Cap-independent Translation of Tobacco Etch Virus Is Conferred by an RNA Pseudoknot in the 5′-Leader

Vladimir Zeenko and Daniel R. Gallie†‡
From the Department of Biochemistry, University of California, Riverside, California 92521-0129

The tobacco etch virus (TEV) 5′-leader promotes cap-independent translation in a 5′-proximal position and promotes internal initiation when present in the intercistronic region of a dicistronic mRNA, indicating that the leader contains an internal ribosome entry site. The TEV 143-nucleotide 5′-leader folds into a structure that contains two domains, each of which contains an RNA pseudoknot. Mutational analysis of the TEV 5′-leader identified pseudoknot (PK) 1 within the 5′-proximal domain and an upstream single-stranded region flanking PK1 as necessary to promote cap-independent translation. Mutations to either stem or to loops 2 or 3 of PK1 substantially disrupted cap-independent translation. The sequence of loop 3 in PK1 is complementary to a region in 18 S rRNA that is conserved throughout eukaryotes. Mutations within L5 that disrupted its potential base pairing with 18 S rRNA reduced cap-independent translation, whereas mutations that maintained the potential for base pairing with 18 S rRNA had little effect. These results indicated that the TEV 5′-leader functionally substitutes for a 5′-cap and promotes cap-independent translation through a 45-nucleotide pseudoknot-containing domain.

Virtually all eukaryotic cellular mRNAs contain a 5′-cap structure (m7G(5′)ppp(5′)N) that, during translation, serves as the binding site for the eukaryotic initiation factor (eIF)4E, a subunit of eIF4F that also contains eIF4A and eIF4G (1, 2). eIF4G promotes binding of 40 S ribosomal subunits through the interaction of eIF4G and eIF3, the latter of which is associated with the 40 S ribosomal subunit. The interaction between eIF4G and eIF3 directs 40 S subunit binding at or close to the 5′ terminus. Following binding, the 40 S subunit scans down the RNA in search of the first AUG codon in a good context at which translation will initiate. The rate of 40 S subunit binding and initiation can be slowed or prevented altogether by the presence of a highly stable secondary structure (3, 4).

Because binding of eIF4E is critical for the subsequent binding of the 40 S subunit and assembly of a competent initiation complex, those mRNAs that naturally lack a 5′-cap must have evolved alternative mechanisms to promote translation. The only known mRNAs that naturally lack a cap are viral in origin. Poliovirus, encephalomyocarditis virus (EMCV), and foot-and-mouth disease virus are examples of animal picornaviruses whose polyadenylated genomic RNA lacks a 5′-cap structure and instead contains a VPg, i.e. a viral protein genome linked to the 5′-terminus that is removed prior to RNA recruitment into polysomes (5–9). The 5′-leader of picornaviruses is long (610 to >1200 nucleotides), highly structured, and contains multiple AUGs upstream of the initiation codon of the polyprotein coding region (10), features that should inhibit normal 40 S ribosomal scanning. Despite these apparent barriers to translation, picornaviral mRNAs are efficiently translated as a consequence of its 5′-leader, which contains an internal ribosome entry site (IRES) to which the 40 S ribosomal subunit is recruited (11–13). For EMCV, the IRES recruits eIF4G which in turn promotes internal binding of the 40 S ribosomal subunit through its interaction with the 40 S subunit-bound eIF3 (14, 15). In addition to canonical initiation factors, several cellular RNA-binding proteins that stimulate viral IRES activity have been identified. Polyuridylic acid–binding protein, poly(rC)-binding protein-2, and unr have been shown to stimulate IRES activity for some but not all IRESs perhaps by acting as RNA chaperones that maintain IRES structure required for 40 S subunit binding (10, 16, 17).

Tobacco etch virus (TEV) is a member of the picornavirus supergroup of positive-strand RNA viruses but uses plants as its host. Its genomic organization is highly similar to picornaviruses in that the genomic RNA functions as a monocistronic mRNA encoding a single polyprotein that, once produced, is processed by viral-encoded proteases into capsid and noncapsid proteins required for the viral life cycle (18, 19). The genomic RNA of TEV is polyadenylated and naturally lacks a 5′-cap but is nevertheless efficiently translated. The 143-nt TEV 5′-leader is sufficient to confer cap-independent translation even in the absence of the 5′-terminal VPg (20, 21). Two regions within the 143-nt TEV 5′-leader were identified previously as necessary to direct cap-independent translation, and their combinatorial effect was approximately multiplicative, suggesting that the two elements are functional parts of a single regulatory region (22). The TEV 5′-leader functionally interacts with the pol(A) tail to promote optimal cap-independent translation and requires eIF4G and poly(A)-binding protein (22, 23). Although the TEV 5′-leader evolved to function in a 5′-proximal position in a monocistronic mRNA, it also promoted translation of a second cistron in a dicistronic mRNA when the TEV leader sequence was introduced into the intercistronic region, indicating that the TEV 5′-leader contains an IRES (22). Although IRES elements from picornaviral RNAs are typically large and highly structured, the TEV element is notable for its small size, representing one of the most compact viral elements identified that can promote cap-independent translation. The precise sequences in the TEV 5′-leader required to promote cap-independent translation have not been identified. Moreover, the struc-
ture of the TEV 5'-leader and its relationship with the function of the 5'-leader in promoting cap-independent translation is unknown. In this study, we show that a 5'-proximal, 45-nt RNA pseudoknot-containing domain within the TEV 5'-leader is required to promote cap-independent translation. Mutations that disrupted the pseudoknot reduced cap-independent translation, including mutations to loop 3 that exhibits complementarity to a conserved region in eukaryotic 18 S rRNAs. Changes to the loop that maintained its potential for base pairing with 18 S rRNA had only a small effect, supporting the possibility that base pairing between the pseudoknot and the 18 S rRNA may be involved in promoting cap-independent translation.

