MOLECULAR AND CELLULAR BIOLOGY, Sept. 1993, p. 5331–5347
0270-7306/93/095331-17$02.00/0

Vol. 13, No. 9

A Phylogenetically Conserved Sequence within Viral 3′ Untranslated RNA Pseudoknots Regulates Translation

VALERIE LEATHERS, ROBERT TANGUAY, MARILYN KOBAYASHI, AND DANIEL R. GALLIE*

Department of Biochemistry, University of California, Riverside, California 92521-0129

Received 1 February 1993/Returned for modification 31 March 1993/Accepted 1 June 1993

Both the 68-base 5′ leader (Ω) and the 205-base 3′ untranslated region (UTR) of tobacco mosaic virus (TMV) promote efficient translation. A 35-base region within Ω is necessary and sufficient for the regulation. Within the 3′ UTR, a 52-base region, composed of two RNA pseudoknots, is required for regulation. These pseudoknots are phylogenetically conserved among seven viruses from two different viral groups and one satellite virus. The pseudoknots contained significant conservation at the secondary and tertiary levels and at several positions at the primary sequence level. Mutational analysis of the sequences determined that the primary sequence in several conserved positions, particularly within the third pseudoknot, was essential for function. The higher-order structure of the pseudoknots was also required. Both the leader and the pseudoknot region were specifically recognized by, and competed for, the same proteins in extracts made from carrot cell suspension cells and wheat germ. Binding of the proteins is much stronger to Ω than to the pseudoknot region. Synergism was observed between the TMV 3′ UTR and the cap and to a lesser extent between Ω and the 3′ UTR. The functional synergism and the protein binding data suggest that the cap, TMV 5′ leader, and 3′ UTR interact to establish an efficient level of translation.

The majority of eukaryotic messages terminate in a poly(A) tail, which is required for efficient translation and message stabilization (reviewed in references 2 and 30). As a regulator of translation in higher eukaryotes, the poly(A) tail requires the cap for function: for uncapped messages, the translational efficiency of poly(A)$^+$ mRNA is not appreciably greater than that of poly(A)$^-$ mRNA (8). Moreover, the degree to which a cap stimulates translation is an order of magnitude greater for poly(A)$^+$ mRNA than for poly(A)$^-$ mRNA. This synergism between a cap and poly(A) tail suggests that these two regulatory elements, in conjunction with associated proteins, communicate to direct efficient translation. Observations made in lower eukaryotes support the idea of an interaction between the termini of an mRNA. In yeast cells, the poly(A)-binding protein gene is an essential gene (38). Suppressor mutants were found to have alterations in either the L46 large ribosomal subunit protein gene or in 25S rRNA processing (36, 37), data suggesting that the poly(A)-binding protein may promote 80S complex formation.

Only a few mRNAs that do not terminate in a poly(A) tail have been found: the cell cycle-regulated animal histone mRNAs (23), several plant virus mRNAs (20, 21), and flavivirus mRNAs (41). Whereas all plant histone mRNAs so far examined are poly(A)$^+$, the 3′-terminal stem-loop structure of the cell cycle-regulated histone mRNAs from animals is involved in 3′ processing and message stabilization (22, 32). A 72-base domain within the 3′ untranslated region (UTR) of tobacco mosaic virus (TMV) genomic mRNA is required for promoting efficient translation in both plant and animal cells (12). How these naturally poly(A)$^-$ mRNAs are efficiently translated is an intriguing but unanswered question.

TMV is an RNA virus whose life cycle is carried out in the cytoplasm. The genomic RNA also serves as an mRNA. Subgenomic mRNAs generated during the viral life cycle are colinear with the genomic mRNA at the 3′ terminus; consequently, all TMV mRNAs contain the same 205-base 3′ UTR. The TMV 3′ UTR is the functional equivalent to a poly(A) tail in that it promotes efficient translation and increases mRNA stability of chimeric mRNA constructs (12). The 3′ UTR is composed of two domains. Located at the 3′ terminus, a 105-base tRNA-like domain mimics the three-dimensional structure of true tRNAs to such an extent that many tRNA-specific enzymes will also recognize and modify it (reviewed in references 20 and 21). Immediately upstream of the tRNA-like domain is a 72-base domain composed solely of three RNA pseudoknots (40). This upstream pseudoknot domain (UPD) is responsible for the regulation of translation associated with the TMV 3′ UTR (12). The fact that the tRNA-like domain, which itself contains two RNA pseudoknots, does not increase translational efficiency (12) demonstrates that the mere presence of RNA pseudoknots within the 3′ UTR is not the basis for enhanced translation.

To precisely determine the sequences within the UPD that are essential for the regulation of translation, mutations were made throughout the region and their effects on translation of chimeric mRNA constructs were measured in vivo. The 5′-proximal pseudoknot was not required for UPD function. The secondary and tertiary structure and primary sequence at several positions within the 3′-proximal pseudoknot were critical for regulatory function. The higher-order structure of the middle pseudoknot of the UPD may be needed as well. We also demonstrate that protein-binding activity from both wheat germ and carrot extracts specifically recognizes the UPD. The same protein activity also recognized the TMV leader (Ω), itself an enhancer of translation.

MATERIALS AND METHODS

Plasmid constructs and site-directed mutagenesis. The T7-based luciferase (*luc*) construct containing the TMV 3′ UTR (Luc-3′TMV) has been described previously (9). The *luc* cartridge contains the entire *luc* 3′ UTR (7) and was used for

* Corresponding author.

the study of the interaction between the TMV 3′ UTR and the cap. For the mutagenesis studies, the native *luc* 3′ UTR was removed before introduction of the TMV 3′ UTR downstream of the *luc* coding region. Site-directed mutagenesis was performed by the method of Kunkel et al. (25). Oligonucleotides were synthesized by using an Applied Biosystems 380B DNA synthesizer.

RNA secondary structure analysis. RNA synthesized from a T7-based vector containing each mutant UPD was dephosphorylated with calf intestinal alkaline phosphatase and 5′ end labeled with [^{32}P]ATP, using T4 polynucleotide kinase. Full-length RNA was purified from a 7 M urea–8% polyacrylamide gel by electroelution. Limited digestion with RNase T_1 or cobra venom (RNase V_1) was carried out at 24°C in 50 mM sodium cacodylate (pH 7.0)–10 mM $MgCl_2$–1 μg of *Escherichia coli* tRNA. Digests with nuclease S1 were carried out in 25 mM sodium acetate (pH 4.5)–5 mM $MgCl_2$–50 mM KCl–1 μg of *E. coli* tRNA. The products were displayed on a 7 M urea–8% polyacrylamide sequencing gel and analyzed by autoradiography. Structure mapping of the mutants was performed at least twice and in some cases three times.

In vitro transcription reaction conditions. All plasmid constructs were linearized with *Nde*I, which cuts immediately downstream of the tRNA-like structure. Concentrations of the construct DNAs were quantitated spectrophotometrically following linearization and brought to a concentration of 0.5 mg/ml. In vitro transcription was carried out as described by Yisraeli et al. (42), using 40 mM Tris-HCl (pH 7.5), 6 mM $MgCl_2$, 2 mM spermidine, 100 μg of bovine serum albumin (BSA) per ml, 0.5 mM each ATP, CTP, and UTP, 160 μM GTP, 1 mM m^7GpppG, 100 mM dithiothreitol (DTT), 0.3 U of RNasin (Promega) per μl, and 0.5 U of T7 RNA polymerase (New England Biolabs) per μl. Under our conditions, 95% of the mRNA is capped. Each RNA construct was synthesized in triplicate in separate transcription reactions so that any variability in RNA yield would be reflected in the standard deviation calculated as part of the expression data for each construct. The integrity and relative quantity of RNA were determined by formaldehyde-agarose gel electrophoresis as described previously (26). Radiolabeled probes were made as uncapped RNAs, using either [^{32}P]ATP or [^{32}P]CTP, and the full-length transcripts were electroeluted from a polyacrylamide gel. Quantitation of RNA yields was done spectrophotometrically.

Preparation and electroporation of protoplasts. Protoplasts were isolated from a tobacco (Xanthi) cell suspension (for the synergism studies) and carrot (RCWC) cell suspension (for the mutational analysis) by digestion with 0.25% CELF cellulase (Worthington Biochemicals), 1% Cytolase M103S (Genencor), 0.05% Pectolyase Y23 (Seishin Pharmaceutical Co.), 0.5% BSA, and 7 mM β-mercaptoethanol in isolation buffer (12 mM sodium acetate [pH 5.8], 50 mM $CaCl_2$, 0.25 M mannitol) for 75 min. Protoplasts were washed once with isolation buffer and once with electroporation buffer (10 mM *N*-1-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid [HEPES; pH 7.2], 130 mM KCl, 10 mM NaCl, 4 mM $CaCl_2$, 0.2 M mannitol) and then resuspended in electroporation buffer to a final concentration of 10^6 cells per ml. Two micrograms of each *luc* mRNA construct was mixed with 0.8 ml of protoplasts immediately before electroporation (500 μF capacitance, 350 V), using an IBI GeneZapper. For time course experiments, aliquots of protoplasts were taken at the time intervals indicated. For endpoint experiments, the protoplasts were incubated overnight. mRNA from each construct was synthesized in triplicate for triplicate electroporations. Luciferase assays were performed in duplicate for each electroporation. For each mutant construct, the mean and standard deviation are reported. The dose response of RNA electroporation versus enzyme activity is linear up to at least 30 μg of input mRNA (10).

In vitro translation. A 160-ng aliquot of each *luc* mRNA construct was translated in wheat germ lysate according to the recommendations of the supplier (Promega) with the exception that all amino acids were supplied at a final concentration of 80 μM in a nonradiolabeled form. Aliquots were removed at time intervals and frozen on dry ice. The extent of translation was determined by assaying each aliquot for luciferase activity.

Luciferase assay. Protoplasts, collected by centrifugation at 100 × *g*, were sonicated for 5 s in 100 mM Tricine (pH 7.8)–2 mM DTT–2 mM 1,2-diaminocyclohexane-*N*,*N*,*N*′,*N*′-tetraacetic acid–10% glycerol–1% Triton X-100, and the cell debris was pelleted. Aliquots of the extract were added to 100 μl of luciferase assay buffer [20 mM Tricine (pH 7.8), 1.07 mM $(MgCo_3)_4Mg(OH)_2 \cdot 5H_2O$, 2.67 mM $MgSO_4$, 0.1 mM EDTA, 33.3 mM DTT, 270 μM coenzyme A, 470 μM luciferin, 500 μM ATP (Promega Biotec)], and the reaction was initiated with the injection of 100 μl of 0.5 mM luciferin in luciferase assay buffer. Photons were counted by using a Monolight 2010 Luminometer (Analytical Luminescence Laboratory). Protein concentration was determined by the method described by Bradford (4).

Wheat germ and carrot extract preparation and gel retardation. For the preparation of wheat germ (Arrowhead Mills) extract, 50 g was ground in a blender for 45 s, and then 150 ml of grinding buffer (5 mM HEPES [pH 6.9], 120 mM potassium acetate, 5 mM magnesium acetate, 1 mM DTT, 1 mM phenylmethylsulfonyl fluoride [PMSF], 1 μg each of leupeptin and pepstatin per ml) was added. Following centrifugation at 30,000 × *g* at 4°C, the supernatant was supplemented with 1/10 volume of 500 mM HEPES (pH 7.6) and centrifuged again for 15 min. The supernatant was dialyzed against 20 mM Tris-HCl (pH 7.5)–50 mM KCl–5 mM $MgCl_2$–1 mM DTT–0.5 mM PMSF–1 μg each of leupeptin and pepstatin per ml.

Protoplasts prepared from carrot cell suspension cells were sonicated for 5 s in 20 mM Tris-HCl (pH 7.5)–50 mM KCl–5 mM $MgCl_2$–1 mM DTT–0.5 mM PMSF–2 μg each of leupeptin and pepstatin per ml; the cell debris was removed by centrifugation and dialyzed against the same buffer.

