

Stem-loop Structure in the 5' Region of Potato Virus X Genome Required for Plus-strand RNA Accumulation

Eric D. Miller¹, Carol A. Plante¹, Kook-Hyung Kim¹, James W. Brown² and Cynthia Hemenway^{1*}

¹Department of Biochemistry and ²Department of Microbiology, North Carolina State University, Raleigh NC 27695, USA

Computer-generated thermodynamic predictions and solution structure probing indicated two stem-loop structures, stem-loop 1 (SL1; nt 32-106) and stem-loop 2 (SL2; nt 143-183), within the 5' 230 nt of potato virus X (PVX) RNA. Because the existence of SL1 was further supported by co-variation analysis of several PVX strains, the functional significance of this structure was investigated by site-directed mutational analysis in a tobacco protoplast system. In general, mutations that reduced genomic plus-strand RNA accumulation similarly affected coat protein accumulation, indicating that subgenomic plus-strand RNA was also affected. In contrast, minus-strand RNA levels remained relatively unchanged. Mutational analysis of the stem C (SC) region of SL1 indicated that pairing was more important than sequence, which was consistent with the covariation analysis. Alterations that increased length and stability of either SC or stem D (SD) were deleterious to plus-strand RNA accumulation. The formation of internal loop C between SC and SD, as well as specific nucleotides within this loop, were also required. Several modifications were made to the terminal GAAA tetraloop, a motif known for enhanced RNA stability. Both GANA and GAAG motifs resulted in wild-type levels of RNA accumulation. However, a UUCG tetraloop was detrimental, indicating that the sequence of this element was important beyond just providing stabilization of the structure. These data indicate that multiple features of SL1 are critical for accumulation of PVX plus-strand RNA.

© 1998 Academic Press

Keywords: RNA structure; RNA virus; RNA replication; potato virus X; stem-loop

*Corresponding author

Introduction

Genome replication for many single-stranded, plus-sense RNA viruses requires production of minus-strand RNA, which in turn serves as the template for genomic and subgenomic (sg) plus-strand RNA species. *Cis*-acting signals that regulate this process have been found at the 5' and 3' termini of the RNA genomes and proximal to the

initiation sites for sgRNAs in both plant and animal viral systems (Duggal *et al.*, 1994; Buck, 1996). Several studies have documented that RNA secondary structures at the 3' terminus are important for RNA replication (Jaspars, 1985; Jacobson *et al.*, 1993; Song & Simon, 1995; Hsue & Masters, 1997; Lauber *et al.*, 1997), and that structural elements can affect host protein binding specificity *in vitro* (Blackwell & Brinton, 1995; Ito & Lai, 1997). Additionally, structures at the 5' terminus have also been found to affect viral RNA accumulation (Andino *et al.*, 1990; Niesters & Strauss, 1990; Pogue & Hall, 1992; Gilmer *et al.*, 1993; Hellendoorn *et al.*, 1997). Although these studies indicate that RNA sequences and/or structures are important during replication, the relationship between structural features and individual aspects of RNA synthesis needs to be further defined.

Abbreviations used: SL, stem-loop; PVX, potato virus X; SC, stem C; SD, stem D; sg, subgenomic; NTR, non-translated region; ORFs, open reading frames; TB, triple gene block; CP, coat protein; OAS, origin of assembly; TL, tetraloop; DMS, dimethyl sulfate; NT-1, *Nicotiana tabacum* strain; SRP, signal recognition particle; AMV, avian myeloblastosis virus; DEPC, diethyl pyrocarbonate; h.p.i, hours post-inoculation; LC, loop C.

E-mail address of the corresponding author: cindy@bchserver.bch.ncsu.edu

Potato virus X (PVX), the type member of the potexvirus group, is a flexuous, rod-shaped virus that contains a positive-sense 6.4 kb RNA genome. The capped and polyadenylated viral RNA is composed of an 84 nt 5' non-translated region (NTR), five open reading frames (ORFs), and a 72 nt 3' NTR (Huisman *et al.*, 1988; Skryabin *et al.*, 1988). ORF1 encodes the viral replicase protein, which is the only viral protein absolutely required for PVX RNA synthesis. Three ORFs (ORFs 2-4), collectively referred to as the triple gene block (TB), encode products necessary for viral cell-to-cell transport (Beck *et al.*, 1991; Angell *et al.*, 1996). The viral coat protein (CP), which is required for viral movement (Chapman *et al.*, 1992; Baulcombe *et al.*, 1995) and encapsidation, is encoded by ORF5. Proteins corresponding to these ORFs are detectable in infected tissue (Price, 1992), as are several RNA products e.g. genomic plus-strand and corresponding minus-strand RNAs, two predominate sgRNA species that initiate upstream of the TB or CP genes, and double-stranded versions of these products (Dolja *et al.*, 1987).

The 5' region of PVX RNA is likely to contain signals for translation, encapsidation and RNA synthesis. Two regions of the PVX 5' NTR, designated α (nt 1-41) and β (nt 42-84), were found to enhance translational efficiency of reporter genes *in vitro* (Smirnyagina *et al.*, 1991) and *in vivo* (Pooggin & Skryabin, 1992; Zelenina *et al.*, 1992). Although the encapsidation signal or origin of assembly (OAS) for PVX has not been defined, the first 47 nt in the 5' NTR of a related potexvirus, papaya mosaic virus, support assembly (Sit *et al.*, 1994). Deletion analysis of the PVX RNA 5' NTR has shown that multiple elements in both of these regions regulate genomic and sg plus-strand RNA accumulation (Kim & Hemenway, 1996). Although multiple elements in the PVX 5' NTR are clearly important for virus replication, very little is known about how specific sequences and/or structures affect different aspects of the infection process.

Given the importance of the 5' region, we initiated studies to define the role of RNA structure in PVX replication. Thermodynamic predictions and analysis of solution structure by chemical modification and RNase digestion indicated two well-defined stem-loop (SL) structures within the 5' 230 nt of PVX RNA, both of which terminate in stable tetraloops. Covarying nucleotides found in the stem regions of SL1 suggested that formation of this structure is important for PVX replication. To understand how individual structural elements in PVX SL1 function in virus replication, we analyzed the effects of site-directed mutations in this SL on accumulation of PVX RNAs in a tobacco protoplast system. Using this approach, we identified key sequence and structural elements required for PVX RNA accumulation, and the data indicate that these elements affect both genomic and sg plus-strand RNA levels.

Results

Thermodynamic predictions and solution structure probing indicate two well-defined, stem-loop structures in the first 230 nt of PVX RNA

Stable secondary structure elements in the 5' region of PVX RNA were first evaluated by computer-generated thermodynamic predictions. The energy dot plot shown in Figure 1A represents all possible structures within 10% of the predicted minimal energy for the 5' 500 nt of PVX RNA. The optimal folding of this region of PVX RNA is depicted in the lower left triangle, where the minimum energy structure is represented by dark dots, and rows of dots denote base-paired regions. Within the upper right triangle, all suboptimal foldings within 4 kcal of the minimum free energy are represented. The optimal folding is again represented by dark dots, and alternative pairings are evident as the dots become lighter in grayness; the least-favored pairings are depicted as the lightest dots. Regions that are cluttered with differently shaded dots represent multiple folding possibilities that are considered less well defined (Jaeger *et al.*, 1989a,b; Zuker, 1989).

Although several SL structures are predicted by the dot plot (Figure 1A) to form in the 5' 500 nt of PVX RNA, we focused on the three SL structures within the first 230 nt. The predicted optimal secondary structure for this region, shown in Figure 1B, indicated that the first 31 nt at the 5' end were unpaired. Of the three SL structures predicted, only the first two were strongly supported by this thermodynamic analysis. SL1 (nt 32-106) was predicted to contain four stem regions (SA, SB, SC and SD), three internal loops (LA, LB and LC), two mismatches, and a terminal tetraloop (TL). Some scatter was observed near the base of SL1, potentially indicating flexibility of pairing in this region. SL2 (nt 143-182) was predicted to contain three stem regions, one mismatch, two bulges, one internal loop, and a terminal tetraloop. SL3, which includes residues 194-229, was less well-defined, as indicated by the scatter of differently shaded dots.

Chemical and enzymatic probing were used to experimentally verify the thermodynamically predicted secondary structures. Transcripts (543 nt) probed in these experiments were generated by runoff transcription of a PVX cDNA clone, pMON8453 (Hemenway *et al.*, 1990). Preliminary experiments indicated that modification and digestion patterns of these truncated transcripts were similar to those for authentic viral RNA (data not shown). Dimethyl sulfate (DMS) modification of unpaired adenosine (A) and cytidine (C) bases was done at several temperatures (4, 25, 37 and 90°C) to determine the relative reactivities under different conditions. Cleavage of unpaired guanosine (G) residues with RNase T₁ and of double-stranded helical or stacked helical regions with nuclease V₁

