The 3’-Terminal Structure Required for Replication of Barley Yellow Dwarf Virus RNA Contains an Embedded 3’ End

Gennadiy Koev,1 Sijun Liu, Randy Beckett, and W. Allen Miller2

Received September 7, 2001; returned to author for revision October 19, 2001; accepted October 31, 2001

We determined the 3’-terminal primary and secondary structures required for replication of Barley yellow dwarf virus (BYDV) RNA in oat protoplasts. Computer predictions, nuclease probing, phylogenetic comparisons, and replication assays of specific mutants and chimeras revealed that the 3’-terminal 109 nucleotides (nt) form a structure with three to four stem-loops followed by a coaxially stacked helix incorporating the last four nt ([A/U]CCC). Sequences upstream of the 109-nt region also contributed to RNA accumulation. The base-pairing but not the sequences or bulges in the stems were essential for replication, but any changes to the 3’-terminal helix destroyed replication. The two 3’-proximal tetraloops tolerated all changes, but the two 3’-distal tetraloops gave most efficient replication if they fit the GNRA consensus. A mutant lacking the 3’-proximal stem-loop produced elevated levels of less-than-full-length minus strands, and no (+) strand. We propose that a “pocket” structure is the origin of (-) strand synthesis, which is negatively regulated by the inaccessible conformation of the 3’-terminus, thus favoring a high (+)/(-) ratio. This 3’ structure and the polymerase homologies suggest that genus Luteovirus is more closely related to the Tombusviridae family than to other Luteoviridae genera.

Key Words: tetraloop; RNA replication; minus-strand promoter; origin of replication; Luteoviridae; Tombusviridae; plant RNA virus.

INTRODUCTION

Initiation of RNA replication of positive-strand RNA viruses takes place at the 3’ terminus of the genomic RNA. Thus, a variety of 3’-terminal structures have evolved that serve as promoters for minus-strand RNA synthesis (Buck, 1996; Dreher, 1999). These include tRNA-like structures (TLSs) in the Bromoviridae family and Tobamovirus, Tymovirus, and Hordeiviridae genera (Dreher, 1999) and other less-characterized structures in other RNA viruses. The TLS is required for replication of Bromoviridae RNA because it contains promoter elements for replicase binding (Chapman and Kao, 1999) and initiation (Dreher and Hall, 1988) of (-)-strand RNA synthesis. The specific structure bound by BMV replicase has been determined at high resolution (Kim et al., 2000).

RNAs of many other viruses have 3’-termini that do not mimic tRNAs, but contain various conserved structural elements that play a role in RNA replication. The 3’-untranslated regions (UTRs) of viruses in such unrelated groups as Tombusviridae (Song and Simon, 1995), Coronavirusidae (Hsue and Masters, 1997; Williams et al., 1999), Pestivirus (Yu et al., 1999), Rubivirus (Chen and Frey, 1999), and others contain stable secondary and tertiary structural elements that are indispensable for replication. Host factors and viral replication proteins can bind specifically to such structures, and often their binding ability correlates with replication competence (reviewed by Lai, 1998).

Many viral RNAs contain poly(A) tails at their 3’ termini. A poly(A) tail is required for stability and translation initiation of nonviral eukaryotic mRNAs (Sachs, 2000). Poly(A) tails on viral RNAs can also play a role in replication. The RNA-dependent RNA polymerase (RdRp) of Zucchini yellow mosaic potyvirus was shown to bind poly(A) binding protein, indicating a direct role for the poly(A) tail in recruiting the replicase to the 3’ end of potyviral RNA (Wang et al., 2000). The poly(A) tail of Bamboo mosaic potexvirus RNA participates in a pseudoknot that serves as a replicase binding site (Huang et al., 2001). The polyadenylated genomes of coronaviruses and potyviruses also contain stem-loops in their 3’-UTRs that are required for replication (Halderman-Cahill et al., 1998; Hsue and Masters, 1997). Specific pseudoknots or kinking stem-loops are necessary in the 3’-UTRs for efficient replication of picornaviruses (Melchers et al., 1997; Mirmomeni et al., 1997; Pilipenko et al., 1996). On the other hand, deletion of entire 3’-UTRs in poliovirus and human rhinovirus, while severely inhib-
iting viral growth, did not abolish replication, suggesting that template recognition by the replicase might not be as stringent as originally believed (Todd et al., 1997). Subcellular compartmentalization of the replicase with the template RNA may also play an important role in specificity of RNA amplification (Todd et al., 1997).

Sequences required for replication initiation are not confined to the 3′ region. The replicase binding site of bacteriophage Qβ is 1.5 kb upstream of the 3′ terminus. Long-distance base-pairing then brings the RdRp to the 3′ end where initiation takes place (Klovins et al., 1998). The intergenic region in the middle of BMV RNA 3 harbors an internal replication enhancer (Sullivan and Ahlquist, 1999). In a very different mechanism, a central stem-loop of poliovirus RNA guides uridylylation of the genome-linked protein (VPg) which then primes (−) strand synthesis at the poly(A) tail (Paul et al., 2000; Rieder et al., 2000).