EXPERIMENTAL PROCEDURES

Secondary Structure Analysis—The TEV 5'-leader sequence used in the experiments represents the 5'-proximal 143-nt of the TEV genome. The predicted secondary structure for the TEV 5'-leader was calculated through minimization of free energy using the Fold program of the GCG package (Genetics Computer Group, Inc., version 7, University of Wisconsin (24)). Prediction of ternary RNA structure employed the parNAs and pknotesRG programs (25), and RNA structures were visualized using PSEUDOVIEWER (26). The experimental data were used to correct the predicted models.

RNA fragments were transcribed in vitro from pT7 constructs containing the appropriate TEV sequence. RNAs were 5'-end-labeled with [γ-32P]ATP, resuspended into probing buffer, and heated rapidly. RNA folding was promoted by slow renaturation under native conditions (i.e. in the presence of 2 mM MgCl₂) in order to refold the molecule into its native conformation and stabilize higher ordered structures. The labeled RNA was subject to partial digestion with enzymic probes in the appropriate buffer, and the RNA fragments were resolved on a denaturing TBE (TrisBorate/EDTA)urea 10% polyacrylamide gel. A ladder generated by alkaline hydrolysis of RNA was run on the same gel to help identify the cleavage site locations.

mRNA Constructs and in Vitro RNA Synthesis—TEV-luc-A₁₄₉ and Con₁₄₄-luc-A₁₄₉ luciferase constructs that contain the 143-nt TEV leader sequence or a 144-nt control sequence and that terminate in a poly(A)₉₀ tract have been described previously (22). The control construct was designed to be 60% AT-rich and nearly the same length as the TEV 5'-leader. The free energy calculated by the FOLD algorithm for the control sequence leader is ΔG = −11.5 kcal/mol, which is approximately equal to the free energy of ΔG = −10.7 kcal/mol of the 5'-leader of the TEV-luc-A₁₄₉ mRNA construct (27).

Following linearization downstream of the poly(A)₉₀ tract, the DNA concentration was quantitated spectrophotometrically and brought to 0.5 mg/ml. Uncapped mRNAs were synthesized in vitro as described previously (28) by using 10 μg of template DNA in 40 mM Tris-HCl, pH 7.5, 6 mM MgCl₂, 100 μg/ml bovine serum albumin, 0.5 μg/ml each of ATP, CTP, UTP, and GTP, 10 mM dithiothreitol, 0.3 units/μl RNasin (Promega), and 0.5 units/μl T7 RNA polymerase. mRNA Delivery to Plant Protoplasts—Protoplasts were isolated from carrot suspension cells used previously in the analysis of the translational regulatory function of viral leaders (21, 29) by digesting with 0.25% CLEF cellulase, 1% cytolase, 0.05% pectolyase Y23, 0.5% bovine serum albumin, and 7 mM β-mercaptoethanol in protoplast isolation buffer (12 mM sodium acetate, pH 5.8, 50 mM CaCl₂, 0.25 mM mannitol) for 90–120 min. Protoplasts were washed with protoplast isolation buffer followed by electroporation buffer (10 mM Hepes, pH 7.2, 130 mM KCl, 10 mM NaCl, 4 mM CaCl₂, 0.2 mM mannitol) and resuspended in electroporation buffer to ~10⁶ cells/ml. Equal amounts of mRNAs (~2.5 μg) were added to 400 μl of cell suspension immediately before electroporation (250 microfarads, 300 V, 0.2-mm electrode) using an IBI Gene-Zapper. The electroporated cells were incubated in protoplast growth media (MS salts, pH 5.8, 30 g/liter sucrose, 100 μg/ml myoinositol, 0.1 mg/ml 2,4-dichlorophenoxyacetic acid, 1.3 mg/liter nicacin, 0.25 mg/liter thiamine, 0.25 mg/liter pyridoxine, 0.25 mg/liter calcium pentoate) supplemented with 20% of cultured medium (protoplast growth medium conditioned with carrot cells for 3 days) overnight prior to assaying for reporter gene activity. For each experiment, an mRNA was delivered to triplicate samples of protoplasts, and the average value of a typical experiment was reported.

