Available online at www.sciencedirect.com

ScienceDirect

Virology 373 (2008) 362–375

VIROLOGY

www.elsevier.com/locate/yviro

The peach latent mosaic viroid replication initiation site is located at a universal position that appears to be defined by a conserved sequence

Julie Motard [a,1], François Bolduc [a,1], Dan Thompson [b], Jean-Pierre Perreault [a,*]

[a] *RNA group/Groupe ARN, Département de biochimie, Faculté de médecine et des sciences de la santé, Université de Sherbrooke, 3001, 12e Avenue Nord, Sherbrooke, Quebec, Canada J1H 5N4*

[b] *Center for Plant Health, Canadian Food Inspection Agency, Sidney, British Columbia, Canada V8L 1H3*

Received 15 October 2007; returned to author for revision 18 November 2007; accepted 10 December 2007

Available online 10 January 2008

Abstract

Viroids replicate through a rolling circle mechanism that is exclusively RNA dependent. In this study, we initially revisited the determination of the replication initiation sites of peach latent mosaic viroid (PLMVd). A universal initiation site for each of the PLMVd polarities (position A_{50}/C_{51} and U_{284} for the plus and minus strands, respectively) that is shared by a relatively wide repertoire of viroid variants was identified, in agreement with a previous report based on a different methodology. Subsequently, an *in vitro* selection procedure based on a model rolling circle replication assay was developed. This latter experiment led to the identification of a highly conserved CAGACG box which is reminiscent of the sequence found in the vicinity of the PLMVd initiation sites. The conserved sequence contributes to delineating the initiation site and provides an explanation for the presence of a specific universal initiation site on the PLMVd molecule.

Keywords: Viroid; RNA replication; SELEX; Initiation site; Molecular evolution

Introduction

Viroids are small (~240–460 nucleotides, nt), single-stranded, circular RNA pathogens that infect higher plants and cause important economical losses in agricultural industries (for a review see Daros et al., 2006). These RNA species replicate through a rolling circle mechanism involving exclusively RNA intermediates, and follow either a symmetric or an asymmetric mode. In the symmetric mode, the infecting circular monomers, which are designated as being of plus (+) polarity by convention, are replicated into linear, multimeric, minus (−) strands (Fig. 1A). The latter are then cleaved and ligated, producing the minus circular monomers that then serve as the templates, through the same three steps (polymerization, cleavage and ligation), for the production of the plus (+) strand. Conversely, in the asymmetric mode, the linear multimeric minus strands serve directly as the template for the synthesis of linear multimeric strands of plus polarity that are then cleaved and ligated in order to produce the progeny.

Viroids have been classified into two families according to both sequence conservation (Flores et al., 2000) and biological hallmarks (Bussière et al., 1999). *Pospiviroidae* members, for which the typical representative species is the *Potato Spindle Tuber Viroid* (PSTVd), share a conserved central region, a nuclear cellular localization and an asymmetric replication that appears to involve host enzymes at each step. In contrast, *Avsunviroidae* members are localized into the chloroplasts and exhibit a symmetric replication mode. *Peach Latent Mosaic Viroid* (PLMVd), which possesses an RNA genome of 336 to 351 nt (Fekih Hassen et al., 2007) and is the causal agent of peach latent mosaic disease (Desvignes, 1986), belongs to the *Avsunviroidae*. The cleavage and ligation steps of the rolling circle replication mechanism of PLMVd have been extensively studied. Some of these studies have shown that the multimeric intermediates in PLMVd replication are processed by a self-catalytic hammerhead motif, while the circularization of the

* Corresponding author. Fax: +1 819 564 5340.
E-mail addresses: Julie.Motard@usherbrooke.ca (J. Motard), Francois.Bolduc@usherbrooke.ca (F. Bolduc), Thompsonda@inspection.gc.ca (D. Thompson), Jean-Pierre.Perreault@USherbrooke.ca (J.-P. Perreault).
[1] These authors contributed equally to this study.

0042-6822/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.virol.2007.12.010

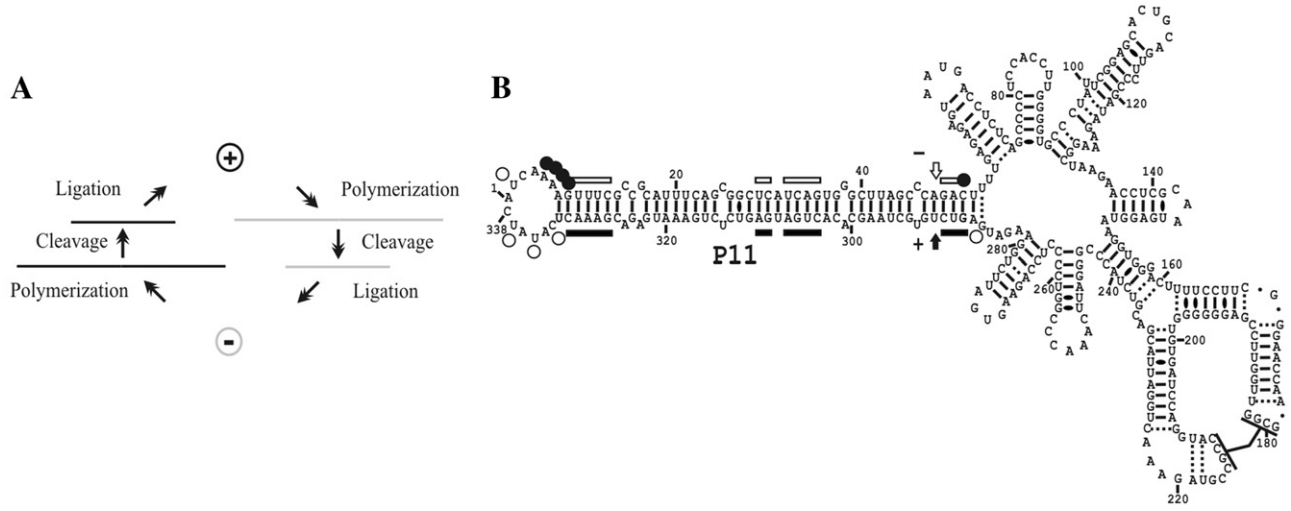

Fig. 1. Rolling circle replication mechanism and structure of PLMVd. (A) Proposed symmetrical rolling circle replication of viroids belonging to the *Avsunviroidae* family. The minus (−) and plus (+) strands are represented by grey and black lines, respectively. (B) Nucleotide sequence and secondary structure of the PLMVd variant used in this work (Siberian C) as defined by Bussière et al. (2000). The arrows indicate the hammerhead self-cleavage sites and the boxes the conserved nucleotides that form the hammerhead motifs. The white objects indicate minus (−) polarity and those in black plus (+) polarity. The black and white circles identify the initiation sites of the plus (+) and minus (−) polarity strands, respectively.

monomeric linear strand is ensured by a non-enzymatic mechanism of self-ligation that produces a characteristic 2′,5′-phosphodiester linkage (Hernandez and Flores, 1992; Beaudry et al., 1995; Côté and Perreault, 1997; Côté et al., 2001). These findings led to the suggestion that the replication of PLMVd relies primarily on a viroid RNA-based mechanism rather than on the involvement of host enzymes as is the case for members of the *Pospiviroidae* family (Côté et al., 2001). The differential subcellular localizations of these two viroid families (chloroplasts versus nucleus) are most likely the main reason for these different replication mechanisms.

The least understood step of PLMVd replication mechanism is the polymerization step in which the circular monomeric RNA strands of both polarities serve as templates. The plastid transcriptional machinery comprises at least two DNA-dependent RNA polymerases (RNAP): the nucleus-encoded polymerase (NEP), which is a single subunit phage-like RNAP; and the plastid-encoded polymerase (PEP), which is a multi-subunit bacterial-like RNAP (Sakai et al., 2004). Recent reports on the identification of the enzymatic complexes that could be implicated in the replication of *Avsunviroidae* genomes indicated that both complexes are implicated in the replication mechanism. The *Avocado sunblotch viroid* (ASBVd) has been shown to replicate in avocado chloroplast extracts in the presence of tagetitoxin, a product known to be a strong inhibitor of both *Escherichia coli* RNAP and the homologous chloroplastic RNAP when they are transcribing DNA templates, suggesting the involvement of the NEP (Navarro et al., 2000). However, when using a commercial preparation of the *E. coli* RNAP holoenzyme, we observed transcription of an RNA template even in the presence of tagetitoxin (Pelchat et al., 2001). Regardless of the polarity of the PLMVd strand used as template, the RNA features triggering the RNAP to initiate specific *in vitro* transcription were shown to be located within the left-hand terminal hairpin (see Fig. 1B) (Pelchat et al., 2002). More specifically, it was shown that the first base-paired uridine was the primary initiation site. This model assay received indirect support from primer extension experiments performed using total RNA samples extracted from PLMVd-infected leaves and a primer derived from the PLMVd sequences located near the other extremity of the P11 stem. In the latter work, the initiation sites found were located at positions A_5, A_6, A_7, and G_8 for the strands of plus polarity, and at positions A_2, A_{337}, A_{335}, and G_{333} for the strands of minus polarity (Pelchat et al., 2001) (Fig. 1B). However, to date, there is no direct evidence for these initiation sites being consistently used in infected cells.

Recently, Delgado and colleagues adapted an RNA-ligase-mediated rapid amplification of cDNA ends (RLM-RACE), developed for the mapping of genuine capped 5′ termini of eukaryotic messenger RNAs (Delgado et al., 2005), to identify viroid initiation sites. They identified several initiation sites based on primer extension assays, but confirmed only a single one possessing a 5′-triphosphate terminus compatible with being a nascent transcript derived from polymerization. Using this elegant strategy, they determined that *in vivo* initiation might take place at only one position, specifically near the hammerhead cleavage site, for transcripts of both polarities: at C_{51} for the plus strand and at U_{286} for the minus strand (Fig. 1B). Both of these positions are located on the right hand side of the P11 stem at the opposite end to the sites reported by the *in vitro* assay.