In contrast to the concept of specific protein–RNA recognition by complex interactions of higher order structures, Dreher and colleagues proposed that, in some viral RNAs, secondary structure in the promoter serves mainly to block initiation at the wrong sites. The promoter within the TLS of Turnip yellow mosaic virus (TYMV) can be reduced to a single-stranded, 3′-terminal CC(A/G) repeat (Singh and Dreher, 1998). Similarly minimal initiation sites function on TCV and Qβ RNAs (Yoshinari et al., 2000). The TLS of TYMV, and the very different secondary structures of TCV and Qβ, were proposed to simply prevent initiation at internal CCA elements and also to ensure that the 3′-terminal CCA is single-stranded (and thus available for recognition by the replicase) by tying up the adjacent sequence in secondary structure.

Many viruses produce subgenomic RNAs (Miller and Koev, 2000). This also requires specific interactions between the replicase and the viral RNA to initiate (+) strand synthesis within the (−) strand (Siegel et al., 1997), or possibly to terminate at a defined site during (−) strand synthesis (Sit et al., 1998). The promoters for subgenomic RNA synthesis also consist of specific structures, yet they rarely resemble the (−) strand promoter, despite RNA synthesis by the same RdRp.

Barley yellow dwarf virus (BYDV) is a single-stranded, positive-sense RNA virus, the type member of genus Luteovirus in the family Luteoviridae (D’Arcy et al., 2000). It has an uncapped, nonpolyadenylated, 5.7-kb genomic RNA (gRNA) that contains six open reading frames (ORFs) (Fig. 1A). During infection, BYDV generates three, 3′-coterminal subgenomic RNAs (sgRNAs). Viral gRNA serves as the mRNA for ORFs 1 and 2; ORF 2 is expressed fused to ORF1 by ribosomal frameshifting (Di et al., 1993). The role of the 39-kDa protein encoded by ORF1 is unclear, whereas the 99-kDa frameshift product of ORFs 1 and 2 is the viral RdRp. Although classified in

FIG. 1. (A) Genome organization of BYDV. Horizontal lines represent genomic and subgenonomic RNAs (sgRNA). Open reading frames (ORFs) are indicated by number (1 through 6), molecular mass (K), and encoded functions (RdRp, RNA-dependent RNA polymerase; CP, coat protein; MP, putative movement protein; AT, aphid transmission protein). (B) Sequence alignment of the 3′ termini of BYDV strains: PAV-Australia (PAV-Aus, GenBank Accession No. X07653), MAV, PAV-Japan (PAV-Jpn, Accession No. D85783), and PAV-129 (Accession No. NC_002160). The PAV-Aus sequence is shown with dots indicating gaps. Hyphens indicate bases in the other isolates that are identical to PAV-Aus. (C) Potential secondary structures of the 3′ termini of PAV-Aus (left) and PAV-129 (right) isolates as predicted by MFOLD (Mathews et al., 1999; Zuker et al., 1999). Nucleotides that vary among non-PAV-129 isolates are in boldface italics on the PAV-Aus structures, with arrow pointing to variant base. Base changes that preserve proposed base-pairing or tetraloop motif are boxed. Nucleotide variations that disrupt secondary structure are circled. On the PAV-129 structure, long blocks of sequences that differ from PAV-Aus are boxed. Other varying bases are in italics.
Luteoviridae, the RdRp gene of BYDV shares significant homology to those of Dianthovirus and Necrovi rus genera of the Tombusviridae, and no homology to the those in the Polerovirus genus of the Luteoviridae (Koonin and Dolja, 1993; Miller et al., 1995). Moreover, similar to other Tombusviridae, the 3’ end of Luteovirus RNA terminates in CCC, while the Polerovirus genomes terminate in GU.

Previous research on origins of replication of Tombusviridae RNAs revealed specific structures (Song and Simon, 1995; Turner and Buck, 1999) and in some cases surprisingly high tolerance of variation (Carpenter and Simon, 1998). Our previous studies have shown that certain sequences in the 869-nt 3′-UTR of the BYDV genome control translation (Miller et al., 1997; Wang et al., 1999), but so far they have been located hundreds of bases upstream of the 3′ terminus. We also mapped the promoters for synthesis of all three sgRNAs (Koev and Miller, 2000; Koev et al., 1999). They show surprisingly little sequence or structural similarity. How BYDV or any virus balances genomic replication with subgenomic RNA synthesis and translation is unclear. One thing certain is that RNA structures play key roles in these functions. Here, we determine the structures necessary for genomic RNA accumulation that presumably participates in (−)-strand synthesis.

RESULTS

Predicted folding of the 3′ terminus of BYDV RNA

To determine the secondary structure of the BYDV 3′ end, we first used computer predictions. Identification of phylogenetically conserved base-pairing despite sequence variation (covariations) is a powerful tool for prediction of functional secondary structure. Therefore, we first aligned the 3′ ends of all known BYDV strains. Most strains show very high conservation throughout their genomes, but one novel strain, BYDVP-PAV-129, varies significantly. All sequenced BYDV isolates other than PAV-129 share >97% identity in ORFs 1 and 2. PAV-129 (GenBank Accession No. NC_002160), which is very se-

eral 88% amino acid sequence identity in ORF2 (the RdRp) and 80% in ORF1 to the sequence in our infectious clone (PAV6), which is a chimera of very similar Illinois and Australian isolates (Di et al., 1993). 3′-end alignment (Fig. 1B) revealed high conservation among all isolates except PAV-129. PAV-129 shows similarity in the 3′-terminal 30-nt bases, but upstream of that many gaps and mismatches were present with only small blocks of similarity. This proved to be a powerful guide for secondary structure prediction.