For the binding reactions, 2 μg (wheat germ extract) or 7 μg (carrot extract) of protein extract and 5 ng of labeled RNA was used in 10 mM Tris (pH 7.5)–35 mM KCl–1.0 mM $MgCl_2$–5% glycerol–1 mM DTT–0.7 mg of total yeast RNA per ml–0.5 U of RNasin per ml. Competitor RNAs, when used, were added to the test RNA before the addition of protein extract. Following 15 min of incubation at 0°C, heparin was added to 5 mg/ml, and the mixture was incubated for an additional 10 min. The RNA-protein complexes were resolved on a native 3.5% polyacrylamide gel, and the dried gels were analyzed by autoradiography. The percentage of shifted RNA was determined by excising and counting the free and shifted bands, using a scintillation counter.

RESULTS

The TMV 3′ UTR is functionally dependent on the cap in vivo. As regulators of translation, the poly(A) tail and the cap interact synergistically (8). Synergism in this context means that the combined effect on translational efficiency of the cap and poly(A) tail is greater than the multiplication of their

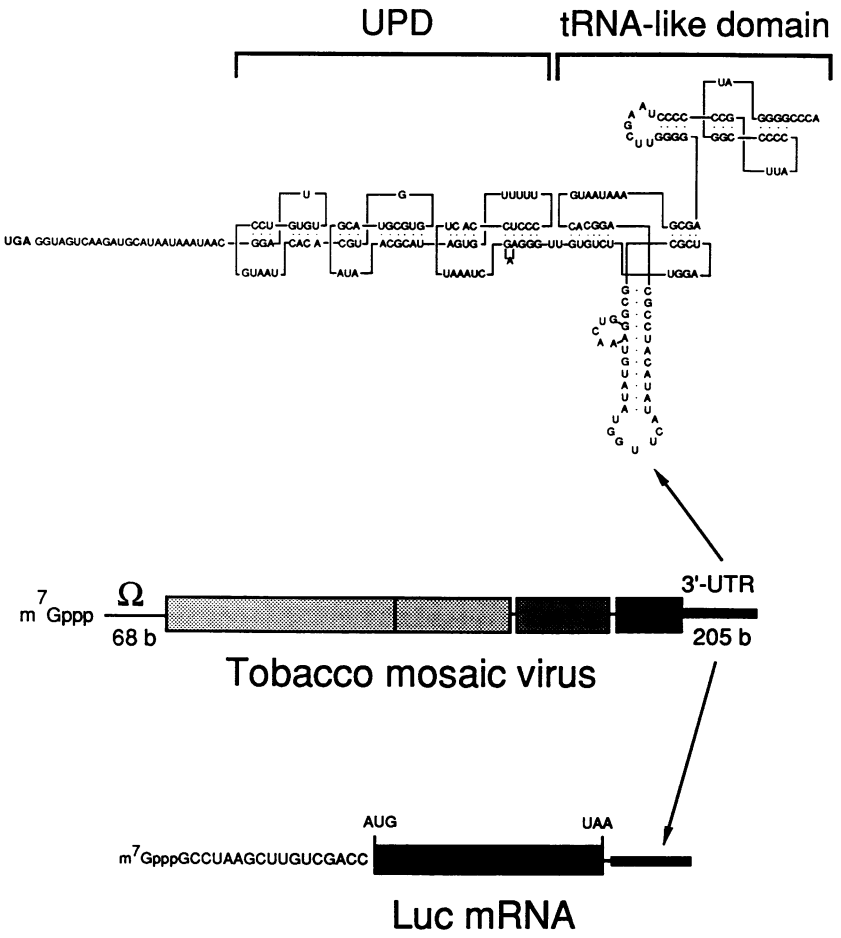
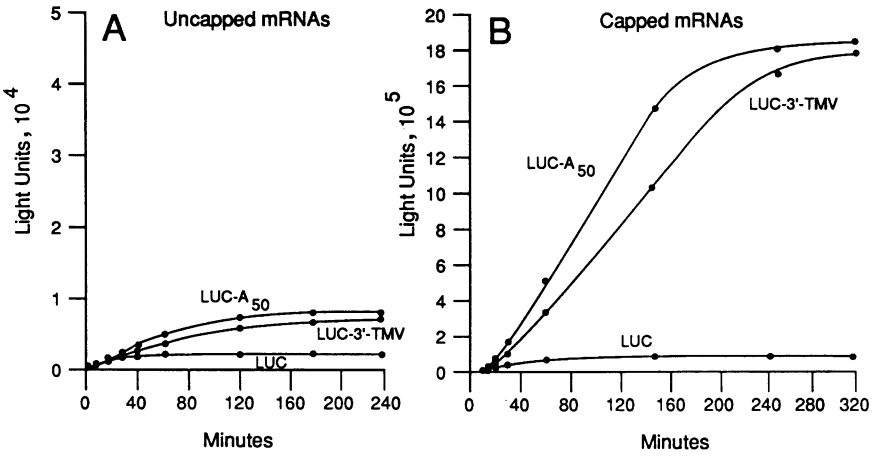


FIG. 1. Primary sequence and higher-order structure of the TMV 3′ UTR. The UPD and tRNA-like domains are indicated by brackets. The organization of the TMV genomic RNA is illustrated; coding regions are indicated by boxes. Each chimeric *luc* plasmid DNA construct was linearized at an *Nde*I site present at the 3′ terminus of the tRNA structure in which the terminal CA dinucleotide sequence is the first two bases (b) of the *Nde*I (CATATG). The *luc* mRNA construct terminating in the TMV 3′ UTR and representative of the in vitro-synthesized mRNA used for the in vivo and in vitro studies is illustrated below.

individual contributions. Previous work demonstrated that the TMV 3′ UTR (Fig. 1) is functionally equivalent to the poly(A) tail (12). Although addition of the TMV 3′ UTR to a reporter mRNA did increase message stability approximately 3-fold (9, 12), this could not account for the 50-fold increase at the protein level. To show directly that the TMV 3′ UTR regulates the rate of translation and determine whether, in this capacity, it requires the cap for function, the effect of the TMV 3′ UTR on the kinetics of reporter mRNA translation was measured in vivo. *luc* mRNA constructs were made in which the transcript terminated in either the TMV 3′ UTR (Fig. 1), a poly(A)$_{50}$ tail, or sequence of equivalent length derived from the T7-based vector. The *luc* mRNA constructs were synthesized with a monomethylated cap (m^{7}GpppG) at the 5′ terminus or as uncapped mRNA. The in vitro-synthesized mRNAs were subsequently delivered to tobacco protoplasts by using electroporation, and the rate of translation was determined by the maximum rate of translation. The functional mRNA half-life, defined as the amount of time to complete a 50% decay in the capacity of the mRNA to synthesize protein (24, 33), was also determined for each mRNA.

Like the poly(A) tail (8), the TMV 3′ UTR did little to increase the translational efficiency of uncapped Luc mRNA. A 3.3-fold increase in the functional mRNA half-life was observed (Fig. 2), which is in good agreement with the 2.8-fold increase in the physical mRNA half-life measured previously (12). In contrast to the uncapped mRNA, the TMV 3′ UTR substantially increased the rate of translation of capped mRNA. The effect of the TMV 3′ UTR on the translation of capped mRNA was similar to that observed for a poly(A)$_{50}$ tail (8). If, in regulating translation, the cap and TMV 3′ UTR function independently, their combined effect would be expected to be the multiplication of their individual effects. However, their combined effect was, in fact, greater than the multiplication of their individual contributions in increasing translational efficiency. These data demonstrate that the TMV 3′ UTR does increase translational efficiency but the mechanism underlying this regulation involves an interaction with the cap. In this respect, the TMV 3′ UTR exhibits functional requirements similar to those for a poly(A) tail.

We observed synergism not only between the cap and the TMV 3′ UTR but also between the TMV 5′ leader (Ω) and the 3′ UTR. In 11 independent experiments, the extent to which the TMV 3′ UTR increased expression was compared



mRNA	Translational Efficiency (Light units/min)	Relative rate of translation	Functional mRNA Half-life (min)	Relative mRNA Half-life	Maximum Accumulation (Light units)	Relative level of expression
Uncapped						
Luc	96	1	15	1	2200	1
Luc-A_{50}	108	1.1	42	2.8	8300	3.8
Luc-3'-TMV	77	0.8	50	3.3	6500	3.0
Capped						
Luc	1150	1	44	1	80,000	1
Luc-A_{50}	13,200	11.5	88	2.0	1,700,000	21
Luc-3'-TMV	9230	8.0	118	2.7	1,640,000	20.5

FIG. 2. Kinetic analysis of the effect of the TMV 3′ UTR on translation in vivo. *luc* mRNA terminating in the TMV 3′ UTR, or as poly(A)$^-$ or poly(A)$^+$ RNA, was synthesized in vitro and translated in electroporated tobacco protoplasts as uncapped (A) or capped (B) mRNA. The translational efficiency and functional stability for mRNA are reported in the table. The maximum rate of translation is determined from the first derivative of each curve. The functional half-life is measured as the amount of time to complete a 50% decay in the capacity of the mRNA to synthesize protein (24, 33).

for capped Luc mRNA with Ω or, as a control, a 17-base leader. On average, addition of the TMV 3′ UTR to Luc mRNA increased expression by 15.6-fold (comparing Luc-3′TMV with Luc) when Ω was absent, but its addition increased expression by 26.6-fold (comparing Ω-Luc-3′TMV with Ω-Luc) when Ω was present (Table 1). Likewise, the enhancing effect of Ω on translation was greater when the Luc mRNA terminated in the TMV 3′ UTR (comparing Ω-Luc with Luc) than when it did not (comparing Ω-Luc-3′TMV with Luc-3′TMV). On average, the synergism between Ω and the 3′ UTR was 1.69-fold. This means that the combined effect of Ω and the TMV 3′ UTR on translation is, on average, 69% greater than would be predicted from the multiplication of their individual effects. This level of synergism is smaller than that observed between the cap and the TMV 3′ UTR, which was approximately 10-fold. However, a paired *t* test demonstrated that the synergism between Ω and the TMV 3′ UTR was highly significant ($P < 0.0001$).

No synergism between the cap and the TMV 3′ UTR is observed in vitro. The synergistic interaction between the cap and the poly(A) tail was observed only in vivo (8), data suggesting that in vitro lysates may be lacking an essential component required for the interaction. To investigate whether the TMV 3′ UTR would function in vitro, wheat germ lysate was programmed with the same batch of *luc* mRNAs used in the in vivo experiment so that the in vitro and in vivo results could be compared directly. Although the in vivo observations were made in a dicotyledonous species and the in vitro experiments were performed with a lysate derived from a monocotyledonous species, we have shown previously that the TMV 3′ UTR functions equally well in both plant types (12). The addition of a cap did stimulate translation in wheat germ lysate, whereas the TMV 3′ UTR failed to increase translation significantly regardless of whether the mRNA was capped or not (Fig. 3). The TMV 3′ UTR also failed to function in lysate derived from rabbit reticulocytes (data not shown), whereas it did function in vivo in Chinese hamster ovary cells (9). Interestingly, the stimulation afforded by a cap in vitro (2.5-fold) was substantially less than that observed in vivo (120-fold). Therefore, as with a cap and a poly(A) tail, the regulatory function of the TMV 3′ UTR is significantly reduced or eliminated in vitro.

Identification of the sequences within the UPD required for regulation. The UPD of the 3′ UTR in TMV strain U1 (Fig.

TABLE 1. Synergism between Ω and the TMV 3′ UTR in tobacco

Expt	Ω			TMV 3′ UTR		
	Fold enhancement		Synergism (A/B)	Fold enhancement		Synergism (C/D)
	In the absence of the TMV 3′ UTR (A)	In the presence of the TMV 3′ UTR (B)		In the absence of Ω (C)	In the presence of Ω (D)	
1	32.4	49.1	1.52	21.1	31.9	1.51
2	35.2	53.9	1.53	15.4	23.5	1.53
3	29.4	54.1	1.84	17.4	32.1	1.84
4	30.9	44.8	1.45	19.0	27.5	1.45
5	31.9	50.8	1.59	10.9	17.4	1.60
6	35.5	53.3	1.50	11.1	16.6	1.50
7	30.6	58.1	1.90	9.7	18.5	1.91
8	35.1	39.6	1.13	15.3	17.2	1.12
9	33.6	61.1	1.82	11.1	20.0	1.80
10	14.2	43.2	3.04	21.4	64.8	3.03
11	14.7	18.2	1.24	19.0	23.0	1.21
Avg	29.4	47.8	1.69	15.6	26.6	1.68

1) contains the information for regulating translation (12). Deletion of the entire tRNA-like domain reduced the regulation associated with the TMV 3′ UTR by only 35% (12). Several viruses related to TMV strain U1 terminate in 3′ UTRs that are similar in higher-order structure (40). A phylogenetic comparison of the UPDs from the eight viruses generated a set of absolutely conserved nucleotides within this domain (Fig. 4). All eight UPDs contain three pseudoknots, although only the structures of UPDs of TMV strains U1 and U2 (also known as tobacco mild green mosaic virus) have been determined by using chemical or enzymatic probes (15, 40). The length of the base-paired region of the first pseudoknot (PK1) varies from 7 to 12 bp, whereas those for pseudoknots 2 (PK2) and 3 (PK3) are conserved at 9 bp each. Moreover, the size of PK2 and PK3 is conserved at 22 and 30 bases, respectively. At the primary level, 5 of the 22 bases in PK2 were absolutely conserved, as were 14 of 30 bases in PK3 (Fig. 4). Absolutely conserved sequence was found in both single-stranded and base-paired regions. In a consensus sequence of seven of eight, an A · U and a U · A base pair present in the 5′ minihelices of PK2 and PK3, respectively, could be included in the conserved sequence, as could a C in the loop that crosses the 5′ minihelix of PK3. Conservation at these positions suggested that the primary sequence was functionally required for the regulation of either translation or viral replication.