In order to resolve this controversy, we decided to revisit the identification of the initiation sites using different approaches. The accumulated evidence indicates that the latter initiation sites appear to be universal, according to the analysis of total RNA obtained from several PLMVd-infected peach cultivars. Moreover, results from an *in vitro* selection procedure based on a model rolling circle replication reaction suggest that the initiation site is fixed by a conserved sequence located within the hammerhead motifs.

Results

Identification of universal PLMVd subgenomic fragments

PLMVd is a circular RNA species, consequently it is possible to imagine that the initiation of replication might occur at any point around this circular molecule. More specifically, once the first hammerhead sequence is synthesized, self-cleavage releases a fragment that should be smaller than 338 nt while the polymerase proceeds with the elongation of the first full-length copy. However, this simplistic view does not take into account that a viroid's sequence could trigger the activity of a host RNA polymerase, that is to say that it must serve as a promoter in addition to defining an initiation site. Consequently, both the promoter and the initiation site should be relatively well-conserved features between PLMVd variants. With this in mind, we initially investigated for the presence of a universal intermediary fragment resulting from the replication of the viroid.

In order to detect the proper fragment, RNA samples isolated from PLMVd-infected peach leaves were analyzed by Northern blot hybridizations using *in vitro* transcribed full-length PLMVd RNA as probe. Examples of autoradiograms are illustrated in Fig. 2 for the detection of both the plus and minus polarities of PLMVd. These autoradiograms confirmed that PLMVd replicates within infected cells, while the RNA extracted from healthy leaves was devoid of any that correspond to PLMVd. The autoradiograms were over-exposed so as to permit the detection of all possible RNA species. Analysis of these three infected RNA samples revealed 13 bands of plus polarity that were smaller than 338 nt (i.e., namely X1 to X13 in Fig. 2A). The bands X1 to X4, X10 and X11 were not considered since they can be formed from PLMVd transcripts synthesized *in vitro* and submitted to the same experiment, indicating that they result from specific degradation during the various treatments (data not shown). Conformers X5 to X7 were found in none of the three cultivars (Fig. 2A). The bands X4 to X6 correspond to RNA transcripts with a size of ~260 to 280 nt, whose length correlates with initiation sites located near the left-hand loop of the P11 stem. Only three out of thirteen conformers were present in each of the three cultivars, specifically conformers X8 (~260 nt), X9 (~240 nt) and X13 (~180 nt). The same experiments were performed with total RNA samples from PLMVd-infected leaves obtained from 8 different peach cultivars (see Supplementary Table 1 for the complete list of cultivars). Whatever the experiment and the cultivar, the only band that was consistently observed in all cases was the one that corresponds to a ~240 nt RNA species (i.e., X9). This species appears to be a unique and universal subgenomic fragment for plus polarity PLMVd. Similar experiments were performed for the identification of a universal fragment of minus polarity. In this case, the autoradiogram in Fig. 2B reveals the presence of only 5 species (X1 to X5) smaller than 338 nt. Analysis of a large collection of RNA samples from the various cultivars revealed that only the smallest species, corresponding to an RNA of ~230 nt (X5), was consistently detected (see Supplementary Table 1).

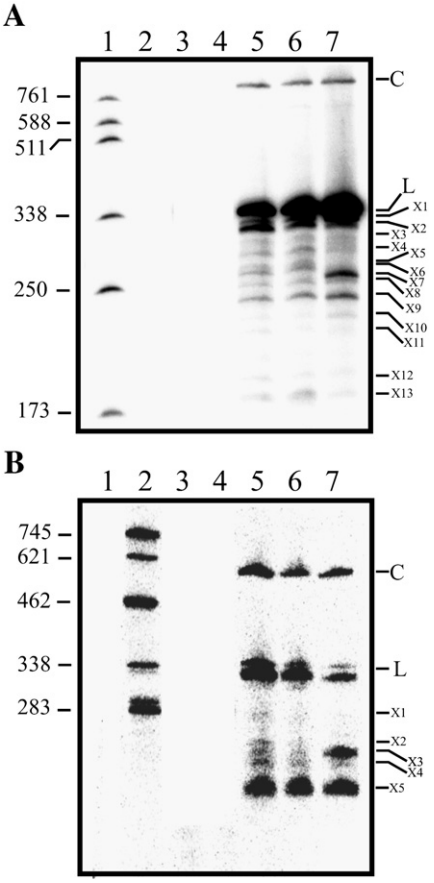

Fig. 2. Detection of PLMVd replication products in three different cultivars. Northern blot hybridizations of total RNA were used to detect plus (+) (A) and minus (−) (B) polarity strands using full-length radioactive PLMVd RNA as a probe. Prior to blotting the samples were fractionated on denaturing 5% PAGE gels. Lanes 1 and 2 are the transcripts and self-cleavage products of plus and minus polarity, respectively, produced by *in vitro* transcription using plasmid pPD1 as DNA template. Lanes 3 and 4 are total RNA extracted from the healthy leaves of the GF305 and Siberian C cultivars, respectively. Lanes 5 to 7 are total RNA samples extracted from the infected leaves of the Redgold, GF305 and Hardired variants, respectively. The numbers on the left refer to the sizes of both the transcripts and the self-cleavage products. The positions of the circular (C), linear (L) and subgenomic (X) conformers are indicated on the right.

Characterization of the universal RNA species

Initially, an RT–PCR amplification strategy was used in order to identify the region in which the 5′ end of the 240-nt fragment of plus polarity was located (i.e., band X9). Total RNA samples were fractionated on denaturing polyacrylamide gels (PAGE), and the gel bands corresponding to both the ~240 and 338 nt fragments excised. The RNA species extracted from the gel fragments were then used as templates for reverse transcription in the presence of either the antisense primer PLM250–275 (for the detection of plus polarity) or of PLM63–86 (for the detection of minus polarity), both of which are located within 20 bases of the self-cleavage site (see Figs. 3A and C). In order to identify the region containing the 5′ end of each RNA species, the resulting mixtures were used as cDNA templates for further amplifications in the presence of the same antisense primers and various sense primers (plus polarity:

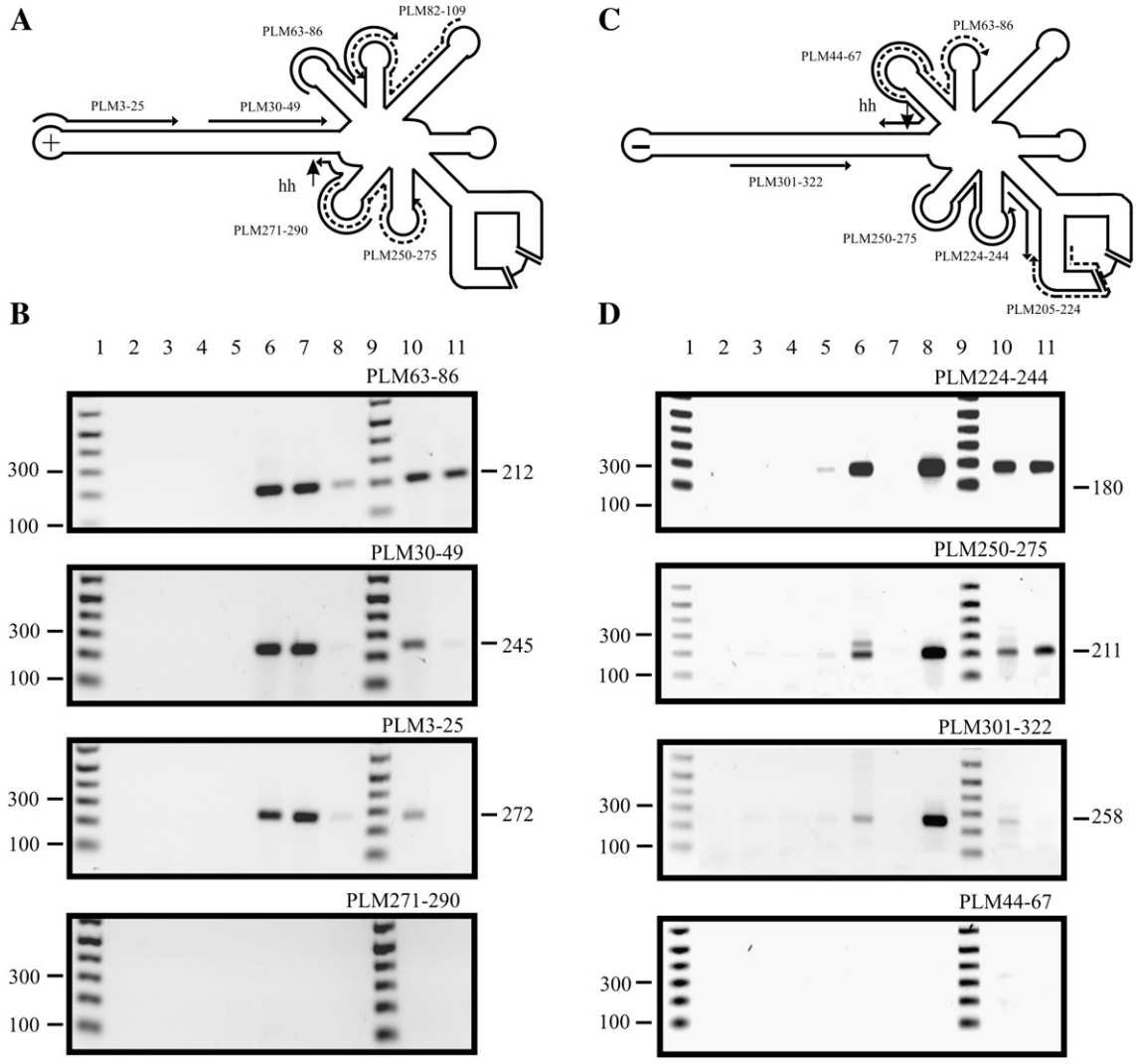

Fig. 3. Characterization of the bands X9 and X5 by RT–PCR. (A, C) Schematic representations of the secondary structure of PLMVd indicating the positions of the different primers used in the RT–PCR reactions for the plus (+) and minus (−) strands, respectively. The dotted lines refer to the reverse primers and the continuous lines to the forward primers. (B, D) Photographs of ethidium bromide stained 1% agarose gels of RT–PCR reactions performed for the plus (+) and minus (−) strands, respectively. Lanes 1 and 9 are aliquots of 1 kb Plus DNA ladder. Lanes 2 and 3 are negative controls that consisted of reverse transcriptase reactions and PCR reactions performed in presence of water instead of the RNA samples. Lane 4 is the pPD1 plasmid containing a head-to-tail dimeric cDNA of PLMVd. Lane 5 contains total RNA isolated from healthy Siberian C leaves. Lane 6 contains total RNA isolated from infected peach leaves from the Hardired variant. Lane 7 is *in vitro*-transcribed full-length PLMVd RNA of plus polarity. Lane 8 is *in vitro*-transcribed full-length PLMVd RNA of minus polarity. Lane 10 is a linear conformer (338 nt) isolated from peach leaves infected by the Hardired variant. Lane 11 is either the subgenomic conformer (X9) or (X5) for the plus (+) and minus (−) strands respectively, isolated from the peach leaves infected by the Hardired variant. Each panel is identified by the forward primer used in the PCR reaction. The numbers on the left refer to the size ladder (in bp). The numbers on the right correspond to the theoretical products obtained.

PLM63–86, PLM30–49, PLM3–25 and PLM271–290; minus polarity: PLM224–244, PLM250–275, PLM301–322 and PLM44–67). For example, in the presence of the primer PLM63–86 that is complementary to the sequence that forms both the P1 and P2 stems, PCR products were detected for both cDNAs indicating that the 5′ end of both species is located upstream of PLM63–86 (Fig. 3B, lanes 10 and 11). The sizes of the PCR products were always in agreement with the positions of the primers around PLMVd. When the experiments were performed using the primers PLM30–49 or PLM3–25, only the cDNA derived from the 338 nt RNA species resulted in the detection of PCR products. The absence of any PCR products with the pool of cDNA derived from the ~240 nt RNA species indicated that the 5′ end of the corresponding RNA species should be approximately located between positions 49 and 63 of PLMVd. As a negative control, we also tested the amplification of the cDNA derived from the 338 nt RNA species in the presence of the primer, PLM271–290, that overlaps with the antisense primer (see Fig. 3A). As expected, no PCR product was detected. Moreover, all of these PCR reactions were also controlled by testing various amplification conditions in order to validate the methods, as well as to establish the expect band sizes when using an *in vitro* synthesized PLMVd molecule (Fig. 3B).

The same strategy was used to estimate the 5′ end position of the ~230 nt species of minus polarity (i.e., band X5). The use of the primers PLM224–244 and PLM250–275 permitted the

detection of PCR products, while that of primer PLM301–322 did not. This indicates that the 5′ end of the ~230 nt species of negative polarity should be approximately located between positions 275 and 301.

In order to precisely determine the 5′ end of the RNA species, we adopted a primer extension strategy using the ^{32}P-5′-end labeled antisense primer PLM82–109 (Fig. 3A) for the detection of that of the plus strand and primer PLM205–224 for the detection of that of the minus strand (Fig. 3C). Gel bands corresponding to RNA conformers of 338 and 240 nt of plus polarity were purified and used as templates for primer extension analyses (Fig. 4A, lanes 4 and 5). We obtained several bands from the primer extension of the 338 nt RNA species. When using a full-length linear PLMVd of plus polarity generated by *in vitro* transcription as template, we observed the same pattern, indicating that each arrest was probably due to the structure of the RNA itself and not to *in vivo* processing (Fig. 4A, compare lanes 2 and 4). When the template was the 240 nt species, we only detected two stops, corresponding to A_{50} and C_{51}, that were not found in the case of the *in vitro*-transcribed RNA species (lane 5). In all of the experiments performed, the species corresponding to C_{51} was always more abundant than that corresponding to A_{50}. A similar strategy was used for the identification of the 5′ terminal nucleotide of the ~230 nt species (i.e., band X5) using the PLM205–224 primer. A unique stop, corresponding to U_{284}, was detected (see Fig. 4B).

Since all primary transcripts produced in chloroplasts are 5′ triphosphated (Vera and Sugiura, 1992; Meng et al., 1991), we attempted to define the nature of some of the subgenomic bands. A series of experiments using the Terminator 5′-monophosphate-dependent exonuclease (Epicentre Biotechnologies), a 5′–3′ exonuclease that digests RNA possessing a 5′-monophosphate, were performed. This enzyme does not digest RNA possessing either a 5′-triphosphate, a 5′-cap or a 5′-hydroxyl group. PLMVd-derived bands were treated with this enzyme either with or without prior T4 polynucleotide kinase treatment to add a 5′-monophosphate to any RNA species possessing a 5′-hydroxyl termini (i.e., thereby producing a substrate for the Terminator exonuclease). There was evidence that the 240-nt RNA species were not hydrolyzed by this exonuclease, suggesting that they probably harbor a 5′-triphosphate, supporting the previous conclusion of Delgado et al. (2005).

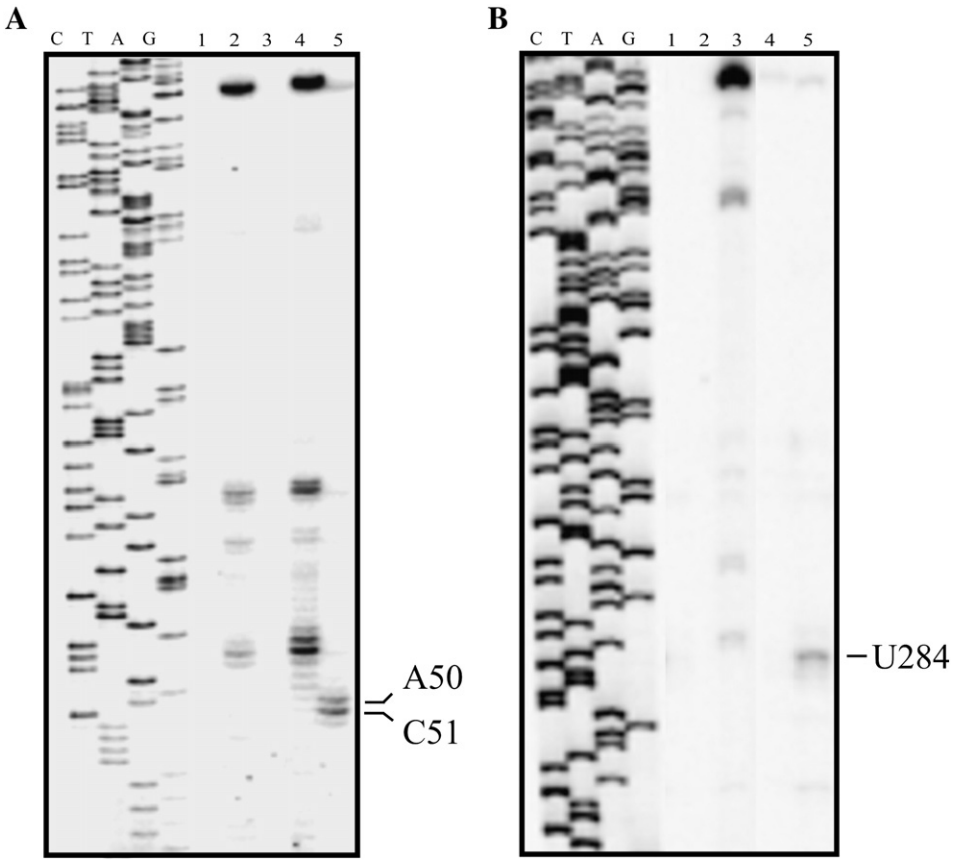

Fig. 4. Identification of the initiation site from the strands for both the plus (+) (conformer X9) and minus (−) polarities (conformer X5). (A) Primer extension using primer PLM82–109 for the detection of the plus (+) polarity. The same primer was used in the sequencing reaction of a plasmid containing PLMVd cDNA. Both reactions were analyzed on the same 6.5% PAGE gel. (B) Primer extension using primer PLM205–224 for the detection of the minus (−) polarity. The same primer was used in a sequencing reaction of a plasmid containing PLMVd cDNA. Both reactions were analyzed on the same PAGE gel. Lane 1 is a negative control performed in the presence of water instead of an RNA sample. Lanes 2 and 3 are the reactions performed in the presence of *in vitro*-transcribed full-length PLMVd RNA of plus (+) and minus (−) polarity, respectively. Lane 4 is the reaction performed in the presence of the linear conformer (338 nt) obtained from the peach leaves infected by the Hardired variant. Lane 5 is the reaction performed with either the X9 (A) or X5 (B) subgenomic conformers of the plus (+) and minus (−) strands, respectively.