RESULTS

Identification of Sequences within the TEV 5'-Leader Required for Cap-Independent Translation—Two regions of the TEV 5'-leader (i.e. nt 28–65 and nt 66–118) that functioned together to promote cap-independent translation were identified previously, suggesting that the region from nucleotides 28 to 118 was required (22). To identify more precisely the functional regions within the TEV 5'-leader involved in promoting cap-independent translation, a more detailed deletion analysis was performed. The full-length TEV 5'-leader and deletion derivatives were introduced upstream of the luc coding region in a T7 promoter-based construct that enabled in vitro synthesis of uncap mRNA terminating in a poly(A)₉₀ tail. The full-length TEV 5'-leader (i.e. TEV-(1–143)) promoted cap-independent translation 285-fold relative to the uncap control luc mRNA containing a 5'-untranslated region of similar length and degree of secondary structure (i.e. Con₁₄₄) (Fig. 1), in good agreement with previous results (21, 22). The truncated TEV-(1–118) and TEV-(1–97) constructs retained 90% of the translational activity of the full-length leader, whereas TEV-
enzymatic and chemical probing data. The AUG at the 3' of the leader containing an alternative structure, similar to deleting the 5' of the full-length TEV leader construct, a level of reduction substantially reduced cap-independent translation to just 9% activity of the full-length TEV leader. Deletion of nt 54–60 of the activity of the full-length leader, whereas deletion of the 5'-proximal 28 nt (i.e. TEV-(28–143)) retained 90% of the translational activity of the full-length leader, demonstrating the importance of the region between nt 29 and 65. Because this analysis suggested the region between nt 29 and 85 was critical to promoting cap-independent translation, a series of deletions within this region was made. Deletion of nt 38–43 resulted in just 20% of the activity of the full-length TEV leader, whereas deletion of nt 47–51 reduced cap-independent translation to 49% of the activity of the full-length TEV leader. Deletion of nt 54–60 substantially reduced cap-independent translation to just 9% of the full-length TEV leader construct, a level of reduction similar to deleting the 5'-proximal 65 nt (i.e. TEV-(66–143)) for each pseudoknot are indicated. PK1 is an H-type pseudoknot with two stems, S1 (6 bp) and S2 (3 bp), that are connected by loops L1 (3 nt), L2 (2 nt), and L3 (7 nt) (Fig. 2A). PK2 is an H-type pseudoknot composed of S1 (5 bp) and S2 (5 bp), that are connected by loops L1 (3 nt), L2 (2 nt), and L3 (7 nt) (Fig. 2B). PK3 is also an H-type pseudoknot composed of S1 (5 bp) and S2 (3 bp) that are connected by loops L1 (2 nt) and L2 (1 nt). In this pseudoknot, S1 and S2 are not separated and thus may coaxially stack.

To determine whether these predicted structures are supported experimentally, the structure of the TEV 5'-leader was examined in solution. The 5'-end-labeled, in vitro-synthesized, 143-nt TEV 5'-leader RNA was heat-denatured and renatured in the presence of 2 mM MgCl₂ to promote folding of the structure. Double-stranded regions were identified following partial digestion with RNase T1, which preferentially cleaves between double-stranded regions RNA or stacked single-stranded regions were identified following partial digestion with RNase V1, which preferentially cleaves between double-stranded regions RNA or stacked single-stranded regions, generating 5'-phosphate fragments. Sin-
nucleotides nonspecifically (34, 35). The TEV 5'-leader was also probed with kethoxal that modifies unpaired G residues (34).

Cleavage of the 5'-proximal 37 nucleotides by RNase A, RNase T2, and nuclease S1 (Figs. 3 and 4) indicated that this region is highly accessible and unstructured, an observation consistent with the predicted structure. Some residues between 7 and 13 nt were less accessible to cleavage by these enzymatic probes, suggesting possible involvement in structure.

RNase V1 cleavage occurred at G38, A39, C42, U58, and C60 (Figs. 3 and 4) supporting the formation of S1 in PK1. RNase T2 cleaved at the bulged C (i.e. C56), indicating that this C was indeed accessible. A57, which lies immediately 3' to the bulged C, was cleaved to a limited extent by RNase T2, indicating some local breathing. G38, which is predicted to be part of the first base pair at the base of the SL1 stem, was accessible as indicated by probing with RNase T1, but G38 and C60 were also cleaved by RNase V1. The accessibility of G38 may be explained by breathing at this end of SL1 where it is cleaved by RNase V1 when it is base-paired with C60 but cleaved by RNase T1 when it is not. G55 on left side of the S1 was largely resistant to RNase T1 cleavage (Figs. 3 and 4), consistent with its internal position in S1.

G47 was largely resistant to cleavage by RNase T1 as were C48 and U51 to RNase A cleavage (Figs. 3 and 4) and C48, A49, A50, and U51 to cleavage by RNase T2 and nuclease S1. The low accessibility G47, C48, A49, A50, and U51 is inconsistent with their presence in a single-stranded loop, which would be expected to be accessible to single-stranded specific RNases and nucleases. The inaccessibility of this region is consistent with cleavage at C44, A45, C48, A49, A50, C52, and A53 by RNase V1 and suggests that the loop is structured or involved in base pairing. A sequence within this region (i.e. \text{GCAAU}^{51}) is complementary to a region downstream of S1 (i.e. \text{AUUGC}^{72}), raising the possibility of base pairing between these two regions to form a pseudoknot. Cleavage at A68, U69, and U70 by RNase V1 supports the putative base pair between these regions and formation of S2 in PK1. The sequence representing L3 (\text{UACUUCUU}^{70}) also exhibited little accessibility to single-stranded specific enzymatic probes (Figs. 3 and 4), suggesting that L3 is sequestered in such a way that is not susceptible to cleavage. These data are consistent with the prediction of an H-type pseudoknot between nucleotides 38 and 72.

The region separating PK1 from the downstream structures (i.e. nucleotides 73–77) showed limited accessibility to RNase A (C75), RNase T1 (G74), RNase T2 (A76 and A77), and nuclease S1 (G74), suggesting that this region is somewhat accessible, but limited cleavage by RNase V1 at A76 and A77 (Figs. 3 and 4) indicated that the region could be structured.

The cleavage data supported the presence of either a pseudoknot or two stem-loops in the region from nt 78 to 104. Strong cleavage by RNase A at C85 and RNase T2 at A86 was observed (Figs. 3D and 4), suggesting that this region is highly accessible and may form loop L1 in PK2 or the loop in SL1 (Fig. 2B). Significant cleavage by RNase V1 was observed at U80 as it was at U91–U93, supporting the presence of S1 in PK2 (Figs. 3D and 4). Cleavage by RNase T2 at A95 and A97 (Figs. 3D and 4) suggests that this region is unstructured, which supports the presence of the loop in SL2 (Fig. 2B). Strong cleavage by RNase T1 at G98 is consistent with this base being present in L3 of PK2 or in the loop of SL2. Cleavage by RNase V1 was observed at A101 and supports the presence of the stem in SL2. Therefore, these data support the formation of PK2 or SL1/SL2 which may be in equilibrium.

