JOURNAL OF VIROLOGY, May 1996, p. 2999–3010 Vol. 70, No. 5
0022-538X/96/$04.00+0

Position-Dependent ATT Initiation during Plant Pararetrovirus Rice Tungro Bacilliform Virus Translation

JOHANNES FÜTTERER,[1]* INGO POTRYKUS,[1] YIMING BAO,[2]† LIU LI,[2] THOMAS M. BURNS,[2] ROGER HULL,[2] AND THOMAS HOHN[3]

Institute for Plant Sciences, ETH Zentrum, CH 8092 Zürich,[1] and Friedrich Miescher Institute, CH-4002 Basel,[3] Switzerland, and John Innes Centre, Norwich NR4 7UH, United Kingdom[2]

Received 7 September 1995/Accepted 14 February 1996

The expression of the rice tungro bacilliform virus open reading frame I was studied in transiently transfected protoplasts. Expression occurs despite the presence of a long leader sequence and the absence of a proper ATG initiation codon. Translation is initiated at an ATT codon. The efficiency of initiation in rice protoplasts depends strongly on the mechanism by which ribosomes reach this codon. From the effects of scanning-inhibiting structures inserted into different leader regions, it can be deduced that this mechanism is related to the ribosome shunt described for cauliflower mosaic virus 35S RNA. The process delivers initiation-competent ribosomes to the region downstream of the leader and is so precise that only the second of two potential start codons only 12 nucleotides apart is recognized. The ATT codon that is used when it is present downstream of the leader is hardly recognized as a start codon by ribosomes that reach it by scanning.

The structure of the pregenomic RNA (Fig. 1A) of the badnavirus rice tungro bacilliform virus (RTBV) closely resembles that of the cauliflower mosaic virus (CaMV) 35S RNA. Both RNAs begin with a leader sequence that is more than 600 nucleotides long harboring a number of short open reading frames (sORFs). Computer-assisted RNA folding predicts strong secondary structure in these leader regions (17, 32). In both cases, the leader ends with the binding site for the tRNA primer for reverse transcription and is followed by several ORFs coding for all of the viral proteins. The ORFs in the coding region are arranged tightly, avoiding long intercistronic distances (32, 57).

Caulimoviruses use the 35S RNA as a polycistronic mRNA (7, 26, 62; for a review, see reference 35). Translation of multiple ORFs from one mRNA is unusual in eukaryotic cells, since the normal mode of translation initiation seems to allow only one initiation event per RNA-bound ribosome and RNA binding occurs almost exclusively at the capped RNA 5′ end (44, 48). Caulimoviruses employ a variety of unusual mechanisms to translate their pregenomic RNA. These mechanisms include transactivation of translation of downstream ORFs on the polycistronic mRNA by a virus-encoded protein (7, 20, 21, 25, 26, 62, 63, 65). Another special feature is nonlinear migration of ribosomes, allowing the bypass of RNA regions in which initiation would otherwise occur: a process that we have called ribosome shunt (19, 22).

The ribosome shunt allows expression of the CaMV ORF VII downstream of the long leader sequence, which contains seven to eight sORFs. In RTBV, the corresponding ORF I is preceded by 12 sORFs with a total of 14 ATG codons (32, 57) (Fig. 1A). ORF I itself lacks an obvious translation initiation codon (Fig. 1B) but is defined by the absence of stop codons for 199 codons. A variety of products with unexpectedly high apparent molecular weights have been detected with an ORF I-specific antiserum in extracts from infected plants (31). To investigate how this ORF is actually expressed, we have analyzed the expression of a reporter gene fused to RTBV ORF I in transfected protoplasts in analogy to experiments previously performed with CaMV (7, 18–22) and figwort mosaic virus (25, 26, 62, 63). Our results indicate that the RTBV ORF I translation is initiated at an ATT codon. The choice of the translation initiation site is determined with precision, probably by the ribosome shunt process which transports ribosomes from the beginning of the leader to this codon. This precise mechanism of ribosome delivery increases the translation efficiency at that specific site, compared with the efficiency at identical initiation regions that are reached by a scanning mechanism. The presence of an ATT codon instead of an ATG codon is crucial for virus viability.