The initiation site is surrounded by a conserved box

The identified initiation sites of both polarities are located in double-stranded regions of the P11 stem that include almost perfectly conserved nucleotides (see Fig. 1) (Ambros et al., 1998; Bussière et al., 2000; Pelchat et al., 2000; Fekih Hassen et al., 2007). Specifically, the plus polarity initiation site is located in front of the conserved $_{288}$GUC$_{291}$ of the hammerhead sequence (C_{51}), while that of the negative polarity is located at the uridine in position 284 of the same region. With this localization, confirmation of the initiation sites by direct mutagenesis of a DNA template and the infection of peach trees would not be informative because mutations in this region would inhibit both the hammerhead self-cleavage and the initiation of replication.

In order to identify a potential RNA promoter for DNA-dependant RNA polymerases, we designed an *in vitro* selection procedure using the PLMVd replication mechanism as a model (Fig. 5). Both the peach chloroplastic NEP and PEP enzymes are not yet commercially available; however, previous experiments using a commercial preparation of *E. coli* RNAP in place of PEP have permitted transcriptional activity from PLMVd-derived fragments (Pelchat et al., 2001, 2002). Similar experiments performed using T7 RNAP instead of NEP did not lead to any observable transcriptional activity. Consequently, we developed a selection procedure using the *E. coli* RNAP as a model bacterial-like enzyme catalyzing the replication of PLMVd.

The selection procedure was inspired by a similar work attempting to identify a single-stranded DNA promoter for the *E. coli* RNAP (Ohmichi et al., 2002). Briefly, circular 93 nt RNA molecules including both a conserved region, composed of sequence complementary to the hammerhead self-cleaving motif, and a region including 30 randomized nucleotides (Fig. 5), was synthesized from a 110 bp DNA template. The remaining nucleotides, which correspond to the sequence complementary to the hammerhead motif (i.e., mainly those composing the stem), differed from those found in PLMVd in order to avoid any bias.

Initially, the steps of the procedure were controlled and their conditions optimized. The templates including a randomized domain were then incubated in the presence of the commercial

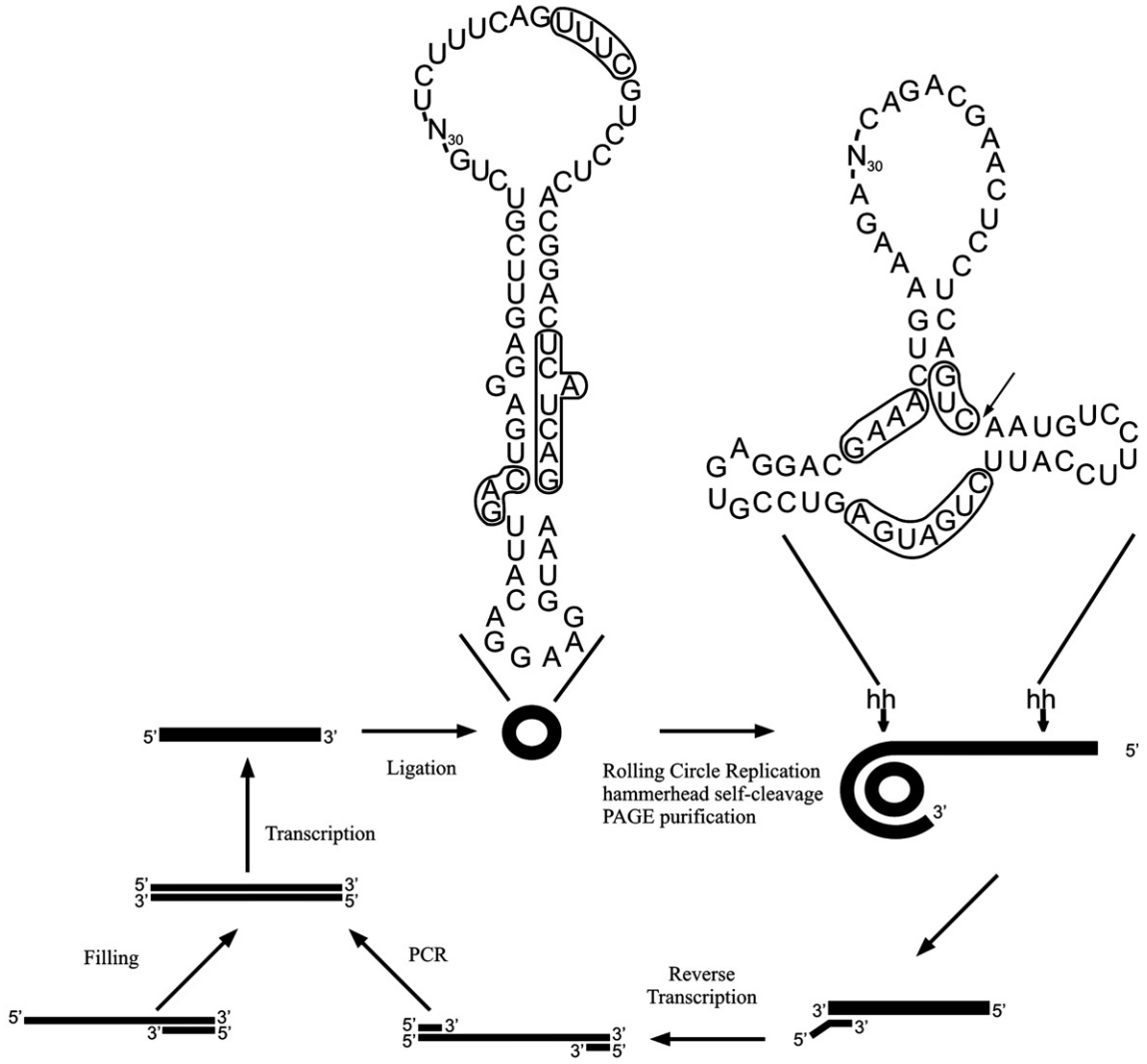

Fig. 5. Schematic representation of the developed *in vitro* cycle and potential secondary structure of both the RNA template and product of the rolling circle replication assay. The boxed nucleotides correspond to the hammerhead conserved positions. The arrow represents the cleavage site. The N_{30} identifies the randomized region.

E. coli RNAP which, theoretically, can initiate the synthesis of a nascent strand at any position. The newly synthesized strand can be of any length up to one unit. The nascent strand that includes the full hammerhead sequence could self-cleave since the replication buffer contained sufficient $MgCl_2$ (Fig. 5). The RNA strand of interest, although always found in tiny amounts, is the one that will be generated from one hammerhead self-cleavage site to the next. These molecules are 93 nt in length and include the sequences that recruited the RNAP. The RNA mixture was fractionated on polyacrylamide gels, and the 93 nt band was cut out and the RNA recovered. Together, the rolling circle replication, the self-cleavage reaction and the purification constituted the positive selection. The desired RNA molecules were then reverse transcribed using an oligonucleotide that contains the T7 RNA promoter, and subsequently PCR amplified in order to produce the DNA template for the next cycle. The last step, the RT–PCR amplification, served as a negative selection removing any 93 nt sequence lacking the appropriate ends, for example, ones produced by spontaneous hydrolysis of the initial RNA templates. As the eleven cycles progressed, the incubation time for the rolling circle replication was reduced from 6 to 1.5 h. Surprisingly, an exponential increase in the amount of RNA produced was not observed. Instead only a negligible increase in the accumulation of the 93 nt fragment was observed. This might reflect the fact that the *E. coli* RNAP is primarily a DNA-dependent RNA polymerase and is not adapted for the use of an RNA template.

Aliquots of DNA from the PCR amplification were cloned after several cycles and a total of 157 clones were sequenced (see Supplementary data). Prior to the beginning of the selection

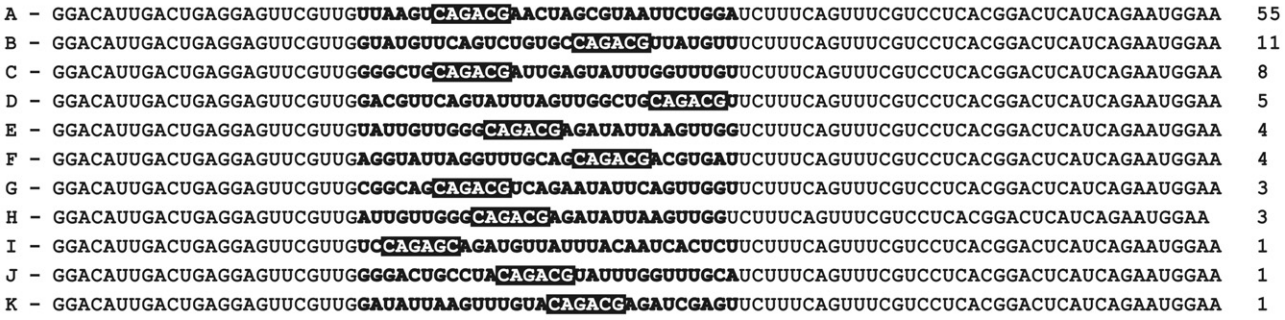

Fig. 6. Example of results obtained from the *in vitro* selection. The upper section shows typical sequences of the selected sequence variants including the 30 nt initially randomized stretch (bold nucleotides). The conserved CAGACG box, or related sequences, are boxed. The attributed letter for each sequence variant and the numbers of times each one was retrieved along the 11 cycles are indicated on the left and right of the sequences, respectively. The lower section includes the most stable structures obtained by bioinformatic prediction for the four most retrieved variants. The attributed letter for each sequence variant is indicated. The conserved CAGACG sequences are boxed. The ΔG values in kcal/mol are in parentheses.

process, 19 clones were sequenced and no bias was detected. From cycles 1 to 4, 18 sequences were determined and found to have nothing in common (as detected by visual inspection). Finally, from cycles 5 to 11, 119 clones were sequenced (see Fig. 6 for typical examples and Supplementary Table 2 for all sequences obtained after cycle 1). Astonishingly, they all included either a CAGACG box or a closely related sequence (one CCGACG, one CAGAGC and two CAGAGG variants were found). Re-analysis of the sequences from cycles 1 to 4 show that three clones including this box were obtained in cycle 4. With the exception of this conserved box, nothing was conserved at the other positions and the deletion of one nucleotide per selected RNA, either in the randomized domain or in the others that include the hammerhead sequence, was frequent. The most abundant sequence was named variant A and was found 55 times. An almost identical sequence (i.e., CAGACT) was obtained in the independent selection for a single-stranded DNA promoter for the *E. coli* RNAP (Ohmichi et al., 2002).

