh}$, we tested if the rod-like viroid structure of Figure 1 could be formed in a 1:1 duplex (lane 8 of Figure 2B). As can be inferred from lane 8 of Figure 2B, however, most PSTVd$_{uh}$ molecules form a dimer only with themselves (D in Figure 2B), whereas a large amount of PSTVd$_{lh}$ still runs as a monomer (M). In addition, minor amounts of even slower migrating bands occur. These bands correspond to heteroduplexes because they show up at stoichiometric amounts of PSTVd$_{uh}$ and PSTVd$_{lh}$ and not for the single fragments alone. The most prominent of these bands (H in Figure 2B) is likely to be a 1:1 complex of PSTVd$_{uh}$ and PSTVd$_{lh}$. Its low electrophoretic mobility indicates a highly branched structure which does not comigrate with the putative cruciform structure (*11*) of circular PSTVd or corresponding full-length transcripts (data not shown).

When a radioactively labeled half-molecule (e.g., PSTVd$_{lh}$) is annealed with the other, nonradioactive transcript (e.g., PSTVd$_{uh}$) and the RNAs are run on nondenaturing gels, the distribution of radioactivity provides information on the composition of the bands, since radioactive bands can only represent multimers of the radioactive RNA fragment or complexes of radioactive and nonradioactive fragments. Different combinations of radioactive and nonradioactive fragments were tested. In one experiment, radioactive PSTVd$_{lh}$ was annealed in HS or Mg^{2+} buffer with an excess of nonradioactive PSTVd$_{uh}$, and most radioactivity was found in the bands marked M and H in lane 8 of Figure 2B, whereas almost no radioactivity (but high amounts of silver-stained

material) was present in the band marked D (data not shown). This means that heating and slow cooling of $PSTVd_{uh}$ and $PSTVd_{lh}$ leads to a mixture of monomeric $PSTVd_{lh}$, dimeric $PSTVd_{uh}$, and highly branched heteroduplex, but not to viroid-like rods. Eventually, a faint band of duplex molecules comigrating with the rod-like viroid conformer could be obtained when radioactive $PSTVd_{lh}$ was annealed with a large excess of unlabeled $PSTVd_{uh}$ (0.2 μg/μL) in HS buffer (but not in Mg^{2+} buffer, data not shown).

Secondary Structure Probing of the Upper Strand of PSTVd (+) RNA. Several of our RNA transcripts were end labeled and subjected to limited nuclease digestion in an effort to find cleavage sites in single-stranded (RNases T1, T2, and A and nuclease S1) or double-stranded regions (nuclease V1). The data from a typical experiment are shown in Figure 3. The half-molecule $PSTVd_{uh}$ apparently folds into four hairpins (Figure 4A). The best fit to the cleavage experiments is obtained with the third best structure (−184 kJ/mol at 37 °C) computed with Zuker's thermodynamic folding program *mfold* (*17*), with only three minor alterations. (i) A small hairpin between nucleotides 22 and 32 is ignored because of extensive nuclease cleavage. (ii) Due to the lack of cleavage, noncanonical G·A and A·A pairs might be present in HP I. (iii) In HP III, a base pair is added to the inaccessible apical loop and an accessible internal loop containing a G·U pair is opened.

To make sure that HP I and III were correctly identified, we also gel purified and sequenced fragments of a $PSTVd_{uh}$ molecule that had been partially digested with RNase T1, which would detect the most prominent cleavage positions. Products generated from monomeric and dimeric $PSTVd_{uh}$ were almost identical. The major fragments consisted of sequences from either HP I, HP III, or both; we found molecules comprising ca. nucleotides 77−115, 125−173, and 126−173 (data not shown), indicating that thermodynamically stable domains are also folding domains and that the structure in Figure 4A is correct.

Secondary Structure Probing of the Lower Strand of PSTVd RNA. In contrast to $PSTVd_{uh}$, the other half-molecule $PSTVd_{lh}$ could not be efficiently labeled at either terminus. Some 3′-labeling could be achieved by ligation to [^{32}P]pCp at 37 °C. To overcome these problems, we probed a shorter 5′-labeled HP II molecule (nucleotides 197−330). The results obtained with the HP II molecule were compatible with those obtained with the entire $PSTVd_{lh}$ RNA (Figure 4B), namely, that a stable HP II dominates the folding of the lower half of PSTVd, as expected from previous experiments (*9*). The structure derived from cleavage experiments corresponds to the fourth best structure calculated with *mfold* (*17*); we had to alter the base pairing of HP I′ at the tip of HP II. Single-strand-specific nicks within putative double-stranded regions might indicate the presence of alternative RNA conformers which can also be inferred from the smear detected on nondenaturing PAGE (Figure 2B). In contrast to what is expected from the computed optimal RNA structure, the U-rich region around position 300 is not involved in base pairing, much like the oligo(A) stretch of $PSTVd_{uh}$ (Figure 4A), which opposes this U run in the native viroid (Figure 1; *9*, *11*) and which is conserved in most viroids (*1*−*3*).