MATERIALS AND METHODS

Construction of expression plasmids. All RTBV sequences were derived from the sequenced infectious clone pJIIS2 (32).

Plasmids containing the CaMV 35S promoter are derivatives of pTZDH (18). The *Bam*HI site in the polylinker between the CaMV 35S promoter and the CaMV polyadenylation signal was converted into a *Cla*I site by filling in with Klenow DNA polymerase. Between this *Cla*I site and the *Pst*I site, we cloned a *Cla*I-*Eco*RI fragment covering positions 7403 to 128 (through 8002/0) of the RTBV genome and an *Xho*I-*Pst*I fragment covering the CAT coding sequence and derived from plasmid pLC15 (17). The *Eco*RI and *Xho*I ends of the two fragments were joined with oligodeoxynucleotides, creating the sequences shown in Fig. 1B for pCIC-21 and pCIC-21ATG. To create CIC-12, a suitable oligodeoxynucleotide was inserted between the *Bst*YI and the *Xho*I sites of plasmid pCIC-21 (Fig. 1B). Deletion of the leader sequence (plasmids diagrammed in panels B and D of Fig. 5 and 6) was accomplished by digestion of pCIC derivates with *Cla*I and *Bst*BI and subsequent religation. Frameshift mutations (see Fig. 3) were introduced into pCIC-21(ATG) by filling in of either the *Bst*BI site at RTBV genome position 18 (plasmids with the extension Bsfi [Fig. 3]) or the *Eco*RI site at position 123 (plasmids with the extension Ecfi [Fig. 3]).

Mutations around the ORF I start codon (plasmids CIC-12A.M1 to -M6) were introduced by first altering the *Bst*YI site in CIC-12 to a *Bam*HI site (CIC-12A.M1 with the same activity as that of CIC-12A) and then by cloning of suitable oligodeoxynucleotides between the *Bam*HI and the *Xho*I sites of CIC-12A.M1, creating the sequences shown in Fig. 5. To produce the plasmids diagrammed in panels C and D of Fig. 5 and 6, the *Xho*I-to-*Eco*RI (site within the chloramphenicol acetyltransferase [CAT] ORF) fragment of CIC-12A.M1 to -M6 or CIC-21A.ATG was replaced by an *Xho*I-*Eco*RI fragment derived from pNRF10Cat (21). This fragment consists of an artificial sequence that adds 14 codons to the 5′ end of ORF I and is followed by a stop codon, an intercistronic sequence of 70 nucleotides, and the 5′ end of the CAT ORF with its own ATG start codon, thus creating the ORF organization shown in Fig. 1, 5, and 6.

* Corresponding author. Mailing address: Institute of Plant Sciences, Universitätstr. 2, CH-8092 Zürich, Switzerland. Phone: 1 632 3866. Fax: 1 632 1044.

† Present address: The Samuel Roberts Noble Foundation, Ardmore, OK 73402.