Analysis of PLMVd sequences led to the detection of a sequence almost identical to the conserved box obtained from the *in vitro* selection. Specifically, PLMVd strands of both plus and minus polarities include the sequences $_{47}$CAGACU$_{52}$ and $_{287}$CAGACA$_{292}$ (i.e., which is $_{292}$CAGACU$_{287}$ in the minus polarity transcript), respectively (see Figs. 1B and 7), both of which are clearly very similar to the conserved CAGAGC box. These sequences are located at the self-cleavage sites of PLMVd and are highly conserved according to the most recent sequence variation study (Fekih Hassen et al., 2007). Moreover, the aforementioned PLMVd initiation sites are located within, or very closely to, these sequences. Clearly, the similarities between both PLMVd RNA templates and the *in vitro* template are impressive considering the fact that the probability of finding a given stretch of 5 nucleotides is 1/1024 and that PLMVd is only 338 nt in size (if only considering perfectly conserved nucleotides).

In order to address the importance of the CAGACG box, rolling circle replication assays were performed using either the isolated variant A (the most abundant sequence retrieved), or derived mutants. The original A template did not support efficient production of one unit-length nascent RNA strands, as was observed during the *in vitro* selection (Fig. 7C). In fact over-exposition of the autoradiograms was required in order to permit detection of the full size product. Conversely, a strong band possessing an electrophoretic mobility corresponding to an RNA species ranging in size from 13 to 17 nt long was detected. This band cannot simply be the result of spontaneous hydrolysis of the template RNA that was non-specifically extended by the RNAP since it was not detected when the reaction was performed in the presence of only two NTPs. Similar observations were reported for transcriptional experiments performed using the *E. coli* RNAP in the presence of 6S RNA; namely that the most abundant products were 14–20 nt in size and corresponded to prematurely aborted syntheses (Wassarman and Saecker, 2006). In order to verify the identity of the nascent transcripts, the band containing the small RNAs was cut out of the gel, the RNA isolated and then cloned using a procedure developed for microRNA sequencing (Lau et al., 2001). All clones were found to possess sequences complementary to the template and, therefore, corresponded to replication products. Moreover, all of the nascent strands corresponding to an initiation in the initially randomized 30 positions contain identical 5′ ends and nearly identical 3′ ends. When using a template in which the conserved box was either completely mutated, or was deleted, the abundance of abortive nascent product was similar to the above (data not shown). However, the sequencing revealed that most originated from

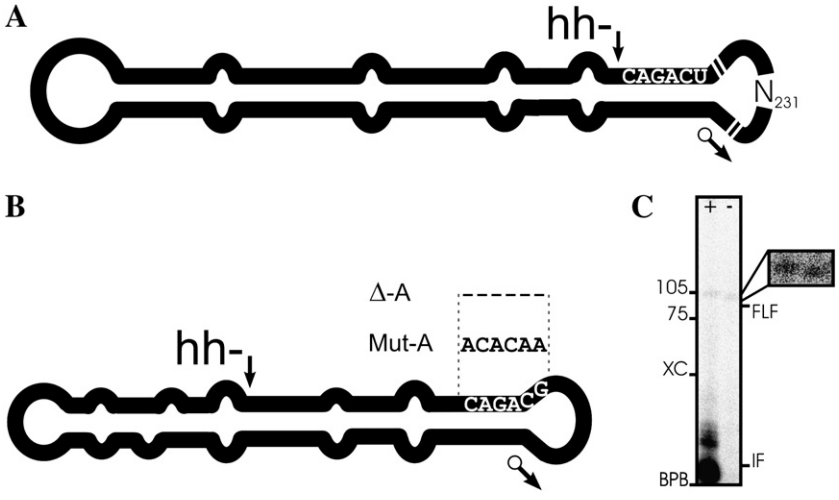

Fig. 7. Similarities between PLMVd and *in vitro* selected molecules in terms of being RNAP templates for the rolling circle replication. (A) Schematic representation of the PLMVd secondary structure. (B) Schematic representation of the original selected template A containing the conserved CAGACG box. In (A) and (B), the hammerhead cleavage sites of the replication products are indicated by the arrows denoted by hh−, while the initiation site and direction of synthesis are indicated by the arrows with circles. The conserved sequence is indicated in both cases. The modifications introduced for both the Mut-A and Δ-A mutated sequences are presented in (B). (C) Typical autoradiogram of a 10% denaturing PAGE gel used for the analysis of the rolling circle replication assays. The experiment was performed either in the presence of 2 nt (−) or 4 nt (+). The migrational positions of the xylene cyanol (XC) and bromophenol blue dyes (BPB), and of the RNA ladder, are indicated. The positions of both the full-length fragments (FLF, 93 nt) and the initiation fragments (IF, 13 to 17 nt) are indicated on the right of the gel. In inset is an overexposure of the region including the full-length fragments.

many other positions around the molecule. Several of these positions were also observed with template A, but to a lesser extent. Together, these results suggest that the presence of the CAGACG box seems to contribute to the positioning of the initiation site on the model template.

In order to further characterize the replication reactions of the circular template A, several experiments were performed (data not shown). For example, we observed that the use of T7 RNA polymerase failed to produce any noticeable products, suggesting that the phenomenon is specific to the *E. coli* RNAP. Moreover, *E. coli* RNAP inhibition experiments performed in the presence of tagetitoxin revealed an efficient inhibition of the synthesis in the presence of a DNA model template when the experiment was performed under various conditions (i.e., different buffers, concentrations, incubation times, etc.). Conversely, we were unable to observe tagetitoxin inhibition (or observed it only at a trace level) in the presence of either the circular template A or the 6S RNA. This indicates that any experiment involving tagetitoxin inhibition of the replication from an RNA template cannot be interpreted as being indicative of the RNAP involved in this reaction (i.e., the NEP or the PEP). The inhibition by tagetitoxin seems to be limited to DNA templates.

Discussion

The original goal of this work was to revisit the determination of the initiation site of PLMVd replication. Initially, Northern blot analyses of a relatively wide repertoire of PLMVd isolates led to the detection of a unique and universal subgenomic RNA strand in the banding patterns for each of the viroid's polarities. The analysis of these fragments, using a combination of RT–PCR amplification and primer extension, led to the identification, with precision, of the universal PLMVd replication initiation sites for both polarities. Specifically, the initiation sites on the plus strand were found to be located at A_{50} and C_{51} (predominant site), while in the minus strands it was located at U_{284}. These initiation sites are in the vicinity of the PLMVd self-cleavage and self-ligation sites. These recent findings are in agreement with the previous report of Delgado and collaborators, using a different methodology, identifying C_{51} as the initiation site for the plus strand and U_{286} as that for the minus strand for another PLMVd variant (Delgado et al., 2005). The minor differences are most likely the result of local primary and secondary structural differences between the PLMVd variants near the initiation sites.

Considering these two reports, it is tempting to suggest that there is only one initiation site per polarity. However, there is neither a direct evidence for discarding the possibility that other sites might exist in some variants nor for the possibility that more than one initiation site exists within the same variant. Moreover, if more than one initiation site is used in a variant, they might be used at different levels for many reasons, including the different structures adopted by the RNA templates. Therefore, the putative initiation sites within the terminal L11 loop previously identified (Pelchat et al., 2001), which are located at the opposite end to those located in this report, might also be used by the RNA polymerase replicating PLMVd, although most likely at a reduced level. RNA polymerase initiation sites located in the terminal stem–loop have been reported for several RNA species, including viroids in both the *Pospiviroidae* and *Avsunviroidae* families, and in the human hepatitis delta virus (Goodman et al., 1984; Navarro and Flores, 2000; Kolonko et al., 2006; Greco-Stewart et al., 2007). Moreover, the possibility of more than one initiation site for each PLMVd strand receives support from the detection of the various profiles of fragments depending on the PLMVd variant analyzed. For example, using RNA isolates from the peach cultivars GF-305, initiation sites located within the L11 loop would correspond to bands X4 to X6 based on their size (~280 nt) (Pelchat et al., 2001; and the present work). While the bands X4 appeared to be a product of degradation, the bands X5 and X6 were retrieved in only 6 out of 8 variants (see Supplementary Table 1). Clearly, this shows that the initiation sites reported for the replication of viroids should be interpreted with care. There are two important requirements to meet prior to reaching any conclusion. First, an initiation site of primary importance for the replication of a viroid should be conserved in all, or almost all, sequence variants. In the present study, the detection of the RNA species in question in a large repertoire of sequence variants should be considered as an indication of its importance. Secondly, the presence of the 5′-triphosphate terminus on the RNA species, which is compatible with it being a nascent transcript produced, should be confirmed. This criterion has been observed for the universal RNA species of both polarities found in all PLMVd variants (Delgado et al., 2005) and has received support from the experiments presented in this work. However, in both cases, the work was initially based on the potential to detect RNA fragments that accumulate in living cells. There is no indication that nascent transcripts are necessarily stable and accumulate sufficiently so as to permit their detection. However, since the two criteria were fulfilled, it appears that a universal initiation site for replication exists on each strand.