To determine which of the pseudoknots were required for UPD-mediated control of translation, the pseudoknots were tested individually and in pairs, using *luc* as the reporter mRNA in carrot protoplasts. The introduction of restriction sites between each pseudoknot to facilitate the deletion of individual pseudoknots resulted in a two-base insertion between PK1 and PK2 (UPD-101; Fig. 5); a four-base insertion between PK2 and PK3 (UPD-100); and a five-base insertion between PK3 and the tRNA-like domain (UPD-107). Because of the coaxial stacking in and between the pseudoknots within the UPD and the tRNA-like domain, a double-stranded RNA helix 35 bp in length runs through both domains (40). The introduction of restriction sites between the pseudoknots, in itself, generated mutants in

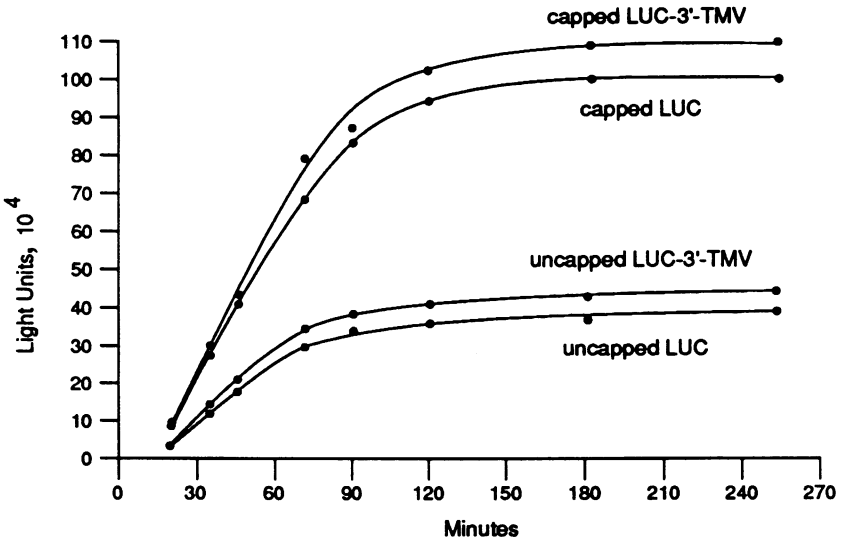


FIG. 3. Analysis of TMV 3′ UTR function in vitro. Wheat germ lysate was programmed with 160 ng of each *luc* mRNA construct. Aliquots were removed at time intervals, assayed for luciferase activity, and plotted as a function of time.

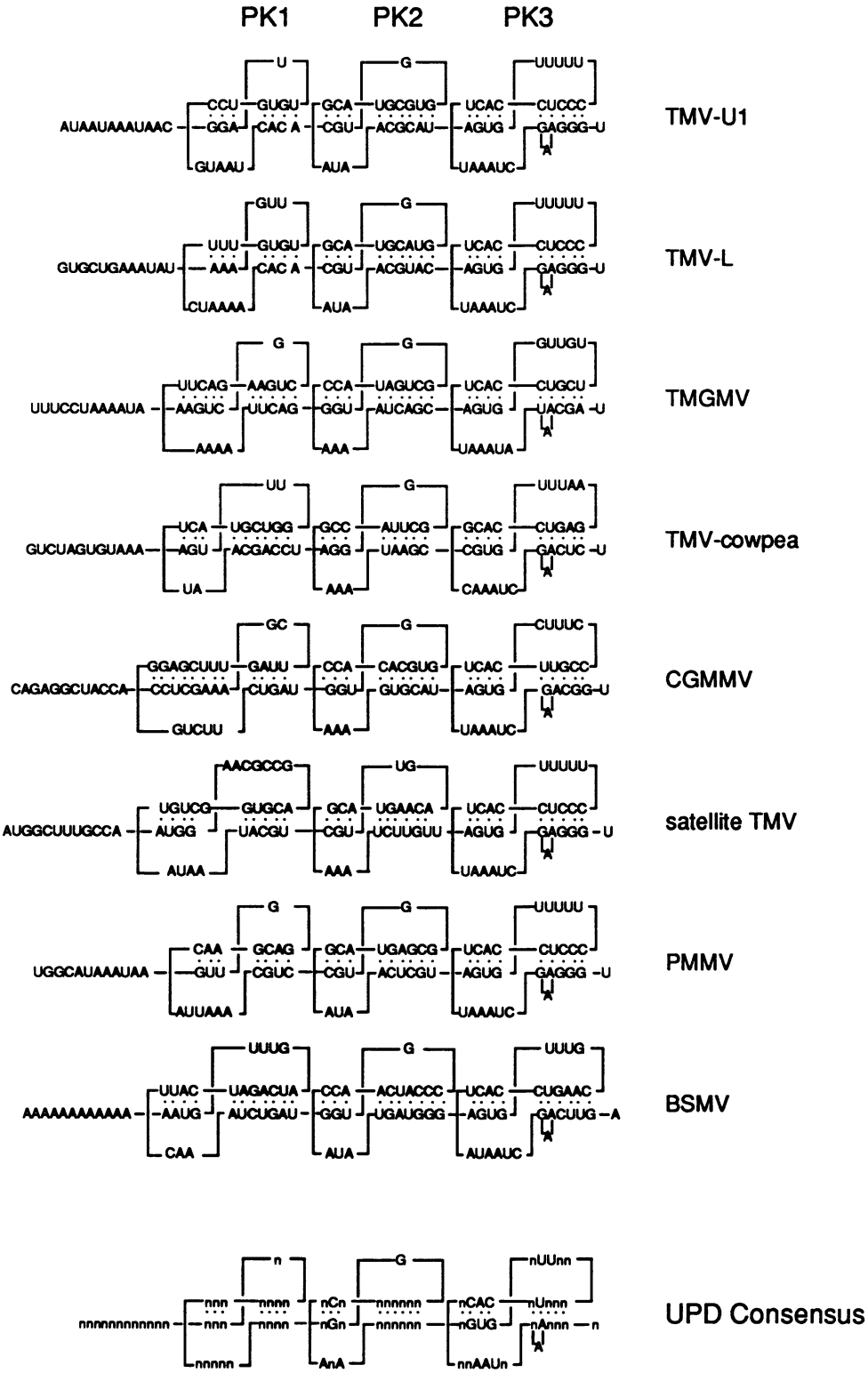


FIG. 4. Phylogenetic analysis of the UPD and its consensus sequence. The UPDs used for the analysis were the U1 (16), L (31), and cowpea (28) strains of TMV, tobacco mild green mosaic virus (TMGMV) (39), cucumber green mottle mosaic virus (CGMMV) (27), and pepper mild mottle virus (PMMV) (3), all of which are tobamoviruses; satellite tobacco mosaic virus (29); and RNAα (18), RNAβ (17, 34), and RNAγ (19) from barley stripe yellow virus (BSMV), a hordeivirus. The absolutely conserved nucleotides in these viruses are shown below. A base is indicated only when present in all UPDs. Positions in which no absolute consensus exists are indicated by n.

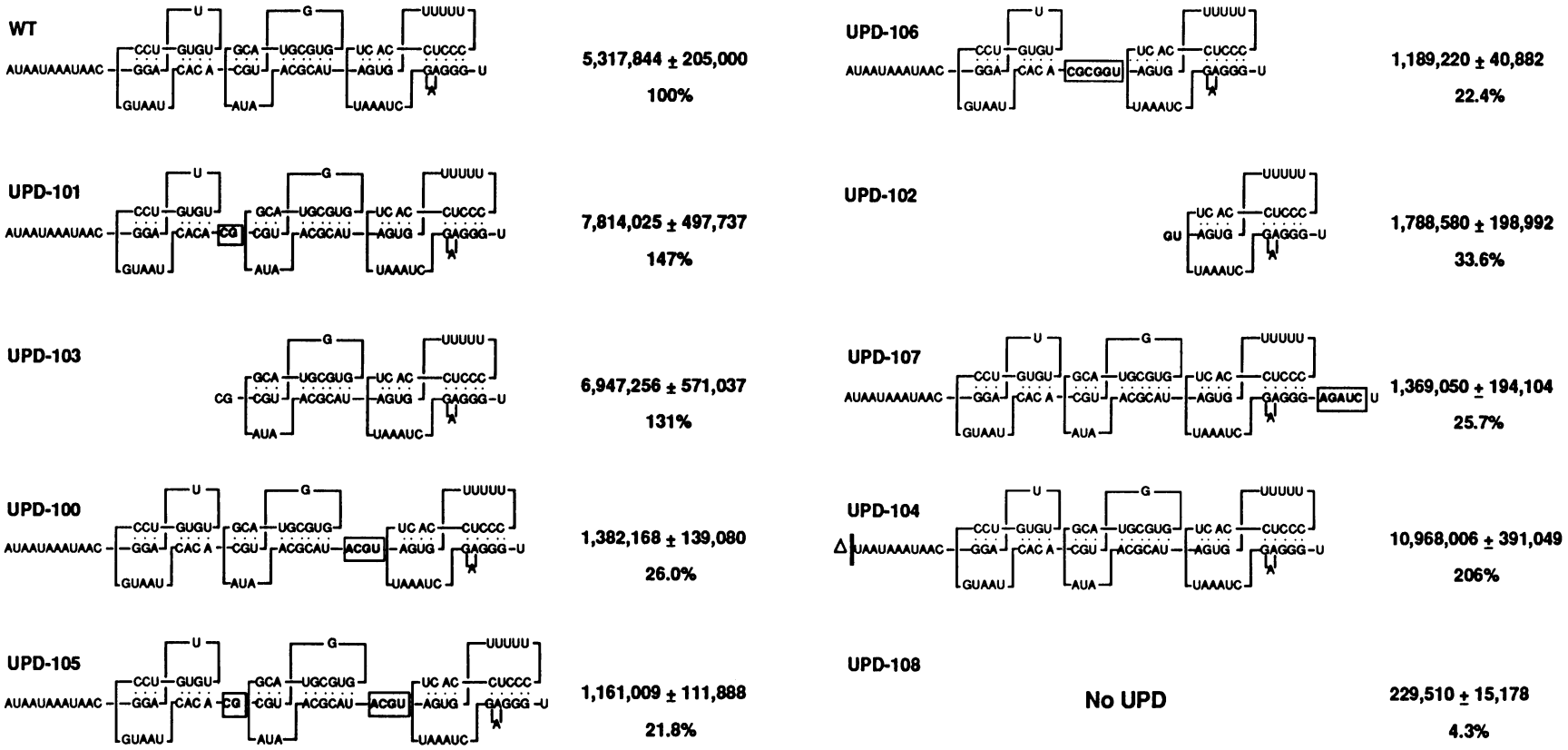


FIG. 5. Effect of pseudoknot deletion on UPD function. Restriction sites engineered between UPD pseudoknots were used to generate constructs in which the effect of deleting one or more pseudoknots on UPD function could be tested. Additional sequences introduced between pseudoknots are boxed. Only the UPD region for each construct is shown. All constructs contain the tRNA-like domain and use the *luc* reporter gene as illustrated in Fig. 1. In vitro-synthesized mRNAs were translated in vivo. Luciferase specific activity (light units per milligram of protein) resulting from each construct is shown to the right of the corresponding construct as the average of triplicate experiments ± standard deviation. Relative values are indicated below each specific activity; the value for the wild-type (WT) construct is 100%.