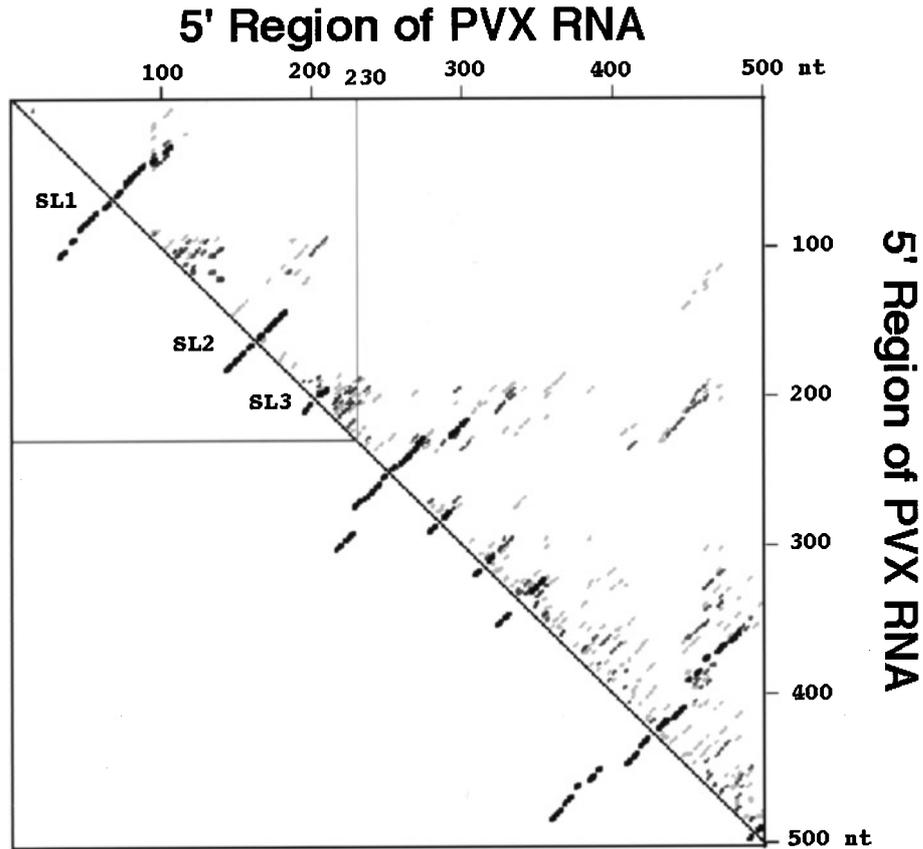
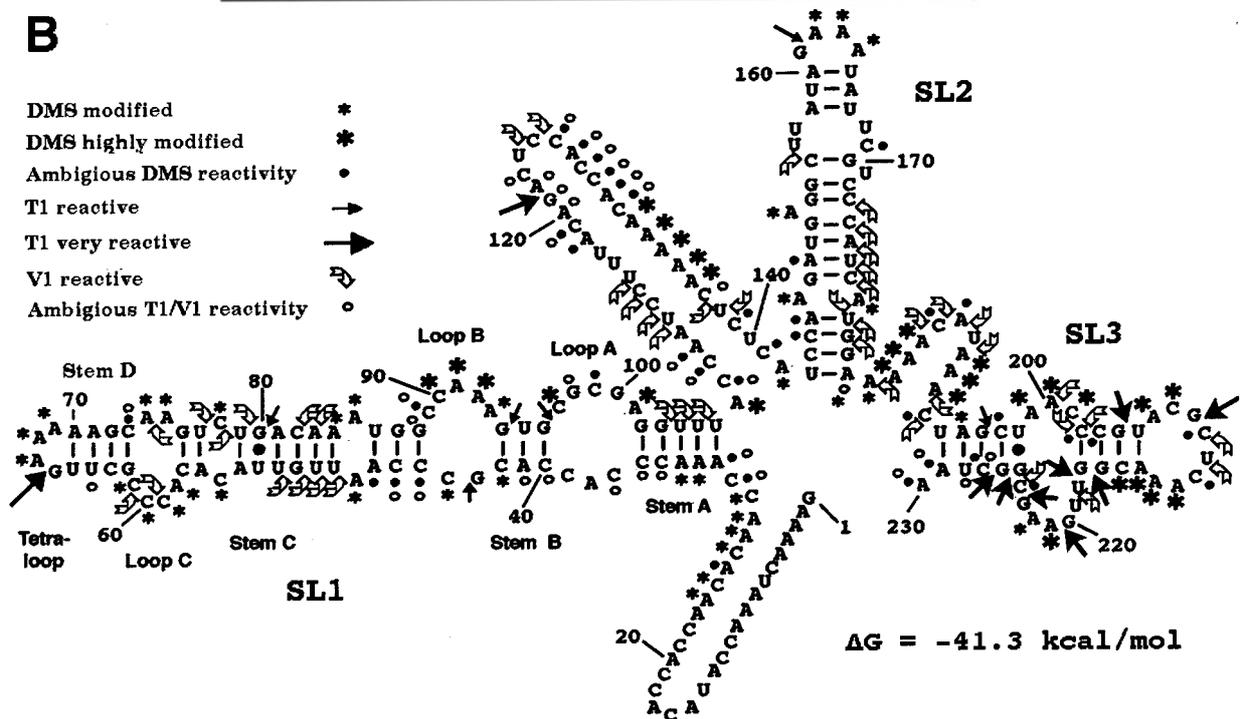
A**B**

Figure 1. Structural analysis of the 5' region of PVX RNA. A, the thermodynamic energy dot plot is shown for all structures within 10% of the most optimal energy (ΔG); a total of seven structures were recovered and plotted. The axes are labeled to indicate the position of each nucleotide in the PVX RNA sequence from nt 1-500. Three stem-loop structures predicted within the 5' 230 nt are boxed and indicated by SL1, SL2 and SL3. B, reactivities in the 5' 230 nt of PVX transcripts to chemical and enzymatic probing are superimposed on the optimal predicted secondary structure model. Positions of DMS reactivity are indicated by asterisks (*), with larger asterisks indicating a greater degree of modification. Filled arrows (\rightarrow) indicate positions of RNase T₁ cleavage, with larger arrows denoting greater reactivity. Positions of RNase V₁ cleavage are marked by open, curved arrows (\curvearrowright). Ambiguous DMS and T₁/V₁ reactivity are indicated by filled and open circles, respectively. Unreactive positions are not marked. Within SL1, four stem regions (SA, SB, SC and SD), three internal asymmetric loops (LA, LB and LC), and a terminal tetraloop (TL) are labeled.

(Lockard & Kumar, 1981) was done at 25°C. Reaction conditions for both DMS modification and RNase digestion experiments were optimized such that less than one nucleotide per RNA molecule was modified. Products of these reactions were evaluated by primer extension with three oligonucleotide primers, eP1, eP2 and eP3, which are complementary to nt 78-97, 148-179 and nt 242-263, respectively.

A summary of the data obtained by DMS modification and RNase digestion of the 5' 230 nt are superimposed on the most stable secondary structure prediction shown in Figure 1B. Primer extension patterns for part of SL1 are depicted in Figure 2. Modification patterns of other regions within the 5' 230 nt were evaluated from additional gels (data not shown). Individual nucleotides were determined to be reactive when: (1) the signal of the extension product was greater in modified than unmodified samples; and (2) the intensity of the extension product was greater at 25 or 37°C than at 90°C, indicating a chemical reactivity at that nucleotide greater than in the random, unfolded molecule. If either of these criteria were not met, residues were considered to be unreactive (not numbered). When considerable background extension products were evident, reactivity was considered to be ambiguous and was noted by filled (DMS) or open (RNases) circles.

The structures predicted for the TL, SD, LC and SC regions of SL1 (nt 45-87), were consistently supported by the modification analyses (Figures 1B and 2). In the TL region, G66 was strongly reactive to RNase T₁ and all three A residues in this region (A67, A68 and A69) were modified by DMS. The A residues of the A·U base-pairs predicted to close this loop (A70, A71) were not accessible to DMS modification, indicating that the predicted terminal loop formed in solution. The other residues in SD also were not reactive to DMS or RNase T₁, indicating pairing in this region. All the residues in LC were accessible to DMS and several exhibited RNase V₁ cleavage (C60, C61, A74 and C78), suggesting that some of these residues may have formed helical, stacked conformations. Within the SC region, most residues predicted to be paired were not susceptible to DMS (Figure 2A) or RNase T₁ (Figure 2B), but did exhibit considerable sensitivity to RNase V₁. The exceptions to this were the A84 modification by DMS, G80 cleavage by RNase T₁, and ambiguous reactivities of A47, C46 and C45 at the base of SC. Residues predicted to be mismatched in SC were accessible to DMS. Given the patterns of reactivity for SD and SC, and accessibility of LC to RNase V₁, it is possible that these regions formed a largely paired and/or helical conformation.

Based on both thermodynamic and modification data shown in Figure 1, it is likely that the portion of SL1 (nt 32-47: nt 86-106) containing SA, LA, SB and LB can form multiple conformations. In the SA region, A32, A33 and A34 were predicted to be paired to U104, U105 and U106. These U residues

were susceptible to RNase V₁, indicating they were paired or in a helical arrangement. However, A33 and A34 were modified by DMS and A32 reactivity was ambiguous, which would argue against pairing in this region of SA. Although residues C35 and C36 exhibited ambiguous reactivity to RNase V₁, they were not accessible to DMS. In addition, the complementary G102 and G103 were not cleaved by RNase T₁, indicating that two predicted base-pairs in this region of SA formed in solution. In the region of asymmetric LA, residues C37, A38 and C39 on the 5' side of this loop and residues G98, C99 and G100 on the 3' side did not exhibit reactivity to DMS or RNase digestion. In contrast, C97 and A101 on the 3' side were reactive to DMS. These data indicate that pairing may occur across the central portion of this loop between C37A38C39 and G98C99G100. Alternatively, these residues may pair with partners elsewhere in the RNA. Because residues in both SB and LB regions were reactive to DMS or RNase T₁, it is possible that SB and LB formed one large loop. One of the suboptimal structures predicted by the thermodynamic analysis exhibited such a conformation (data not shown).

Considerable RNase V₁ digestion and the absence of DMS modification to most C and A residues in the predicted stem region of SL2 (Figure 1B) indicated base-pairing or stacking. The A152 residue predicted to be bulged and the A147/A178 mismatch were modified by DMS. Accessibility of G161 to RNase T₁ and A162, A163 and A164 to DMS supported the terminal TL in SL2. As was observed for SL1, some residues near the base of SL2 exhibited ambiguous reactivities. The region separating SL1 and SL2, which was predicted to have several alternative structures or to be unstructured (Figure 1), exhibited many ambiguous reactivities. However, some residues in this area were clearly unreactive to DMS (C113, C114 and C125) and were also digested by RNase V₁, suggesting that they either paired with residues in this or other regions or were in stacked/helical conformations. Also, A132, A133, A134, A135, A136 and C137 were highly reactive to DMS, indicating that they were unpaired.

In support of the thermodynamic predictions, the modification studies indicated that SL3 was the least likely of the predicted structures to form in the 5' 230 nt of PVX RNA. Within the SL3 region (nt 194-229) and the region between SL2 and SL3 (nt 183-193), most C and A residues were modified by DMS and all G residues were accessible to RNase T₁ (Figure 1B). Although some scattered RNase V₁ digestions were observed, they did not correlate with regions predicted to be paired.

Covariation analysis of sequences from various strains of PVX supports the existence of SL1

Six strains of PVX (Figure 3A), which belong to three out of four strain groups and exhibit 85–98%

A

Group 2	PVX CP	CACGCCCAAUU AA CACACACCCCGCUUG AAAAAG AA AGUCUG UU AAAUGGCC
		40 51 62 67 73 78 80 90
Group 3	PVX X3	CACGCCCAAUU GUU ACACACCCCGCUUG G AAAAG C AAAGUCU AA CAAAUGGCC
	PVX R	CACGCCCAAUU GUU ACACACCCCGCUUG G AAAAG U AAAGUCU AA CAAAUGGCC
	PVX UK3	CACGCCCAAUU GUU ACACACCCCGCUUG AAAAAG AA AGU U AA CAAAUGGCC
	PVX WA	CACGCCCAAUU GUU ACACACCCCGCUUG AAAAAG CA AGUCU G ACAAAUGGCC
Group 4	PVX HB	CACGCCCAAUU AA CACACACCC A CUUG AAAAAG U AAAGUCUG UU AAAUGGCC

B

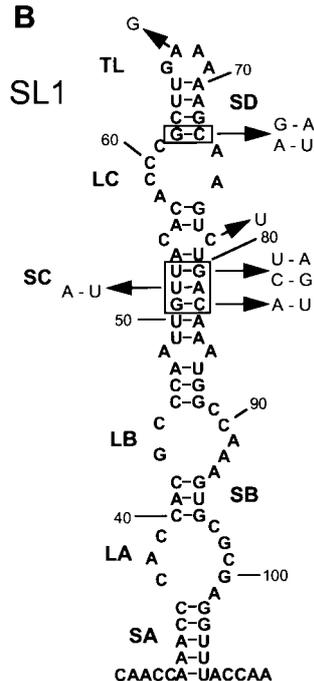


Figure 3. Sites of RNA sequence covariation in the SL1 region of PVX RNA. A, all six PVX sequences between nt 40 and 90 are shown and grouped according to guidelines established by Cockerham (1955). Nucleotide positions involved in compensatory changes or covariation that are different from our isolate (PVX-WA) are shown in bold face and underlined. Also shown are nucleotide changes that conserve the GNRA tetraloop motif (PVX-X3 and PVX-R) and the mismatch in stem C of SL1 (PVX-UK3). B, the optimal secondary structure of SL1 is labeled as in Figure 1. Covarying positions are marked by boxes, and arrows indicate changes observed in the six PVX strains.