Ultimately, the determination of the initiation sites must receive support from an experiment showing that mutation at, or near, the site either inhibits or modulates the replication level during infection experiments. In the case of PLMVd, this experiment is difficult to envisage because the initiation sites are located in the essential region that contains both the self-cleavage and the self-ligation sequences. Any mutation in this region would attenuate not only the initiation of replication, but also the processing of the viroid. Consequently, the interpretation would be irrelevant. This situation led us to develop an *in vitro* selection procedure for a model replication reaction in order to obtain more support for the identification of the universal initiation sites. This procedure was based on the use of the *E. coli* RNAP which has already been used as a model of the PEP enzyme, although there is no solid evidence that this polymerase is responsible for PLMVd replication. The identification of either the NEP or the PEP chloroplastic polymerase as being responsible for the PLMVd replication remains elusive. Initial work using a protein extract supporting ASBVd replication has shown that the polymerization step was unaffected by high levels of α-amanitin (Navarro et al., 1999). Subsequently,

it has been suggested, based on the fact that the replicating ASBVd species was inhibited by the presence of tagetitoxin, that the NEP enzyme is more likely the polymerase responsible for ASBVd replication (Navarro et al., 2000). As both ASBVd and PLMVd belong to *Avsunviroidae* and all their replicational intermediates were detected in the chloroplasts (Bussière et al., 1999; Navarro et al., 1999), it was proposed that the NEP enzyme was the polymerase involved in this biological process for this family (Navarro et al., 2000). The tagetitoxin is an antibiotic known to strongly inhibit both *E. coli* RNAP and the homologous chloroplastic PEP enzyme in the presence of a DNA template (Mathews and Durbin, 1990; Pelchat et al., 2001; and the present work). However, to our knowledge, there is no report in the literature supporting tagetitoxin inhibiting a polymerase using an RNA template. In fact, only limited inhibition (if any at all) was observed for the *E. coli* RNAP in the presence of either PLMVd–derived fragments (Pelchat et al., 2001), the circular template A or the 6S RNA. This indicates that the inhibition of the *E. coli* RNAP by the tagetitoxin is observed only in the presence of a DNA template and is of no help for the identification of the polymerase replicating a viroid species. In the case of the *Pospiviroidae*, which are localized in the nucleus, it is well established that the RNA polymerase II is responsible for the replication (Rackwitz et al., 1981; Goodman et al., 1984; Warrilow and Symons, 1999; Fels et al., 2001; Kolonko et al., 2006). Interestingly, the hepatitis delta virus RNA, an infectious RNA reminiscent of the viroid that is also localized in the nucleus, is proposed to be replicated by the RNA polymerase II (Taylor, 2003; Lai, 2005; Greco-Stewart et al., 2007). The multisubunit RNA polymerase as the plant and human RNA polymerase II, and the bacterial RNAP, behave in an analogous fashion for the recognition of their promoters (Steitz, 2006; Lehmann et al., 2007). On that basis, it is tempting to suggest that the plastid-encoded polymerase might be the enzyme responsible of the *Avsunviroidae* replication. This hypothesis receives physical support from the demonstration that the homolog *E. coli* RNA initiated the replication of PLMVd-derived RNA molecules as templates (Pelchat et al., 2001, 2002; Pelchat and Perreault, 2004). The finding of a conserved sequence between the model template used in the present work and the PLMVd is also reinforcing this hypothesis, although it does not constitute a direct demonstration. An unambiguous identification of the peach polymerase involved in PLMVd replication remains a key issue for the field and experiments currently in progress might lead to this highly suitable information.

Nevertheless, the experiment performed in this study led to the identification of a CAGACG box that was found to be conserved in almost all selected sequence variants. Since only the conserved box appears to be shared by the selected sequence variants, we wondered if this RNA species might fold into an unique secondary structure. The most stable secondary structures, in terms of free energy (ΔG), were predicted for several of the selected variants (see Fig. 6 for examples). Overall, the predictions resulted in 1 to 3 secondary structures for each sequence variant. The ΔG values varied less than 20% between the structures of a given variant. The predicted structures for the different variants are relatively heterogenous. However, we observed that in most cases, the CAGACG box was most often found in double-stranded regions that are followed by a single-stranded region. In several cases, the last nucleotide of the box is already single-stranded. In PLMVd natural variants, the corresponding sequence is retrieved in a perfectly base-paired region surrounded by mismatches and single-stranded regions on strands of both polarities on the left and right hands, respectively. This could be an indication that the box has to be in a double-stranded domain to contribute as helper (or promoter) for the *E. coli* RNA polymerase, while the single-stranded part would favor the initiation of the synthesis.

In the course of these bioinformatics analyses, we also verified if both polarities of the RNA species with the most stable secondary structures in the frequently retrieved variants can also rearrange in order to include the hammerhead structure (or the equivalent structure for the complementary polarity). In order to do so, the secondary structures of the variants were predicted using constraints that corresponded to the three helices surrounding the catalytic core of the hammerhead structure. For most of the variants, a single secondary structure with a ΔG 10% to 30% higher than the most stable one (i.e., without the hammerhead structure) was obtained. In other words, the structures with the hammerhead were only slightly less stable than the most stable ones. Moreover, we noticed that the CAGACG sequence was less base-paired in the structures including the hammerhead motif. When the prediction was performed using the sequence of the P11–L11 stem–loop of PLMVd of plus polarity, a single most stable structure with a ΔG of −61 kcal/mol was obtained. The structure including the hammerhead motif for the plus polarity strand and a corresponding structure for the opposite strand had a ΔG of −32 kcal/mol. This is a ΔG value 50% smaller. A similar difference was obtained when predicting the structures for the minus polarity strands. This difference between the stability of the PLMVd and the small RNA templates used in this work might be an explanation as to why no exponential increase in the amount of RNA produced was observed. Instead, only a negligible increase in the accumulation of the 93 nt fragment was detected. This might reflect that following the self-cleavage, the RNA templates folded into the most stable structure. These can easily be rearranged into various structures because the differences in terms of ΔG values are not significant. Since the CAGACG sequence is not necessarily in double-stranded regions in these structures, they might not be used as promoter sequences for the RNAP. Conversely, the PLMVd strands self-cleaved, and following the completion of its synthesis, the nascent strands folded into a stable structure that cannot be easily rearranged because of the important difference between the ΔG values. Under the most stable secondary structure, the CAGACG box is double-stranded. If this box plays the same role in the viroid itself, it can be used as promoter for PLMVd recruitment of the host polymerase.

The CAGACG sequence was shown to determine the position of the initiation site on small RNA templates. Similar boxes are found in PLMVd located adjacent to the region responsible not only for the initiation of replication but also for

the self-cleavage and the self-ligation activities. If this box, which is highly conserved in all known PLMVd variants (Fekih Hassen et al., 2007), serves the same function in the viroid as it does for the small templates, it could provide an explanation for the existence of a unique, universal initiation site in all of the PLMVd variants studied. In other words, a specific sequence triggering the host polymerase provides an explanation for the replication starting at a precise position on the circular genome. The model found by *in vitro* selection in this report should not be considered as a formal indication that the peach chloroplastic PEP enzyme is responsible for PLMVd replication. However, the finding of the reminiscent box in both the small model templates and PLMVd variants is probably not just simple coincidence, but rather an indication that points in this direction. The appropriate RNAP should select molecules sharing common essential features. Clearly, once the host RNA polymerase is unambiguously identified, it will be instructive to verify whether or not the conserved CAGACG box plays the same role, as well as to define the PLMVd promoter sequence. Such a promoter should be located within the P11 stem since, to date, everything else involved in PLMVd replication is located there.

Materials and methods

RNA extraction

Total RNA (400 mg) were extracted from both infected and healthy peach leaves grown on the west coast of Canada (Sidney, British Columbia) using the Qiazol method (Qiagen) as recommended by the manufacturer. These RNA samples were quantified by UV spectroscopy and their quality verified by 1% agarose gel electrophoresis.

Northern blot hybridizations

Northern blot hybridizations were performed as reported previously (Bussière et al., 1999). Briefly, RNA samples (5 μg) were heated at 70 °C for 2 min in formamide dye (98% formamide, 10 mM EDTA, 0.05% bromophenol blue and 0.05% xylene cyanol) and fractionated by denaturing (8 M urea) 5% polyacrylamide gel electrophoresis (PAGE) using 1× TBE buffer (50 mM Tris–borate, pH 8.3, and 1 mM EDTA) prior to being transferred overnight, in 10× SSC (1.5 M NaCl, 150 mM sodium citrate, pH 7.0), onto nylon membranes (Hybond-N; GE Healthcare). The samples were then UV cross-linked to the membranes at 254 nm using a Stratalinker (Stratagene), and pre-hybridized for 4 h at 60 °C in a solution containing 50% formamide, 5× SSC, 1% sodium dodecyl sulfate (SDS), 5% Dernhardt's solution and 100 μg/mL of salmon sperm DNA. Hybridizations were performed overnight at 60 °C with fresh pre-hybridization solution containing the probe. After hybridization, the membranes were washed 3 × 15 min in 0.1× SSC/0.1% SDS solution at 60 °C and then analyzed by autoradiography on a Phosphoscreen (GE Healthcare). ^{32}P-Radiolabeled probes of full-length linear PLMVd (338 nt) of both plus and minus polarities were synthesized by *in vitro* by run-off transcription from the plasmid pPD1 as described previously (Beaudry et al., 1995). The resulting transcripts were purified by denaturing 5% PAGE electrophoresis. The bands corresponding to full-length PLMVd were excised, the RNA eluted overnight and ethanol-precipitated prior to being added to the pre-hybridization solution.