Secondary Structure Probing of Hairpin I Molecules. The inability of single-strand-specific nucleases to cleave within the internal loops of HP I (Figure 4A) was confirmed with short transcripts spanning the HP I regions in the upper C domains of PSTVd and grapevine yellow speckle viroid (GYSVd), a member of a different viroid group (*18*). Both HP I molecules contained unpaired G nucleotides at the 5′-end for efficient transcription and 5′-labeling (Figure 5; *16*). All known HP I's harbor a palindromic apical loop and unpaired G's opposed by A's (*3*). At the low concentrations

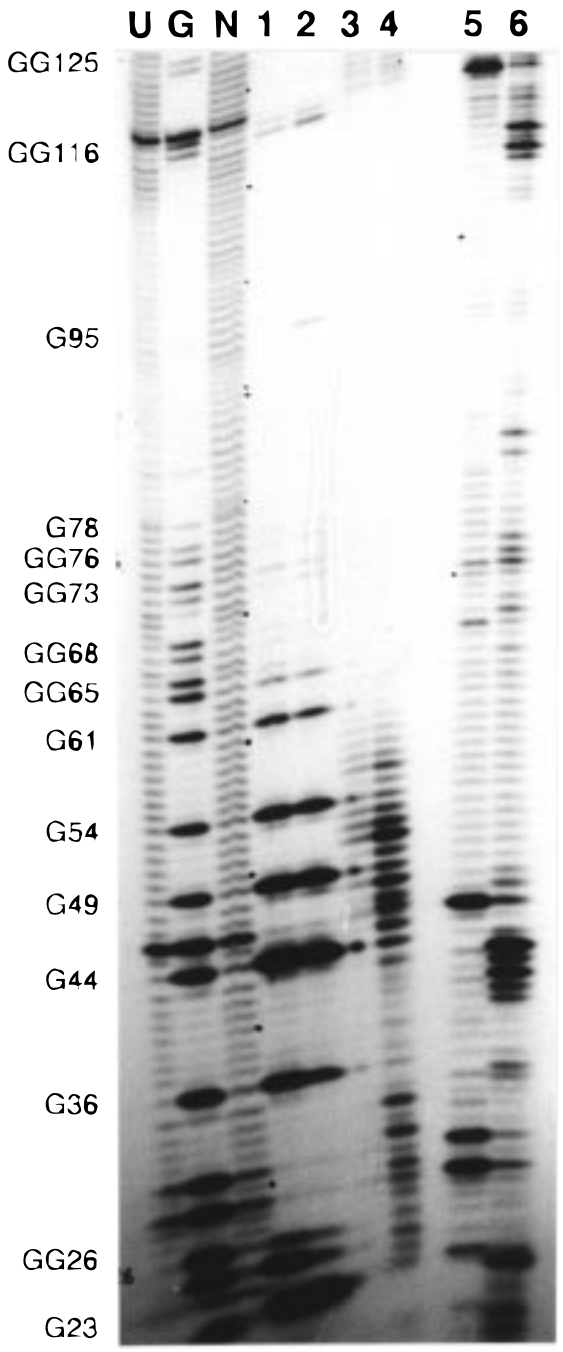

FIGURE 3: Part of a 6% sequencing gel representing the enzymatic probing of the LS form of 5′-labeled $PSTVd_{uh}$. The lanes are labeled as follows: U, untreated sample; G, G-specific sequencing with RNase T1 at 56 °C; N, OH^- cleavage; 1, RNase T1 cleavage at room temperature in the absence of salt; 2, RNase T1 cleavage at room temperature in the presence of 5 mM $MgCl_2$; 3, S1 cleavage at 37 °C; 4, nuclease S1 cleavage at neutral pH in the presence of 20 mM $MgCl_2$ (see ref *11*); 5, RNase A cleavage; and 6, nuclease V1 cleavage. The numbering of the G nucleotides according to the standard PSTVd numbering (*11*, *15*; see Figure 1) is indicated on the gel margin.