Four sets of nucleotides were identified to covary: G51-C82, U52-A81, U53-G80, and G62-C73 (Figure 3B). Changes in residues G51 to A, U52 to A and U53 to C were compensated by changes of C82 to U, A81 to U, and G/A80 to G, strongly supporting the existence of SC. The G62-C73 pair also covaries in strains PVX-CP, PVX-UK3, PVX-R and PVX-HB, which would indicate that pairing at the base of SD is also critical for PVX multiplication. For two of the strains (PVX-UK3 and PVX-CP), a G-A base-pair is predicted in this region, suggesting that these nucleotides may form non-canonical base-pairs to maintain the structure. Thus, the covariation analysis provided additional evidence for the paired central portion of the SL1 structure and its significance to survival of the virus.

When the corresponding 3' 230 nt, minus-strand RNA sequences from the different PVX strains were analyzed, no covariation was apparent. Although thermodynamic predictions indicated secondary structures in this region of the minus-strand RNA (data not shown), these structures were not supported by sequence covariations. However, given the limited sample size for these studies, the information may not be sufficient to

completely rule out this possibility. Thermodynamic structural comparisons and covariation analysis among different potexviruses were not extensively analyzed because low sequence similarity (25–40%) and diversity in leader length (72–107 nt) made the identification of homologous residues impossible.

Structure near the base of stem C in SL1 is important for PVX plus-strand RNA accumulation in tobacco protoplasts

The significance of SL1 structural features to viral RNA accumulation was determined by inoculation of transcripts from modified or wild-type PVX cDNA clones into *Nicotiana tabacum* (NT-1) protoplasts. Six mutations (SC1-SC6) were introduced into the SC region of SL1, as shown in Figure 4D. The SC1 (C46 to U), SC2 (A47 to C), SC4 (A47A48 to CC), SC5 (C46A47A48 to AU) and SC6 (U86 to A) mutations were predicted to change the optimal, wild-type secondary structure near the base of SL1 and alter the free energy of the SL1 structure. In addition, the optimum predicted structures for the 5' 230 nt of these mutants indicated disruption of SA and formation of new

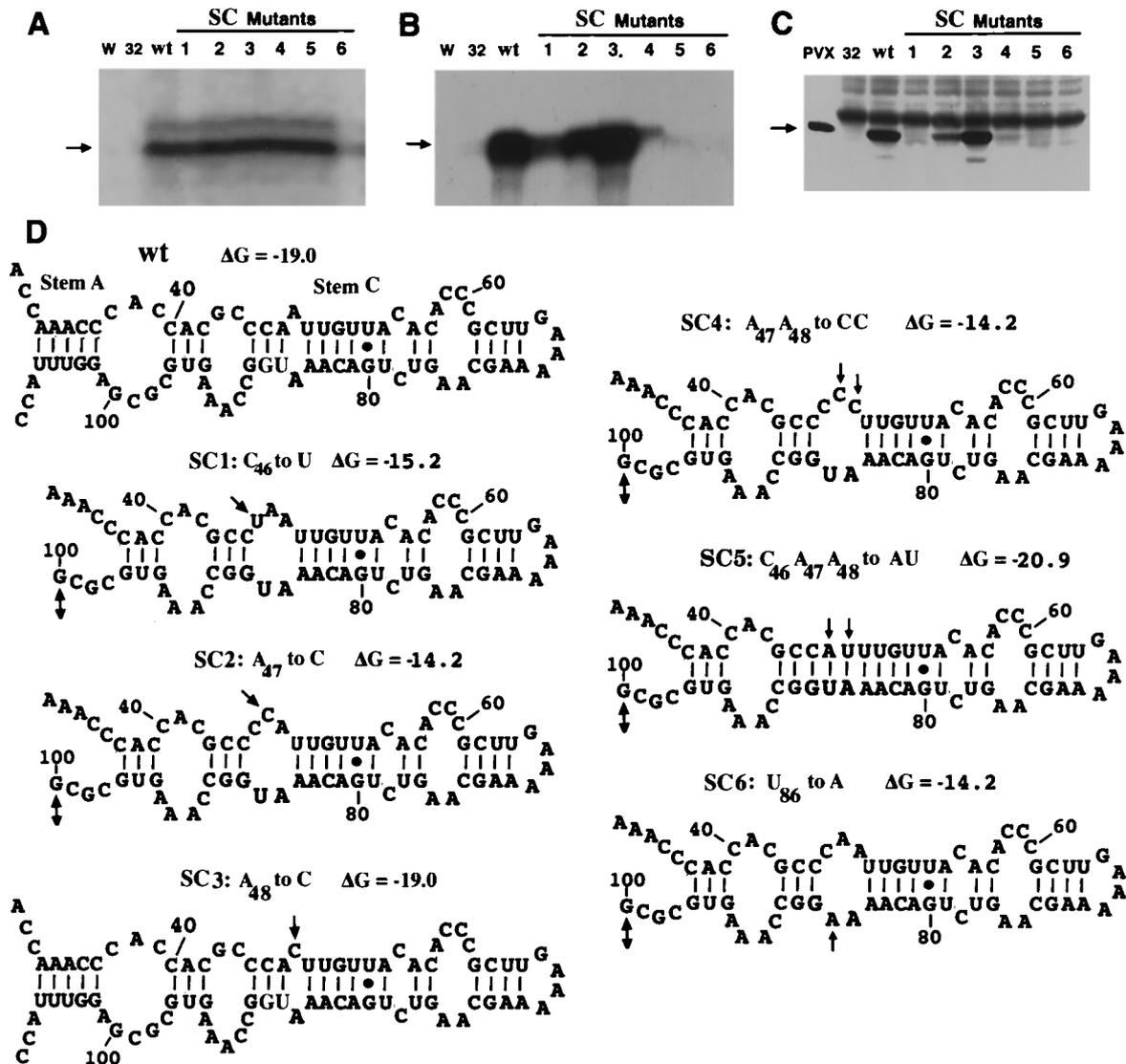


Figure 4. Effects of mutations near the base of stem C on accumulation of PVX RNAs and coat protein. Protoplasts inoculated with water (W), replication-defective transcripts derived from p32 (32), wild-type transcripts derived from pMON8453 (wt), or mutant transcripts SC1-SC6 were analyzed at 48 hours post-inoculation (h.p.i.). S₁ nuclease protection was utilized for detection of minus (A) and plus-strand (B) RNA accumulation. C, The relative coat protein accumulation as detected by Western blotting. A comparison of the predicted optimal secondary structures and free energy (ΔG) values for SL1 of wild-type and mutant RNAs is illustrated in D. Single-headed arrows indicate positions of mutations; double-headed arrows mark positions that pair with nt downstream to form stem-loop structures between SL1 and SL2. The positions (nt 85-87) corresponding to the initiation codon for the replicase protein are marked by enlarged type in all of the structures.

SL structures positioned between SL1 and SL2 (data not shown). Although these SLs appeared in all structures predicted to be within 10% of the optimum folded structure, there was considerable scatter in the energy dot plots in this region, indicating less-favored structures. Although one might predict that changing C46 to U in mutant SC1 would still enable pairing to G87, as shown in the wild-type structure, the thermodynamic prediction for the optimum structure of mutant SC1 indicates restructuring at the base of SL1. In addition, this G-U pair did not show up in any of the nine structures predicted to be within 10% of the optimum

structure for SC1 (data not shown). The SC6 mutation (U86 to A) additionally changed the initiation codon of the replicase gene. In contrast to all other mutants, conversion of the the A-A mismatch to a C-A mismatch in mutant SC3 (A48 to C) did not result in any other alterations to the predicted structure or stability of SL1.

Protoplasts electroporated with transcripts containing these mutations were analyzed for minus-strand RNA (Figure 4A), plus-strand RNA (Figure 4B), and CP (Figure 4C) accumulation. Minus-strand RNA accumulation in protoplasts inoculated with SC1, SC2, SC3, SC4 or SC5 tran-

Table 1. Plus and minus-strand RNA accumulation in NT-1 protoplasts

Construct	Relative RNA levels (%)		Construct	Relative RNA levels (%)	
	(+)	(-)		(+)	(-)
w.t.	100 ± 0	100 ± 0	LC1	7 ± 4	78 ± 9
p32	0 ± 0	0 ± 0	LC2	66 ± 13	85 ± 9
			LC3	48 ± 8	88 ± 8
SC1	16 ± 5	86 ± 4	LC4	7 ± 4	85 ± 12
SC2	36 ± 7	106 ± 9	LC5	0 ± 0	70 ± 11
SC3	159 ± 16	108 ± 8	LC6	9 ± 3	81 ± 13
SC4	3 ± 3	90 ± 10	LC7	0 ± 0	85 ± 11
SC5	0 ± 0	73 ± 11	TL1	40 ± 10	86 ± 10
SC6	0 ± 0	4 ± 2	TL2	19 ± 10	83 ± 14
			TL3	83 ± 15	84 ± 14
SC7	0 ± 0	65 ± 10	TL4	215 ± 55	86 ± 16
SC8	0 ± 0	120 ± 20	TL5	152 ± 22	96 ± 11
SC9	26 ± 15	96 ± 14	TL6	22 ± 2	93 ± 17
SC10	1 ± 0	88 ± 12	TL7	138 ± 27	90 ± 17
SC11	111 ± 10	98 ± 8	TL8	2 ± 1	95 ± 18
SC12	0 ± 0	88 ± 16			
SD1	0 ± 0	83 ± 14			

Construct designations are as noted in the text and Figure legends. Plus and minus-strand RNA levels were measured with a Molecular Dynamics Phosphorimager. The average value and standard error for each mutant were calculated from three or more independent experiments. Relative percentage values shown were normalized relative to RNA levels observed in protoplasts inoculated with transcripts derived from the wild-type (w.t.) PVX cDNA clone, pMON8453.

scripts was similar to levels in wild-type inoculated protoplasts, although the small reductions observed for SC1 (86 ± 4%) and SC5 (73 ± 11%) were significantly different (Figure 4A; Table 1). Because the predicted structure for SC5 has a higher stability and exhibits pairing of the nucleotides for the initiation codon, the reduction in minus-strand RNA levels may reflect reduced translation of the replicase gene. Protoplasts inoculated with SC6 transcripts were greatly reduced in minus-strand RNA levels (4 ± 2%). This reduction may be due to altered RNA structure, but more likely reflects the modified replicase initiation codon. Although there is another in-frame methionine codon 31 amino acids downstream of the initiation codon, the product translated from this codon may have decreased conformational stability and/or activity.

In contrast to their effects on minus-strand RNA accumulation, all mutations near the base of SC substantially altered plus-strand RNA (Figure 4B; Table 1). In addition, changes in the levels of CP accumulation mirrored the corresponding changes in genomic plus-strand RNA (Figure 4C), suggesting that sgRNA levels were similarly affected. The reduction of plus-strand RNA accumulation to 16 ± 5% in protoplasts inoculated with the SC1 mutant indicated a requirement for a C at position 46 and/or pairing at the base of SC. Likewise, alteration of A47 to C in the SC2 mutant resulted in a reduction of plus-strand RNA to 36 ± 7%. In contrast, conversion of the mismatch A48 to C in mutant SC3, which did not alter the predicted structure for SL1, enhanced plus-strand RNA levels to 159 ± 16%. When this A48 to C change was combined with A47 to C in mutant SC4, the structure at the base of SC was altered and plus-strand RNA levels were reduced to 3 ± 3% of wild-type levels.