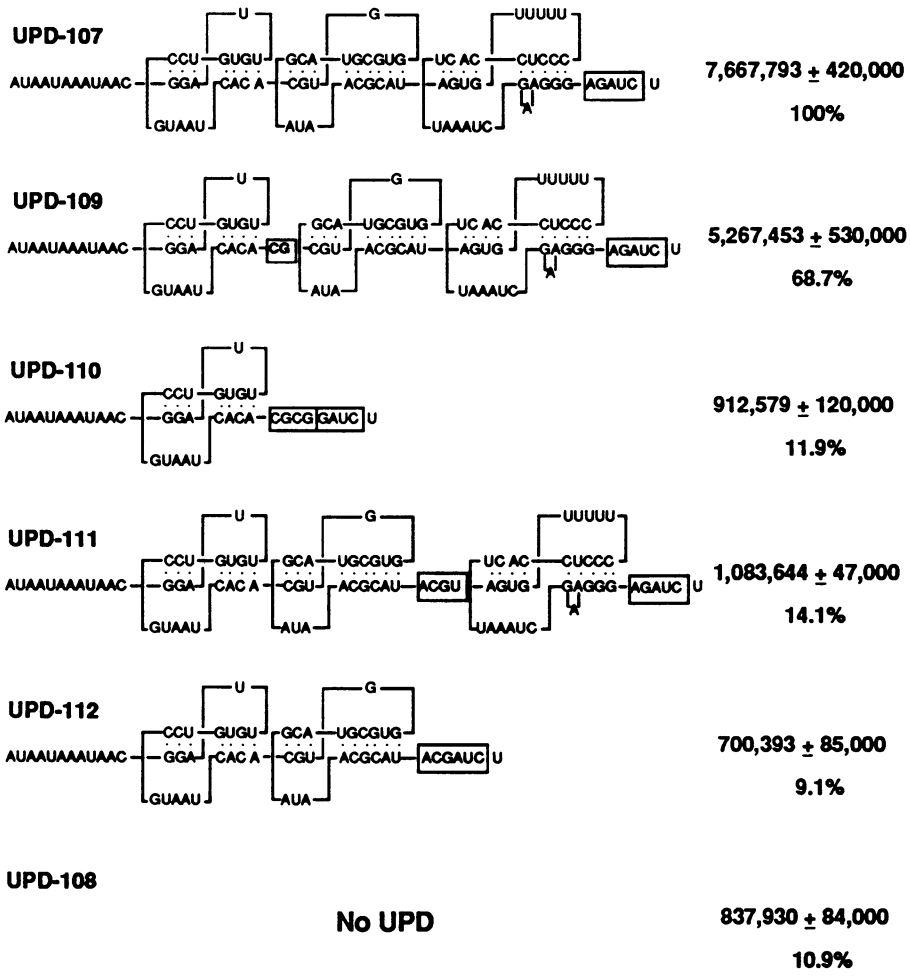


FIG. 6. Additional deletion mutation analysis of the UPD. Additional sequences introduced between pseudoknots as a result of the introduction of restriction sites are boxed. Only the UPD region for each construct is shown. All constructs contain the tRNA-like domain and use the *luc* reporter gene as illustrated in Fig. 1. In vitro-synthesized mRNAs were translated in vivo. Luciferase specific activity (light units per milligram of protein) resulting from each construct is shown to the right of the corresponding construct as the average of triplicate experiments ± standard deviation. Relative values are indicated below each specific activity; the value for UPD-107 is 100%.

which the effect of eliminating the coaxial stacking between pairs of pseudoknots could be examined.

Deletion of the UPD domain, leaving only the tRNA-like domain (UPD-108), resulted in a level of expression that was only 4% of that measured for the wild-type TMV UPD construct (Fig. 5). This is representative of the 25- to 50-fold increase in expression previously observed when the UPD is added to a reporter mRNA (12). For the following mutations of the UPD, those that result in at least 75% of the wild-type activity we have designated as having little effect on UPD function, those that result in approximately 40 to 75% of the wild-type activity are designated as having a moderate effect, and those that result in less than 40% of the wild-type activity are designated as having a severe effect. Neither the deletion of PK1 (UPD-103) nor the two-base insertion between PK1 and PK2 (UPD-101) reduced UPD function (Fig. 5). Deletion of 58 bases of sequence upstream of the UPD (UPD-104) also had no detrimental effect on regulation. In contrast, the mutant in which PK2 was deleted (UPD-106) was only 22.4% as active as the wild-type UPD. Interestingly, a four-base insertion between PK2 and PK3 (UPD-100) had nearly the same effect as did deletion of PK2, resulting in a level of activity just 26.0% of the wild-type level. No additional reduction in activity was noted for the double-insertion mutant, UPD-105. Deletion of both PK1 and PK2 (UPD-102) had approximately the same effect on UPD function (33.6% of wild-type activity) as did the deletion of just PK2 alone. Introduction of five bases between PK3 and the tRNA-like domain (UPD-107) reduced UPD function to 25.7% of wild-type activity.

In a separate experiment in which UPD-107 was used as the reference construct as well as to create a second set of deletion mutants, the deletion of both PK2 and PK3 (UPD-110; Fig. 6) or just PK3 alone (UPD-112) reduced UPD function to 11.9 and 9.1%, respectively. Separation of PK3 from coaxial stacking with the pseudoknots on either side (UPD-111) had an effect similar to that of the deletion of PK3. Even the two-base insertion between PK1 and PK2 that had no negative effect in UPD-101 had a moderately detrimental effect in UPD-109. These data suggest that both PK2 and PK3 are required for function but PK3 is more important for UPD function. As isolation of the regulatory pseudoknots had an effect similar to that of pseudoknot

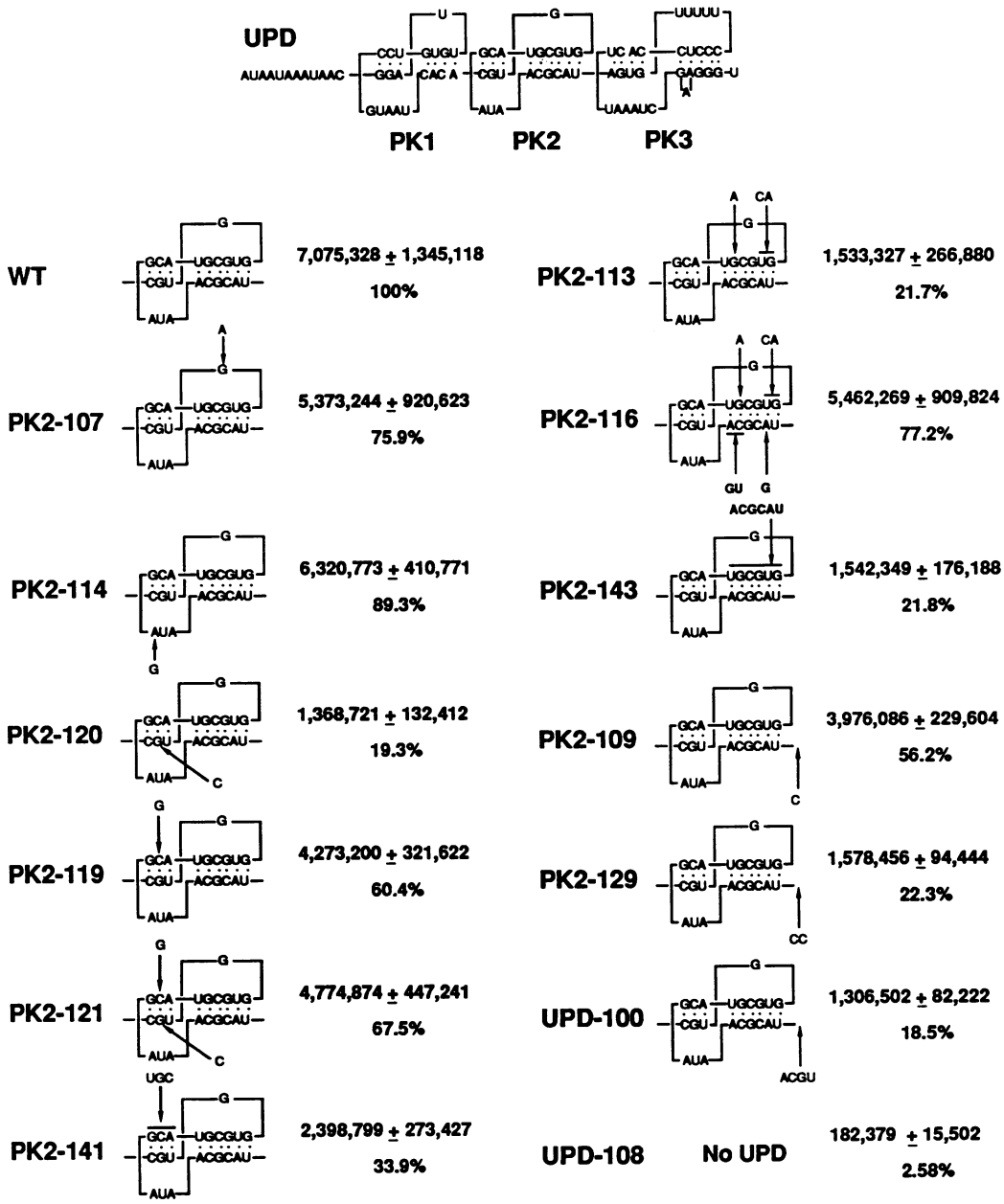


FIG. 7. Analysis of the role of primary sequence and higher-order structure of PK2 function. The complete UPD is shown at the top for reference. Only PK2 is shown for each point and compensatory base change mutant. All constructs contain the tRNA-like domain. In vitro-synthesized mRNAs were translated in vivo. Luciferase specific activity (light units per milligram of protein) resulting from each construct is shown to the right of the corresponding construct as the average of triplicate experiments ± standard deviation. Relative values are indicated below each specific activity; the value for the wild-type (WT) construct is 100%.

deletion, the position of the pseudoknots with respect to each other is also essential for UPD function.

To more precisely determine which sequences within PK2 and PK3 were required for UPD function, we made a series of point mutations and tested their effects in vivo. The effects of mutations within PK2 are illustrated in Fig. 7. As noted above, in all constructs, the tRNA-like domain is present. In this experiment, deletion of the entire UPD (UPD-108) resulted in a level of expression only 2.58% of that measured for an intact UPD. Mutation of the conserved G in the one-base loop (PK2-107) or the A in the three-base loop (PK2-114) had little effect. Changes to the single phylogenetically conserved base pair in PK2 were more significant. Mutation of the G to a C, creating a C · C base pair (PK2-120), resulted in a drop to 19.3% of the wild-type level. The effect of this mutation is similar to that observed when PK2 was deleted (UPD-106; Fig. 5). The C→G mutant (PK2-119), which created a G · G base pair, retained 60.4% of wild-type activity, whereas full disruption of this 3-bp minihelix (PK2-141) resulted in a drop to 33.9% of wild-type activity (Fig. 7). The double mutant (PK2-121) retained 67.5% of wild-type activity. Moreover, the higher-order