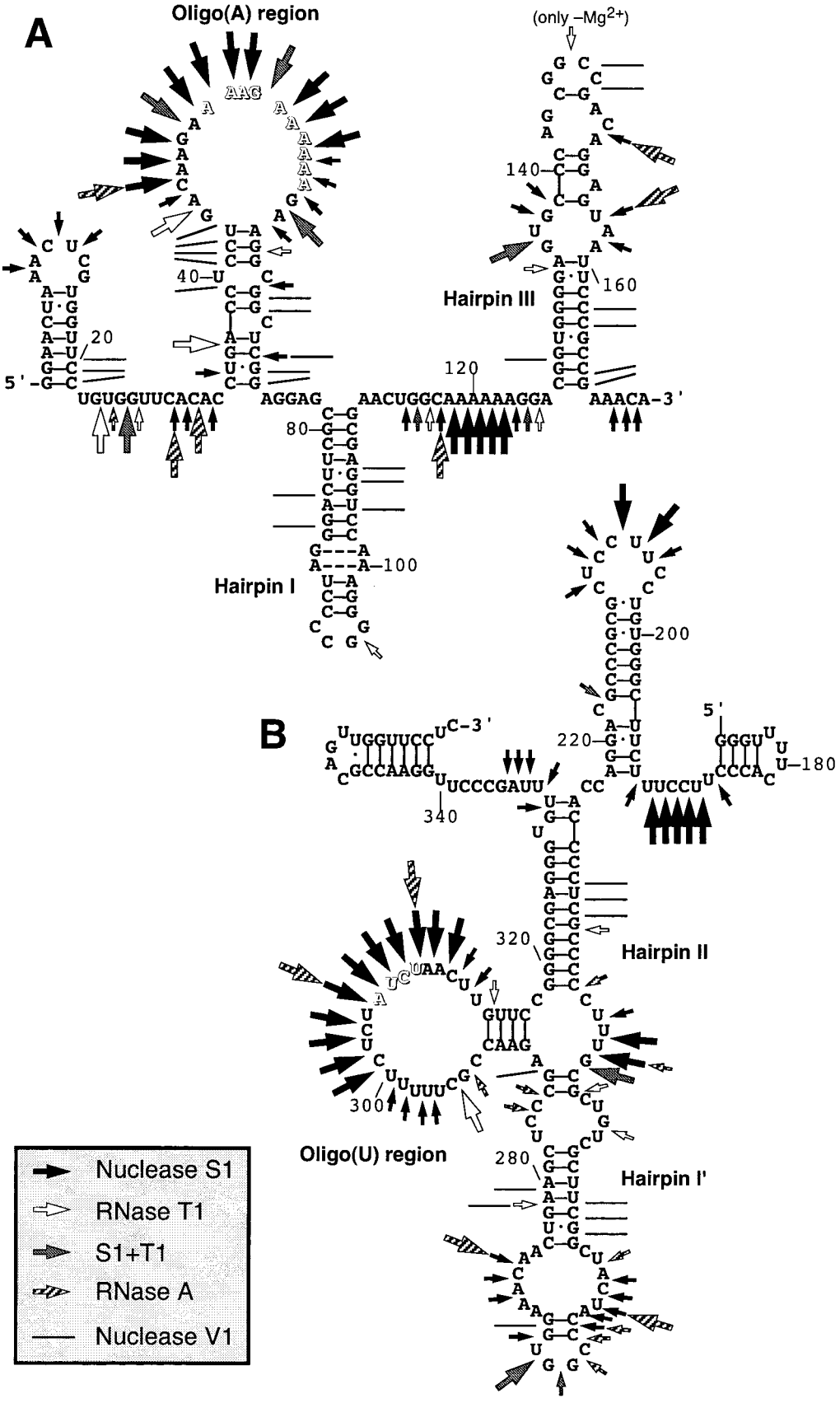

FIGURE 4: Structures of subgenomic RNA transcripts of (+) PSTVd as deduced from nuclease cleavage (arrows explained in the legend). (A) Probing of monomeric 5′- and 3′-labeled $PSTVd_{uh}$. Nucleotides within the oligo(A) loop that are invariant in all known PSTVd isolates are drawn outlined. (B) Combined results of 3′-labeled $PSTVd_{lh}$ and 5′-labeled HP II experiments (nucleotides 197−330; 3′-labeling was impossible). The invariant sequence AUCU in the oligo(U) loop is shown outlined. $PSTVd_{lh}$ is drawn in the same orientation as in the rod-like model of Figure 1; i.e., the 5′-end is at the right side.