RT–PCR amplifications

Total RNA (100 μg) from both healthy and infected Hardired cultivars was incubated in a total volume of 20 μL containing 3 U of RQ1 DNase (Promega), 40 mM Tris–HCl pH 8.0, 10 mM $MgSO_4$ and 1 mM $CaCl_2$ at 37 °C for 1 h. Following phenol/chloroform extraction and ethanol precipitation, the RNAs were then separated on denaturing 5% PAGE gels. The gel bands corresponding to the region of the electrophoretic mobilities of full-length linear PLMVd and the universal subgenomic fragment (i.e., 338 and ~240 nt, respectively) were cut out, the RNA eluted and then used as template. Reverse transcription reactions of PLMVd RNA were carried out using SuperScript II (Invitrogen) in a 20-μL reaction as recommended by the manufacturer. After the reverse transcription reaction, 10 μg of RNase A was added and the mixtures were incubated for 10 min at 37 °C. Two different antisense primers, PLM250–275 (5′-ATCACTTCTGGAGGGGACCGGGTTTG-3′) and PLM63–86 (5′-TAATGACCTCTCAGCCCCTCCACC-3′), were used for the detection of the plus and minus strands, respectively.

Three microliters of the reverse transcription reactions were used in PCR amplification using purified PFU DNA polymerase. The reactions were carried out in a final volume of 50 μL containing 2 μM of each DNA primer, 1 mM $MgSO_4$, 0.2 mM of each dNTP, 10 mM Tris–HCl pH 8.8, 25 mM KCl and 5 mM $(NH_4)_2SO_4$. The reactions were incubated at 94 °C for 1 min prior to the addition of 2 μL of the enzyme and subjection to 25 amplification cycles (30 s at 94 °C, 30 s at 50–56 °C and 30 s at 72 °C). For the amplification of the plus strand, the antisense primer PLM250–275 (see above) was used in combination with one of the different sense primers: PLM63–86 (see above), PLM30–49 (5′-CATCAGTGGGCTTAGCCCAG-3′), PLM3–25 (5′-CAAAAGTTTCGCCGCATTTCAGC-3′) or PLM271–290 (5′-GTG ATTCTGGAAGATGAGTC-3′). For the amplification of the minus strand, the antisense primer PLM63–86 (see above) was used in combination with one of the four different sense primers PLM224–244 (5′-CGGGTAGACG-TCGTAATCCAG-3′), PLM250–275, PLM301–322 (5′-CATTTCAGAACTCATCAGTGTGC-3′) or PLM44–67 (5′-CATTACTCTCTCAAAAGTCTGGGC-3′). The amplifications were analyzed on 1% agarose gels.

Primer extensions

The isolated linear 338 nt and ~240 nt RNA species of both polarities were used as templates for primer extension reactions in order to determine their 5′ ends. The reverse transcription reactions were performed as described above with the exception that we used 2 pmol of 5′-^{32}P-labeled primers PLM82–109 (5′-GTGCTCCGAATAGGGCACCCCAAGGTGG-3′) and PLM205–224 (5′-CCAGGTACCGCCGTAGAAAC-3′) for

the RNA species of plus and minus polarities, respectively. Four microliters of the resulting mixtures were analyzed on 6.5% PAGE sequencing gels. The same primers were also used for performing sequencing reactions using the pPD1 plasmid, which contains two consecutive copies of PLMVd (Beaudry et al., 1995), that were analyzed on the same gel.

In vitro selection

Initially, a DNA library of transcription templates containing 63 constant nucleotides and 30 randomized nucleotides was generated by the hybridization of 20 nmol of each of 2 oligonucleotides (sense primer 5′-TAATACGACTCACTATAGGACATTGACTGAGGAGTTCGTCTG N_{30}TCTTTCAGTTTCGTCCTCACGGACTCA-3′ in which the T7 RNA promoter sequence is underlined, and the antisense primer 5′-TTCCATTCTGATGAGTCCGTGAGGAC GAAACTGAAAGA-3′) followed by 5 cycles of PCR amplification. The amplifications were performed in a final volume of 100 μL containing 10 mM Tris–HCl, pH 8.8, 25 mM KCl, 5 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$, 0.2 mM of each dNTP and 2 μL of purified PFU using a hybridization temperature of 55 °C. The oligonucleotide containing the randomized sequence was manually synthesized, thereby ensuring an equal distribution of each of the 4 deoxyribonucleotides at each position (IDT).

The first step of the *in vitro* selection was the run-off transcription. The transcriptions were performed at 37 °C for 90 min in a total volume of 50 μL containing 80 mM Hepes–KOH pH 7.5, 24 mM $MgCl_2$, 2 mM spermidine, 40 mM dithiothreitol, 4 mM of each NTP, 15 U of RNAguard (Amersham Biosciences), 0.01 U of yeast pyrophosphatase (Roche Diagnostic), 2 μL of purified T7 RNA polymerase and DNA from either the preliminary filling step or the previous PCR amplification. The reactions were stopped by the addition of 5 U of RQ1 RNase-free DNase I (Promega) and a subsequent incubation at 37 °C for 30 min. The resulting products were extracted with phenol/chloroform and ethanol-precipitated prior to being fractionated on a denaturing 8% PAGE gel. The band corresponding to the full-length RNAs was excised, the RNA eluted and then ethanol-precipitated. From cycle 1 to 5, aliquots of 500 pmol or less, depending on the cycle, were dephosphorylated in a final volume of 50 μL containing 250 mM Tris–HCl, pH 8.0, 15 U RNAguard and 4 U of alkaline phosphatase (Roche Diagnostic) at 37 °C for 30 min. The phenol/chloroform-extracted nucleic acids were ethanol-precipitated prior to phosphorylation in the presence of 4 mM ATP and 12 U of T4 polynucleotide kinase (USB) at 37 °C for 90 min according to the manufacturer's conditions. The resulting 5′-monophosphate RNA molecules were recovered by phenol/chloroform extraction and ethanol precipitation. For the residual cycles (6th to the 11th), the transcriptions were supplemented with 32 mM GMP so as to avoid the need to dephosphorylate and phosphorylate before the ligation, as guanosine is the first nucleotide of the transcripts. The second step is the circularization of the RNA template. Circularization of the RNA was achieved in a final volume of 30 μL per phosphorylated reaction containing 50 mM Tris–HCl pH 7.8, 10 mM $MgCl_2$, 1 mM ATP, 10 mM dithiothreitol, 15 U RNAguard (Amersham Biosciences) and purified T4 RNA ligase. The circular products were fractionated by 8% denaturing PAGE gel electrophoresis in TBE buffer. The band corresponding to the full-length circular RNAs was excised, the RNA eluted and then ethanol-precipitated. The third step was the rolling circle replication (RCR). The circular transcripts (aliquots of 100 pmol or less) were used as templates in a final volume of 15 μL containing 40 mM Tris–HCl pH 8.0, 50 mM KCl, 10 mM $MgCl_2$, 50 μg/mL BSA, 5 mM dithiothreitol, 0.5 mM of each NTP, 15 U RNAguard (Amersham Biosciences) and 2 U *E. coli* RNA polymerase (USB). The RCR reaction, which is the selection step, was performed at 37 °C for times decreasing slowly from 6 h (cycle 1) to 90 min (cycles 10 and 11). The resulting mixtures were fractionated on 8% denaturing PAGE gels and the bands corresponding to the full-length RNAs excised, the RNA eluted and then ethanol-precipitated. Step 4 is a reverse transcription coupled to a PCR amplification. The linear full-length products were reverse transcribed using SuperScript III (Invitrogen) for 1 h at 52 °C according to the manufacturer's protocol in the presence of an oligonucleotide that reproduced the 5′ portion of the original template (5′-TAATACGACTCACTATAGGACATTGACTGAGGAGTTCGTCTG-3′, the T7 RNA promoter sequence is underlined). The samples were then incubated for 10 min at 37 °C in the presence of 10 μg of RNase A, and the nucleic acids were then ethanol-precipitated. Products from the reverse transcription step were PCR-amplified in 4 reaction tubes under the same conditions as described for the preliminary filling step, but using specific oligonucleotides (sense primer, 5′-TAATACGACTCACTATAGG-3′, the T7 RNA promoter sequence is underlined; and, antisense primer, 5′-TTCCATTCTGATGAGTCCGTG-3′) and a hybridization temperature of 52°C. Aliquots of the PCR reaction were ligated into pGEM-T vector (Promega) and then transformed into *E. coli* XL1-Blue. Plasmid DNA was isolated and sequenced either manually, using the T7 sequencing kit (GE Healthcare), or sent to McGill University and Genome Quebec Innovation Center for automated sequencing. The remaining products were phenol/chloroform-extracted and ethanol-precipitated prior to undergoing another *in vitro* selection cycle.

RCR on selected RNA templates and the cloning of nascent transcripts

Three selected mutants of the RNA templates were synthesized as described above except that initially the DNA templates were obtained from the filling reaction using different sense oligonucleotides (A-ori 5′-TAATACGACTCACTATAGGACATTGACTGAGGAGTTCGT TGTTAAGT**CAGACG**AACTAGCGTAATTCTGGATCTTTCAGTTTCGTCCTCACGGACTCATC-3′; A-Δ 5′-TAATACGACTCACTATAGGACATTGACTGAGGAGTTCGT TGTTAAGT**Δ**AACTAGCGTAATTCTGGATCTTTCAGTTTCGTCCTCACGGACTCATC-3′; or A-mut 5′-TAATACGACTCACTATAGGACATTGACTGAGGAGTTCGTTGTT AAGT**ACACAA**AACTAGCGTAATTCTGGATCTTTCAGTTTCGTCCTCACGGACTCATC-3′; the underlined sequences correspond to the T7 RNA promoter, the bold ones the

conserved box position and the delta (Δ) the site of deletion). The rolling circle replication was then performed as described previously except that the reactions were performed either in the absence or the presence of 6 pmol of α-^{32}P UTP (3000 Ci/mmol; Amersham Biosciences) and the non-radioactive UTP concentration was reduced to 10 μM. The resulting products were analyzed on denatured 10% PAGE gels. The bands corresponding to the small RNAs (sRNA) were excised, the RNA eluted and then ethanol-precipitated. Linkers were then ligated to both ends of the sRNA according to a procedure reported previously (Landry and Perreault, 2005; Lau et al., 2001). The 3′-linker (5′PO_4-**CUGUA**GGATCCATCAA-3′N-3′, the bold letters represent ribonucleotides while the other positions are deoxyribonucleotides) includes a 5′ phosphate (PO_4) and a 3′NH2 in order to avoid multimerization between the monomers, while the 5′-linker (5′ ATCGTAGGAUCCUGAAA-3′) did not have any special requirement. These oligonucleotides were purchased from Dharmacon. The resulting products were reverse transcribed, PCR amplified and then were ligated into pGEM-T vector (Promega), transformed into *E. coli* XL1-Blue bacteria and sequenced.