Replacement of C46A47A48 with AU in mutant SC5 disrupted the structure in SC, with little change in stability of SL1, and resulted in no detectable accumulation of plus-strand RNA. Although the 27% decrease in minus-strand RNA levels observed for SC5 may have accounted for some decrease in plus-strand RNA levels, it is likely that this mutation directly affected plus-strand RNA accumulation. In contrast, the absence of plus-strand RNA accumulation in protoplasts inoculated with SC6 likely reflects the reduction in minus-strand RNA levels (4 ± 2%) observed for this mutant.

These data indicate that accumulation of PVX plus-strand RNA requires either specific bases at positions 46 and 47, or pairing of these positions at the base of SC. The positive effect achieved by replacement of A at position 48 with a C suggests a requirement for a mismatch in this region, which may be maintained by pairing of C46A47 with U86G87.

Plus-strand RNA levels are affected more by pairing and stability than sequence in the central region of stem C

Five mutations (SC7-SC11) were introduced to evaluate the importance of pairing and sequence in the central region of stem C to RNA accumulation. As shown in Figure 5D, deletion of the 5' side of the stem (nt U50G51U52U53) in mutant SC7 and deletion of the 3' side of this stem in mutant SC8 (nt G80A81C82A83) resulted in significantly increased ΔG values. The optimum predicted structures for both SC7 and SC8 resulted in rearrangement of SA, LA, SB and LB, but more of the pairing in SC was maintained in the SC8 mutant. Mutant SC9 includes both of the deletions in SC7 and SC8. Although the overall predicted

structure for SL1 in SC9 was similar to wild-type, the removal of four base-pairs in the upper region of stem C resulted in a significantly higher ΔG

value. Mutants SC10 and SC11 were designed to further address the issue of sequence *versus* pairing in this region. Alteration of the sequence on the 5'

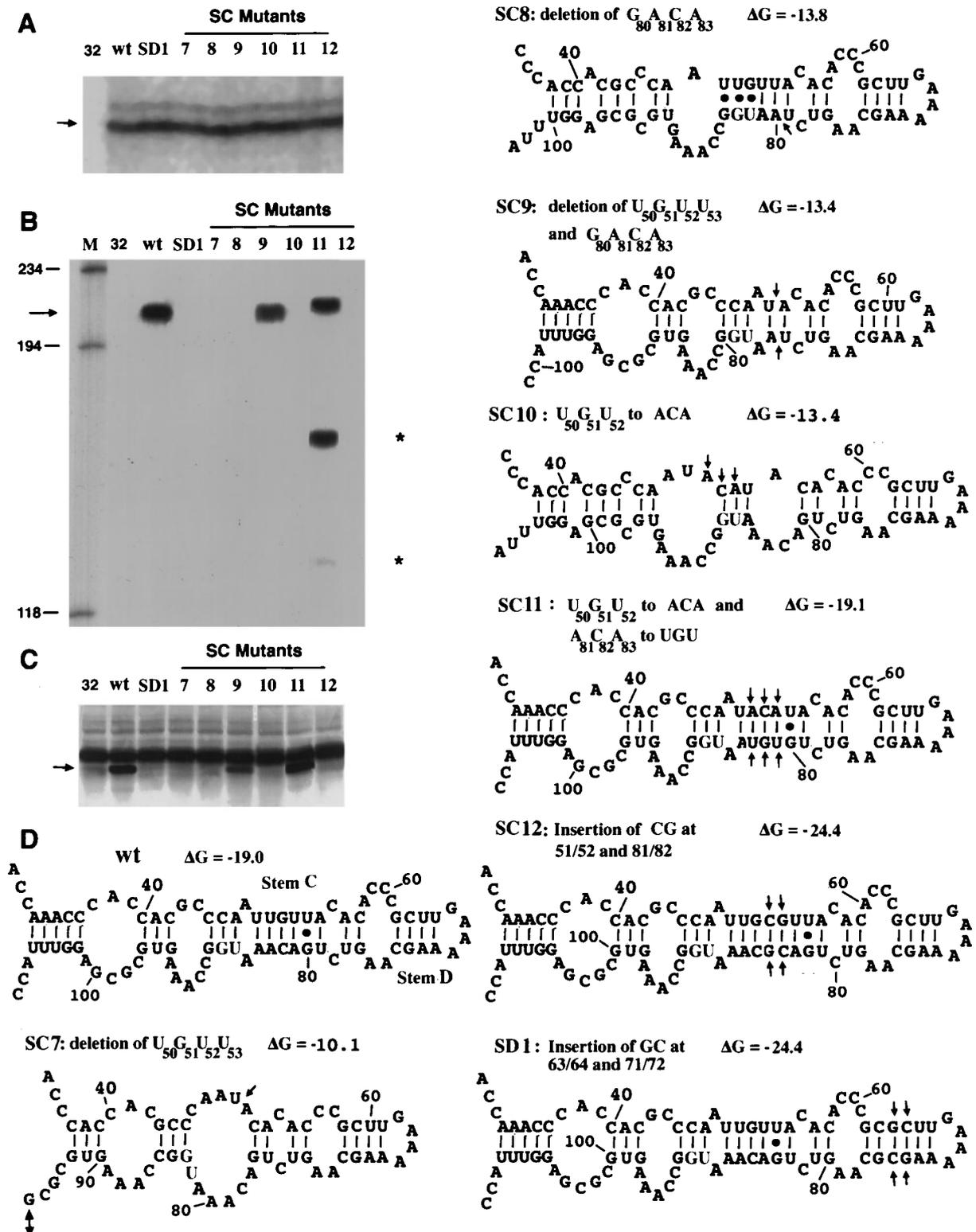


Figure 5. Effects of mutations in the central region of stem C and in stem D on accumulation of PVX RNAs and coat protein. Protoplasts inoculated with replication-defective transcripts derived from p32 (32), wild-type transcripts derived from pMON8453 (wt), or mutant transcripts SC7-SC12 and SD1 were analyzed at 48 h.p.i. Panels and labeling are as described in the legend to Figure 4.

side of the stem (U50G51U52 to ACA) in mutant SC10 resulted in an increased ΔG value that was similar to values for the deletion mutants, and the optimum predicted structure for this mutant was substantially altered below the C-C mismatch in stem C. When compensating changes were introduced into the 3' side of SC in mutant SC10 to generate SC11 (U50G51U52 to ACA:A81C82A83 to UGU), the optimum structure and ΔG value of SL1 were not predicted to change from the wild-type RNA.

Minus-strand RNA production in protoplasts inoculated with these mutants was not significantly different from levels in protoplasts inoculated with wild-type transcripts, except for SC7 ($65 \pm 10\%$; Figure 5A; Table 1). The overall reduced pairing and stability of SL1 in mutant SC7 may contribute to decreased translation of the replicase gene and thus account for some reduction in minus-strand RNA accumulation. In contrast, plus-strand RNA accumulation and CP levels were decreased for most of these mutants (Figure 5B, C; Table 1). Protoplasts inoculated with either SC7 or SC8 exhibited no detectable accumulation of genomic plus-strand RNA or CP. The substantially altered structures and stabilities predicted for SL1, or the removal of critical nucleotides in these mutants may account for the reduction in plus-strand RNA accumulation. Because protoplasts inoculated with the corresponding double mutant, SC9, did accumulate plus-strand RNA ($26 \pm 15\%$) and CP, identity of the deleted sequences may not be as critical as stability. Also, the reduced stability due to reduced stem length or altered spacing in the stem C region of SC9 may account for the reduction in plus-strand RNA levels compared to those in protoplasts inoculated with wild-type transcripts.

The pattern of plus-strand RNA and CP accumulation (Figure 5B, C; Table 1) in protoplasts inoculated with mutants SC10 and SC11 further supports the importance of pairing, rather than sequence, in the central region of SC. Conversion of the sequence on the 5' side of the stem in mutant SC10 (U50G51U52 to ACA), which substantially altered the predicted structure and stability, resulted in no detectable accumulation of plus-strand RNA or CP. In contrast, when wild-type structure and stability were restored in SC11 (U50G51U52 to ACA and A81C82A83 to UGU) by introduction of the complementary sequence on the 3' side of SC in mutant SC10, plus-strand accumulation was at wild-type levels. Thus, the sequence in this region was not a factor in RNA accumulation. As indicated in Figure 5B, the S_1 nuclease digestion pattern of RNA isolated from protoplasts inoculated with SC11 resulted in three bands, each corresponding to the region of mismatch with the probe, P1. From top to bottom, these bands correspond to the full-length protected fragment, the product cleaved at nt 51-53, and the product cleaved at nt 81-83. These products were also observed when purified transcripts of SC11

were similarly treated (data not shown), further verifying that the multiple digestion products were generated by probe mismatch with this mutant. Although not evident in Figure 5B, much longer exposures indicated similar sized smaller bands for the mutant SC9. It is likely that these bands were less evident for SC9 because plus-strand RNA accumulation was significantly lower than observed for SC11. Similarly, such bands were not observed for SC7, SC8 and SC10 because plus-strand RNA did not accumulate. These data suggest that pairing, stability and/or spacing, but not precise sequence, in the central region of SC are important for plus-strand RNA accumulation.

Increased length and stability in stems C and D result in reduced plus-strand RNA levels

To further analyze the effects of stem length and stability, we inserted two G·C pairs into the upper region of SC and into SD to create mutants SC12 and SD1, respectively. For both of these mutants, the optimum structures were predicted to be similar to wild-type, but with lower ΔG values. Although minus-strand RNA levels were unaffected in protoplasts inoculated with SC12 or SD1 (Figure 5A; Table 1), this increase in length in either stem proved to be deleterious to plus-strand RNA and CP accumulation (Figure 5B, C; Table 1). Thus, as was observed for mutant SC9 described above, our data indicate that stem length and/or optimum stability in SC and SD are critical for plus-strand RNA accumulation.

Sequence and formation of internal loop C are important for plus-strand RNA accumulation

Several mutants were constructed to determine if the presence of asymmetric LC and/or its sequence are important for RNA synthesis. As shown in Figure 6D, modifications of the sequence on the 5' side of this loop in mutants LC1 (C59 to A), LC2 (C60 to A), LC3 (C61 to A), and LC4 (C59C60C61 to AAA) and on the 3' side in mutant LC6 (A74A75 to CC) did not change the predicted secondary structures or ΔG values from those of wild-type PVX RNA. In contrast, conversion of A58C59C60C61 to UU in the 5' side of this loop in mutant LC5 resulted in a loss of LC and increased stability of SL1. Similarly, conversion of A74A75 to GG on the 3' side of LC in mutant LC7 resulted in a more stable predicted structure, with loss of the internal loop.