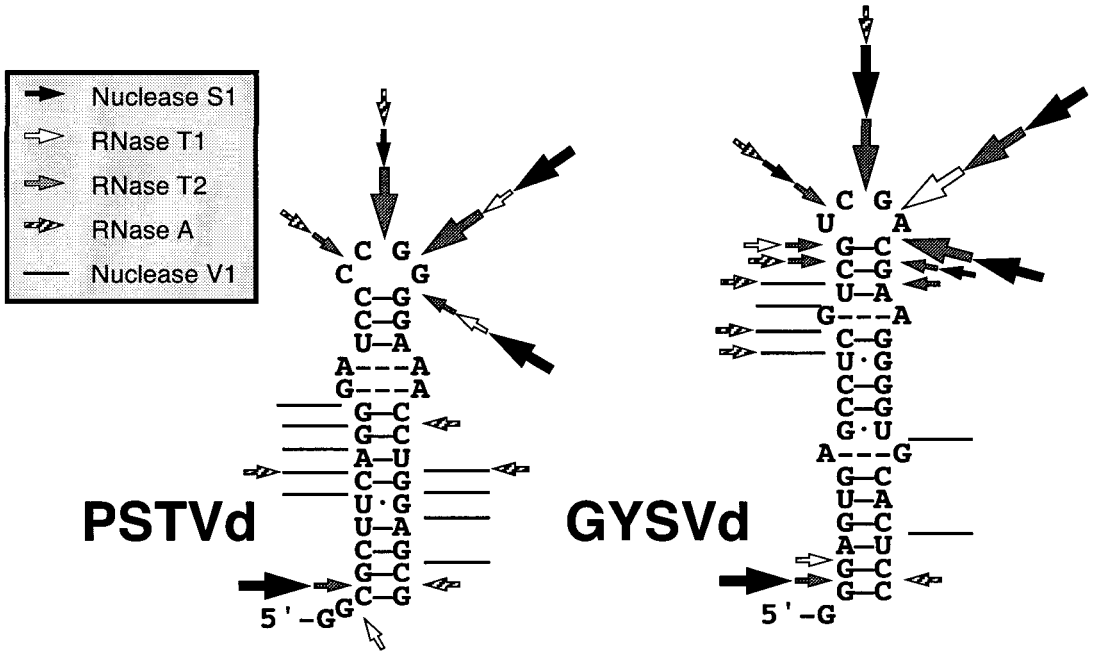

FIGURE 5: Nuclease probing of 5′-labeled HP I molecules of PSTVd and GYSVd. The base-paired structure corresponds to nucleotides 79−110 in PSTVd (variant KF 440-2; GenBank file PSTVD440) and nucleotides 87−122 in GYSVd (the variant from ref *18*; GenBank file GYSVXX). Additional G nucleotides at the 5′-end are used for efficient transcription (*16*) and labeling; this results in a G which is not present in the PSTVd sequence, whereas all G's of GYSVd are present in the complete viroid sequence. Arrows are explained in the legend.

used, only the monomeric hairpins form. Figure 5 shows the nuclease positions determined from 5′-labeled molecules. Except for a few unspecific RNase A sites under low-salt conditions and strong S1 signals at the base of the major helix, probably caused by fraying, single-strand-specific nicks occur only in the apical loop, with PSTVd being more stable than GYSVd, whereas double-strand specific nicks are distributed all over the rest of the hairpin, including the internal loops. Thus, the internal loops stabilize the stem of the hairpin, probably by forming noncanonical base pairs.

Determination of the Dimerization Site by Nuclease Digestion of the PSTVd$_{uh}$ Dimer and by Binding of Antisense Oligodeoxynucleotides. Nonradioactive monomers and dimers of PSTVd$_{uh}$ were produced (by heating and slow cooling in either the absence or presence of Mg^{2+}), and these complexes were digested (both in the presence of the same amount of Mg^{2+}) with high concentrations (20 units/μL) of RNase T1. After cleavage for 16 h at room temperature, only two slowly migrating bands from the dimer and two faster migrating bands from the monomer could be detected in a denaturing gel. They were excised, 5′-labeled, and sequenced. The legible sequences from the monomer were UUCAGG-GAUCCCCG and GAAACNUGGAGC; these fragments correspond to the proximal and distal part of HP I. That the most stable hairpin is cleaved in its terminal loop and not in an internal loop is another indication of G·A and A·A base pairs in HP I. Sequencing of the slower of the two bands from the T1 digest of the dimer is shown in Figure 6A. The longer fragment encompasses ca. nucleotides 79−115, and the shorter one ca. nucleotides 79−110. Thus, HP I is the dimerization domain. A possible pairing scheme of the dimeric HP I is shown in Figure 6B.

This result was confirmed by competition with antisense oligodeoxynucleotides, as used previously for the determination of the dimerization site of retroviruses (*19*, *20*). Only the oligonucleotide directed against the entire HP I region (i.e., that used for HP I transcription, nucleotides 79−110 and an antisense T7 promoter) inhibited dimerization of PSTVd$_{uh}$ in Mg^{2+} buffer, whereas oligonucleotides directed against other regions, such as the left quarter of HP I and the adjacent region (nucleotides 68−84) or the right half of HP III (nucleotides 152−173), bound but did not prevent dimerization (data not shown). This could be reproduced in HS and Mg^{2+} buffer with two full-length PSTVd transcripts (nucleotides 2−359/1 and 174−359/1−173). Figure 2A shows that the block of dimerization by the antisense oligodeoxynucleotide against HP I is also exerted on circular PSTVd molecules (lane 8). The resulting monomeric RNA migrates like the branched monomer (BM) of circular PSTVd and of PSTVd-specific transcripts. When end-labeled oligonucleotides were used, they were shifted to the position(s) of the RNA transcripts on the gel. The oligodeoxynucleotide complementary to nucleotides 68−84, which did not inhibit dimerization, bound equally well to the monomeric and dimeric form of both PSTVd$_{uh}$ and the full-length PSTVd transcript (nucleotides 174−359/1−173). In contrast, the oligonucleotide directed against the entire HP I (nucleotides 79−110) bound almost exclusively to the monomeric RNAs (data not shown), thus confirming the conclusions that have been drawn.