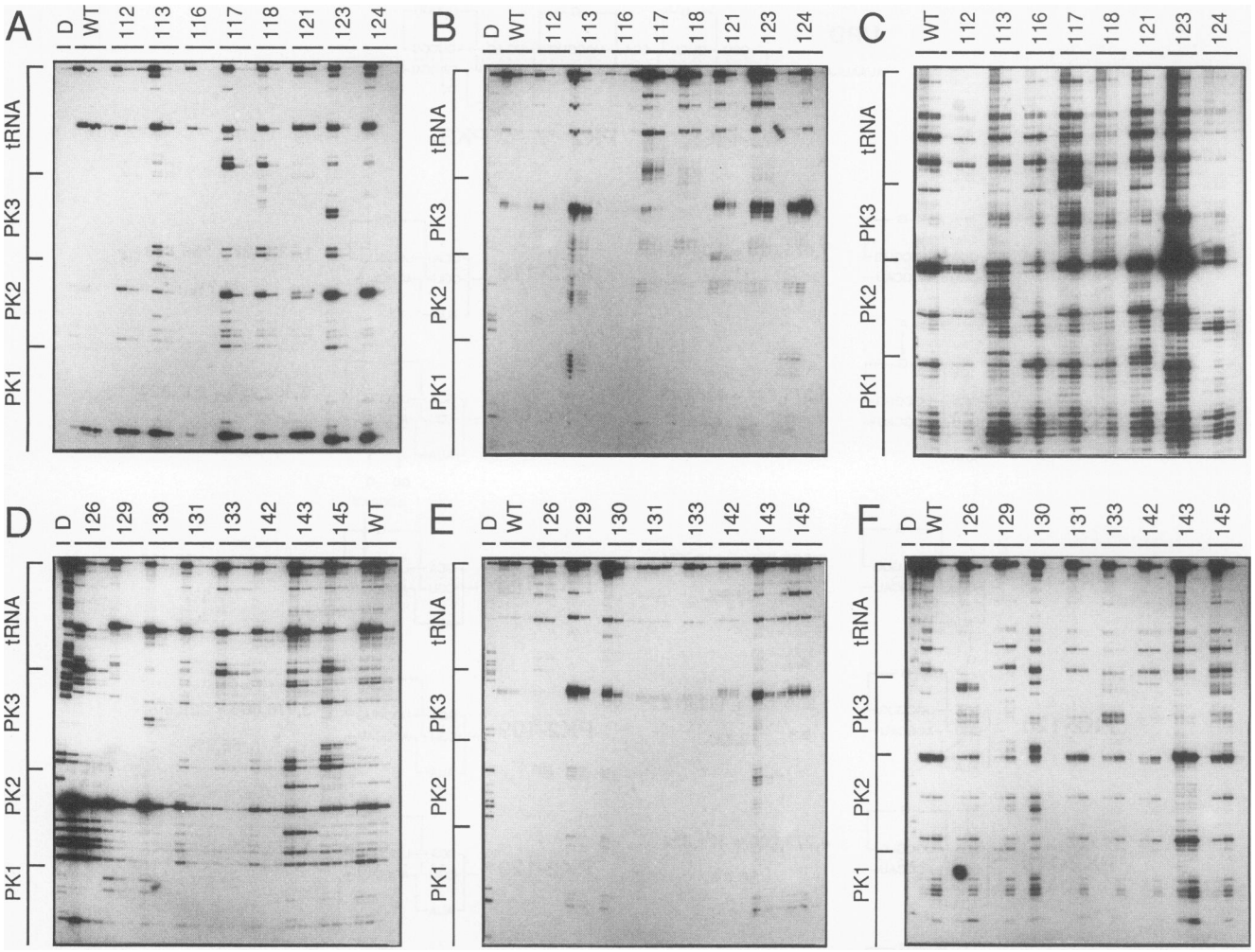


FIG. 8. Structural analysis of selected mutants. RNAs containing only the UPD and tRNA-like domains were synthesized in vitro, gel purified, and end labeled at the 5′ terminus. Limited digests using RNase T_1 (A and D), nuclease S1 (B and E), or RNase V_1 were performed at three enzyme concentrations, and the products of the reactions were resolved on a 7 M urea–8% polyacrylamide gel. The mutant tested is indicated above each set of three lanes. D, a limited digest (using RNase T_1) of fully denatured wild-type sequence; WT, wild-type UPD. For each gel, positions corresponding to PK1, PK2, PK3, and the tRNA-like domain are indicated at the left.

structure, as determined by using the enzymatic probes RNase T_1, RNase V_1, and nuclease S1, was largely restored in PK2-121 (Fig. 8). Because the effect on function due to the disruption of this 3-bp minihelix was not as great in PK2-119 or PK2-141 as it was in PK2-120, which was not completely reversed by the compensatory mutant PK2-121, these data suggest that the G of the conserved G · C base pair may be important but that base pairing within this region is also required.

The 6-bp 3′-proximal minihelix of PK2 was rotated 180° in two steps. First, the top six bases were replaced with the bottom six (PK2-113; Fig. 7). As this sequence is partially palindromic, this mutation resulted in disruption of only 3 bp. Structural mapping of this mutant (PK2-113) confirmed that the entire PK2 had been disrupted (Fig. 8). This mutation reduced activity to 21.7% (Fig. 7). Making the compensatory base change mutant (PK2-116) in which the bottom six bases of PK2-113 were replaced with the original top six bases resulted in restoration of the higher-order structure (Fig. 8) and near restoration of function (Fig. 7), data suggesting that the secondary structure is of greater importance than the primary sequence within this region.

The effect of separating PK2 and PK3 was examined in greater detail. An insert of a single C between the two pseudoknots (PK2-109) reduced activity to 56.2% (Fig. 7). Activity was further reduced to 22.3% when two C's were introduced (PK2-129) and to 18.5% with a four-base insertion (UPD-100). Structural mapping revealed that the insertions between PK2 and PK3 did reduce the stability of PK2, with no effect on PK1 or PK3 (PK2-129; Fig. 8). The deleterious effect on UPD function due to separation of PK2 and PK3 correlates with the localized disruption of the higher-order structure.

The effects of mutations within PK3 are illustrated in Fig. 9. When the top-strand 5′-proximal 4-bp minihelix of PK3 was replaced with the bottom strand (PK3-123), resulting in disruption of at least 2 bp, activity dropped to 11.4%, compared with 3.52% when the entire UPD was deleted (UPD-108). This mutation (PK3-123) caused disruption of the higher-order structure of the four-base mini-helix as well

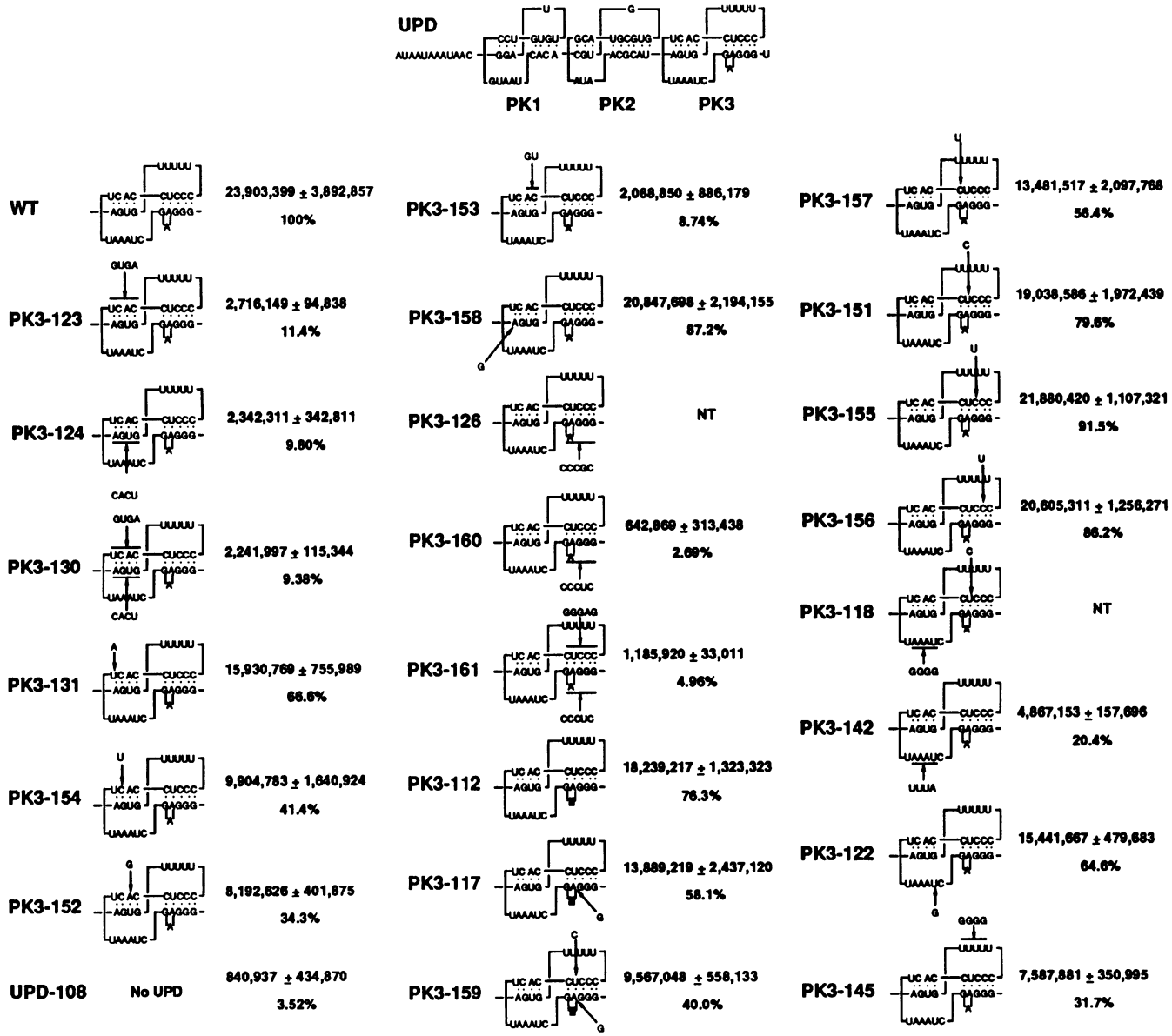


FIG. 9. Analysis of the role of primary sequence and higher-order structure of PK3 function. The complete UPD is shown at the top for reference. Only PK3 is shown for each point and compensatory base change mutant. All constructs contain the tRNA-like domain. In vitro-synthesized mRNAs were translated in vivo. Luciferase specific activity (light units per milligram of protein) resulting from each construct is shown to the right of the corresponding construct as the average of triplicate experiments ± standard deviation. Relative values are indicated below each specific activity; the value for the wild-type (WT) construct is 100%.

as a reduction in the stability of PK2 and the 5′-proximal pseudoknot of the tRNA-like domain (Fig. 8). Replacement of the bottom strand with the top strand (PK3-124) reduced activity to 9.8% (Fig. 9). Again, the stability of PK3 and PK2 was reduced, with no significant changes to the tRNA-like domain (Fig. 8). The double mutant PK3-130, in which wild-type base pairing for this region is maintained but the 4-bp minihelix has been rotated 180°, did not restore activity (9.38% of wild-type activity; Fig. 9), data suggesting that primary sequence within this minihelix is essential for UPD function. This conclusion was confirmed by the structural mapping: the higher-order structure of PK2 and PK3 in PK3-130 was restored (Fig. 8). Single base mutations within the top strand of the four-base minihelix suggest that the three conserved base pairs are critical for function. A U→A mutation (PK3-131) that created an A · A base pair had only a moderate effect on function (66.6% as active as the wild type; Fig. 9). The effect of this mutation on the structure of PK3 was also minimal (Fig. 8). Single base changes which maintained base pairing (PK3-154 and PK3-152) reduced UPD function to 41.4 and 34.3%, respectively (Fig. 9). Interestingly, these mutations may have reduced the stability of PK2 and PK3. An AC→GU double mutant (PK3-153) that maintained base pairing was, nevertheless, only 8.74% as active as the wild type (Fig. 9), a level of activity close to that observed when the entire UPD was deleted (UPD-108).

The only mutation possible in the bottom strand of this minihelix that maintained base pairing was an A→G mutation (PK3-158) that had only a minor effect on activity (87.2% of wild-type activity).

Mutation of the bottom strand of the distal 5-bp minihelix (PK3-160) eliminated UPD function completely (2.69% of wild-type activity; Fig. 9). The function of the UPD of a second mutant in this same region, PK3-126, although not included in this data set, was affected to the same extent. The higher-order structure of the UPD and the 5′-proximal pseudoknot of the tRNA-like domain of PK3-160 (data not shown) and PK3-126 (Fig. 8) were disrupted in a way that suggested a rearrangement in base pairing. A 180° rotation of this minihelix (PK3-161) that maintained the base pairing within PK3 also eliminated UPD function (4.96% of wild-type activity; Fig. 9). Although the pseudoknot structure of PK3 was maintained in PK3-161, PK2 was less stable. Removal of the bulged A (PK3-112) had a small negative effect on function (76.3% of wild-type activity; Fig. 9) and little effect on UPD structure. Changing the single base-paired A to a G (PK3-117), which maintains the base pairing with the U, decreased UPD function even further (58.1% of wild-type activity). The 5′-proximal pseudoknot of the tRNA-like domain was disrupted by this mutation. Changing the U to a C (PK3-159), which also maintains base pairing, reduced activity to 40.0% (Fig. 9). Curiously, in this mutant, no change in the structure, including the tRNA-like domain pseudoknot, was observed. Changing just the U of this base pair to C (PK3-151), which disrupts base pairing, had a small effect on UPD function (79.6% of wild-type activity; Fig. 9). A C→U change in the first C · G base pair of this minihelix (PK3-157) reduced UPD function to 56.4% of the wild-type level even though base pairing and the higher-order structure were maintained. Other single base changes in the top strand of this minihelix that maintained base pairing had little effect on UPD function: changing either of the central C's to a U (PK3-155 and PK3-156) resulted in a level of activity similar to the wild-type level (Fig. 9).