Although small, but significant reductions in minus-strand RNA levels were observed for protoplasts inoculated with mutants LC1 ($78 \pm 9\%$) and LC5 ($70 \pm 11\%$), all other LC mutants supported similar to wild-type levels of minus-strand RNA accumulation (Figure 6A; Table 1). Again, this was in contrast to accumulation patterns for plus-strand RNA and CP (Figure 6B, C; Table 1). Protoplasts inoculated with mutants LC1 (C59 to A), LC2 (C60 to A) and LC3 (C61 to A) exhibited reduced plus-

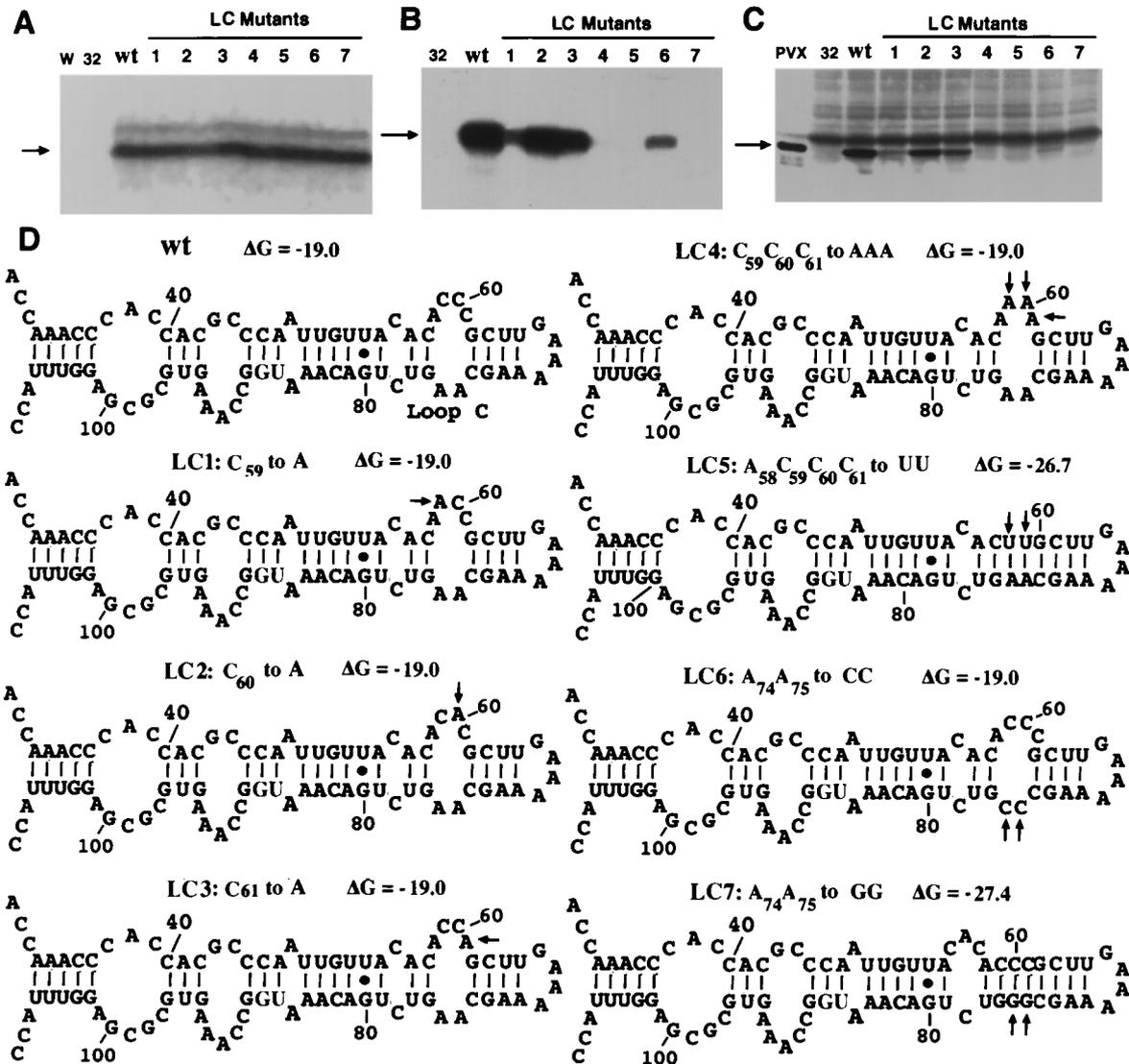


Figure 6. Effects of mutations in loop C on accumulation of PVX RNAs and coat protein. Protoplasts inoculated with replication-defective transcripts derived from p32 (32), wild-type transcripts derived from pMON8453 (wt), or mutant transcripts LC1-LC7 were analyzed at 48 h.p.i. Panels and labeling are as described in the legend to Figure 4.

strand RNA levels of 7 ± 4 , 66 ± 3 , and $48 \pm 8\%$, indicating that sequence of this portion of the loop, and position 59 in particular, were important for optimum RNA accumulation. Similarly, protoplasts inoculated with mutant LC4 (C₅₉C₆₀C₆₁ to AAA) contained plus-strand RNA levels ($7 \pm 4\%$) similar to those observed for LC1, again indicating the importance of sequence for the 5' side of LC. When protoplasts were inoculated with mutant LC6 (A₇₄A₇₅ to CC), containing changes on the 3' side of LC, plus-strand RNA accumulation was reduced to $9 \pm 3\%$ of wild-type levels. Thus, sequences on both sides of LC are important for plus-strand RNA accumulation.

Protoplasts inoculated with mutants LC5 (A₅₈C₅₉C₆₀C₆₁ to UU) and LC7 (A₇₄A₇₅ to GG), which were constructed to close the internal loop and increase stability of the predicted structures,

contained no detectable plus-strand RNA or CP. Because levels of plus-strand RNA in protoplasts inoculated with LC7 were significantly lower than in those inoculated with LC6, these data suggest that both the presence of LC and the sequence of this loop were required for RNA accumulation.

GNRA, GANA or GAAG motifs in the terminal tetraloop support efficient plus-strand RNA accumulation

GNRA tetraloop motifs, like the GAAA loop found in SL1 of PVX RNA, have been shown to be unusually stable loops that serve important structural and/or functional roles in many biological systems. Eight mutants, TL1-TL8 (Figure 7D), were constructed to analyze the significance of the GNRA tetraloop element in PVX RNA synthesis;

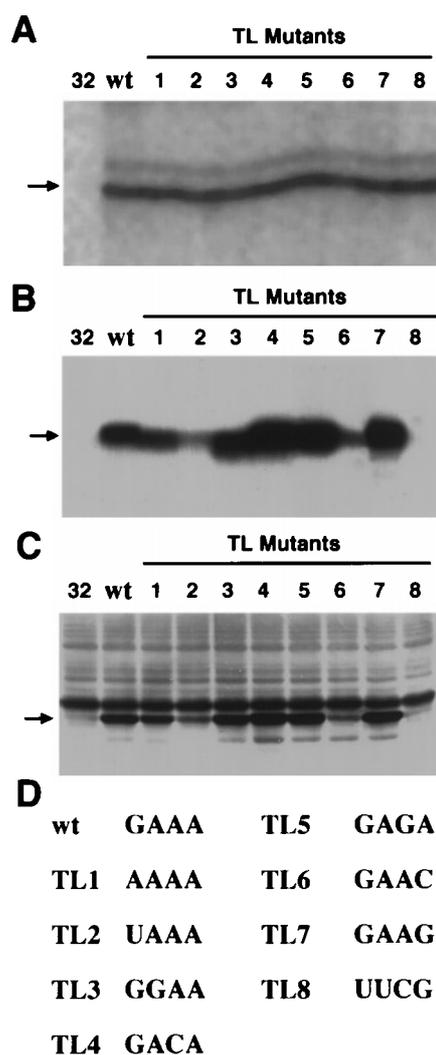


Figure 7. Effects of mutations in the tetraloop of SL1 on accumulation of PVX RNAs and coat protein. Protoplasts inoculated with replication-defective transcripts derived from p32 (32), wild-type transcripts derived from pMON8453 (wt), or mutant transcripts TL1-8 were analyzed at 48 h.p.i. A to C, and labeling, are as described in the legend to Figure 4. Tetraloop sequences for mutants TL1-8 are illustrated in D. Predicted structures for each of these mutant RNAs are not shown because they are identical to those of the wild-type PVX RNA.

none of these altered the predicted structure of SL1. Once again, protoplasts inoculated with these mutants were not significantly altered in minus-strand RNA accumulation (Figure 7A; Table 1), but exhibited a spectrum of levels of plus-strand RNA (Figure 7B; Table 1) that was mirrored by CP levels (Figure 7C). TL1 (G66 to A) and TL2 (G66 to U), which change the first position of the GNRA motif, reduced plus-strand RNA accumulation to $40 \pm 10\%$ and $19 \pm 10\%$, respectively. Thus, a G at position 66 was preferred over an A or U. Alteration of the second TL position in mutant TL3 (A67 to G) did not significantly affect plus-strand RNA

accumulation ($83 \pm 15\%$). In addition, replacement of A67 with U also resulted in wild-type levels of plus-strand RNA (data not shown), indicating that the second position of the SL1 tetraloop can vary without affecting RNA accumulation.

Discrepancies from the GNRA motif occurred when the last two positions of the terminal TL were changed. Protoplasts inoculated with either TL4 (A68 to C) or TL5 (A68 to G) exhibited plus-strand RNA and CP levels that were greater than wild-type (Figure 7B,C; Table 1), indicating that a purine in the third position was not required. Alteration of the fourth position to a pyrimidine in mutant TL6 (A69 to C) caused a decrease in plus-strand RNA accumulation to $22 \pm 2\%$ of wild-type levels. In contrast, when A69 was changed to a G in mutant TL7, greater than wild-type levels of plus-strand RNA ($138 \pm 27\%$) were produced. Thus, an A was not required in the fourth position, but a purine was preferred at this site. When protoplasts were inoculated with TL8, which contains another stable TL motif, UUCG, plus-strand RNA accumulation was reduced to $2 \pm 1\%$. In addition, plus-strand RNA levels were similarly reduced in protoplasts inoculated with a mutant containing a UUCG TL closed by a C-G pair (U65G66A67A68A69A70 to CUUCGG; data not shown). These data indicate that there are specific sequence requirements in this terminal loop, which could result in a unique structural architecture similar to the GNRA motif. Although GANA and GAAG can substitute for GNRA, any four nucleotides will not suffice, even if stability of the loop is maintained.

Discussion

Given that the 5' region of PVX RNA is likely to contain *cis*-acting signals important for translation, RNA synthesis, and encapsidation during virus replication, we analyzed the propensity of this region to form structural elements that may function in one or more of these processes. Thermodynamic predictions and solution structure analyses indicated that two thermodynamically favored stem-loop structures, SL1 and SL2, can form in the 5' region of PVX RNA, and covariation data further supported formation of SL1. Site-directed mutational studies were used to further discern the significance of the SL1 structure and affiliated sequences to PVX RNA accumulation *in vivo*. Although our data indicate that both the sequence and structure of SL1 are required for one or more aspects of plus-strand RNA accumulation, it is possible that elements in the corresponding 3' end of minus-strand RNA are also required. Andino *et al.* (1990) and Pogue & Hall (1992) have demonstrated that structures in the 5' ends of poliovirus and BMV RNAs, respectively, and not those at the 3' end of corresponding minus-sense RNAs, are critical for RNA replication. In contrast, Shi *et al.* (1996) illustrated that structures in the 3' end of the

minus-strand in West Nile virus were involved in cellular host protein interactions, and may be important for plus-strand RNA synthesis. Although covariation analysis of minus-strand RNAs from different PVX strains did not identify any covarying nucleotides that would support the existence of a particular secondary structure, additional experiments will be required to address this issue.