DISCUSSION

How viroids replicate and cause pathogenicity is still an unresolved question (*1*, *2*). Because RNA polymerase II in the nucleoplasma is responsible for viroid replication (*4*), it is conceivable that accumulation of the mature, circular (+) strand in the nucleoli (*21*) is the final and not the primary event of their replication cycle. Circularity prevents degradation by exonucleases, and the base-paired and rod-like structure confers stability against single-strand-specific en-

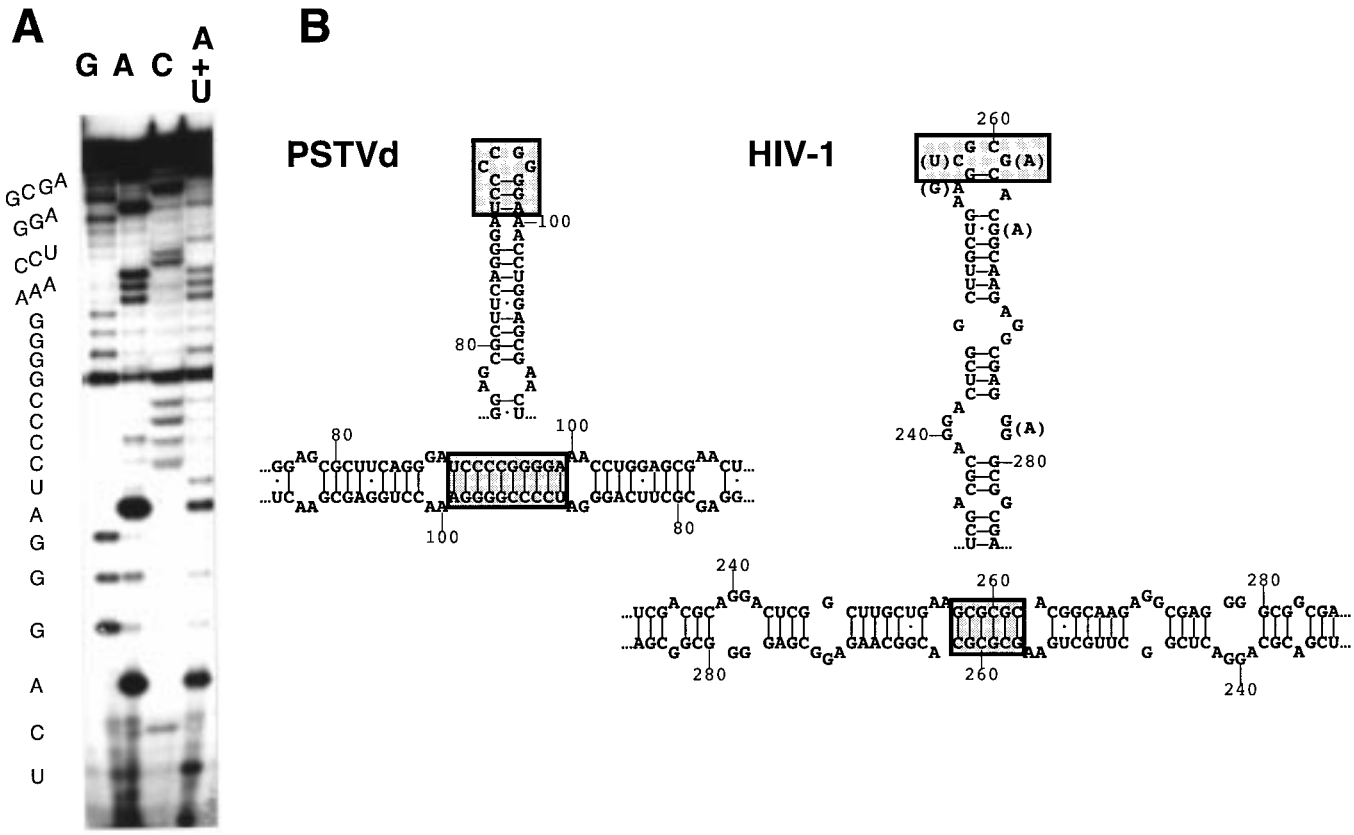

FIGURE 6: Identification of the dimerization site of PSTVd$_{uh}$ by RNase T1 digestion. (A) A 20% gel of the sequencing of the larger of the two T1-resistant cores of the dimer. The lanes are labeled according to the specificity of the RNases: G (T1), A (U2), C (CL3), and A + U (Phy M). The pyrimidine-specific *Bacillus cereus* RNase did not cleave in this region due to extensive secondary structure. The sequence of the smaller T1-resistant fragment is identical, but truncated at its 3′-end (data not shown). (B) Proposed structures of the dimerization sites of transcript PSTVd$_{uh}$ (left, this work) and HIV-1 (from ref *27*; compensatory mutations in parentheses). Hairpins are shown at the top and dimeric structures at the bottom. The palindromic tips of the hairpins are boxed. It is possible (and likely for the full-length viroid RNA molecules) that viroid dimerization involves only the boxed palindromic region, leading to a kissing loop complex (*31*).