Cleavage by RNase V1 was observed at A106, A107, U108, U109, U110, A117, and U120 (Figs. 3D and 4), supporting the

![Figure 3: Structural analysis of the TEV 5'-leader.](Image)
The enzymatic and chemical probing data of the TEV 5'-leader from Fig. 3 are summarized. Bases that were highly or moderately accessible to the probes are indicated by large or small symbols, respectively. Unmarked bases represent those that were not accessible. Tertiary base pairing is indicated by the gray lines.

FIG. 4. Summary of the enzymatic and chemical probing data of the TEV 5'-leader. The enzymatic and chemical probing data of the TEV 5'-leader is shown. The regions representing PK1, SL1/SL2, and PK3 and the stems and loops within each are indicated.

Because PK1 is present in the region between nt 28 and 75 identified in Fig. 1 as functionally critical in promoting cap-independent translation, the involvement of tertiary base pairing within PK1 was investigated further. The structure of TEV-(1–85), in which the entire PK1 is present, and TEV-(1–65), in which the absence of nt 66–75 abolishes the potential for tertiary base pairing in PK1, were compared with the full-length TEV-(1–143). Probing TEV-(1–143) with RNase T1 in 5 M urea, which disrupts base pairing within the structure thus rendering the RNA single-stranded, resulted in increased cleavage at G47, G55, and particularly at G71 (Fig. 6A, compare lane $T_1^*$ with $T_1$), supporting the notion that the guanosines at each of these positions are involved in base pairing. In contrast, no increase in cleavage was observed at G74, suggesting that it is not involved in base pairing, consistent with its position just downstream of PK1. No increase in cleavage was observed at G38 (Fig. 6A, compare lane $T_1^*$ with $T_1$), which in the wild-type structure is already subject to

FIG. 5. The effect of Mg$^{2+}$ on the stability of structures within the TEV 5'-leader. Representative PAGE autoradiogram of the full-length TEV 5'-leader following partial enzymatic digestion with RNase T1 in the presence of EDTA (E) or in the presence of increasing amounts of Mg$^{2+}$. Arrows indicate cleavage at G residues by RNase T1. The regions representing PK1, SL1/SL2, and PK3 and the stems and loops within each are indicated. Lane C, undigested RNA included as a control. Partially hydrolyzed RNA (lane L) served as an RNA ladder to identify base positions.
extensive cleavage perhaps due to breathing at the base of S1 in PK1. At G98, a decrease in cleavage was observed when probed with RNase T1 in urea, consistent with the notion that this guanosine is present in a highly exposed position within PK2 (or SL2). An increase in cleavage was observed at G104 when probed in urea, consistent with its involvement in a base pair in S2 of PK2 or the stem of SL2. No increase in cleavage was observed at G114 (Fig. 6A, compare lane T1* with T1), consistent with the possibility that the tertiary base pairing in PK3 may be weak.

Similar increases in accessibility were observed for G47, G55, and G71 when TEV-(1–85) was probed with RNase T1 in urea with a small increase at G38 observed as well. The observed increase in cleavage at G74 when probed with RNase T1 in urea may be a result of the termination of the RNA at nt 85, which disrupts PK2 or SL1 and may allow ectopic base pairing to G74. A reduction in cleavage by RNase V1 was observed at A68, U69, and U70 in TEV-(1–85) compared with the extent of cleavage in TEV-(1–143). As cleavage at these positions in TEV-(1–143) supported the presence of tertiary interactions in PK1, the decreased cleavage in TEV-(1–85) may indicate some reduction of stability of PK1. In contrast to TEV-(1–143) and TEV-(1–85), when TEV-(1–65) was probed with RNase T1 in urea, no increase in cleavage was observed at G47 and only a slight increase in cleavage was observed at G55 (Fig. 6A, compare lane T1* with T1), consistent with the notion that, in the absence of the distal side of S2 in PK1, G47 is no longer base-paired, and the stability of the C42-G55 base pair is reduced.

Direct comparison of cleavage at G38, G47, and G55 by RNase T1 under conditions to promote formation of structure (i.e. in the absence of urea and presence of Mg$^{2+}$) revealed substantially more cleavage at G47 in TEV-(1–65) than in TEV-(1–143) or TEV-(1–85) (Fig. 6B) with little change in cleavage at G38 or G55, consistent with the involvement of G47 in base pairing in TEV-(1–143) or TEV-(1–85) but not in TEV-(1–65). A substantial increase in cleavage by nuclease S1 at G47 and a smaller increase in cleavage by nucleases S1 at G55 support the notion that G47 is no longer base-paired and a reduction in the stability of the C42-G55 base pair in TEV-(1–65). Several additional changes in accessibility of TEV-(1–65) were observed, most notably an increase in cleavage at A57,
U58, and U59 by RNase V1 and a decrease in cleavage by RNase T2 at these same positions relative to TEV-(1–143) (Fig. 6A), suggesting an increase in the stability of these base pairs of S1 in PK1 or that RNase V1 has greater access to S1 when not present in a pseudoknot.