The mutational studies described herein not only indicate the importance of SL1 to RNA accumulation, but also identify key features within this structure. Both the presence of internal LC and the sequence on both sides of this loop are important for PVX RNA accumulation. Although alteration of C59, C60 or C61 reduced plus-strand RNA levels, modification of C59 had the greatest effect. Hellendoorn *et al.* (1997) found that protonatable cytosine bases in single-stranded internal loop regions of turnip yellow mosaic virus are important for infection. Given that loop regions of several types of RNAs appear to be important for protein binding (Draper, 1995), LC may be required for binding of virus and/or host proteins during PVX infection. Structure probing data of PVX RNA show that LC residues are accessible to DMS and to RNase V₁, indicating that the LC region of SL1 contains some helical or stacked residues. Because stacking within loops has been found to be important for protein binding in other systems (Scripture & Huber, 1995; Gubser & Varani, 1996), such interactions within LC may additionally be needed for recognition by viral and/or host factors. Alternatively, the sequence requirements in LC may not represent residues critical for contacting protein, but rather be important for establishment of a critical loop structure. Analysis of a crystal structure for the P4-P6 domain of the *Tetrahymena thermophila* self-splicing intron by Cate *et al.* (1996b) indicated planar "A-platform" structures involving two adenosine residues located in internal loops. These observations supported previous modification studies showing that the 5' most A of the pair is less accessible to modifying agents (Murphy & Cech, 1993). In addition, mutation of the A-A platform resulted in 70% less activity of the ribozyme (Murphy & Cech, 1994; Cate *et al.*, 1996b). Modification of A74A75 to CC in LC of PVX RNA, which reduced RNA accumulation, may have disrupted such a platform. Although the DMS modification data for PVX RNA showed that both A74 and A75 of LC are accessible to DMS, this does not rule out potential interactions across this loop or that protein binding might alter the structure within LC.

Certain tetraloop sequences, such as the GAAA motif in SL1, occur in many RNAs and have been shown to increase stability of RNA molecules (Woese *et al.*, 1990), participate in tertiary contacts (Jaeger *et al.*, 1994; Murphy & Cech, 1994; Cate *et al.*, 1996a), and serve as protein recognition sites (Gluck *et al.*, 1992; Zwieb, 1992; De Guzman *et al.*, 1998). Although the GNRA motif is conserved in

all PVX strains, our data indicate that several changes in the GAAA tetraloop of SL1 are tolerated and that a GNRA tetraloop is not absolutely required for RNA accumulation. In contrast, conversion to another highly stable, common tetraloop motif, UUCG (Cheong *et al.*, 1990), caused the greatest reduction in plus-strand RNA accumulation. Thus, any stabilizing loop and any four membered loop are not sufficient for plus-strand RNA accumulation. Surprisingly, plus-strand RNA accumulation was increased for mutants containing either GANA or GAAG motifs. Preference for a purine in the fourth position of a tetraloop (GNAR) has also been observed in helix 6 of signal recognition particle (SRP) RNA (Zwieb, 1992), which is recognized by SRP19 protein (p19). When the tetraloop was changed from GNAR to UNCG, p19 was unable to interact (Zwieb, 1992). Additional experiments will be required to determine if the TL in PVX SL1 affects RNA accumulation by binding viral and/or host proteins, by stabilizing SL1 directly or through tertiary interactions, or a combination of these.

The covariation data and mutational analysis indicate that pairing rather than sequence in SC is important for RNA accumulation. Compensatory site-directed changes in this region (mutant SC11) do not affect RNA accumulation, whereas modifications to one side of SC (mutants SC7, SC8 and SC10) drastically reduced RNA levels. In addition, most modifications near the base of SC in SL1, which substantially alter the structure of this region and/or affect the A-A mismatch, significantly reduce plus-strand RNA accumulation. Only one of the SC mutants (SC3), in which the A48-A85 mismatch was changed to a C-A mismatch without any further alteration to the predicted wild-type SL1 structure and stability, exhibits significantly higher levels of plus-strand RNA than those observed for wild-type. Comparatively, C-A mismatches are more commonly found in other RNAs, suggesting that they are less disruptive or more thermodynamically favorable than other types of mismatches (Gutell, 1996). It is possible that pairing in SC may be needed for correct positioning of internal loop C and the mismatches.

In addition to a role for pairing in the central region of SL1, our data also indicate that stability and/or length of stem regions affect RNA accumulation. Extension of SC or SD by two G-C base-pairs (mutants SC12 and SD1, respectively) abolishes plus-strand RNA accumulation. In contrast, shortening of the SC region (mutant SC9) by four base-pairs decreases RNA accumulation, but to a lesser extent. Similarly, Olsthoorn *et al.* (1994) observed that increased hairpin stability was more detrimental than reduced stability for bacteriophage MS2 gene expression. They also reported that MS2 with mutations in a hairpin containing the start codon of the CP gene reverted to a wild-type stability but not to wild-type sequence. In addition to stability, an optimal helix length may be required for PVX RNA accumulation. Selinger

et al. (1993) have shown that the length of the helix adjoining a tetraloop element in the RNA component of SRP is important for interaction with the SRP54 protein. Thus, stem length and/or stability in different regions of PVX SL1 may function to position other elements for interaction with viral and/or host proteins. Analysis of reversion pathways in plants inoculated with some of the PVX mutants described herein should help to distinguish the relative importance of length and/or stability to PVX RNA accumulation.

It was somewhat surprising that most mutations in SL1 did not significantly affect minus-strand RNA accumulation. Although minus-strand RNA synthesis is typically regulated by signals residing in the 3' ends of viral genomes (Buck, 1996), alterations in the 5' region might also affect translation of replicase and subsequent RNA synthesis. Of the α (nt 1-41) and β (nt 42-84) elements in the PVX leader, which have been shown to enhance translation of reporter gene products (Smirnyagina *et al.*, 1991), the β region is contained within the SL1 structure. Given that altered sequences and structures on the 5' side of SL1 in mutants SC1, SC5, SC7, LC1 and LC5 resulted in small, but significant declines in minus-strand RNA accumulation, the 5' side of SL1 may be important for recognition by translation factors and/or ribosomes. Carberry *et al.* (1992) have shown that in addition to steric accessibility of the cap structure, other sequence and secondary structural features of mRNA can enhance binding of the cap-binding protein (eIF-4E). Appropriately placed structures may also facilitate initiation of the ribosome at the correct site (Merrick & Hershey, 1996). It may be that the site-directed changes made to PVX SL1 are not as deleterious to translation (and subsequent minus-strand RNA synthesis) as are deletions within this region (Tomashevskaya *et al.*, 1993; Kim & Hemenway, 1996). Alternatively, levels of replicase required to support minus-strand synthesis in the absence of plus-strand genomic and sgRNA synthesis may be lower than those required during a normal replication cycle. Thus, assessment of minus-strand accumulation may not be an accurate measure of translational efficiency. A better understanding of translational signals in the PVX leader will require development of techniques to quantify replicase levels *in vivo*.

Although reductions in plus-strand RNA accumulation likely reflect a decline in plus-strand RNA synthesis, alterations in RNA stability and/or RNA packaging cannot be ruled out. If one or more sequence and/or structural elements in SL1 are important for binding of CP, packaging may be inhibited and may lead to reduced stability of the genomic RNA. The OAS for PVX has not been defined, although the 5' most 47 nt of a related potexvirus, papaya mosaic virus, are sufficient for packaging *in vitro* (Sit *et al.*, 1994). If the 5' most region of PVX RNA that is predicted to be unstructured contains the OAS, then mutations in SL1 may not directly affect packaging. Alternatively,

the OAS for PVX may include SL1 or formation of the SL1 structure may be important for positioning during assembly. Requirements for secondary structures in capsid protein binding have been demonstrated for several other plus-strand RNA viruses (Houwing & Jaspars, 1982; Turner *et al.*, 1988; Fujimura *et al.*, 1990; Fosmire *et al.*, 1992; Wei *et al.*, 1992; Zhong *et al.*, 1992). For the retrovirus HIV-1, the encapsidation signal (Ψ) consists of four SL structures, SL1-SL4 (Clever *et al.*, 1995). Although the HIV-1 nucleocapsid protein binds to the tetraloop sequence of HIV-1 SL3 *in vitro* (De Guzman *et al.*, 1998), both SL1 and SL3 are position-dependent and contribute to RNA encapsidation *in vivo* (McBride & Panganiban, 1997). Because all mutations in PVX SL1 that reduced plus-strand RNA levels also lowered CP levels (and presumably sgRNA synthesis), additional studies will be required to determine if SL1 is required for assembly and to distinguish the effects of this structure on translation, assembly and/or RNA synthesis.

Our previous deletion studies of the PVX 5' NTR (Kim & Hemenway, 1996) indicate that multiple elements throughout this region affect both genomic and subgenomic plus-strand RNA accumulation. In addition, critical elements near the 5' terminus, in the region predicted to be unstructured, exhibit complementarity to conserved regions upstream of the sgRNAs, and this complementarity may be important for sgRNA accumulation (Kim & Hemenway, 1997; K.-H.K. & C.H., unpublished data). Given that SL1 is also required for these processes, it appears that both unstructured and structured elements function locally and/or at a distance to affect plus-strand RNA accumulation. A localized role for SL1 may be to pull apart the two strands after completion of minus-strand RNA in order to facilitate initiation of plus-strand RNA synthesis, as was suggested for the role of secondary structure formation in bacteriophage Q β RNA (Priano *et al.*, 1987). Such a role for SL1 could also affect sgRNA levels if all plus-strand RNA synthesis is dependent upon initiation events at the 3' end of the minus-strand RNA. Future experiments will determine if one or more of these elements serves to bind or localize components of the replication machinery or to provide an appropriate architecture that facilitates plus-strand RNA synthesis.

Methods and Materials

Materials

Restriction enzymes, modifying enzymes, polymerases and m⁷GpppG cap analog were purchased from New England BioLabs. Avian myeloblastosis virus (AMV) reverse transcriptase was from Life Sciences. RNasin and RQ1 DNase were obtained from Promega. RNase T₁ and V₁ were purchased from Pharmacia Biotech and dimethyl sulfate was from Fluka. Boehringer Mannheim was the source for deoxy, dideoxy- and ribonucleotides, and oligonucleotides were synthesized by Genosys. Radiolabeled materials were products of Amersham. Cel-

lulase "Onozuka" R-10 was purchased from Yakult Honsha Co., Ltd and pectolyase Y23 is a product of Karlan Research Products.

Generation of thermodynamic predictions for PVX RNA structure

The energy dot plot and secondary structure folding predictions for the 5' region of PVX RNA were obtained with the Mfold program (Jaeger *et al.*, 1989a,b; Zuker, 1989) on Michael Zuker's internet location (www.ibc.wustl.edu/~zucker/rna/form1.cgi). The thermodynamic parameters consist of a folding temperature of 37°C and a per cent suboptimality number of ten.

Covariation analysis of PVX strains

RNA sequences from six different PVX strains were aligned and analyzed by the computer program, COVARIATION (Brown, 1991), to identify regions within the 5' 230 nt that may covary in order to maintain structural integrity.

Construction of site-directed mutants

Mutations were introduced in the SL1 region of the PVX cDNA clone, pMON8453 (Hemenway *et al.*, 1990), by the method of Kunkel (1985). Subsequent to verification by sequencing, mutant clones were digested with *MfeI* (nt 46) and *BsiWI* (nt 203), and the resulting mutant fragments were resected back into a wild-type pMON8453 clone that was similarly digested. Site-directed mutants that altered the *MfeI* restriction site were resected using *BsaAI* (nt 9436) and the *BsiWI* site. The resulting clones were sequenced through the entire resected fragment and flanking regions to ensure that only the mutations introduced were present.

Synthesis of PVX transcripts

PVX transcripts (543 nt) for structural studies were synthesized by *in vitro* transcription of the plasmid, pMON8453 (full-length PVX cDNA), linearized with *BamHI*. RNAs utilized for protoplast inoculations were derived from pMON8453 and mutant clones digested with *SpeI*. Transcription reaction conditions and product isolation were as described by Kim & Hemenway (1996). The quality and relative concentrations of transcripts were checked by electrophoresis on 1% (w/v) agarose gels at 4°C and visualized by ethidium bromide staining.