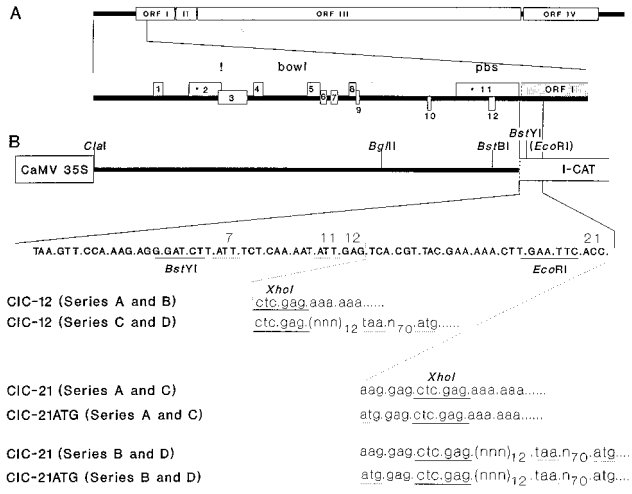

FIG. 1. (A) Schematic presentation of the pregenomic RNA of RTBV. ORFs are indicated by numbered open boxes; the 5′- and 3′-untranslated regions are indicated by thick lines. At bottom, the leader region of this RNA is shown enlarged with the sORFs presented as numbered boxes aligned in their respective reading phases. ORF-internal ATG codons are indicated by asterisks. ORF I is shaded, and the absence of an ATG start codon is indicated by the stippled 5′ border. The locations of the polyadenylation signal (!), a conserved sequence element without known function (bowl), and the primer binding site for reverse transcription (pbs) are indicated. (B) The basal structure of the expression plasmids used, consisting of a CaMV 35S promoter, the leader sequence of the RTBV pregenomic RNA, and a CAT reporter ORF fused to the RTBV ORF I. Also shown is the sequence of the beginning of the ORF I, starting with the TAA stop codon of the preceding sORF 11. Relevant codons are indicated by numbers, and potential start and stop codons are underlined with dots. Fusions to the RTBV sequence are appended in lowercase letters. Relevant restriction sites are indicated in the plasmid map and (by underlining) in the sequence. The plasmid names used in this and the following figures indicate the basal CaMV 35S promoter-RTBV ORF I (codon 12 or 21)-CAT fusion by CIC-12 or -21, respectively. Plasmids of series A and C all contain the complete leader sequence; the leader is deleted from plasmids of series B and D. In series A and B, the CAT ORF starts with a codon in the RTBV ORF I; in series C and D it starts with its own ATG start codon downstream of a termination codon in RTBV ORF I.

The mutant CIC-12A.M3-Xhfi was produced by filling in the *Xho*I site of CIC-12M3 with Klenow DNA polymerase.

Construct CIC-12A.sd is a derivative of a construct which contains four additional restriction sites in the 5′ part of the leader (CIC-L^m). We introduced a *Hin*dIII site by a G-to-A mutation at genome position 7452, a *Bam*HI site and a *Nco*I site by a C-to-G mutation at position 7489 and a A-to-C mutation at position 7494, and a *Spe*I site by a G-to-C mutation at position 7512. These mutations had no effect on any investigated aspect of gene expression (not shown). An oligodeoxynucleotide which contains an AGGT-to-AAAT mutation of the splice donor site was inserted between the *Bam*HI and the *Spe*I sites to produce CIC-12A.sd.

CIC-12A.Δ89 is a *Cla*I-to-*Bam*HI deletion derivative of CIC-L^m. CIC-12-St1 was produced by insertion of a palindromic oligodeoxynucleotide into the *Bam*HI site of CIC-L^m. The same oligonucleotide, which has also been used before to analyze ribosome migration on CaMV-derived reporter RNAs (20, 22), was inserted into the *Bgl*II site in the RTBV leader (CIC-12A.St2) and into a *Bam*HI site in the region between the truncated ORF I and the CAT ORF in CIC-21C to yield CIC-21C.St3.

Relevant portions of plasmid constructs were analyzed by DNA sequencing with a Sequenase kit according to the protocols provided by the supplier (U.S. Biochemicals).

Construction of plasmids for agroinoculation. DNA manipulations leading to the construction of mutant clones pJIIS2-ATG698 and pJIIS2-ATG686 were done with the full-length viral DNA clone pJIIS2 (14). Site-directed mutagenesis was performed by using the pALTER-1 mutagenesis vector (Promega) as described by the manufacturer. Construction of the one-and-a-bitmer clones from the full-length constructs pJIIS2-ATG698 and pJIIS2-ATG686 into the *Agrobacterium tumefaciens-Escherichia coli* vector pBIN (5) to produce pRTRB-ATG698 and pRTRB-ATG686, respectively, was as described by Dasgupta et al. (14). Transfer of pRTRB-ATG698 and pRTRB-ATG686 to *Agrobacterium tumefaciens* A4 (14) was by electroporation (53).