Whether deletions within PK1 would affect the stability of the pseudoknot was then examined. Deletion of L3 resulted in decreased RNase T1 cleavage at G47 and G71, suggesting an increase in the stability of S2 (Fig. 7A). Note that deletion of L3 (i.e. nt 61–67) results in a 7-nucleotide shift in the gel at what was A68 in the wild-type so that it runs at nt 61 in the mutant. To make comparisons of specific nucleotides between mutants without confusion, the wild-type position designation is used for all mutants as well despite the shift in migration following deletion of sequence in the mutants. Thus, the position of G74 and G71 has shifted seven nucleotides, but their designation as G74 and G71 has been maintained in reference to the wild-type sequence. For ΔG71,72, the deletion results in loss of G71 and a shift in migration of G74 by two nucleotides.

The effect on the stability of PK1 that reversing the entire S2 (construct S2-3) or reversing just the two CG base pairs within S2 (construct S2-6), which would maintain base pairing within S2, was then examined (Fig. 7B). Structural probing confirmed that base pairing in S2 was maintained as revealed by the poor accessibility of G72 (Fig. 7B). The accessibility of G48 was also reduced in construct S2-3, suggesting that not only was base pairing maintained within S2 but that the stability of S2 was increased. In contrast, an increase in accessibility of S1 was observed by the increase in RNase T2 cleavage in S1 of S2-3 and RNase T2 and nuclease S1 cleavage of S1 in S2-6 (Fig. 7B), indicating a decrease in the stability of S1. These results support the conclusion that a pseudoknot is present between G38 and C72 and that the structure of the pseudoknot is maintained by its primary sequence and stability of the two stems.

PK1 Is Required to Confer Cap-independent Translation—

The results of the structural analysis provided a framework to understand the observation from the data in Fig. 1 that multiple discrete regions within the TEV 5’-leader were required to promote cap-independent translation. The sequence from nt 38 to 43 and from nt 54 to 60 that was identified as functionally important represents S1 of PK1, whereas nt 61–67 corresponds to L3. The sequence from nt 47 to 51 and from nt 67 to 75 that was also identified includes S2 of PK1.

In order to examine the primary sequence and structural requirements of PK1 and its flanking regions involved in promoting cap-independent translation, mutational analysis was performed, and the effect on cap-independent translation was examined. Deletion of the 10-nt flanking region (i.e. 25°CAAAACAAAC57) immediately upstream of PK1 reduced cap-independent translation to 31% of that provided by the wild-type TEV leader (Fig. 8), despite the fact that a similar
but not identical sequence (i.e. \(^{18}\text{ACACCAUAUA}^{27}\)) now flanked PK1 in the deletion mutant. The spacing between PK1 and the 5' terminus could be shortened without loss of function as deletion of the first 27 nt (i.e. TEV 5'-untranslated region nt 28–143) did not substantially reduce cap-independent translation (90% of the full-length TEV leader) (Fig. 1). Changing the native sequence of nt 28–36 to poly(U) reduced cap-independent translation to just 20% of that provided by the wild-type TEV leader (Fig. 8, CAA→UUU). Changing C28 and C33 to adenosines (i.e. \(^{28}\text{AAAAA}^{37}\)) reduced cap-independent translation to 66% of the wild-type leader (Fig. 8, C→A), whereas changing these same bases to guanosines (i.e. \(^{28}\text{GAAAG}^{37}\)) reduced cap-independent translation to 5.5% of the wild-type leader (Fig. 8, C→G). Substitution of this region with its complementary sequence (i.e. \(^{28}\text{GUUUUGUUUC}^{37}\)) resulted in a reduction in cap-independent translation to 3.5% of the wild-type leader (Fig. 8, CAA→GUU). These data suggest that the primary sequence of nt 28–37 and not just an unstructured region immediately upstream of PK1 or the spacing between PK1 and the 5' terminus was required for cap-independent translation.

Deletion of nt 38–43 reduced cap-independent translation to 20% of the full-length TEV leader (Fig. 1). However, as this represents the 5'-proximal side of S1 in PK1, its deletion may have altered the spatial relationship between the upstream flanking sequence and PK1. Therefore, to disrupt S1 while maintaining the upstream flanking region in its immediate proximal position, the 3'-half of S1 was substituted for the 5'-half of S1, which resulted in a reduction in cap-independent translation to 14% of the wild-type TEV leader (Fig. 9, S1-1). Substituting the 5'-half of S1 for the 3'-half of S1 resulted in a reduction in cap-independent translation to 27% of the wild-type TEV leader (Fig. 9, S1-2). Completely reversing S1 in PK1 such that the 5'-side of S1 became the 3'-side and the 3'-side became the 5'-side altered the primary sequence of S1, whereas it maintained its potential to base pair (Fig. 9, S1-3). Reversing S1 reduced cap-independent translation to 7.4% of the wild-type TEV leader, suggesting that restoring base pairing within S1 of PK1 was not sufficient to restore cap-independent translation.

Deletion of nt 67–75, which included deletion of one nucleotide from L3, the 3'-half of S2, and three nucleotides of the downstream flanking region reduced cap-independent translation to 40% of the full-length TEV leader (Fig. 1). To examine whether base pairing within S2 of PK1 was required, the 5'-half of S2 was substituted for the 3'-half of S2 which reduced cap-independent translation to 39% of the wild-type TEV leader (Fig. 9, S2-1). Substituting the 3'-half of S2 for the 5'-half of S2 also reduced cap-independent translation (18% of the wild-type TEV leader) (Fig. 9, S2-2). Completely reversing S2 in PK1 (construct S2-3) altered the primary sequence of S2 while maintaining its base pairing as shown in Fig. 7B. Reversing S2 reduced cap-independent translation to 34% of the wild-type TEV leader (Fig. 9, S2-3), suggesting that restoring base pairing within S2 of PK1 was not sufficient to restore cap-independent translation.