Chemical and enzymatic probing

Modification of truncated transcripts with DMS was in a total of 50 µl of buffer A (80 mM potassium cacodylate (pH 7.2), 300 mM potassium chloride, 20 mM magnesium acetate), as described by Moazed *et al.* (1986). Reactions were initiated by the addition of DMS to a final concentration of 5 mM in the presence of 2 µg RNA transcripts. Unmodified control reactions were treated similarly, except water was substituted for DMS. Reactions were incubated for 60 minutes at 4°C, 40 minutes at 25°C, 20 minutes at 37°C, or two minutes at 90°C, and terminated by addition of 12.5 µl of stop solution (1 M Tris-acetate (pH 7.5), 1.5 M sodium acetate, 0.1 mM EDTA, 1.42 M 2-mercaptoethanol). After incubation on ice for ten minutes, samples were precipitated with 2.5 vol. ethanol. Pellets were washed with cold 80% (v/

v) ethanol and resuspended in 10 µl diethyl pyrocarbonate (DEPC)-treated water.

For enzymatic probing, reactions contained 2 µg PVX truncated transcripts in 50 µl buffer B (100 mM Tris-HCl (pH 7.6), 100 mM sodium chloride, 20 mM magnesium chloride). Upon addition of RNase T₁ (0.04 unit/µl) or RNase V₁ (1.4 × 10⁻⁴ unit/µl), reactions were incubated at 25°C for 15 minutes or 10 minutes, respectively, and then stopped by extraction once with phenol/chloroform and subsequently with chloroform alone. The partially digested RNA products were precipitated with ethanol, washed with cold 80% ethanol and the pellets were resuspended in 10 µl DEPC-treated water.

Primer extension analyses of modified RNA

DMS-modified and RNase-digested PVX RNA samples were analyzed by primer extension (Moazed *et al.*, 1986). To evaluate extension products from nt 1-230, three primers were utilized: eP1 (complementary to nt 78-97), eP2 (complementary to nt 148-179), and eP3 (complementary to nt 242-263). After end-labeling with polynucleotide kinase and [³²P]ATP, primers (10 pmol) were combined with 5 µl modified PVX RNAs. Samples were dried and resuspended in 1 µl R-loop buffer (80% (v/v) formamide, 400 mM NaCl, 40 mM piperazine-*N,N'*-bis(3-ethanesulfonic acid) (Pipes, pH 4.6), 1 mM EDTA, 10 µg *Escherichia coli* tRNA), and incubated for 15 minutes at 45°C. After hybridization, 20 µl reverse transcriptase mix (40 mM KCl, 6 mM MgCl₂, 50 mM Tris-HCl (pH 8.3), 0.5 mM each dNTP, 10 mM DTT, 0.25 unit/µl AMV reverse transcriptase, 0.25 unit/µl RNasin) was added and the reaction was incubated at 45°C for an additional 10 minutes. Extension reactions were stopped by the addition of EDTA and SDS to final concentrations of 30 mM and 0.1% (w/v), respectively. Products were purified by phenol/chloroform extraction and ethanol precipitation. Pellets were washed and resuspended in 10 µl of sequencing stop solution (95% formamide, 20 mM EDTA, 0.05% (v/v) xylene cyanol FF, 0.05% (w/v) bromophenol blue), and analyzed by electrophoresis through 8% acrylamide/7.5 M urea gels.

Preparation and inoculation of protoplasts

Protoplasts utilized for functional analyses were prepared from a rapidly dividing *N. tabacum* suspension cell line, NT-1, maintained in NT-1 medium containing MS salts (0.43%), sucrose (88 mM), B1-inositol stock (55 mM myo-inositol, 0.3 mM Thiamine), Miller's I stock (440 mM KH₂PO₄), 0.2 mg/ml 2,4-dichlorophenoxyacetic acid (An, 1985). Preparative conditions were similar to those described by Kim & Hemenway (1997). Briefly, cells were digested in NT-1 protoplast basic solution (0.4 M mannitol, 20 mM Mes (pH 5.5)) containing 0.75% (w/v) cellulase Onozuka R-10 and 0.1% (w/v) pectolyase Y23 for 50 minutes with gentle shaking (60 rpm) at room temperature. After centrifugation and washing, protoplasts were resuspended in electroporation buffer (0.8% (w/v) NaCl, 0.02% (w/v) KCl, 0.02% (w/v) KH₂PO₄, 0.11% (w/v) Na₂HPO₄, 0.4 M mannitol, pH 6.5) and adjusted to a final concentration of 5 × 10⁶ cells/ml.

For inoculation, 5 µg transcripts, in a total volume of 400 µl electroporation buffer, were added to 400 µl protoplasts (2 × 10⁶) and subjected to 250 V at a capacitance of 500 µF with a BioRad Gene Pulser. After electroporation, cells were placed on ice for ten minutes and subsequently plated in 8 ml NT-1 culture medium

supplemented with 0.4 mg/ml 2,4-dichlorophenoxyacetic acid, 0.4 M mannitol, 10 µg/ml RNase A. Inoculated protoplasts were then allowed to incubate for 48 hours at 25°C on a 14 hours light and ten hours dark cycle. After incubation, cells were washed and incubated with an additional 10 µg/ml RNase for 60 minutes to digest the remaining input RNA. Inoculated protoplast cells were then used for CP detection and total RNA isolation.

Analysis of PVX RNA accumulation in protoplasts

At 48 h.p.i., total RNA was purified from 3.2×10^6 inoculated protoplasts using Trizol (Life Technologies) reagent and resuspended in DEPC-treated water. S₁ nuclease protection assays were used for detection of minus- and plus-strand PVX RNAs with 5' end-labeled probes P3 and P1, respectively, following a method described by Kim & Hemenway (1996). Protected products were separated on 8% sequencing gels for plus-sense detection or 6% gels for minus-sense detection, and visualized by autoradiography or by a Molecular Dynamics phosphorimager. All mutants were inoculated into NT-1 protoplasts in three or more independent experiments to validate the levels of RNA and CP accumulation. The ImageQuant 1.0 program was utilized for quantification of RNA levels; differences in RNA accumulation for both plus and minus-strand were subjected to student *t*-tests and were considered significant if within a confidence interval of 95%.

For detection of CP, 8×10^5 protoplasts were added to an equal volume of 2 × SDS loading buffer (125 mM Tris-HCl (pH 6.8), 4% (w/v) SDS, 20% (v/v) glycerol, 1.42 M 2-mercaptoethanol, 0.1% (w/v) bromophenol blue, 0.1% (w/v) xylene cyanol FF). After vortexing, samples were heated to 95°C for ten minutes and subsequently centrifuged for 20 minutes at 14,000 rpm. Fifty per cent of the supernatant was loaded onto a 12% Laemmli gel (Laemmli, 1970). After electrophoresis, the gel was transferred to nitrocellulose membrane and CP was detected using antisera to PVX and a biotinylated/streptavidin colorimetric detection system (Amersham).

Acknowledgments

We are grateful to P. Wollenzien for discussions, technical advice and for critical reading of the manuscript. This research was supported by Public Health Service grant GM 49841 to C.H.

References

- An, G. (1985). High efficiency transformation of cultured tobacco cells. *Plant Physiol.* **79**, 568–570.
- Andino, R., Rieckhof, G. E. & Baltimore, D. (1990). A functional ribonucleoprotein complex forms around the 5' end of poliovirus RNA. *Cell*, **63**, 369–380.
- Angell, S. M., Davies, C. & Baulcombe, D. C. (1996). Cell-to-cell movement of potato virus X is associated with a change in the size-exclusion limit of plasmodesmata in trichome cells of *Nicotiana glauca*. *Virology*, **216**, 197–201.
- Baulcombe, D. C., Chapman, S. & Santa, Cruz S. (1995). Jellyfish green fluorescent protein as a reporter for virus infections. *Plant J.* **7**, 1045–1053.
- Beck, D. L., Guilford, P. J., Voot, D. M., Andersen, M. T. & Forster, R. L. (1991). Triple gene block proteins of white clover mosaic potexvirus are required for transport. *Virology*, **183**, 695–702.
- Blackwell, J. L. & Brinton, M. A. (1995). BHK cell proteins that bind to the 3' stem-loop structure of the West Nile virus genome RNA. *J. Virol.* **69**, 5650–5658.
- Brown, J. W. (1991). Phylogenetic comparative analysis of RNA structure on Macintosh computers. *Cabios Commun.* **7**, 391–393.
- Buck, K. W. (1996). Comparison of the replication of positive-stranded RNA viruses of plants and animals. *Advan. Virus Res.* **47**, 159–251.
- Carberry, S. E., Friedland, D. E., Rhoads, R. E. & Goss, D. J. (1992). Binding of protein synthesis initiation factor 4E to oligoribonucleotides: effects of cap accessibility and secondary structure. *Biochemistry*, **31**, 1427–1432.
- Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Kundrot, C. E., Cech, T. R. & Doudna, J. A. (1996a). Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science*, **273**, 1678–1685.
- Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Szewczak, A. A., Kundrot, C. E., Cech, T. R. & Doudna, J. A. (1996b). RNA tertiary structure mediation by adenosine platforms. *Science*, **273**, 1696–1699.
- Chapman, S., Kavanagh, T. & Baulcombe, D. (1992). Potato virus X as a vector for gene expression in plants. *Plant J.* **2**, 549–557.
- Cheong, C., Varani, G. & Tinoco, I. (1990). Solution structure of an unusually stable RNA hairpin, 5'GGAC(UUCG)GUCC. *Nature*, **346**, 680–682.
- Clever, J., Sassetti, C. & Parslow, T. G. (1995). RNA secondary structure and binding sites for gag gene products in the 5' packaging signal of human immunodeficiency virus type 1. *J. Virol.* **69**, 2101–2109.
- Cockerham, G. (1955). *Proceedings of the Second Conference on Potato Virus Diseases*, Lisse-Wageningen.
- De Guzman, R. N., Wu, Z. R., Stalling, C. C., Pappalardo, L., Borer, P. N. & Summers, M. F. (1998). Structure of the HIV-1 nucleocapsid protein bound to the SL3 Ψ-RNA recognition element. *Science*, **279**, 384–388.
- Dolja, V. V., Grama, D. P., Morozov, S. Y. & Atabekov, J. G. (1987). Potato virus X related single- and double-stranded RNAs. *FEBS Letters*, **214**, 308–312.
- Draper, D. E. (1995). Protein-RNA recognition. *Annu. Rev. Biochem.* **64**, 593–620.
- Duggal, R., Lahser, F. C. & Hall, T. C. (1994). *Cis*-acting sequences in the replication of plant viruses with plus-sense RNA genomes. *Annu. Rev. Phytopathol.* **32**, 287–309.
- Fosmire, J., Hwang, K. & Makino, S. (1992). Identification and characterization of a coronavirus packaging signal. *J. Virol.* **66**, 3522–3530.
- Fujimura, T., Esteban, R., Esteban, L. M. & Wickner, R. B. (1990). Portable encapsidation signal of the L-A double-stranded RNA virus of *S. cerevisiae*. *Cell*, **62**, 819–828.
- Gilmer, D., Allmang, C., Ehresmann, C., Guilley, H., Richards, K., Jonard, G. & Ehresmann, B. (1993). The secondary structure of the 5' noncoding region of beet necrotic yellow vein virus RNA 3: evidence for a role in viral RNA replication. *Nucl. Acids Res.* **21**, 1389–1395.
- Gluck, A., Endo, Y. & Wool, I. G. (1992). Ribosomal RNA identity elements for ricin A-chain recognition