Mutation of the conserved 5′-AAAU-3′ to 5′-UUUA-3′ within the loop that crosses the minor groove of PK3 (PK3-142) reduced activity to just 20.4% of wild-type activity (Fig. 9). No significant impact on the higher-order structure of PK3 was observed for PK3-142 even in an exposure longer than that shown in Fig. 8. A U→G change within this same sequence (PK3-122) resulted in a drop to 64.6% of wild-type activity. Changing the 5′-AAAU-3′ to 5′-GGGG-3′ (PK3-118) reduced UPD function to the same extent as did PK3-142 (PK3-118 was not included in the data set in Fig. 9) but had an impact on PK3 structure that may be due to the disruption of a base pair within the 3′ minihelix (Fig. 8). Changing four of the U's within the loop that crosses the major groove to G's (PK3-145) caused changes in the structure of PK3 and reduced activity to 31.7% of wild-type activity (Fig. 9). Considering the effects of the PK3 mutations on both function and structure, we conclude that base pairing within either minihelix is not sufficient for function and that the three conserved base pairs within the 5′-proximal minihelix are necessary for UPD function. Primary sequence within the 3′-proximal minihelix, including the A of the conserved A · U base pair and the C of the first C · G base pair within this minihelix, is also important. In addition to primary sequence within the base-paired regions, the conserved 5′-AAAU-3′ loop sequence is also required for function.

The cumulative effect of multiple mutations within the single-stranded regions was next examined (Fig. 10). As demonstrated above, changes within the single-stranded regions of PK3 had a greater impact on UPD function than did changes in the loops of PK2. Individually, most mutations within the PK2 and PK3 loops had, at most, a moderate effect on UPD function (PK3-111, PK3-112, and PK2-114 in Fig. 10; PK2-107 in Fig. 7). Exceptions to this are the 5′-AAAU-3′→5′-UUUA-3′ (PK3-142) and 5′-UUUU-3′→5′-GGGG-3′ (PK3-145) mutants, which, as shown above, reduced UPD function significantly (Fig. 10). However, as mutations that have little effect individually are combined, their effects on UPD function become increasingly significant (Fig. 10). A correlation between the accumulation of mutations within the PK2 and PK3 loops and changes in the stability of these pseudoknots was observed. Not only was the structure of PK2 and PK3 increasingly disrupted with the increase in the number of loop mutations, but the 5′-proximal pseudoknot of the tRNA-like domain was disrupted as well (data not shown). These data suggest that the primary sequence of the loops is important mainly for structural reasons (with the exception of the 5′-AAAU-3′ sequence of PK3) and probably plays no direct role as a recognition sequence for any UPD-binding proteins.

Proteins specifically bind to TMV UPD and the 5′ leader. The work described above identified PK3 as comprising the regulatory core in the UPD, with perhaps PK2 playing a supporting role. Previous work identified a 35-base region within the 68-base TMV leader (Ω) as the core enhancing element (13). That Ω and the TMV 3′ UTR interact synergistically in regulating translational efficiency (Table 1) suggests an interaction between these two regulatory elements. We postulate that these regulatory sequences do not exert their control on translation as naked RNA but rather function as binding sites for cellular proteins that mediate the regulation. As a first step in characterizing Ω/UPD-binding proteins, extracts were prepared from both wheat germ and carrot suspension cells and used to generate RNA-protein complexes with Ω and UPD RNA that could be detected as electrophoretically retarded complexes. Ideally, the RNA used for protein binding to Ω or the UPD should be a full-length, capped mRNA containing either Ω or the UPD. However, the size of such an mRNA prevents the resolution of the complexed and free RNAs. Therefore, short RNAs, an 83-base RNA that contained Ω and a 91-base RNA that contained the UPD, were used to obtain the necessary resolution. Moreover, the RNAs used in these studies were uncapped to avoid complications due to binding of eIF-4F to the cap structure. A single shifted complex was observed with labeled, uncapped UPD RNA in extract made from wheat germ (Fig. 11A, lane 2). The percentage of RNA present in the complex was determined by excising each part of the gel containing the free and shifted bands and then counting in a scintillation counter. Competition with a 1:1.4 molar ratio of cold to hot UPD RNA resulted in a 50% inhibition of complex formation. Cold Ω RNA also acted as an efficient competitor against UPD: 50% inhibition was achieved by a 1:0.3 molar ratio of cold Ω to hot UPD RNA. To eliminate nonspecific protein binding to the target RNAs, all assays were carried out in an 800-fold excess of total yeast RNA and 5 mg of heparin per ml. UPD complex formation was not affected by competition with a 20-fold molar excess of a random sequence or a 1,000-fold molar excess of poly(C). UPD complex formation was also observed when the binding reaction was performed under conditions approximating those for the RNA structural mapping experiments (i.e., 5 mM $MgCl_2$, 24°C), suggesting that

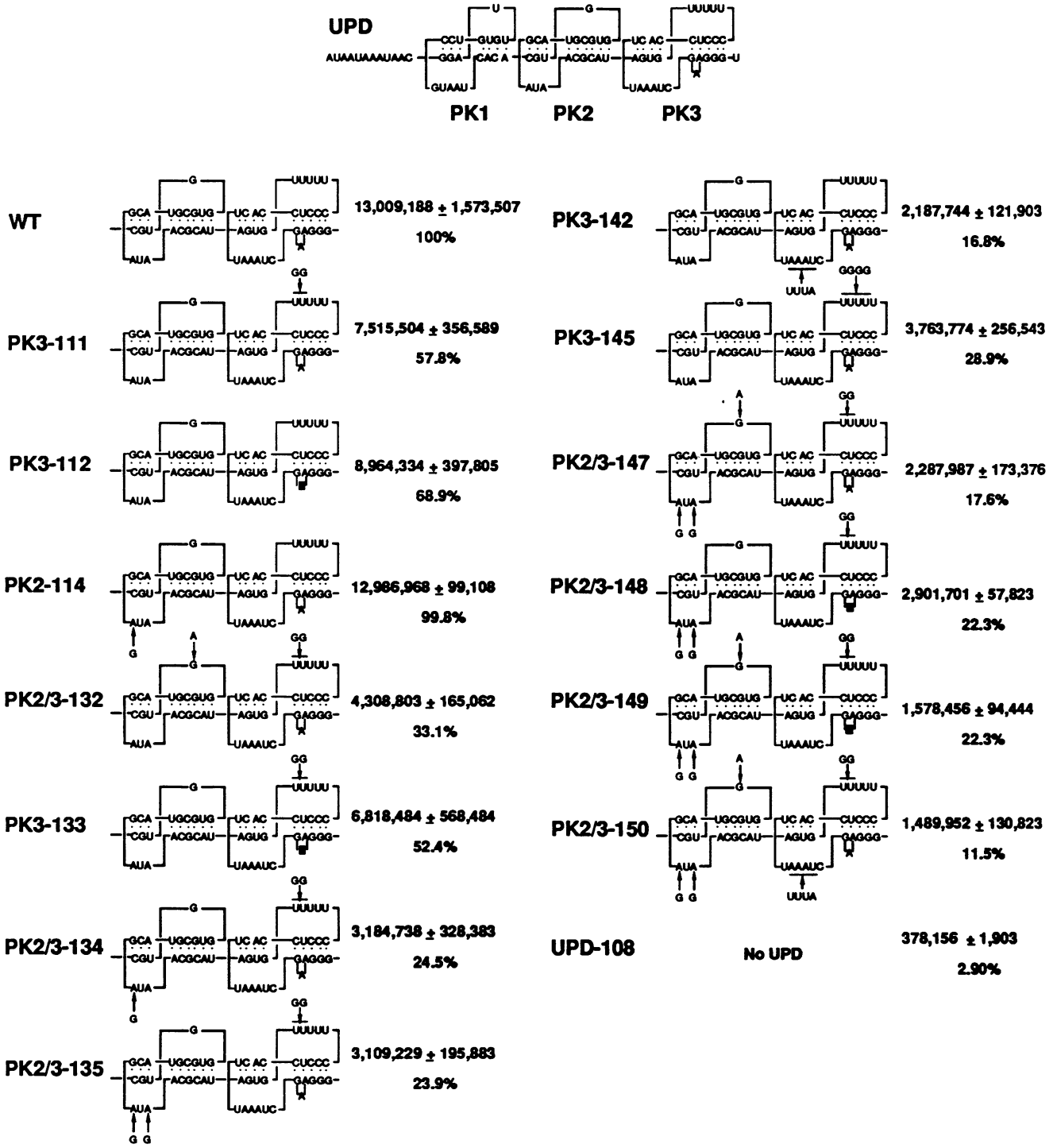


FIG. 10. Analysis of the role of primary sequence within the loops of PK2 and PK3 for UPD function. The complete UPD is shown at the top for reference. Only PK2 and PK3 are shown for each mutant. All constructs contain the tRNA-like domain. In vitro-synthesized mRNAs were translated in vivo. Luciferase specific activity (light units per milligram of protein) resulting from each construct is shown to the right of the corresponding construct as the average of triplicate experiments ± standard deviation. Relative values are indicated below each specific activity; the value for the wild-type (WT) construct is 100%.

the UPD is present in the same conformation for both types of analyses.

Although the UPD is not functionally active in in vitro translation lysates made from wheat germ, Ω is functionally active (11, 14). Therefore, we also tested complex formation in wheat germ extract by using Ω. A shifted complex was observed when labeled, uncapped Ω RNA was used in the wheat germ extract binding assay (Fig. 11B, lane 2). In this

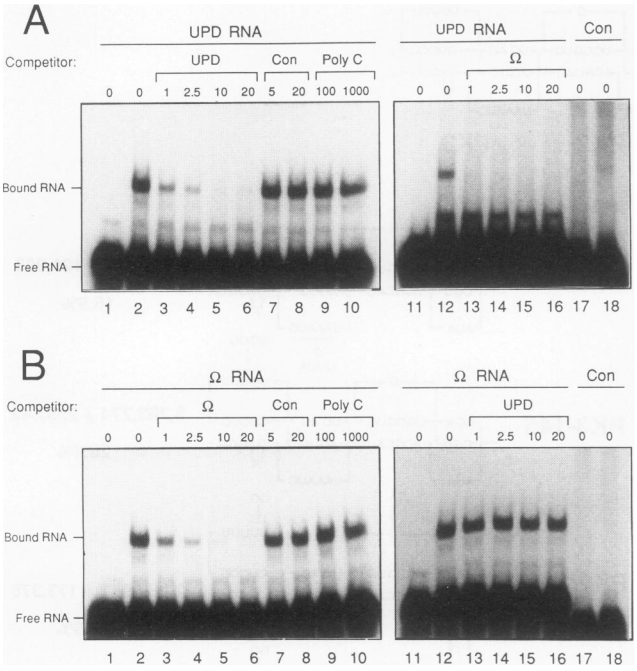


FIG. 11. Ω and the UPD are specifically recognized by RNA-binding protein present in wheat germ extract. Radiolabeled UPD (A) and Ω (B) RNAs were incubated in the absence (lanes 1 and 11) or presence (lanes 2 to 16) of extract. No competitor was added in lanes 1, 2, 11, 12, 17, and 18. (A) Cold UPD was added as a competitor in lanes 3 to 6, a 58-base control random sequence (Con) was added in lanes 7 and 8, poly(C) was added in lanes 9 and 10, and Ω was added in lanes 13 to 16. (B) Cold Ω RNA was added as a competitor in lanes 3 to 6, control random sequence was added in lanes 7 and 8, poly(C) was added in lanes 9 and 10, and UPD RNA was added in lanes 13 to 16. The fold molar excess of competitor RNA added is indicated above each lane. In both panels, control random sequence was incubated in the absence (lanes 17) or presence (lanes 18) of extract.

experiment using 2 μg of protein extract, only one band was observed. However, when 6 μg or more of protein extract was used, we observed a second, more retarded band that was not seen when UPD RNA was tested with this amount of protein extract (data not shown). This finding suggests that a second protein species may bind either directly to Ω or to the species already bound to Ω. Alternatively, multiple binding sites for a single protein species may be present in Ω. Fifty percent inhibition of Ω complex formation was achieved by a 1:1.25-fold molar ratio of cold to hot, uncapped Ω RNA. Cold UPD competed with Ω but only poorly: 29% inhibition was achieved when a 20-fold molar excess of cold UPD RNA was present. Neither a 20-fold molar excess of control random sequence nor a 1,000-fold molar excess of poly(C) competed with Ω. No shifted complex was observed with the control random sequence. Preliminary mapping of the binding site within Ω indicated that the region representing the core enhancing element responsible for the translational enhancement (13) also serves as the binding site (8a).