In a more conservative change to S2, changing G71 and C72 to the complementary nucleotides and thus disrupting base pairing with C48 and G47, respectively, reduced cap-independent translation to 14% of the wild-type TEV leader (Fig. 9, S2-4), whereas changing C48 and G47 to the complementary nucleotides reduced cap-independent translation to 17% of the wild-type TEV leader (Fig. 9, S2-5). Completely reversing these 2 bp (i.e. C48→G71 and G47→C72) restored base pairing as shown in Fig. 7B, and partial restoration of cap-independent translation was observed, although the level was still just 40% of the wild-type TEV leader (Fig. 9, S2-6). The lack of complete restoration of cap-independent translation may be due to partial destabilization of S1 as observed in Fig. 7B and/or due to a change in the primary sequence at these positions.

To examine further how mutations affecting tertiary base pairing within PK1 impacted cap-independent translation, additional mutations were made to S2 of PK1, and their effect on translation was examined. Deleting the last two nucleotides of the 3'-half of S2 (ΔGC71,72) reduced cap-independent translation to 24% of the wild-type level, whereas deletion of G71 or C72 reduced cap-independent translation to 33 or 10% of the wild-type level, respectively, and a G71C mutation reduced translation to 14% (Fig. 10A). Deletion of A49 and A50 in the 5'-half of S2 (ΔAA49,50) reduced translation to 12% of the wild-type level, whereas deletion of A49 alone only reduced translation to 48% (Fig. 10A). Deletion of G47 and C48 in this 5'-half of S2 (ΔGC47,48) reduced translation to 62% of the wild-type level, whereas deletion of G47 or C48 alone reduced translation to 51 or 37%, respectively (Fig. 10A). A C48U mutation, which should maintain base pairing with G71, reduced...
translation to 22%. These results demonstrate that changes to the 3'-half of S2 are more disruptive than most changes to the 5'-half and that alterations to S2 significantly affect cap-independent translation.

Deletion of L2 (ΔCA52–53) reduced cap-independent translation to 8% of the full-length TEV leader (Fig. 10B), whereas mutating L2 to AU reduced cap-independent translation to 13% of the full-length TEV leader (Fig. 10B, CA52,53AU), data suggesting that L2 was important for PK1 activity. Deletion of L1 (nt 44–46) reduced cap-independent translation only moderately (Fig. 10B, ΔCAAA44–46), whereas changing the sequence of L1 to ACU reduced cap-independent translation to 45% of the full-length TEV leader (Fig. 10B, CAAA44–46ACU), indicating that L1 contributed only moderately to PK1 activity.

Deletion of L3 in PK1 (nt 61–67) reduced cap-independent translation (Fig. 1), suggesting that L3 contributes to PK1 activity. Most interestingly, the sequence of loop 3 is complementary to a region in 18 S rRNA that is conserved throughout eukaryotes. Although it is not known whether L3 base pairs with 18 S rRNA, mutations were made in L3 that would either disrupt, maintain, or extend its potential base pairing with 18 S rRNA. The effect of each mutation on base pairing with 18 S rRNA is shown in Fig. 11. Mutating the last five nucleotides of the L3 sequence (i.e. 61UACUUCU67) to the complementary sequence (i.e. 61UACUCUG67) reduced cap-independent translation to 14% of the wild-type level (Fig. 11, ΔL3-5). Extending the potential for base pairing by changing the sequence of loop 3 to its complementary sequence (i.e. 61UACUCG67) increased cap-independent translation to 22%. These results demonstrate that changes to the 5’-end of L3 sequence that was complementary to the 18 S rRNA at either the 5’ or 3’ end of the L3 sequence reduced cap-independent translation to just 17 or 11.5%, respectively, of the wild-type leader (Fig. 11, 61UACUUCG67 and 61UACUCG67, respectively).

Sequence Distal to PK1 May Affect Its Function in Promoting Translation—To examine whether specific sequences downstream of PK1 are important, mutations were made that altered the primary structure and disrupted base pairing of the structured elements distal to PK1 and their effect on cap-independent translation examined. Mutating nt 84–91 (i.e. 84UCAUUUCU91) to the complementary sequence (i.e. 84AGUAAGAG91) to disrupt base pairing in SL1 (or the alternative pseudoknot structure) resulted in an increase in cap-independent translation to 119% of the wild-type level (Fig. 12). Changing nt 92–97 (i.e. 92UUUAAA97) to the complementary sequence (i.e. 92AAAUU97) to disrupt base pairing in SL2 or PK2 also increased cap-independent translation to 156% of the wild-type level, whereas mutating nt 100–103 (i.e. 100AAAA103) of SL2 to its complementary sequence (i.e. 100UUUAAA103) had only a small effect if any (Fig. 12). Altering nt 104–108 (i.e. 104GCAAAT108) to its complementary sequence (i.e. 104GUUA108), which might affect the stability of SL2 and disrupt S1 of PK3, increased cap-independent translation to 116% of the wild-type level (Fig. 12). Finally, changing nt 114–116 (i.e. 114GAA116) to its complementary sequence (i.e. 114CU116), which would disrupt S2 in PK3, increased cap-independent translation to 212% of the wild-type level (Fig. 12). These results indicate that changes to the primary sequence and disruption of the structural elements distal to PK1 do not reduce cap-independent translation and, in several cases, increase translation moderately.