Using MFOLD (Mathews et al., 1999; Zuker et al., 1999), we predicted two alternative conformations of the BYDV (PAV6) 3′ terminus (PAV-Aus, Fig. 1C). Both structures harbor four stem-loops (SL1 through SL4 numbered from the 3′ end) that contain terminal tetraloops. The tetraloops of SL1, SL3, and SL4 fit the GNRA consensus; the SL2 terminal loop is a UNCG tetraloop. Both GNRA and UNCG tetraloops confer unusually high stability to their stem-loop structures due to non-Watson–Crick interactions between the nucleotides in the loops (Cheong et al., 1990; Heus and Pardi, 1991). The structure with predicted ΔG = −46.9 kcal/mol differs from the “alternative” structure (ΔG = −45.4) by its shorter SL2 and base-pairing of the 3′-terminal four bases (ACCC5674–5677) in a helix that stacks coaxially with SL3 (Fig. 1C).

The slight variation among the BYDV-MAV and BYDV-PAV-Jpn strains supported existence of some stems and tetraloops (Fig. 1C). The base differences in tetraloops of SL1 and SL3 both maintain the GNRA consensus. The variable nucleotides in other single-stranded regions do not change their single-stranded nature. Covariations in SL3 and SL4 support existence of the base-paired regions, with variations in the internal bulge in SL4. The closely related sequences did not allow us to distinguish between the two predicted structures. In contrast, PAV-129 3′ end can be folded only into the structure with the embedded 3′ terminus (Fig. 1C). The PAV-129 structure lacks SL2, which perturbed alignment upstream (Fig. 1A), has a bulge in SL3, and a longer stem and longer loop at the SL4 position. Except for the proximal end, SL4 bears no primary sequence similarity to that of PAV6. The PAV-Aus and PAV-129 structures, from the beginning of SL4 to the 3′ end, end up being exactly the same length.

Nuclease probing of the 3′-terminal RNA secondary structure

To determine the solution structure of the 3′ end of BYDV RNA, we probed a transcript comprising the 3′-terminal 109 bases with the single-strand-specific, sequence-nonspecific chemical nuclease imidazole and with ribonuclease T1 (cuts single-stranded guanosine nucleotides). Most predicted single-stranded bases were cut strongly with imidazole and RNase T1 (Fig. 2A). Imidazole failed to cleave the predicted tetraloop nucleotides probably due to the non-Watson–Crick interactions within the loops (Cheong et al., 1990; Heus and Pardi, 1991). Imidazole is thought to require flexibility at the phosphodiester bonds to catalyze cleavage (Vlassov et al., 1995). The GGGU5619–5622 sequence was insensitive to imidazole and T1 (Fig. 2), as predicted by both possible structures (Fig. 1C). ACUC5638–5641 was very sensitive to imidazole cleavage, which is inconsistent with the alternative structure in which this region base pairs with GGGU5674–5677. Therefore, the structure probing data sup-
ported formation of the most stable calculated structure, which is the conformation shared with PAV-129 RNA.

The 3'-terminal stem-loops are required for viral RNA replication

To test the importance of the 3'-terminal stem-loop structures for viral RNA replication, we constructed mutants in the full-length infectious clone of BYDV (PAV6) that contained deletions that spanned each stem-loop including the larger version of SL2 (Fig. 3A). Also, to test whether sequence upstream of the four stem-loops is necessary to support viral RNA replication, we designed an additional mutant, 3'SL5D, that contained a deletion of nts 5503–5567 upstream of SL4 (Fig. 3A). Because portions of the 3'-terminal region of BYDV genomic RNA have been implicated in translational regulation (Miller et al., 1997; Wang et al., 1999), we first tested these mutants for translation in a wheat germ extract. All five mutant transcripts translated similarly to wild-type (PAV6) RNA in vitro, producing abundant 39-kDa product of ORF1 and low levels of the 99-kDa frameshift product, the RdRp (Fig. 3B). None of the full-length viral in vitro transcripts containing any of the stem-loop deletions replicated (Fig. 3C) in protoplasts. The faint genomic RNA bands in all lanes are residual inoculum. Absence of subgenomic RNAs indicates absence of replication. Mutant 3'SL5D replicated at a low level (Fig. 3C), suggesting that all four 3'-terminal stem-loops are necessary and sufficient for basal-level replication, but that upstream sequence (bases 5503–5567) is required for wild-type activity.

To further assess effects of mutations on replication, we probed for minus-strand RNA, which is present only if the origin of replication is functional. In cells infected with the wild-type (PAV6) RNA, low levels of (−)-gRNA and sgRNA were detected with the minus-strand-specific probe (Fig. 3C). However, we found that the minus-strand-specific probe can anneal to BYDV plus-strand RNAs, albeit very inefficiently (data not shown). Thus, we do not know if the subgenomic-sized RNA species are minus-strand sgRNAs or represent inefficient detection of plus-strand sgRNAs. Interestingly, the mutant 3'SL1D, which lacked SL1, consistently produced very high amounts of negative-strand RNA with the molecular mass lower than that of the full-length gRNA (Fig. 3C). Owing to the absence of such a pattern with the plus-strand-specific probe, and the very strong intensity of the signal, this is not artifactual detection of plus strand. A smaller amount of this low molecular weight RNA was detected in cells inoculated with the deletion mutants 3'SL2D through 3'SL5D as well (Fig. 3C).