To determine whether complex formation occurred in the same cell type used for the functional assays, gel shift analysis of Ω and the UPD was performed in extracts made from carrot suspension cells. Both Ω and the UPD formed retarded complexes with 7 μg of carrot extract protein (Fig. 12A, lanes 2 and 7). The shifted bands were consistently not as sharp as those observed with wheat germ extract, which may be a result of protease activity. Although the two retarded bands observed with Ω and 7 μg of carrot protein are also observed with 6 μg of wheat germ protein, UPD RNA forms two retarded complexes only with carrot extract. This finding suggests that carrot extract may contain a binding activity that is not present in wheat germ. The difference is probably not due to differences in activity between the two extracts, as even highly purified preparations from wheat germ that exhibit a high level of binding activity do not give more than one retarded complex with UPD RNA (39a). A third, more faint complex can be observed migrating below the two main bands. This band, also present in the experiment with wheat germ extract, was not consistently observed.

As with wheat germ extract, cold Ω RNA efficiently competed with both labeled Ω and UPD RNA (Fig. 12B). Although cold UPD competed with labeled UPD, just as observed wheat germ extract, it was a less efficient competitor with labeled Ω. Neither poly(C) nor the random control RNA competed with either Ω or UPD RNA (Fig. 12C). A poly(CAA)$_n$ RNA reduced complex formation of either Ω or UPD by approximately 50%. The poly(CAA)$_n$ constitutes part of the core element within Ω responsible for the enhancement associated with this leader. The mutual competition between Ω and the UPD suggests that both RNAs are specifically recognized by the same protein factor(s) present in both carrot and wheat germ extract. However, as Ω is a more efficient competitor than the UPD RNA, the interaction of the binding proteins with Ω appears to be tighter than that with the UPD.

DISCUSSION

In this study, we have examined the role of the 3′ UTR in regulating translational efficiency. The 72-base UPD, present within the 205-base 3′ UTR of TMV, had previously been identified as responsible for increasing the translational efficiency of chimeric reporter mRNAs (8, 12). We demonstrated here that the TMV 3′ UTR functioned in vivo but not in vitro, possibly because in vitro lysates lack an essential component. Alternatively, if the TMV 3′ UTR is involved in reinitiation, a low rate of reinitiation in wheat germ lysate may account for the poor activity of the TMV 3′ UTR in vitro. A similar difference in regulatory activity for the poly(A) tail has been observed in vivo and in vitro (8).

Using gel shift analysis, we have demonstrated that Ω and the UPD are specifically recognized by protein(s) from wheat germ and carrot extracts. Ω and the UPD serve as mutual competitors in the in vitro binding assay, suggesting that they may bind the same protein(s). This conclusion is supported by the fact that poly(C) or RNA of random sequence does not bind the protein(s). The gel shift data are also supported by the synergistic effect that Ω and the TMV 3′ UTR have on the translation of *luc* mRNA in vivo. The fact that the combined effect of these two regulatory elements was greater than the multiplication of their individual contributions suggests that Ω and the UPD interact in a positive manner to enhance the efficiency of translation. As there is no obvious sequence similarity between Ω and the UPD and they differ considerably to the extent in which secondary structure is present, binding of each RNA element may not occur at the same site within a single protein and may involve different subunits of a larger complex. Initial characterization of the protein species by UV cross-

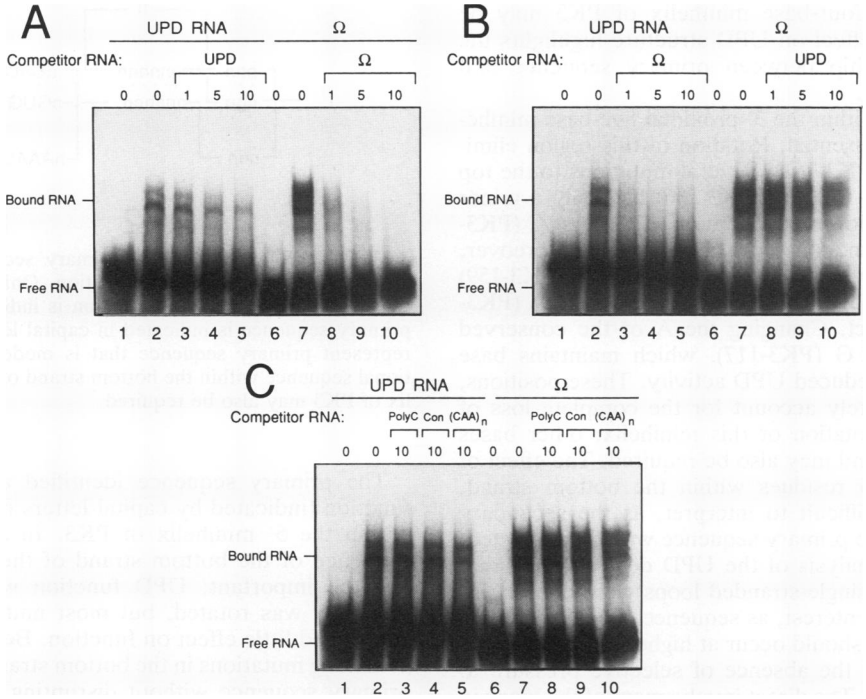


FIG. 12. Analysis of the binding activities present in extract made from carrot cell suspension that specifically recognize Ω and the UPD. (A) Radiolabeled UPD or Ω RNA is competed against by cold UPD or Ω RNA, respectively. (B) UPD and Ω act as mutual competitors. (C) The UPD and Ω binding activities are not affected by competition with poly(C) or the control (Con) random sequence. A poly(CAA)$_n$ RNA was also used as a competitor in lanes 5 and 10. In all panels, radiolabeled UPD and Ω RNA were incubated in the absence (lanes 1 and 6, respectively) or presence (lanes 2 to 5 and 7 to 10, respectively) of extract. Competitor RNA was added in a molar ratio indicated above the appropriate lane.

linking has identified three proteins that cross-link to both Ω and the UPD (39a). It is interesting to note that although Ω functions in wheat germ lysate, the UPD does not, yet both are capable of forming retardation complexes. Similar observations have been made with the poly(A) tail in that it also does not function in vitro (8) yet is capable of binding the poly(A)-binding protein. The binding of specific factors, therefore, may not be sufficient for function for these 3′ regulatory elements. Moreover, the additional retardation complex present when carrot extract was used may represent an activity that is necessary for UPD function. Whether any of these binding activities are important for either Ω or UPD function remains to be determined.

A phylogenetic analysis of the UPD from seven viruses as well as a satellite virus generated a consensus sequence in which several positions within the primary sequence of PK2 and PK3 are absolutely conserved. Deletion analysis of the UPD pseudoknots demonstrated that PK2 and PK3 were sufficient and necessary for UPD-mediated regulation. The complete removal of PK2 reduced UPD function to approximately 22% of the wild-type level, whereas deletion of PK3 eliminated UPD function altogether.

Changes within the primary sequence of PK2 that disrupted either the secondary or tertiary structure virtually eliminated the contribution made by this pseudoknot to UPD function. Compensatory base mutations that restore the higher-orde: structure largely restored PK2 function. Mutations that altered the primary sequence but did not disrupt the secondary structure had only a small effect on the regulation. We conclude that the higher-order structure is more essential than is the primary sequence of PK2 for UPD function. However, a minor role for some bases within PK2 may well exist. Mutation of the G in the C · G base pair within the 5′-proximal three-base minihelix had a larger effect on PK2 function than did any other mutation within this pseudoknot. As the functional activity of the compensatory base change mutant was not fully restored, a G at this position may play a small role.

In contrast to PK2, primary sequence within PK3 was essential. The positions identified correspond in large part to the phylogenetically conserved bases in both the single and doubled regions of the pseudoknot. Rotation of the 5′-proximal four-base minihelix in PK3 altered the primary sequence but maintained base pairing. Nevertheless, UPD function was virtually eliminated. That the mutation had no impact on UPD structure suggests that primary sequence within this region is essential. Changes to either base of the nonconserved U · A base pair (PK3-131 and PK3-158) within this same region had no more than a moderate effect, whereas changes to the 5′-CAC-3′ sequence (PK3-123, PK3-152, PK3-153, and PK3-154) significantly impaired UPD function. A mutation as small as two nucleotides that altered the primary sequence but not the base pairing (PK3-153) resulted in the virtual elimination of UPD regulatory function. Conclusions about the relative importance of each base within the 5′-CAC-3′ sequence are somewhat complicated by the effects that these mutations had on UPD structure. Although base pairing was maintained (PK3-152, PK3-153, and PK3-154), some reduction in the stability of PK2, PK3, or the 5′-proximal pseudoknot of the tRNA-like domain was observed. The actual base composition within the loops and stacked regions of a pseudoknot can influence stability (5,

35). That the entire four-base minihelix of PK3 may be rotated 180° without effect on UPD structure highlights the complicated relationship between primary sequence and pseudoknot formation.

Primary sequence within the 3'-proximal five-base minihelix of PK3 was also essential. Rotation of this region eliminated UPD function (PK3-161). Several mutations to the top strand (PK3-151, PK3-155, and PK3-156) had only a minor effect on regulation. However, mutation of the first C (PK3-157) reduced UPD function more significantly. Moreover, mutation of the conserved A · U base pair to G · C (PK3-159) also reduced UPD function. Removal of the bulged A (PK3-112) had a small effect. Changing the A of the conserved U · A base pair to a G (PK3-117), which maintains base pairing, significantly reduced UPD activity. These positions, alone, do not completely account for the complete loss of regulation following rotation of this minihelix; other bases within the bottom strand may also be required. The effect of altering any of the G residues within the bottom strand, however, would be difficult to interpret, as the secondary structure as well as the primary sequence would be affected.

The phylogenetic analysis of the UPD demonstrated that sequences within the single-stranded loops were conserved. This was of particular interest, as sequence changes within a single-stranded region should occur at higher rate than in the base-paired regions in the absence of selective pressure to maintain the sequence. The direct involvement of the loops in UPD function was one possibility. This hypothesis was partially supported in the case of the sequence 5'-AAAU-3' within PK3. The effect of mutations within this sequence (PK3-142 and PK3-122) suggested a role for this loop sequence that, if not as essential as primary sequence within the base-paired regions of PK3, nevertheless made a contribution to UPD function as important as that made by PK2. No change in the higher-order structure was observed for mutations confined solely within this loop. Although conserved sequence is present within the remaining three loops, mutations in these loops were minor when made individually. Only when mutations were made in several loops was a significant impact on UPD function seen. As UPD structure was increasingly compromised, we conclude that the sequence of three of the loops is important but only as far as pseudoknot formation is concerned. In this regard, it is important to note that all mutations made in PK2 and PK3 reduced UPD function to some extent, although many mutations that did not disrupt the higher-order structure had only a small effect. Nevertheless, these results may indicate that most of the primary sequence present in both the single- and double-stranded regions is optimal for UPD structure and function.