Protoplast transformation and CAT assays. Protoplasts were isolated from suspension cultures of *Orychophragmus violaceus* or rice (*Oryza sativa* cv. *Nipponbare*) and were transfected with equimolar amounts of circular plasmid DNA by electroporation (*O. violaceus*) or by the polyethylene glycol method (rice) as described elsewhere (15, 18). In many experiments, a β-glucuronidase (GUS) expression plasmid was cotransfected to serve as an internal standard for transfection efficiency. Routinely, about 10 μg of DNA was used to transfect either 2×10^6 (*O. violaceus*) or 0.6×10^6 (*O. sativa*) protoplasts. In cotransfection experiments, control samples were supplemented with equal amounts of calf thymus DNA instead of the cotransfected test plasmid. Crude protein extracts were prepared after overnight incubation of protoplasts and were analyzed for CAT activity as described elsewhere (18).

Agroinoculations. All agroinoculations were performed as described by Dasgupta et al. (14).

Western blotting (immunoblotting). Proteins from infected plants were analyzed as described by Hay et al. (31).

Sequence analysis of RTBV RNA and DNA. RTBV RNA from virus-infected rice plants was analyzed by 5′ rapid amplification of cDNA ends-PCR as described before (2) with oligodeoxynucleotide V305 (5′ 127-AATTCAAGTTTT TCGTAA-110 3′) as the reverse transcription primer. The PCR product was gel purified, and 0.3 pmol of the product was sequenced with 10 pmol of oligodeoxynucleotide V361 (5′ 7939-GCACTAGTAGGTAACTAA-7954 3′) as the primer (36). For determination of the detection limits of sequence variants in the original RNA preparation, different mixtures of a genomic clone of RTBV (pJIIS2 [32]) and mutant forms of this DNA with T-to-G exchanges at positions 89 and 101 (by courtesy of L. Liu) were sequenced with primer V361.

RNase protection assays. Total RNA from transfected protoplasts was isolated and analyzed by RNase A/T_1 mapping as described before (7). As an internal standard, in all samples pC4CΔBB (23) was cotransfected. In this plasmid, the CAT ORF is fused to the RTBV ORF4. The anti-sense RNA for probing was produced by in vitro transcription by T7 RNA polymerase of a plasmid containing an *Eco*RI-to-*Pst*I fragment consisting of the 216 nucleotides of the CAT ORF, a short linker, and adjacent RTBV ORF III sequences (23). Expected protected fragments have sizes of 299 and 350 nucleotides for spliced and unspliced RNAs, respectively, from the internal standard and of about 222 nucleotides for ORF I fusions. RNA levels were quantified with a PhosphorImager (Molecular Dynamics).

RESULTS

RTBV ORF I RNA sequence. Viral RNA from RTBV-infected rice plants was directly sequenced to determine whether an RNA subpopulation which contains a conventional start codon for ORF I might exist. Such RNAs might have been produced by some kind of editing and for unknown reasons might have been excluded from reverse transcription or packaging and thus would not be represented as reverse-transcribed viral DNA. In contrast to this speculation, the RNA sequence was found to be identical to the DNA sequence (Fig. 2A). By addition of traces of an RNA containing an ATG codon, it was shown that the detection limit for an edited RNA was approximately 5% in our assay (Fig. 2B). We conclude that even on the RNA level, ORF I lacks an ATG codon that could be used as a translational start site. The RNA sequence is also unambiguous in the region containing the stop codon that separates the in-frame sORF 11 from ORF I (Fig. 2A). This excludes the existence of any significant amount of an RNA processed to give translational continuity with an ATG of a leader sORF.

Expression plasmids. To study the expression of RTBV ORF I directly, we transfected protoplasts with constructs containing a CAT reporter ORF fused to the viral ORF I (Fig. 1B). Transcription was controlled by the CaMV 35S promoter, since the RTBV promoter is poorly active in the protoplast systems used (11). Downstream of the 35S promoter, the plasmids contain RTBV sequences starting at a *Cla*I site located at RTBV genome position 7403 (RTBV sequence numbering is according to that of Hay et al. [32]). This position is close to the actual transcription start site of the RTBV pregenomic RNA, which was mapped to position 7404 or 7405 (2, 64). Numbering of positions in the different constructs for convenience refers to position 7404 as +1. All constructs contain the CaMV polyadenylation signal. The system for naming plasmids is given in the legend to Fig. 1.