**DISCUSSION**

Although a member of the picornavirus supergroup, the TEV 5'-leader differs considerably from those of picornaviruses in that it is considerably shorter in length, less structured, and
contains no upstream AUG triplets. Deletion analysis indicated that several disparate regions in the 5'-proximal 75 nucleotides were required for cap-independent translation and corresponded to PK1 and its upstream flanking region. Mutations in the leader outside these regions had little to no effect on cap-independent translation, demonstrating that the 5'-proximal pseudoknot is specifically required for mediating cap-independent translation. Disruption of S1 or S2 of PK1 resulted in a substantial loss of cap-independent translation, whereas reversing either S1 or S2 to restore base pairing did not restore translation. However, the deleterious effect that disrupting the distal two GC base pairs in S2 had on PK1 activity could be partially reversed when base pairing was restored, suggesting that the primary sequence at these positions was not absolutely essential. As some alterations in pseudoknot stability could be detected even when the base pairing in the stems was maintained, the primary sequence of PK1 may be optimal for the correct formation of the pseudoknot. Of the three loops in PK1, L1 was the least critical to PK1 activity, whereas alteration of L2 reduced cap-independent translation substantially as did some changes to L3. The observation that changes to L1 and some changes to L3 had little effect on cap-independent translation demonstrated that the importance of the PK1 primary sequence was confined to specific regions within the structure. The flanking sequence immediately upstream of PK1 was also required for cap-independent translation. This sequence shares similarity with the translational enhancer sequence identified in O, the 5'-leader of tobacco mosaic virus genomic mRNA (29, 40, 41).
Mutations within structures in the TEV leader were not always correlated with reduced expression as we observed that disrupting the structures downstream of PK1 did not impair expression, suggesting that unfolding these structures does not reduce mRNA stability consistent with previous work showing that deletion of the entire TEV leader did not affect mRNA stability (21). The observation that mutations affecting the region distal to PK1 resulted in increased expression suggests that the sequence downstream of PK1 may serve as a negative regulator. The presence of sequences that promote and inhibit IRES activity within the same leader has been demonstrated previously, e.g. in the cold stress-induced Rbm3 mRNA (42, 43). Whether the structured nature of the region downstream of PK1 serves to slow 40 S ribosomal subunit scanning will depend on whether 40 S subunits are recruited directly to the initiator AUG or scan through at least a portion of the leader. Whether the structured nature of the region downstream of PK1 serves to slow 40 S ribosomal subunit scanning will depend on whether 40 S subunits are recruited directly to the initiator AUG or scan through at least a portion of the leader. Deletion of nearly the entire region distal to PK1 (e.g. TEV-(1–85)) reduced cap-independent translation moderately, correlating with some reduction in the stability of PK1. The deletion also changed the spacing between PK1 and the initiator AUG, although replacing nt 66–143 with an unrelated sequence of the same length increased the activity of nt 1–65 only marginally relative to nt 1–65 without the spacer sequence (22). It is also possible that there is a requirement for other cis-acting elements such as polypyrimidine tracts that bind polypyrimidine tract-binding protein and stimulate IRES activity in some viral RNAs such as rhinovirus (17). Several polypyrimidine tracts are present downstream of PK1, and at least one remained in the deletions introduced downstream of PK1 (Fig. 12), whereas all would have been absent following deletion of the entire region distal to PK1.

The role of pseudoknots in facilitating translation initiation from mRNAs from eubacteria and eukaryotes has been reported. An RNA pseudoknot in hepatitis C and classical swine fever virus is required to promote 40 S ribosomal subunit binding to the IRES and thus is required for IRES-mediated translation (44–47). Appropriate spacing between the pseudoknot and the initiator AUG was required for translation (46, 47). Up to three pseudoknots in the leader of the cricket paralysis-like virus, Plateau stall intestine virus (PPS), are required for IRES-mediated translation at a non-methionine initiation codon (48). In these cases, the pseudoknot may largely play a structural role in positioning the 40 S ribosomal subunit at the initiator AUG.

The sequence of L3 in PK1 is notable because it is complementary to nt 1117–1123 of 18 S rRNA that is conserved throughout eukaryotes. L3 is part of a region (i.e. 57–75) encompassing the distal half of S1, L3, the distal half of S2, and flanking sequence that is complementary to 16 of 19 bases of nt 1109–1127 of eukaryotic 18 S rRNA. Mutating the two cytidines to adenosines in L3 or changing the three uridines to adenosines that would disrupt base pairing with 18 S rRNA reduced cap-independent translation up to 96%. In contrast, mutating two uridines to cytidines that would maintain potential base pairing with 18 S rRNA had the least effect on PK1 activity, consistent with the possibility of an interaction with 18 S rRNA. Increasing the length of L3 that is complementary to 18 S rRNA reduced cap-independent translation, demonstrating that the length and primary sequence of L3 are optimal for PK1 activity.

Although a direct interaction between L3 and 18 S rRNA remains to be demonstrated, base pairing between other sequence elements that function as translational enhancers and 18 S rRNA has been proposed. A 40-nt segment of the leader from the Gtx homeodomain mRNA cross-linked to 18 S rRNA, following 40 S ribosomal subunit binding to the mRNA and a 9-nt sequence within the 40-nt segment binds 18 S rRNA, which was complementary to nt 1123–1131 within helix 26 of 18 S rRNAs and enhanced translation when the length of the complementary match was shortened to 7 nt (49–51). A sequence that was complementary to the same region of 18 S rRNAs was selected from random oligonucleotides on the basis of its observed IRES activity (52). A second sequence that was complementary to nt 68–76 of rat 18 S rRNA was also selected in the same screen for its IRES activity (52). A similar interaction was proposed for a 22-nt sequence within the leader of the cold stress-induced Rbm3 mRNA, which is complementary to nt 808–819 of 18 S rRNA (42, 43).
nt 1112–1121 of plant 18 S rRNA was identified previously by using murine 18 S rRNA as “clanger” segment number 10 (nt 1163–1172), so named for its potential involvement in stable intermolecular base pairing (53). A subsequent study showed that nt 1115–1124 of plant 18 S rRNA is exposed and accessible for intermolecular base pairing (54). Moreover, mRNAs containing a region with perfect complementarity to this region exhibited a high affinity for 40 S ribosomal subunits and substantially enhanced expression (54). The L3 of TEV PK1 has the potential to base pair to the same sequence of plant 18 S rRNA to which this sequence binds. Similar observations were made for a 10-nucleotide sequence complementary to nt 1105–1114 (54), suggesting that the region from at least 1105–1124 is available for stable intermolecular base pairing. These observations suggest that sequences within the leader of an mRNA that are able to interact within this region (i.e. nt 1105–1124) of 18 S rRNA may function to enhance translation through direct base pairing with the 18 S rRNA.