To summarize the mutational analysis of the UPD, Fig. 13 illustrates the sequence identified as important for UPD function. We conclude that the higher-order structure of the two pseudoknots is required. Virtually no primary sequence within PK2 was required for UPD function; however, disruption of the higher-order structure did affect UPD activity. This may be interpreted in two ways: the higher-order structure of PK2 plays a direct role in the mechanism by which the UPD functions, or it simply serves to maintain the stability of PK3. Either the disruption of PK2 or the separation of PK2 and PK3 would result in the loss of the coaxial stacking between the two pseudoknots and therefore reduce the stability of PK3. Disruption of the coaxial stacking on both sides of PK3 (UPD-111; Fig. 6) virtually eliminated UPD function, further supporting the idea that the pseudoknots on either side of PK3 serve to maintain PK3 structure.

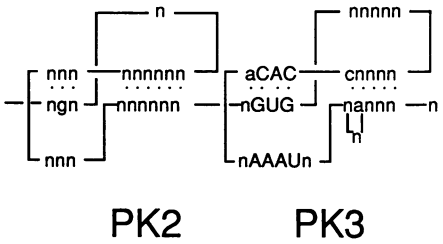


FIG. 13. Summary of the primary sequence and higher-order structure required for UPD function. Only the primary sequence that is important for UPD function is indicated. The most critical primary sequence is indicated in capital letters. Lowercase letters represent primary sequence that is moderately important. Additional sequence within the bottom strand of the 3'-proximal minihelix of PK3 may also be required.

The primary sequence identified as essential for UPD function (indicated by capital letters in Fig. 13) is clustered within the 5' minihelix of PK3. In addition, the primary sequence of the bottom strand of the 3' minihelix of PK3 may be important: UPD function was lost when the 3' minihelix was rotated, but most mutations within the top strand had little effect on function. Because of the difficulty in making mutations in the bottom strand that would alter the primary sequence without disrupting the base pairing, the possibility remains that the primary sequence in the bottom strand is required for UPD function. Other positions in which the primary sequence has a moderate impact on UPD function (lowercase letters in Fig. 13) are grouped about the essential core region. Decisions on which bases should be included in this second group of less important primary sequence are complicated by their moderate impact. Although the primary sequence of other positions within both PK2 and PK3 had a minor effect on both UPD function and structure (e.g., PK2-121 and PK3-112), it is not clear to what extent minor changes in the stability of the UPD are responsible for the minor changes in UPD function.

Interestingly, the sequence identified by this analysis is found in pseudoknots in the 3' UTRs of a diverse group of viral RNAs. In addition to the six tobamoviruses, most or all of these sequences are found in pseudoknots in a satellite virus, a hordeivirus (barley stripe mosaic virus [17–19]), and tobraviruses (tobacco rattle virus [1, 6]) as well as being present in as many as three copies in odontoglossum ringspot viral RNA (6a). These pseudoknots have evolved and have been maintained in viral evolution as an alternative to poly(A)-dependent translation. The elucidation of the mechanism underlying UPD function and the proteins that mediate the regulation will provide an opportunity to explore the versatility of the eukaryotic translational machinery.

ACKNOWLEDGMENT

This work was supported by grant NRICGP 91-37301-6459 from the U.S. Department of Agriculture.

REFERENCES

1. **Angenent, G. C., H. J. M. Linthorst, A. F. van Belkum, B. J. C. Cornelissen, and J. F. Bol.** 1986. RNA 2 of tobacco rattle virus strain TCM encodes an unexpected gene. Nucleic Acids Res. **14**:4673–4682.
2. **Atwater, J. A., R. Wisdom, and I. M. Verma.** 1990. Regulated mRNA stability. Annu. Rev. Genet. **24**:519–541.
3. **Avila-Rincon, M. J., M. L. Ferrero, E. Alonso, I. Garcia-Luque,**

and J. R. Diaz-Ruiz. 1989. Nucleotide sequences of 5′ and 3′ non-coding regions of pepper mild mottle virus strain S RNA. J. Gen. Virol. **70**:3025–3031.
4. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72**:248–254.
5. **Cheong, C., G. Varani, and I. Tinoco, Jr.** 1990. Solution structure of an unusually stable RNA hairpin, 5′GGAC(UUCG)GUCC. Nature (London) **346**:680–682.
6. **Cornelissen, B. J. C., H. J. M. Linthorst, F. T. Brederode, and J. F. Bol.** 1986. Analysis of the genome structure of tobacco rattle virus strain PSG. Nucleic Acids Res. **14**:2157–2169.

6a. **Dawson, W.** Personal communication.

7. **de Wet, J. R., K. V. Wood, M. DeLuca, D. R. Helinski, and S. Subramani.** 1987. Firefly luciferase gene: structure and expression in mammalian cells. Mol. Cell. Biol. **7**:725–737.
8. **Gallie, D. R.** 1991. The cap and poly(A) tail function synergistically to regulate mRNA translational efficiency. Genes Dev. **5**:2108–2116.

8a. **Gallie, D. R.** Unpublished data.

9. **Gallie, D. R., N. Feder, R. T. Schimke, and V. Walbot.** 1991. Post-transcriptional regulation in higher eukaryotes: the role of the reporter gene in controlling expression. Mol. Gen. Genet. **228**:258–264.
10. **Gallie, D. R., W. J. Lucas, and V. Walbot.** 1989. Visualizing mRNA expression in plant protoplasts: factors influencing efficient mRNA uptake and translation. Plant Cell **1**:301–311.
11. **Gallie, D. R., D. E. Sleat, J. W. Watts, P. C. Turner, and T. M. A. Wilson.** 1987. The 5′-leader sequence of tobacco mosaic virus RNA enhances the expression of foreign gene transcripts *in vitro* and *in vivo*. Nucleic Acids Res. **15**:3257–3273.
12. **Gallie, D. R., and V. Walbot.** 1990. RNA pseudoknot domain of tobacco mosaic virus can functionally substitute for a poly(A) tail in plant and animal cells. Genes Dev. **4**:1149–1157.
13. **Gallie, D. R., and V. Walbot.** 1992. Identification of the motifs within the tobacco mosaic virus 5′-leader responsible for enhancing translation. Nucleic Acids Res. **20**:4631–4638.
14. **Gallie, D. R., V. Walbot, and J. W. B. Hershey.** 1988. The ribosomal fraction mediates the translational enhancement associated with the 5′-leader of tobacco mosaic virus. Nucleic Acids Res. **16**:8675–8694.
15. **Garcia-Arenal, F.** 1988. Sequence and structure at the genome 3′ end of the U2-strain of tobacco mosaic virus, a histidine-accepting tobamovirus. Virology **167**:201–206.
16. **Goelet, P., G. P. Lomonossoff, P. J. G. Butler, M. E. Akam, M. J. Gait, and J. Karn.** 1982. Nucleotide sequence of tobacco mosaic virus RNA. Proc. Natl. Acad. Sci. USA **79**:5818–5822.
17. **Gustafson, G., and S. L. Armour.** 1986. The complete nucleotide sequence of RNAβ from the type strain of barley stripe mosaic virus. Nucleic Acids Res. **14**:3895–3909.
18. **Gustafson, G., S. L. Armour, G. C. Gamboa, S. G. Burgett, and J. W. Shepherd.** 1989. Nucleotide sequence of barley stripe mosaic virus RNAα: RNAα encodes a single polypeptide with homology to corresponding proteins from other viruses. Virology **170**:370–377.
19. **Gustafson, G., B. Hunter, R. Hanau, S. L. Armour, and A. O. Jackson.** 1987. Nucleotide sequence and genetic organization of barley stripe mosaic virus RNAγ. Virology **158**:394–406.
20. **Hall, T. C.** 1979. Transfer RNA-like structures in viral genomes. Int. Rev. Cytol. **60**:1–26.
21. **Haenni, A.-L., S. Joshi, and F. Chapeville.** 1982. tRNA-like structures in the genomes of RNA viruses. Prog. Nucleic Acid Res. Mol. Biol. **27**:85–104.
22. **Harris, M. E., R. Bohni, M. H. Schneiderman, L. Ramamurthy, D. Schumperli, and W. F. Marzluff.** 1991. Regulation of histone mRNA in the unperturbed cell cycle: evidence suggesting control at two posttranscriptional steps. Mol. Cell. Biol. **11**: 2416–2424.
23. **Hentschel, C. C., and M. L. Birnsteil.** 1981. The organization and expression of histone gene families. Cell **25**:301–313.
24. **Kepes, A.** 1963. Kinetics of induced enzyme synthesis determination of the mean life of galactosidase-specific messenger RNA. Biochim. Biophys. Acta **76**:293–309.
25. **Kunkel, T. A., J. D. Roberts, and R. A. Zakour.** 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods Enzymol. **154**:367–382.
26. **Melton, D. A., P. A. Krieg, M. R. Rebagliati, T. Maniatis, K. Zinn, and M. R. Green.** 1984. Efficient *in vitro* synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucleic Acids Res. **12**:7035–7056.
27. **Meshi, T., R. Kiyama, T. Ohno, and Y. Okada.** 1983. Nucleotide sequence of the coat protein cistron and the 3′ noncoding region of cucumber green mottle mosaic virus (watermelon strain) RNA. Virology **127**:54–64.
28. **Meshi, T., T. Ohno, H. Iba, and Y. Okada.** 1981. Nucleotide sequence of a cloned cDNA copy of TMV (cowpea strain) RNA, including the assembly origin, the coat protein cistron, and the 3′ noncoding region. Mol. Gen. Genet. **184**:20–25.
29. **Mirkov, T. E., D. M. Mathews, D. H. Du Plessis, and J. A. Dodds.** 1989. Nucleotide sequence and translation of satellite tobacco mosaic virus RNA. Virology **170**:139–146.
30. **Munroe, D., and A. Jacobson.** 1990. Tales of poly(A): a review. Gene **91**:151–158.
31. **Ohno, T., M. Aoyagi, Y. Yamanashi, H. Saito, S. Ikawa, T. Meshi, and Y. Okada.** 1984. Nucleotide sequence of the tobacco mosaic virus (tomato strain) genome and comparison with the common strain genome. J. Biochem. **96**:1915–1923.
32. **Pandey, N. B., and W. F. Marzluff.** 1987. The stem-loop structure at the 3′ end of histone mRNA is necessary and sufficient for regulation of histone mRNA stability. Mol. Cell. Biol. **7**:4557–4559.
33. **Pedersen, S., S. Reeh, and J. D. Friesen.** 1978. Functional mRNA half lives in *E. coli*. Mol. Gen. Genet. **166**:329–336.
34. **Pleij, C. W. A., J. P. Abrahams, A. van Belkum, K. Rietveld, and L. Bosch.** 1987. The spatial folding of the 3′ noncoding region of aminoacylatable plant viral RNAs, p. 299–316. *In* M. A. Brinton and R. R. Rueckert (ed.), Positive strand RNA viruses. Alan R. Liss, Inc. New York.
35. **Puglisi, J. D., J. R. Wyatt, and I. Tinoco.** 1990. Conformation of an RNA pseudoknot. J. Mol. Biol. **214**:437–453.
36. **Sachs, A. B., and R. W. Davis.** 1989. The poly(A) binding protein is required for poly(A) shortening and 60S ribosomal subunit-dependent translation initiation. Cell **58**:857–867.
37. **Sachs, A. B., and R. W. Davis.** 1990. Translation initiation and ribosomal biogenesis: involvement of a putative rRNA helicase and RPL46. Science **247**:1077–1079.
38. **Sachs, A. B., R. W. Davis, and R. D. Kornberg.** 1987. A single domain of yeast poly(A)-binding protein is necessary and sufficient for RNA binding and cell viability. Mol. Cell. Biol. **7**:3268–3276.
39. **Solis, I., and F. Garcia-Arenal.** 1990. The complete nucleotide sequence of the genomic RNA of the tobamovirus tobacco mild green mosaic virus. Virology **177**:553–558.

39a. **Tanguay, R., and D. R. Gallie.** Unpublished data.

40. **van Belkum, A., J. P. Abrahams, C. W. A. Pleij, and L. Bosch.** 1985. Five pseudoknots are present at the 204 nucleotides long 3′ noncoding region of tobacco mosaic virus RNA. Nucleic Acids Res. **13**:7673–7686.
41. **Wengler, G., and E. Castle.** 1986. Analysis of structural properties which possibly are characteristic for the 3′-terminal sequence of the genome RNA of flaviviruses. J. Gen. Virol. **67**:1183–1188.
42. **Yisraeli, J. K., and D. A. Melton.** 1989. Synthesis of long, capped transcripts *in vitro* by SP6 and T7 RNA polymerases. Methods Enzymol. **180**:42–50.