donucleases. On the other hand, the rod-like structure is incompletely base-paired, making viroids a poor substrate for double-strand-specific RNases and allowing some melting and refolding during the life cycle. When total RNA is quickly phenol extracted from PSTVd-infected tomato leaves and separated on nondenaturing gels, (+) PSTVd RNA comigrates with the rod-like conformer formed in vitro (unpublished data). This extraction method is prone to artifacts, but since the different conformers of viroid RNA are stable against phenol and SDS (unpublished data), there is a good chance that the rod-like (+) strand conformation is the major viroid RNA form in planta.

During replication, however, the compact viroid structure must rearrange into at least a multi-hairpin intermediate, which is also observed during in vitro transcription of viroid-specific RNA transcripts (*22*) and which possibly resembles the structures of the isolated upper and lower strand (Figure 4). Formation of such intermediates, with some of the hairpins conserved (Figure 1; refs *8*, *9*, and *23* and our own unpublished sequence alignments), appears to be the key event in replication. Since the folding of the RNAs appears to be kinetically controlled, the exact nature of the (+) as well as the (−) RNA intermediates depends on the start position of the replication, which is unknown.

Diener (*12*) has suggested that dimerization of the palindromic HP I of an oligomeric (+) strand, which could be visualized for in vitro transcripts by electron microscopy (*24*), forms the site of processing of the oligomeric (+) strands into circular monomers. Results from experiments using plant nuclear extracts (*6*) seemed to support this view. However, the processing of model RNA transcripts by newly characterized specific endonucleases from nuclear extracts (*7*, *8*) or from plant nucleoli (our own unpublished work) does not occur in the trihelical region, but in such RNA conformers which either adopt a rod-like structure or contain a GAAA tetraloop. Whatever the function of this region, it seems to be fundamental. A conserved hairpin whose loop is perfectly palindromic in all viroid groups, although their central conserved regions differ (*3*), is unlikely to have evolved by chance.

Given the existence of a palindromic hairpin, we expected viroid RNA to dimerize also intermolecularly, which we found. This event could be enhanced by using subgenomic RNA fragments of PSTVd in which the complementarity between the upper and lower strand was disrupted. We found that the dimerization signal is located in the upper strand of the viroid RNA. Dimerization does not depend on K^+ ions and appears to be more specific in the presence of Mg^{2+} than in the presence of monovalent cations.

In our experiments, annealing of the upper and lower strands to a viroid-like duplex, which is the viroid RNA structure found in vitro (*11*), was less likely than dimerization of two molecules of the upper strand (Figure 2B). The monomers of the isolated upper and lower PSTVd strands adopt stable folds whose fine structures were probed enzymatically. Almost the entire lower region is involved

in the formation of HP II (*9*, *23*), whereas in the upper region several smaller hairpins form: one comes from the left terminal branch (Figure 1), one spans the entire oligo(A) region, and the others are the well-known hairpins I and III (Figure 4). Under certain circumstances, e.g., during synthesis of the (+) RNA strand in vivo, the folding of such half-molecules may be fast enough to compete with the stable pairing to the upper PSTVd strand as in the rod-like model of Figure 1. When upper and lower viroid strands are folded into their hairpin structures (Figure 4), an interaction with one another which leads to a heteroduplex would be possible only along very short single-stranded stretches. In fact, annealing of equimolar amounts of the upper and lower strand of PSTVd in vitro resulted only in highly branched heteroduplexes (lane 8 of Figure 2B), indicating that the folding of the single molecules is quicker than duplex formation.

That such metastable structures are not exotic is illustrated by the fact that cDNAs, random primed and reverse transcribed from purified circular viroids, are restricted to loop regions of large stable hairpins (like HP II) that can form at high temperatures (*25*). Formation of intramolecularly folded multi-hairpin structures instead of viroid-like rod molecules might also explain our inability to induce viroid replication by coinoculation of the subgenomic fragments $PSTVd_{uh}$ and $PSTVd_{lh}$ into host plants (unpublished data).

Dimerization of RNA genomes also plays a role in other viral systems. HIV-1, as an example of a retrovirus, is diploid and has a dimerization domain which is thought to be important for its life cycle, e.g., by facilitating encapsidation or recombination (*19*, *26*−*31*, and references therein). The region involved has been debated, but now appears to be structurally similar to the region within the central conserved region of PSTVd and other viroids (cf. Figures 5 and 6B), because it consists of a hairpin with a palindromic loop that can dimerize via "kissing hairpin loops" (*19*, *26*−*31*).