RTBV ORF I downstream of the leader sequence is expressed. To verify expression of an ORF I product, initially two

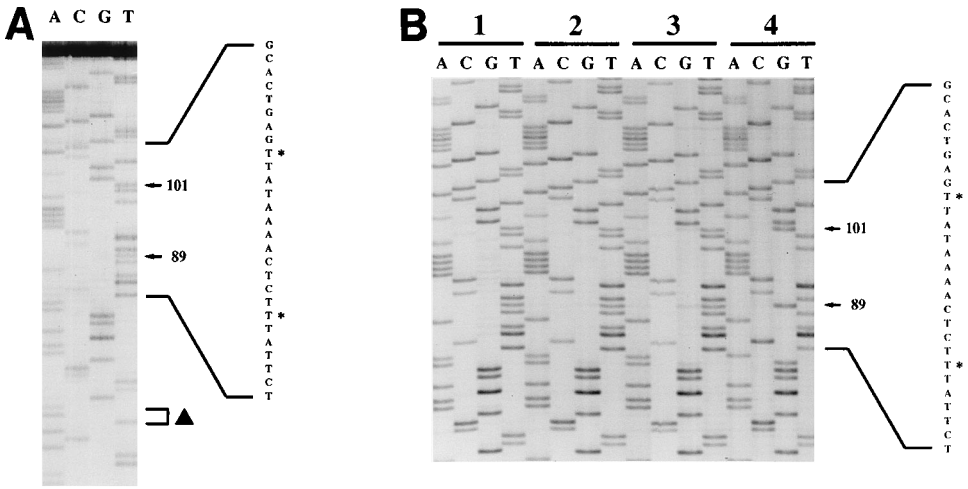

FIG. 2. (A) Sequence analysis of RTBV RNAs from infected plants. The region encompassing the end of sORF 11 (▲, stop codon) and the beginning of ORF I is shown. The positions of the second T's in the two ATT codons are marked according to the sequence numbering used by Hay et al., (32). (B) Sequence analysis of mixtures of RTBV wild-type DNA and mutant DNAs. 1, wild-type DNA; 2, mixture of 95% wild-type DNA and 5% mutant DNA; 3, mixture of 90% wild-type DNA and 10% mutant DNA; 4, DNA with ATT-to-ATG mutations in both ATT codons at the positions indicated. In lane 3G, additional bands are clearly visible at the positions marked by arrowheads. In lane 2G, a very faint signal can still be detected.

CAT fusions were made to position +727, i.e., to codon 21 of ORF I (counted from the stop codon of the preceding sORF 11 [Fig. 1B]). In one construct, the CAT ORF retained its own ATG start codon (pCIC-21A.ATG; nucleotides are referred to in their DNA form), while in the other construct, this ATG was changed to AAG (pCIC-21A) to eliminate any potential start codon within the reporter gene and to ensure that translation initiation at this ORF is controlled by RTBV sequences. Upon transfection of these constructs into protoplasts from cell suspensions of *O. violaceus* or rice (*O. sativa*), CAT expression was obtained for both plasmids (Fig. 3). In rice, the expression level of the ATG-less construct was usually approximately 10% of that of the ATG-containing control, and in *O. violaceus* the relative expression level of the ATG-less construct was around