FIG. 2. Nuclease probing of the 3’ terminus of BYDV RNA. (A) Autoradiograph of an 8% polyacrylamide, 7 M urea gel showing partial digests of 5’ end-labeled transcript comprising the 3’-terminal 109 nt of BYDV (PAV6) RNA. Lane 1 (T1 denatur.), sequencing ladder generated by partial digestion with the T1 ribonuclease (0.2 u) under denaturing conditions. Lanes 2–4, partial digests with imidazole (concentration shown above each lane). Lane 5 (T1 native), partial digest with T1 ribonuclease (0.01 u) under native conditions. Numbers on the left indicate genomic positions of guanosine residues. Numbers on the right indicate genomic positions of nucleotides sensitive to T1 or imidazole. (B) Predicted secondary structure of the 3’ end of BYDV genome with imidazole-sensitive nucleotides indicated by arrows, and T1-sensitive nucleotides indicated by open triangles.
single-strand mutant disrupted base-pairing and the double mutant restored base-pairing (Fig. 4A).

Mutants with the stem structures disrupted in SL1, SL3, and SL4 failed to replicate, whereas restoration of these stems in the double mutants restored virus replication to the wild-type level in SL3 and to levels lower than the wild-type in SL1 and SL4 (Fig. 4B). Interestingly, deletion of all four bulged nucleotides in SL1 (Fig. 4A, 3′SL14) had no deleterious effect on replication, indicating that unpaired bases in SL1 are not required for replication (Fig. 4B). These data showed that the base-pairing of the stems in SL1, SL3, and SL4, and not their primary RNA sequences, are required for replication.

We designed a set of mutants in an attempt to determine which conformation involving SL2 or the 3′ terminus is necessary. One group (3′SL2U1, 3′SL2U2, 3′SL2U3) disrupted and restored the distal end of SL2 that is present in both predicted structures. The results clearly supported a requirement for secondary but not primary structure. The second set disrupted and restored either the bottom of the longer SL2 that is present only in the alternative conformation (GGGU5619–5622:ACUC5638–5641, mutants 3′SL21, 3′SL22, 3′SL23, Fig. 4A inset) or the base-pairing of GGGU5619–5622 to the 3′ terminus (ACCC5674–5677, 3′SL22, 3′SL24, 3′SL25, Fig. 4A). To alter the 3′-terminal ACCC, we could not transcribe from plasmid DNA, because the CCC forms part of the SmaI site required for plasmid linearization prior to run-off transcription. Instead, full-length DNA was prepared by PCR using a 3′ primer containing the appropriate mutation, and a 5′ primer that included the T7 promoter. Full-length genomic RNA was transcribed directly off the PCR product. To allow fair comparison, all constructs involving potential base-pairing with the GGGU sequence (3′SL21, 3′SL22, 3′SL24, 3′SL25, Fig. 4A), as well as wild-type RNA, were prepared this way. As seen in Fig. 4C, the RNA from protoplasts inoculated with transcripts from PCR products appeared to contain more degraded viral RNA. However, it is clear that all mutants containing alterations in the GGGU5619–5622 (3′SL22, 3′SL23, 3′SL25) sequence or the 3′-terminal ACCC (3′SL24, 3′SL25) were dead. The single ACUC5638–5641 mutant (3′SL21) replicated at a reduced level. Thus, ACUC5638–5641 does not have to be base-paired to GGGU for replication, nor is its primary sequence essential. Either the primary sequence or secondary structures (or both) of GGGU5619–5622 and ACCC5674–5677 are essential. The data cannot distinguish between these possibilities. However, the data support the absence of a requirement for base-pairing of GGGU5619–5622 to ACUC5638–5641, consistent with the most stable computer-predicted structure and that determined by probing.

Helical regions of the 3′-terminal stem-loops are necessary for replication

To test the role of the primary sequence vs the secondary structure of the stems in SL1 through SL4, we designed a series of mutations that disrupted and restored these helices. Three mutants were generated for each stem-loop: the 3′-proximal strand, the 5′-proximal strand, or both strands were mutated such that each
Requirements for the terminal tetraloops

Because of the previously reported importance of GNRA and UNCG tetraloops in RNA structure and function (Jaeger et al., 1994; Murphy and Cech, 1994), we elucidated the necessity of the terminal loop sequences in SL1, SL2, SL3, and SL4 for viral RNA accumulation. In a series of mutants (Fig. 5A), the sequence of the loop in SL1 (GAAA) was altered to a UNCG consensus tetraloop (UUCG), another GNRA tetraloop (GAGA), and a nontetraloop sequence (GACA). All these mutants replicated at the wild-type level, as did the mutant with the SL2 tetraloop (UUCG) changed to the complementary sequence (AAGC) (Fig. 5B), indicating that the sequences of the terminal loops in SL1 and SL2 are not important for virus RNA replication in our assay.

Alteration of the SL3 tetraloop sequence (GCAA) to its complement (CGUU) abolished virus replication (Fig. 5B). Changing it to a UNCG tetraloop (UUCG), a different GNRA tetraloop (GCGA), and a nontetraloop sequence (GACA). All these mutants replicated at the wild-type level, as did the mutant with the SL2 tetraloop (UUCG) changed to the complementary sequence (AAGC) (Fig. 5B), indicating that the sequences of the terminal loops in SL1 and SL2 are not important for virus RNA replication in our assay.