Limited nuclease digestion of the dimer of the upper PSTVd strand and competition with antisense oligonucleotides indicate that viroids can indeed dimerize via the palindromic HP I. This process may involve the entire sequence of HP I (Figure 6B), possibly including noncanonical G·A and A·A pairs whose high frequencies in ribosomal RNA have been outlined recently (*32*). Since higher-order complexes of the upper viroid RNA strand form at high RNA concentrations, other regions must also be able to dimerize. We suggest that HP III (which contains in its loop the palindromic sequence CGGCCG) can bind to a second molecule, and two dimerization sites cause oligomerization. This is supported by the molecule containing nucleotides 2−146 which lacks one-half of HP III; it can dimerize via HP I, but does not form oligomers (lanes 5 and 6 of Figure 2B).

Instead of the viroid pairing across the entire trihelical region (as shown in Figure 6B), viroid dimerization might occur via a kissing loop, i.e., encompassing only the 10 central nucleotides, which would be essentially similar to the HIV-1 dimerization loop (*31*). At the moment, our data cannot discriminate between both possibilities. However, whereas both schemes are equally likely in linear viroid RNA transcripts, kissing hairpin loops are the only way for an interaction between two covalently closed circular viroid molecules to occur, since it seems to be impossible to wrap three helical RNA turns around each other as in Figure 6B without opening the circle in another position.

It is feasible that even the viroid half-molecule interacts with itself via a kissing hairpin loop. The first indication for this may be that an oligodeoxynucleotide complementary to nucleotides 68−84 (see Figure 6B for numbering) is unable to prevent dimerization of $PSTVd_{uh}$ even though it binds to a large extent to HP I. The occurrence of dimeric molecules during transcription, i.e., far below the monomer−dimer transition temperature of ∼60 °C (data not shown), might also be due to base pairing of only a short RNA stretch, such as the 10 nt palindromic loop of HP I. Finally, two bands representing dimers of the upper strand of PSTVd are visible after heating and cooling in HS buffer (lane 2 of Figure 2B). This might correspond to the two possible dimeric states of HP I, namely, a fully base-paired duplex of ∼32 base pairs (as shown in Figure 6B) and a kissing loop complex of 10 base pairs. Heating and slow cooling in Mg^{2+}-containing buffer results in the formation of the lower of the two dimer bands, whereas heating and snap-cooling in Mg^{2+} buffer or heating and cooling in LS buffer increases the fraction of the upper of the two dimer bands. Because intermolecular base pairing should be less likely under low-salt conditions (due to insufficient counterions) or upon snap-cooling (due to the kinetic barrier), it is tempting to speculate that the lower band contains a full duplex dimer and the upper band contains a kissing loop complex. Of course, further studies are needed to clarify this point.

If there is a functional similarity of the dimerization domain of viroids to that of HIV-1, HP I might not be important for processing, but a signal for packaging, replication, or recombination. Not much is known about these processes in the viroid system; it has been shown that PSTVd can be encapsidated (*33*; our own unpublished results), but it is unclear what triggers this event. Dimerization via palindromic hairpin loops has been observed with RNA regions at the 5′-end of several retroviruses, including HIV-1 (*19*, *26*−*31*), Moloney murine leukemia virus (*20*), Harvey sarcoma virus (*34*), human T cell leukemia virus type 1 (*35*), and avian sarcoma-leukosis virus (*36*), whereas dimerization of the Rous sarcoma virus does not rely on a palindrome (*37*). In light of all these results, it might well be that dimerization is a common feature of infectious RNA replicons, their different hosts and replication strategies notwithstanding. Since we have found a process in a viral system from plants similar to those others have found in viral systems from animals, one must now ask what these systems have in common, and this warrants the further clarification of these events wherever they are observed.

ACKNOWLEDGMENT

The kind gift of the T7 RNA polymerase-overproducing strain by Dr. F. W. Studier (Brookhaven National Laboratory, Upton, NY) and the critical reading of the manuscript by Drs. U. Pieper and A. Pingoud (Justus-Liebig-Universität Giessen, Giessen, Germany) are gratefully acknowledged. We also thank the reviewers for their detailed comments and corrections.