Previous work had demonstrated that the TEV 5′-leader specifically required eIF4G to promote cap-independent translation that was enhanced further by eIF4A, eIF4B, and the poly(A)-binding protein. That the TEV leader promotes cap-independent translation through direct base pairing with the 18 S rRNA.

REFERENCES
1. Gallie, D. R. (2002) Plant Mol. Biol. 50, 949–970
2. Hershey, J. W. B., and Merck, W. C. (2000) in Translational Control of Gene Expression (Sonenberg, N., Hershey, J. W. B., and Mathews, M. B., eds) pp. 33–88, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
3. Kozak, M. (1986) Cold Spring Harbor Symp. Quant. Biol. 51, 5345–5349
4. Pelletier, J., and Sonenberg, N. (1985) J. Biol. Chem. 260, 12484–12487
5. Blackburn, P. (1979) J. Biol. Chem. 254, 12484–12487
6. Draper, D. E. (2004) RNA (N.Y.) 10, 335–343
7. Cole, P. E., Yang, S., and Crothers, D. M. (1972) Biochemistry 11, 4558–4568
8. Romer, R., and Hach, R. (1975) Eur. J. Biochem. 55, 271–284
9. Stein, A., and Crothers, D. M. (1976) Biochemistry 15, 160–168
10. Gallie, D. R., Sleat, D. E., Watts, J. W., Turner, P. C., and Wilson, T. M. A. (1987) Nucleic Acids Res. 15, 2827–2843
11. Wells, D. R., Tanguay, R. L., Le, H., and Gallie, D. R. (1998) Genes Dev. 12, 3236–3251
12. Chappell, S. A., and Edelman, G. M., and Mauro, V. P. (2000) Proc. Natl. Acad. Sci. U. S. A. 97, 1536–1541
13. Chappell, S. A., Edelman, G. M., and Mauro, V. P. (2000) Proc. Natl. Acad. Sci. U. S. A. 97, 1536–1541
14. Chappell, S. A., Edelman, G. M., and Mauro, V. P. (2000) Proc. Natl. Acad. Sci. U. S. A. 97, 1536–1541
15. Pestova, T. V., Shatsky, I. N., and Hellen, C. U. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 9109–9128
16. Allison, R. F., Johnston, R. E., and Dougherty, W. G. (1986) Virology 154, 9–20
17. Allison, R. F., Sorensen, J. C., Kelly, M. E., Armstrong, F. B., and Dougherty, W. G. (1985) Proc. Natl. Acad. Sci. U. S. A. 82, 3969–3972
18. Carrington, J. C., and Freed, D. D. (1990) J. Virol. 64, 1590–1597
19. Allison, R. F., Sorensen, J. C., Kelly, M. E., Armstrong, F. B., and Dougherty, W. G. (1985) Proc. Natl. Acad. Sci. U. S. A. 82, 3969–3972
20. Cattoretti, F., and Carrington, J. C. (1990) J. Virol. 64, 1590–1597
21. Allison, R. F., Sorensen, J. C., Kelly, M. E., Armstrong, F. B., and Dougherty, W. G. (1985) Proc. Natl. Acad. Sci. U. S. A. 82, 3969–3972
22. Carrington, J. C., and Freed, D. D. (1990) J. Virol. 64, 1590–1597
23. Kuhn, R., Luz, N., and Beck, E. (1990) J. Virol. 64, 4625–4631
24. Jang, S. K., Krausslich, H. G., Nicklin, M. J., Duke, G. M., Palmenberg, A. C., and Nicklin, M. J. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 7706–7710
25. Sczyrba, A., Kruger, J., Mersch, H., Kurtz, S., and Giegerich, R. (2003) RNA (N.Y.) 9, 133–148
26. Han, K., and Byun, Y. (2003) Nucleic Acids Res. 31, 3767–3770
27. Zakeri, I., and Stiegler, P. (1981) Nucleic Acids Res. 9, 33–88, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
28. Kozak, M. (1986) Cold Spring Harbor Symp. Quant. Biol. 51, 5345–5349
29. Pelletier, J., and Sonenberg, N. (1985) J. Biol. Chem. 260, 12484–12487
30. Chappell, S. A., and Mauro, V. P. (2003) Nucleic Acids Res. 31, 3767–3770
Cap-independent Translation of Tobacco Etch Virus Is Conferred by an RNA Pseudoknot in the 5′-Leader
Vladimir Zeenko and Daniel R. Gallie

J. Biol. Chem. 2005, 280:26813-26824.
doi: 10.1074/jbc.M503576200 originally published online May 23, 2005

Access the most updated version of this article at doi: 10.1074/jbc.M503576200

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts

This article cites 50 references, 24 of which can be accessed free at http://www.jbc.org/content/280/29/26813.full.html#ref-list-1