FIG. 4. Effect of stem-loop disruption and restoration on virus RNA replication. (A) Secondary structure with mutants indicated (boldface italics). Disrupted and restored helical regions shown in boxes. The mutants based on the alternative SL2 structure are in the shaded boxes. Mutants 3′SL21 and 3′SL22 are shown on both structures. (B) Accumulation of positive-sense viral RNAs in protoplasts inoculated with wild-type (PAV6) or mutant transcripts from Smal-linearized plasmid was detected by Northern blot hybridization of total RNA isolated from infected oat protoplasts. (C) Accumulation of viral RNAs in protoplasts inoculated with transcripts derived directly from uncloned, full-length PCR products with the indicated mutations. Far right lane was from protoplasts inoculated with plasmid-derived RNA. Ethidium bromide staining of the gel (below blot) prior to blotting shows approximately equal loading as revealed by ribosomal RNA abundance in each lane.
sequence with only one base different from the wild-type (GCCA) restored replication to various degrees but never to the wild-type level (Fig. 5B). The mutant 3′SL3GCCA, which had one base different from the wild-type sequence and fit the GNR4 consensus, replicated at the highest level of the mutants. Similar results were obtained with the SL4 tetraloop mutants designed in the same fashion as those of SL3 (Fig. 5A). Although none of the mutations abolished virus replication, they all reduced RNA accumulation by various degrees (Fig. 5B), indicating that the original wild-type sequence GAAA is optimal.

Replication of chimeric PAV6-PAV129

To test a natural variant that represents extreme alteration of the 3′ end, while conserving key structures, we tested the ability of the 3′ end from PAV-129 to support replication of a chimeric RNA. This construct contains a natural deletion of the distal end of SL2 that is conserved in the two alternative PAV6 structures. The 5′ half of the chimeric construct (5′ UTR and ORFs 1 and 2) is derived from PAV6, while the 3′ half is derived from PAV-129. In the coding regions, the isolates diverge most in ORF1 (80% amino acid sequence identity) and ORF 2 (88% identity). It was possible that the divergent 3′ structure of PAV-129 reflects a different template specificity of the replicase conferred by its different amino acid sequence. In this case, the PAV6 replicase of the chimera might not recognize the PAV-129-derived 3′-terminus. However, it is clear that the chimera replicated, albeit at about 10% of the efficiency of full-length PAV6 RNA (Fig. 6). Thus, while SL2 may contribute to viral RNA accumulation, it is not absolutely necessary. Moreover, the sequences of stem 3, and stem and loop of SL4 can vary radically compared to PAV6, while still allowing significant RNA replication.

FIG. 5. Effect of the terminal tetraloop sequence changes on BYDV RNA replication. (A) Diagram of the mutants with altered loop sequences shown in boxes. (B) Northern blot showing accumulation of viral RNAs, ~24 hpi, in protoplasts inoculated with indicated transcripts.

FIG. 6. Replication of chimeric PAV6-129 transcript in oat protoplasts. (A) Map of PAV6-129 showing PAV6 (filled) and PAV-129 (open)-derived portions, and the common restriction site at which they were joined. (B) Oat protoplasts were inoculated with transcripts from Smal-linearized pPAV6 or pPAV6-129. Viral replication products were detected as in previous experiments. The two PAV6-129 lanes were from different sets of protoplasts inoculated with transcripts from independent clones of the same construct. Ethidium bromide staining of the gel (right) reveals relative loading of each lane.
DISCUSSION

Role(s) of the 3′-terminal stem-loop structures in RNA accumulation

Most, if not all, RNA viruses have a specific structure at the 3′ end of the genome that is required for initiation of (−) strand synthesis. Here, we identified structures in the 3′ end of the BYDV genome that are required for accumulation of viral RNA. In addition to (−) strand synthesis, the structures may contribute to BYDV RNA stability or translation in vivo. Although translation of the deletion mutants in wheat germ extract showed no apparent defects (Fig. 3B), some as-yet unmapped BYDV 3′-UTR sequences not required for translation in vitro are required for translation in vivo (Guo et al., 2000; Wang et al., 1999). These may include sequences that mimic the poly(A) tail function, functionally resembling the pseudoknot-rich domains of TMV (Gallie and Walbot, 1990), because wheat germ extracts are insensitive to the presence of poly(A) tails. Involvement of the 3′-terminal structures exclusively in conferring (+) strand stability is unlikely because intact inoculum RNA is still detectable in cells 24 h after inoculation with replication-defective mutants (Figs. 3–5).

Features likely to participate directly in (−) strand synthesis are SL1 and the terminal ACCC sequence. Similar stem-loops that tolerate variation in primary sequence and are located a few bases upstream of a 3′-terminal CCC sequence have been identified as origins of (−) strand synthesis for Turnip crinkle virus (TCV) (Song and Simon, 1995), Cymbidium ringspot virus (CymRSV) (Havelda and Burgyan, 1995), and Red clover necrotic mosaic virus (RCNMV) (Turner and Buck, 1999). In the polymerase genes, these viruses, all members of the Tombusviridae family, are closely related to BYDV (Koonin and Dolja, 1993), so similar replication origins would be predicted in BYDV RNA.

Stem-loops with a relaxed primary sequence requirement in 3′ replication elements have also been identified in Tobacco etch virus (TEV) (Haldeman-Cahill et al., 1998), Bovine viral diarrhea virus (BVDV) (Yu et al., 1999), Alfalfa mosaic virus (AMV) (van Rossum et al., 1997), and others. On the other hand, Beet necrotic yellow vein virus and Satellite tobacco necrosis virus both require specific primary and the secondary structures in the 3′-helical regions for replication (Bringloe et al., 1999; Lauber et al., 1997). The secondary structures of 3′-terminal stem-loops of RNA viruses have been shown to be required for protein binding (Lai, 1998). Perhaps failure to bind replication factors rendered the stem disruption mutants nonviable.