REFERENCES

1. Diener, T. O., Ed. (1987) *The Viroids*, Plenum Press, New York.
2. Symons, R. H. (1991) *Mol. Plant-Microbe Interact. 4*, 111−121.
3. McInnes, J. L., and Symons, R. H. (1991) in *Viroids and Satellites: Molecular Parasites at the Frontier of Life* (Maramorosch, K., Ed.) pp 21−58, CRC Press, Boca Raton, FL.
4. Rackwitz, H.-R., Rohde, W., and Sänger, H. L. (1981) *Nature 291*, 297−301.
5. Branch, A. D., and Robertson, H. D. (1984) *Science 223*, 450−455.
6. Tsagris, M., Tabler, M., Mühlbach, H.-P., and Sänger, H. L. (1987) *EMBO J. 6*, 2173−2183.
7. Baumstark, T., and Riesner, D. (1995) *Nucleic Acids Res. 23*, 4246−4254.
8. Baumstark, T., Schröder, A. R. W., and Riesner, D. (1997) *EMBO J. 16*, 599−610.
9. Riesner, D. (1990) *Semin. Virol. 1*, 83−99.
10. Wassenegger, M., Heimes, S., and Sänger, H. L. (1994) *EMBO J. 13*, 6172−6177.
11. Gast, F.-U., Kempe, D., Spieker, R. L., and Sänger, H. L. (1996) *J. Mol. Biol. 262*, 652−670.
12. Diener, T. O. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 58−62.
13. Pace, U., Branch, A. D., and Robertson, H. D. (1992) *Nucleic Acids Res. 20*, 6681−6686.
14. Gast, F.-U., and Sänger, H. L. (1994) *Electrophoresis 15*, 1493−1498.
15. Tabler, M., and Sänger, H. L. (1985) *EMBO J. 4*, 2191−2199.
16. Milligan, J. F., and Uhlenbeck, O. C. (1989) *Methods Enzymol. 180*, 51−62.
17. Zuker, M. (1989) *Science 244*, 48−52.
18. Koltunow, A. M., and Rezaian, M. A. (1988) *Nucleic Acids Res. 16*, 849−864.
19. Muriaux, D., Girard, P.-M., Bonnet-Mathonière, B., and Paoletti, J. (1995) *J. Biol. Chem. 270*, 8209−8216.
20. Girard, P.-M., Bonnet-Mathonière, B., Muriaux, D., and Paoletti, J. (1995) *Biochemistry 34*, 9785−9794.
21. Harders, J., Lukács, N., Robert-Nicoud, M., Jovin, T. M., and Riesner, D. (1989) *EMBO J. 8*, 3941−3949.
22. Matousek, J., Schröder, A. R. W., Trnená, L., Reimers, M., Baumstark, T., Dedic, P., Vlasák, J., Becker, I., Kreuzaler, F., Fladung, M., and Riesner, D. (1994) *Biol. Chem. Hoppe-Seyler 375*, 765−777.
23. Qu, F., Heinrich, C., Loss, P., Steger, G., Tien, P., and Riesner, D. (1993) *EMBO J. 12*, 2129−2139.
24. Steger, G., Tabler, M., Brüggemann, W., Colpan, M., Klotz, G., Sänger, H. L., and Riesner, D. (1986) *Nucleic Acids Res. 14*, 9613−9630.
25. Francis, M. I., Szychowski, J. A., and Semancik, J. S. (1995) *J. Gen. Virol. 76*, 1081−1089.
26. Skripkin, E., Paillart, J.-C., Marquet, R., Ehresmann, B., and Ehresmann, C. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 4945−4949.
27. Laughrea, M., and Jetté, L. (1994) *Biochemistry 33*, 13464−13474.
28. Paillart, J.-C., Marquet, R., Skripkin, E., Ehresmann, B., and Ehresmann, C. (1994) *J. Biol. Chem. 269*, 27486−27493.
29. Clever, J. L., Wong, M. L., and Parslow, T. G. (1996) *J. Virol. 70*, 5902−5908.
30. Laughrea, M., Jetté, L., Mak, J., Kleiman, L., Liang, C., and Wainberg, M. A. (1997) *J. Virol. 71*, 3397−3406.
31. Paillart, J.-C., Westhof, E., Ehresmann, C., Ehresmann, B., and Marquet, R. (1997) *J. Mol. Biol. 270*, 36−49.
32. Gautheret, D., Konings, D., and Gutell, R. R. (1994) *J. Mol. Biol. 242*, 1−8.
33. Querci, M., Owens, R. A., Bartolini, I., Lazarte, V., and Salazar, L. F. (1997) *J. Gen. Virol. 78*, 1207−1211.
34. Feng, Y. X., Fu, W., Winter, A. J., Levin, J. G., and Rein, A. (1995) *J. Virol. 69*, 2486−2490.
35. Greatorex, J. S., Laisse, V., Dokhelar, M.-C., and Lever, A. M. L. (1996) *Nucleic Acids Res. 24*, 2919−2923.
36. Fosse, P., Motte, N., Roumier, A., Gabus, C., Muriaux, D., Darlix, J. L., and Paoletti, J. (1996) *Biochemistry 35*, 16601−16609.
37. Lear, A. L., Haddrick, M., and Heaphy, S. (1995) *Virology 212*, 47−57.

BI980830D