- and catalysis. Analysis with tetraloop mutants. *J. Mol. Biol.* **226**, 411–424.
- Gubser, C. C. & Varani, G. (1996). Structure of the polyadenylation regulatory element of the human U1A pre-mRNA 3'-untranslated region and interaction with the U1A protein. *Biochemistry*, **35**, 2253–2267.
- Gutell, R. R. (1996). Comparative sequence analysis and the structure of 16S and 23S RNA. In *Ribosomal RNA: Structure, Evolution, Processing and Function in Protein Biosynthesis* (Zimmerman, R. A. & Dahlberg, A. E., eds), pp. 111–128, CRC Press, Boca Raton.
- Hellendoorn, K., Verlaan, P. W. G. & Pleij, C. W. A. (1997). A functional role for the conserved protonatable hairpins in the 5' untranslated region of turnip yellow mosaic virus RNA. *J. Virol.* **71**, 8774–8779.
- Hemenway, C., Fang, R.-X., Kaniewski, W. K., Chua, N.-H. & Tumer, N. E. (1988). Analysis of the mechanism of protection in transgenic plants expressing the potato virus X coat protein or its antisense RNA. *EMBO J.* **7**, 1273–1280.
- Hemenway, C., Weiss, J., O'Connell, K. & Tumer, N. E. (1990). Characterization of infectious transcripts from a potato virus X cDNA clone. *Virology*, **175**, 365–371.
- Houwing, C. & Jaspars, E. (1982). Protein binding sites in nucleation complexes of alfalfa mosaic virus RNA 4. *Biochemistry*, **21**, 3408–3414.
- Hsue, B. & Masters, P. S. (1997). A bulged stem-loop structure in the 3' untranslated region of the genome of the coronavirus mouse hepatitis virus is essential for replication. *J. Virol.* **71**, 7567–7578.
- Huisman, M. J., Linthorst, H. J. M., Bol, J. F. & Cornelissen, B. J. C. (1988). The complete nucleotide sequence of potato virus X and its homologies at the amino acid level with various plus-stranded RNA viruses. *J. Gen. Virol.* **69**, 1789–1798.
- Ito, T. & Lai, M. M. C. (1997). Determination of the secondary structure of and cellular protein binding to the 3'-untranslated region of the Hepatitis C virus RNA genome. *J. Virol.* **71**, 8698–8706.
- Jacobson, S. J., Konings, D. A. M. & Sarnow, P. (1993). Biochemical and genetic evidence for a pseudoknot structure at the 3' terminus of the poliovirus RNA genome and its role in viral RNA amplification. *J. Virol.* **67**, 2961–2971.
- Jaeger, J. A., Turner, D. H. & Zuker, M. (1989a). Improved predictions of secondary structures for RNA. *Proc. Natl Acad. Sci. USA*, **86**, 7706–7710.
- Jaeger, J. A., Turner, D. H. & Zuker, M. (1989b). Predicting optimal and suboptimal secondary structure for RNA. *Methods Enzymol.* **183**, 281–306.
- Jaeger, L., Michel, F. & Westhof, E. (1994). Involvement of a GNRA tetraloop in long-range RNA tertiary interactions. *J. Mol. Biol.* **236**, 1271–1276.
- Jaspars, E. M. J. (1985). Interaction of alfalfa mosaic virus nucleic acid and protein. In *Molecular Plant Virology* (Davies, J. W., ed.), pp. 155–221, CRC Press, Boca Raton.
- Kavanagh, T., Goulden, M., Santa, Cruz S., Chapman, S., Barker, I. & Baulcombe, D. (1992). Molecular analysis of a resistance-breaking strain of potato virus X. *Virology*, **189**, 609–617.
- Kim, K.-H. & Hemenway, C. (1996). The 5' nontranslated region of potato virus X RNA affects both genomic and subgenomic RNA synthesis. *J. Virol.* **70**, 5533–5540.
- Kim, K.-H. & Hemenway, C. (1997). Mutations that alter a conserved element upstream of the potato virus X triple block and coat protein genes affect subgenomic RNA accumulation. *Virology*, **232**, 187–197.
- Kunkel, T. A. (1985). Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl Acad. Sci. USA*, **82**, 488–492.
- Laemmli, U. K. (1970). Cleavage of structural protein during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680–685.
- Lauber, E., Guilley, H., Richards, K., Jonard, G. & Gilmer, D. (1997). Conformation of the 3' end of beet necrotic yellow vein benevirus RNA 3 analyzed by chemical and enzymatic probing and mutagenesis. *Nucl. Acids Res.* **25**, 4723–4729.
- Lockard, R. E. & Kumar, A. (1981). Mapping tRNA structure in solution using double-strand-specific ribonuclease V1 from cobra venom. *Nucl. Acids Res.* **9**, 5125–5140.
- McBride, M. S. & Panganiban, A. T. (1997). Position dependence of functional hairpins important for human immunodeficiency virus type 1 RNA encapsidation *in vivo*. *J. Virol.* **71**, 2050–2058.
- Merrick, W. C. & Hershey, J. W. B. (1996). The pathway and mechanism of eukaryotic protein synthesis. In *Translational Control* (Hershey, J. W. B., Mathews, M. B. & Sonenberg, N., eds), pp. 31–69, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Moazed, D., Stern, S. & Noller, H. F. (1986). Rapid chemical probing of conformation in 16 S ribosomal RNA and 30 S ribosomal subunits using primer extension. *J. Mol. Biol.* **187**, 399–416.
- Murphy, F. L. & Cech, T. R. (1993). An independently folding domain of RNA tertiary structure within the tetrahymena ribozyme. *Biochemistry*, **32**, 5291–5300.
- Murphy, F. L. & Cech, T. R. (1994). GAAA tetraloop and conserved bulge stabilize tertiary structure of a group I intron domain. *J. Mol. Biol.* **236**, 49–63.
- Niesters, H. G. & Strauss, J. H. (1990). Defined mutations in the 5' nontranslated sequence of Sindbis virus RNA. *J. Virol.* **64**, 4162–4168.
- Olsthoorn, R. C. L., Licis, N. & van Duin, J. (1994). Leeway and constraints in the forced evolution of a regulatory RNA helix. *EMBO J.* **13**, 2660–2668.
- Orman, B. O., Celnik, R. M., Mandel, A. M., Torres, H. N. & Mentaberry, A. N. (1990). Complete cDNA sequence of a South American isolate of potato virus X. *Virus Res.* **16**, 293–305.
- Pogue, G. P. & Hall, T. C. (1992). The requirement of a 5' stem-loop structure in brome mosaic virus replication supports a new model for viral positive-strand RNA initiation. *J. Virol.* **66**, 674–684.
- Pooggin, M. M. & Skryabin, K. G. (1992). The 5'-untranslated leader sequence of potato virus X RNA enhances the expression of a heterologous gene *in vivo*. *Mol. Gen. Genet.* **234**, 329–331.
- Priano, C., Kramer, F. R. & Mills, D. R. (1987). Evolution of the RNA coliphages: the role of secondary structures during RNA replication. *Cold Spring Harbor Symp. Quant. Biol.* **52**, 321–330.
- Price, M. (1992). Examination of potato virus X proteins synthesized in infected tobacco plants. *J. Virol.* **66**, 5658–5661.
- Querci, M., van der Vlugt, R., Goldbach, R. & Salazar, L. F. (1993). RNA sequence of potato virus X strain HB. *J. Gen. Virol.* **74**, 2251–2255.
- Scripture, J. B. & Huber, P. W. (1995). Analysis of the binding of *Xenopus* ribosomal protein L5 to oocyte 5 S rRNA. *J. Biol. Chem.* **270**, 27358–27365.

- Selinger, D., Brennwald, P., Liao, X. & Wise, J. A. (1993). Identification of RNA sequences and structural elements required for assembly of fission yeast SRP54 protein with signal recognition particle RNA. *Mol. Cell. Biol.* **13**, 1353–1362.
- Shi, P.-Y., Li, W. & Brinton, M. A. (1996). Cell proteins bind specifically to West Nile virus minus-strand 3' stem-loop. *J. Virol.* **70**, 6278–6287.
- Sit, T. L., Leclerc, D. & AbouHaidar, M. G. (1994). The minimal 5' sequence for *in vitro* initiation of papaya mosaic potexvirus assembly. *Virology*, **199**, 238–242.
- Skryabin, K. G., Kraev, A. S., Morozov, S. Y., Rozanov, M. N., Chernov, B. K., Lukasheva, L. I. & Atabekov, J. G. (1988). The nucleotide sequence of potato virus X RNA. *Nucl. Acids Res.* **16**, 10929–10930.
- Smirnyagina, E. V., Morozov, S. Y., Rodionova, N. P., Miroshnichenko, N. A., Solovev, A. G., Fedorkin, O. N. & Atabekov, J. G. (1991). Translational efficiency and competitive ability of mRNAs with 5'-untranslated alpha beta-leader of potato virus X RNA. *Biochimie*, **73**, 587–598.
- Song, C. & Simon, A. E. (1995). Requirement of a 3'-terminal stem-loop in *in vitro* transcription by an RNA-dependent RNA polymerase. *J. Mol. Biol.* **254**, 6–14.
- Tomashevskaya, O. L., Solovyev, A. G., Karpova, O. V., Fedorkin, O. N., Rodionova, N. P., Morozov, S. Y. & Atabekov, J. G. (1993). Effects of sequence elements in the potato virus X RNA 5' non-translated alpha beta-leader on its translation enhancing activity. *J. Gen. Virol.* **74**, 2717–2724.
- Turner, D. R., Joyce, L. E. & Butler, P. J. (1988). The tobacco mosaic virus assembly origin RNA. Functional characteristics defined by directed mutagenesis. *J. Mol. Biol.* **203**, 531–547.
- Wei, N., Hacker, D. L. & Morris, T. J. (1992). Characterization of an internal element in turnip crinkle virus RNA involved in both coat protein binding and replication. *Virology*, **190**, 346–355.
- Woese, C. R., Winker, S. & Gutell, R. R. (1990). Architecture of ribosomal RNA: constraints on the sequence of "tetra-loops". *Proc. Natl Acad. Sci. USA*, **87**, 8467–8471.
- Zelenina, D. A., Kulaeva, O. I., Smirnyagina, E. V., Solovyev, A. G., Miroshnichenko, N. A., Fedorkin, O. N., Rodionova, N. P., Morozov, S. & Atabekov, J. G. (1992). Translation enhancing properties of the 5'-leader of potato virus X genomic RNA. *FEBS Letters*, **296**, 267–270.
- Zhong, W., Dasgupta, R. & Rueckert, R. (1992). Evidence that the packaging signal for nodaviral RNA2 is a bulged stem-loop. *Proc. Natl Acad. Sci. USA*, **89**, 11146–11150.
- Zuker, M. (1989). On finding all suboptimal foldings of an RNA molecule. *Science*, **244**, 48–52.
- Zwieb, C. (1992). Recognition of a tetranucleotide loop of signal recognition particle RNA by protein SRP19. *J. Biol. Chem.* **267**, 15650–15656.

Edited by D. E. Draper

(Received 6 April 1998; received in revised form 6 August 1998; accepted 14 August 1998)