An intriguing feature of the comparison of the conserved structures formed by PAV6 and PAV-129 termini (Fig. 1) is that most of the conserved primary sequence is in a “pocket” comprising all of SL1 and immediately adjacent bases, the putative GGG(U/A)5619–5622(U/A)UCCC5674–5677 helix, the proximal ends of SL’s 3 and 4, and the bases connecting them. Although the two-dimensional rendition does not reflect the true 3D structure, most of these bases must be in fairly close proximity. We propose that this area of conserved bases comprises a replicase recognition site. This leads to the question of why the sequence in SL1 is conserved when it can be altered drastically without reducing replication, as long as it can form a stem-loop.

The answer to the above question may lie in the important caveat for all of these experiments: the functional mutant structures detected here may not be optimal. This is because we are examining replication only at 24 h postinoculation (hi) in plant cells that have been inundated with large quantities of inoculum. In these conditions, the inoculum RNA has an advantage in sheer numbers in the competition with host RNAs for host proteins. Second, the inoculum has no competition with more fit viral variant RNAs that would normally be present in a mixed natural population. Third, we are viewing only a 24 h time point. After longer periods of infection and many more rounds of replication that take place in whole plant infections, more fit structures have time to evolve. Thus, the wild-type SL1 sequence may replicate more efficiently in the long run than the mutant constructs that replicate in protoplasts at 24 days p.i., as long as the stem can form.

Roles of tetraloops

The one conserved sequence (between PAV6 and PAV-129) that is not in the conserved pocket is the tetraloop at the distal end of SL3 that fits the highly stable GNRA consensus. GNRA and UNCG tetraloops have intraloop interactions that greatly stabilize them and the helix. Tetraloops have also been implicated in protein binding. Thus, this element could also be a potential binding site for replication proteins (reviewed in Lai, 1998). A third interaction for GNRA tetraloops involves interaction with a “receptor” in a specific RNA helix, as in the group I intron of Tetrahymena thermophila and the td intron of the bacteriophage T4 (Butcher et al., 1997; Costa and Michel, 1997; Pley et al., 1994). Such interactions have not been reported in viral RNAs.

Interestingly, sequences of the SL1 and the SL2 terminal tetraloops were not important for replication in our assay. The lack of a requirement for specific sequences in the loops of the essential 3′-proximal stem-loops has been demonstrated in vitro in TCV (Song and Simon, 1995) and CymRSV, which also contains a 3′-terminal GNRA tetraloop (Havelda and Burgyan, 1995). On the other hand, RNA2 of RCNMV, which is closely related to BYDV, was unable to replicate with an alteration in the three base loop of the 3′-terminal stem-loop structure.
(Turner and Buck, 1999). Perhaps the more stable tetraloops provide an advantage only in more competitive or longer term replication conditions. Such was the case for a GNRA tetraloop in the 5’-proximal stem-loop of Potato virus X (PVX) (Miller et al., 1998). A preference for a GNRA tetraloop was obvious in SL3 and SL4 of PAV6, but the PAV-129 3’ end permits replication with a much larger loop 4. This may explain the reduced accumulation of chimeric RNA or perhaps the longer stem 4 of PAV-129 compensates for loss of the stabilizing tetraloop.

Possible roles of upstream sequence elements

The 3’-terminal 109 nts that contain the four stem-loops were not sufficient for the wild-type level of RNA accumulation in protoplasts. Sequence upstream (bases 5503–5567) appear to play a supporting role. They may be involved in functions other than (−)-strand synthesis, such as (+)-strand synthesis. In TCV and RCNMV genomes, elements located hundreds of nucleotides upstream of the 3’ end are needed for efficient replication (Carpenter et al., 1995; Turner and Buck, 1999). Elements located upstream of the 3’-terminal TLS stimulate replication of BMV and TMV RNAs (Lahser et al., 1993; Sullivan and Ahlquist, 1999; Takamatsu et al., 1990). In the bacteriophage Qβ RNA, an internal region 1.5 kb upstream of the 3’ end is the actual replicase binding site (Brown and Gold, 1996). Long-distance base-pairing delivers the replicase complex to the 3’ end (Klovins et al., 1998). Thus, it is possible that other upstream elements of the BYDV genome may play a role in the (−)-strand synthesis. We can rule out bases 2789–4515, 5016–6045, and 5183–5205 because they can be deleted with little effect on BYDV RNA replication (Mohan et al., 1995; Paul et al., 2001).

Regulation of replication by an embedded 3’ end?

The computer modeling and phylogenetic comparisons predict that the 3’ end of BYDV RNA is embedded in a helix that is coaxially stacked with stem 3 (Fig. 1). All changes that disrupted and restored the GGGU5619–5622:ACCC5674–5677 helix destroyed RNA replication (except for the natural U/A covariation in PAV-129 which supports a role for this helix). Given that all Tombusviridae and all viruses with related polymerases terminate in essential CCC or CCA sequences, it is highly likely that the alterations to CCC5675–5677 prevented (−)-strand synthesis, even if the base-pairing were restored. The combination of the covariation mutagenesis data that support functional roles for SL1, SL3, SL4, and the short, four-base SL2, with the probing data that indicate GGGU5619–5622 is base-paired but not to ACUC5638–5641, leave little choice but for the GGGU5619–5622 to base pair to the 3’-terminal ACCC.