Acknowledgments

The authors would like to thank Julien Forest-Parent for the technical assistance with the Northern blot hybridization experiments. This work was supported by a grant from the Natural Sciences and Engineering Research Council (NSERC, grant 155219-07) of Canada. The RNA group is also supported by grants from both the Université de Sherbrooke and Canadian Institutes for Health Research. J. M. was the recipient of a NSERC graduate student fellowship. J.P.P. holds the Canada Research Chair in Genomics and Catalytic RNA.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.virol.2007.12.010.

References

Ambros, S., Hernandez, C., Desvignes, J.C., Flores, R., 1998. Genomic structure of three phenotypically different isolates of peach latent mosaic viroid: implications of the existence of constraints limiting the heterogeneity of viroid quasispecies. J. Virol. 72, 7397–7406.

Beaudry, D., Bussière, F., Lareau, F., Lessard, C., Perreault, J.P., 1995. The RNA of both polarities of the peach latent mosaic viroid self-cleaves in vitro solely by single hammerhead structures. Nucleic Acids Res. 23, 745–752.

Bussière, F., Lehoux, J., Thompson, D.A., Skrzeczkowski, L.J., Perreault, J.P., 1999. Subcellular localization and rolling circle replication of peach latent mosaic viroid: hallmarks of group A viroid. J. Virol. 73, 6353–6360.

Bussière, F., Ouellet, J., Côté, F., Lévesque, D., Perreault, J.P., 2000. Mapping in solution shows the peach latent mosaic viroid to possess a new pseudoknot in a complex, branched secondary structure. J. Virol. 74, 2647–2654.

Côté, F., Perreault, J.P., 1997. Peach latent mosaic viroid is locked by a 2′,5′-phosphodiester bond produced by in vitro self-ligation. J. Mol. Biol. 273, 533–543.

Côté, F., Lévesque, D., Perreault, J.P., 2001. Natural 2′,5′-phosphodiester bonds found a the ligation sites of peach latent mosaic viroid. J. Virol. 75, 19–25.

Daros, J.A., Elena, S.F., Flores, R., 2006. Viroids: an Ariadne's thread into the RNA labyrinth. EMBO Rep. 7, 593–598.

Delgado, S., Martinez de Alba, A.E., Hernandez, C., Flores, R., 2005. A short double-stranded RNA motif of peach latent mosaic viroid contains the initiation and the self-cleavage sites of both polarity strands. J. Virol. 79, 12934–12943.

Desvignes, J.C., 1986. Peach latent mosaic and its relation to peach mosaic and peach yellow virus disease. Acta Hortic. 193, 51–57.

Fekih Hassen, I., Massart, S., Motard, J., Roussel, S., Parisi, O., Kummert, J., Fakhfakh, H., Marrakchi, M., Perreault, J.P., Jijakli, M.H., 2007. Molecular features of new peach latent mosaic viroid variants suggest that recombination may have contributed to the evolution of this infectious RNA. Virology 360, 50–57.

Fels, A., Hu, K., Reisner, D., 2001. Transcription of potato spindle tuber viroid by RNA polymerase II starts predominantly at two specific sites. Nucleic Acids Res. 29, 4589–4597.

Flores, R., Randles, J.W., Bar-Joseph, M., Diener, T.O., 2000. Virus taxonomy: the classification and nomenclature of viruses. The Seventh Report of the International Committee on taxonomy of Viruses. In: van Regenmortel, M.H.V., Fauquet, C.M., Bishop, D.H.L., Carsten, E.B., Estes, M.K., Lemon, S.M., Maniloff, J., Mayo, M.A., McGeoch, D.J., Pringle, C.R., Wickner, R.B. (Eds.), Viroids. Academic Press, San Diego, Calif, pp. 1009–1024.

Goodman, T.C., Nagel, L., Rappold, W., Klotz, G., Reisner, D., 1984. Viroid replication: equilibrium association constant and comparative activity measurements for the viroid polymerase interaction. Nucleic Acids Res. 12, 6231–6246.

Greco-Stewart, V.S., Miron, P., Abrahem, A., Pelchat, M., 2007. The human RNA polymerase II interacts with the terminal stem–loop regions of the hepatitis delta virus RNA genome. Virology 357, 68–78.

Hernandez, C., Flores, R., 1992. Plus and minus RNAs of peach latent mosaic viroid self-cleave in vitro via hammerhead structures. Proc. Natl. Acad. Sci. U. S. A. 89, 3711–3715.

Kolonko, N., Bannach, O., Aschermann, K., Hu, K.H., Moors, M., Schmitz, M., Steger, G., Reisner, D., 2006. Transcription of potato spindle tuber viroid by RNA polymerase II starts in the left terminal loop. Virology 347, 392–404.

Landry, P., Perreault, J.P., 2005. Identification of a peach latent mosaic viroid hairpin able to act as a Dicer-like substrate. J. Virol. 79, 6540–6543.

Lai, M.M., 2005. RNA replication without RNA-dependent RNA polymerase: surprises from hepatitis delta virus. J. Virol. 79, 7951–7958.

Lau, N.C., Lim, L.P., Weinstein, E.G., Bartel, D.P., 2001. An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. Science 294, 797–799.

Lehmann, E., Brueckner, F., Cramer, P., 2007. Molecular basis of RNA-dependent RNA polymerase II activity. Nature 450, 445–449.

Mathews, D.E., Durbin, R.D., 1990. Tagetitoxin inhibits RNA synthesis directed by RNA polymerase from chloroplasts and *Escherichia coli*. J. Biol. Chem. 265, 493–498.

Meng, B.Y., Wakasugi, T., Sugiura, M., 1991. Two promoters within the *psbK–psbI–trnG* gene cluster in tobacco chloroplast DNA. Curr Genet. 20, 250–264.

Navarro, J.A., Flores, R., 2000. Characterization of the initiation sites of both polarity strands of a viroid reveals a motif conserved in sequence and structure. EMBO J. 19, 2662–2670.

Navarro, J.A., Daros, J.A., Flores, R., 1999. Complexes containing both polarity strands of avocado sunblotch viroid: identification in chloroplasts and characterization. Virology 253, 77–85.

Navarro, J.A., Vera, A., Flores, R., 2000. A chloroplastic RNA polymerase resistant to tagetitoxin is involved in replication of avocado sunblotch viroid. Virology 268, 218–225.

Ohmichi, T., Maki, A., Kool, E.T., 2002. Efficient bacterial transcription of DNA nanocircle vertors with optimized single-stranded promoters. Proc. Natl. Acad. Sci. U. S. A. 99, 54–59.

Pelchat, M., Perreault, J.P., 2004. Binding site of *Escherichia coli* RNA polymerase to an RNA promoter. Biochem. Biophys. Res. Commun. 319, 636–642.

Pelchat, M., Lévesque, D., Ouellet, J., Laurendeau, S., Lévesque, S., Lehoux, J., Thompson, D.A., Eastwell, K.C., Skrzeckowski, L.J., Perreault, J.P., 2000. Sequencing of peach latent mosaic viroid variants from nine North American peach cultivars shows that this RNA folds into a complex secondary structure. Virology 271, 37–45.

Pelchat, M., Côté, F., Perreault, J.P., 2001. Study of the polymerization step of the rolling circle replication of peach latent mosaic viroid. Arch. Virol. 146, 1753–1763.

Pelchat, M., Grenier, C., Perreault, J.P., 2002. Characterization of a viroid-derived RNA promoter for the DNA-dependent RNA polymerase from *Escherichia coli*. Biochemistry 41, 6561–6571.

Rackwitz, H.R., Rohde, W., Sanger, H.L., 1981. DNA-dependent RNA polymerase II of plant origin transcribes viroid RNA into full-length copies. Nature 291, 297–301.

Sakai, A., Takano, H., Kuroiwa, T., 2004. Organelle nuclei in higher plants: structure, composition, function, and evolution. Int. Rev. Cytol. 238, 59–118.

Steitz, T.A., 2006. Visualizing polynucleotide polymerase machines at work. EMBO J. 25, 3458–3468.

Taylor, J.M., 2003. Replication of human hepatitis delta virus: recent developments. Trends Microbiol. 11, 185–190.

Vera, A., Sugiura, M., 1992. Combination of in vitro capping and ribonuclease protection improves the detection of transcription start sites in chloroplasts. Plant Mol. Biol. 19, 309–311.

Warrilow, D., Symons, 1999. Citrus exocortis viroid RNA is associated with the largest subunit of RNA polymerase II in tomato in vivo. Arch. Virol. 144, 2367–2375.

Wassarman, K.M., Saecker, R.M., 2006. Synthesis-mediated release of a small RNA inhibitor of RNA polymerase. Science 314, 1601–1603.