The embedded nature of the 3’-terminus is intriguing, given that a free CCA or CCC 3’ end is required for (−)-strand initiation in other viruses. Base-paired 3’-termini was suggested to be a feature peculiar to RNA bacteriophages (Klovins et al., 1998). It is tempting to speculate that the putatively base-paired 3’ end of BYDV RNA negatively regulates minus-strand synthesis. Base-pairing has been implicated in preventing initiation of minus-strand RNA synthesis from promoter-like elements within the TLS in TYMV (Singh and Dreher, 1998). Possible regulatory conformational changes in viral 3’-UTRs have been suggested for AMV (Olsthoorn et al., 1999) and flaviviruses (Khromykh et al., 2001). Conceivably, a conformational switch between the embedded 3’-end structure, and the alternative structure with the free ACCC5674–5677 mediate a switch between negative (free) and positive (embedded) strand synthesis in BYDV. The increase in (−)-strand accumulation resulting from deletion of SL1, and to a lesser extent from deletion of SL2 and SL4, and the concomitant loss of (+)-strand accumulation is consistent with a role of these structures in inhibiting (−)-strand synthesis. (−)-Strand synthesis could be inhibited directly, by the structures themselves, or indirectly by their recruitment of a protein factor. Translation factor EF-1α has been proposed to negatively regulate (−)-strand synthesis of TYMV RNA by binding the TLS (Dreher, 1999). Blocking (−)-strand promoter activity also could facilitate translation of genomic RNA, since translation and replication are incompatible on the same RNA molecule (Gamarnik and Andino, 1998).

Such regulation would predict that the (+)-strand promoters would be different from the (−)-strand promoter. Indeed the 5’ ends of BYDV genomic (+) strand and subgenomic RNAs 1 and 2 begin with an essential GUGAAG sequence that is absent in the 5’ end of the (−) strand, and they show no obvious similarity to the (−)-strand promoter.

Implications for BYDV taxonomy

Isolate PAV-129 differs substantially in sequence from other PAV and MAV isolates in ORFs 1, 2, and in the long 3’-untranslated region. It also has very different symptomatology (Chay et al., 1996). Thus, it might be considered a different virus than BYDV. However, the fact that a chimera containing the replicase from PAV6 and the 3’ origin of replication from PAV129 can replicate in protoplasts (Fig. 6) and in plants (S. Liu, unpublished observations) leads us to conclude that they are merely divergent isolates of one virus species.

The polymerase, translational control signals, and 3’-terminal cis-acting replication structures of BYDV (genus Luteovirus, family Luteoviridae) and those of family Tombusviridae are very similar. In contrast, the polymerase genes and 3’-termini of the genus Polerovirus (ending in GU), family Luteoviridae, bear no resemblance to those
of BYDV (Koonin and Dolja, 1993; Miller et al., 1995). Thus genus *Luteovirus* should perhaps be reassigned to the Tombusviridae.

**MATERIALS AND METHODS**

**Plasmids**

The mutant constructs were derived from the full-length infectious clone of BYDV, pPAV6 (Di et al., 1993), by PCR-mediated site-directed mutagenesis using Vent DNA polymerase (New England Biolabs). The mutants p3'SL1D, p3'SL1UUCG, p3'SL1GAGA, p3'SL1GACA, p3'SL11, p3'SL12, p3'SL13, p3'SL14, and p3'SL2D were constructed by PCR amplification of the 3’ portion of the BYDV genome with the downstream mutagenic primers and the upstream primer CB0416 (GGTCTAGATATACGACTCACTACTATAGAGTGAAAGATTGACCATCTCACAAAAGC). The product was digested with the restriction endonucleases *KpnI* and *SmaI*, purified by 0.8% low-melt agarose gel electrophoresis, and inserted into pPAV6 cut with the same enzymes. The remaining mutants were generated by two-step PCR (Landt et al., 1990). In the first step, upstream mutagenic primers containing the mutant bases flanked by appropriate homologous sequence were used with the downstream primer 3'wt (ATACTACCTATAGGTCACAAAAACAGCGAAAT). The product was gel-purified and used as a downstream primer in the second-step PCR with the upstream primer CB0416. The second-step PCR product was cut with *KpnI* and *SmaI*, gel-purified, and placed in pPAV6 cut with the same enzymes. The construct p3' used for RNA structural probing was generated by subcloning a *BglII–SalI* fragment of pgl016, containing the 3' end of BYDV, into pSS1 (Miller and Silver, 1991) cut with *BamHI* and *SalI*. Plasmid preparations were performed using the QuantumPrep DNA purification kit (Bio-Rad). All mutants were confirmed by sequencing.

pPAV6-129 was derived from pPAV129-CPRT, a plasmid derived from pPAV6, in which ORFs 3, 4, and 5 of pAV6 were replaced by the counterparts from PAV129. To replace the 3'-end sequence of pAV6 in pPAV129-CPRT, a *KpnI–SmaI* fragment (nts 4156–5677) was amplified from two subclones (p129–10 and p129–12) of pPAV129 and joined by three PCR reactions. The first PCR amplified an 832-bp fragment from pPAV129–10, corresponding to positions 4054–4885. The primers used for PCR were 5'-CTGACGCCGTTTTCTACCTTGCG-3' and 5'-CCAATTACCCGAGCTTAGAACC-3' for the 5' and 3' ends, respectively. The second PCR generated a fragment corresponding to bases 4863–5677 in pPAV129, using primers 5'-GGTTGCTAGCCTGGGTTAATTGG-3' (complementary to the above 3' primer) and 5'-GGGATGTGCGCGGCTTCCTTTTC-3' and p129-12 as template. The two PCR products were agarose gel purified and about 50 ng of each used as templates for the third PCR using the 5’ primer of the first PCR with the 3’ primer of the second PCR to generate a single product of 1522 bases. This product was digested with *KpnI* and ligated into the corresponding sites (*KpnI* and *SmaI*) of pPAV129CPRT. The resulting plasmid was confirmed by partial sequencing and restriction mapping.

**Genome-length PCR**

To allow direct in vitro transcription of correctly terminated, full-length viral RNA from uncloned PCR products, full-length cDNA templates were generated by PCR using VENT DNA polymerase (New England Biolabs). An upstream primer, annealing to the 5’ end of the viral genome, 5’-ACGCGGCGCTAATACGACTACTATAAGAGTGAAAGATTGACCATCTCACAAAAGC-3’ (T7 promoter in italics, transcription initiation site underlined, viral sequence in bold), was paired with either of two 3’-end primers, 5’-GGGTTCGGCAAATCGCTTTTGGATGTCGATTAGTTAAGATGTCACC-3’ (wild-type) or 5’-CTCATGCGAAGCTTTCGGAATTTTCGCGTTAAGATGTCACC-3’ (for mutants 3’SLL24 and 3’SLL25). The mutant 3’ primer was used to amplify 3’SLL22 template to generate mutant 3’SLL24. The same primer was used with pAV6 template to generate 3’SLL25. As controls, the wild-type primer was used on the same templates to allow transcription of RNA from PCR products rather than linearized plasmids for comparison purposes. Full-length transcripts were obtained using Ambion’s T7 Megascript kit.

**Protoplast infection and Northern blot analysis**

Wild-type and mutant infectious transcripts were generated by in vitro transcription of *SmaI*-linearized plasmids using Megascript T7 RNA polymerase system (Ambion, Austin, TX). We used 10–15 µg RNA for electroporation of oat protoplasts prepared as described in Dinesh-Kumar and Miller, 1993. Total RNA was extracted from protoplasts ~24 hpi using RNeasy plant RNA isolation kit (Qiagen, Los Angeles, CA). For minus-strand detection, RNA was isolated following the procedure described in Seeley et al., 1992, using aurin tricarboxylic acid as RNase inhibitor. RNA (5–10 mg) was analyzed by Northern blot hybridization essentially as described in Seeley et al., 1992. Equal amounts of total RNA were loaded in each lane as verified by ethidium bromide staining. A 32P-labeled riboprobe complementary to the 3'-terminus of BYDV was used to detect viral gRNA and sgRNAs. For positive-strand detection, plasmid pSP10 (Dinesh-Kumar et al., 1992) was linearized with HindIII and transcribed in vitro with T7 RNA polymerase (New England Biolabs). For negative-strand detection, the same plasmid was linearized with *SmaI* and transcribed in vitro with SP6 RNA polymerase (Promega, Madison, WI). GeneScreen nylon membranes (DuPont) were hybridized with the probes and exposed to phosphorim-
ager screens (Molecular Dynamics, Sunnyvale, CA) for 5–24 h (positive-strand detection) or 2–4 days (negative-strand detection).

Transcripts of the stem-loop deletion mutants were tested for translation in wheat germ extract translation system (Promega) following manufacturer’s instructions, using 35S-labeled methionine, as described in Wang et al., 1999. Translation products were analyzed by SDS–PAGE (10% acrylamide) as described in Wang et al., 1999. All experiments were performed at least twice.

RNA sequence and structural analysis

Sequence alignments of BYDV isolates were performed using the GCG software package (Madison, WI). RNA secondary structure was predicted using MFOLD, version 3.0, at the MFOLD web site (http://bioinfo.math.rpi.edu/~mfold/) (Mathews et al., 1999; Zuker et al., 1999). RNA secondary structure probing was performed with imidazole (Vlassov et al., 1999) and with T1 nuclease (Miller and Silver, 1991) on in vitro transcripts 5′ end-labeled with 32P essentially as described in Koev et al., 1999. Imidazole cleavage was performed at 25°C for 1–2 h. The transcripts were derived from p31991999. Translation products were analyzed by SDS–polyacrylamide gel electrophoresis (PAGE) for detection of minus-strand viral RNA, Mark Young and Sergei Filichkin for providing the NRI. We thank Kay Scheets for advice on RNA extraction for detection using the GCG software package (Promega) following manufacturer’s instructions, tested for translation in wheat germ extract translation system (Promega) following manufacturer’s instructions, using 35S-labeled methionine, as described in Wang et al., 1999. Translation products were analyzed by SDS–PAGE (10% acrylamide) as described in Wang et al., 1999. All experiments were performed at least twice.

ACKNOWLEDGMENTS

This work was supported by Grant 2001-35319-10011 from USDA-NRI. We thank Kay Scheets for advice on RNA extraction for detection of minus-strand viral RNA, Mark Young and Sergei Filichkin for providing unpublished sequence of the 3′ end of BYDV-MAV, Theo Dreher for useful discussion, and Jacquelyn Jackson for advice on full-length RT-PCR. This is paper J-18711 of the Iowa Agriculture and Home Economics Experiment Station, Project 3545.

REFERENCES


---

**BYDV RNA 3′ END STRUCTURE**

125

---