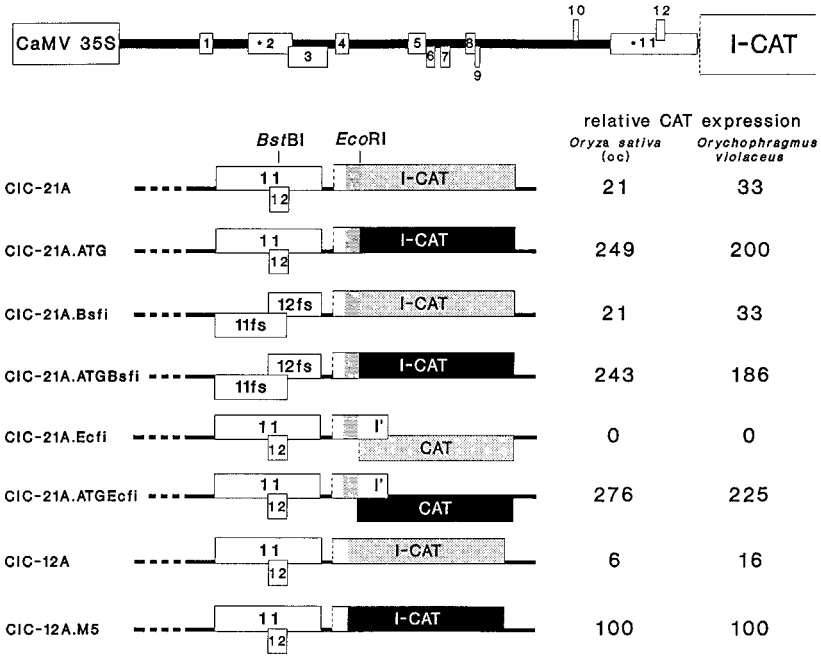

FIG. 3. Expression of RTBV ORF I. The basal expression unit is shown in Fig. 1. For the different expression plasmids, only the regions surrounding the start of ORF I, including the sORFs 11 and 12 and their derivatives, are shown enlarged. The fusions of the CAT ORF (shaded or black) either in frame or out of frame are indicated by the positionings of the respective boxes. The presence of an ATG start codon is indicated by a strong vertical line at the 5′ end of an ORF. CAT ORFs initiated at an ATG codon are indicated by black boxes; CAT ORFs without an ATG codon are indicated by lightly shaded boxes. The region between codons 12 and 21 of the RTBV ORF I is indicated by darker shading. Average expression levels, obtained in at least five independent transfection experiments, are shown for rice and *O. violaceus* protoplasts. Expression values for CIC-12A.M5 (see Fig. 5) were set to 100% for both systems. Standard deviations are omitted for clarity in this and the following figures. As described in the text, values for ATG-less constructs varied by about 50% of the indicated values, whereas values for ATG-containing constructs varied by only 20% (see Fig. 4).

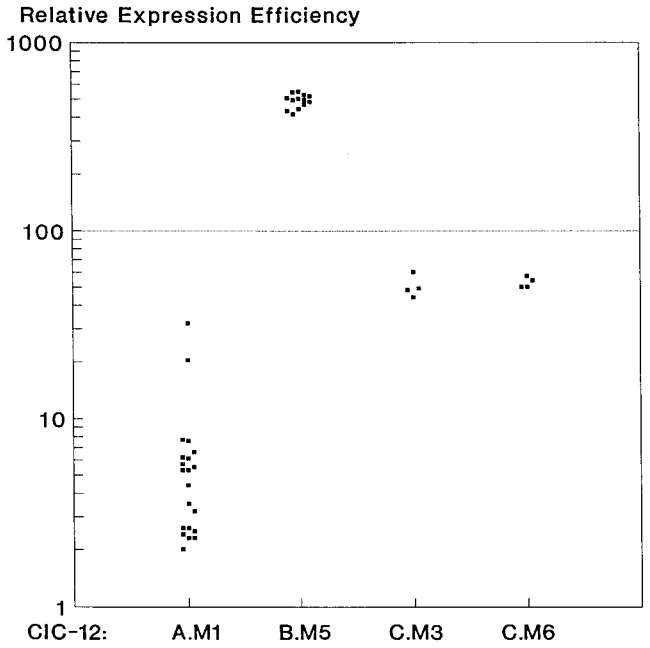

FIG. 4. Examples of the variation in relative expression levels in different experiments. Values obtained in transfected rice protoplasts with constructs CIC-12A.M1 (no ATG), -B.M5, -C.M3, and -C.M6 (all with an ATG start codon; see Fig. 5) relative to CIC-12A.M5 (taken as 100%) are shown. Each dot represents an independent transfection experiment.

20%. In both protoplast systems, these relative expression levels varied much more (from 2 to 30% compared with the ATG-containing construct [Fig. 4]) than is usual in this type of experiment (deviations of ± 20% of a measured value). These data demonstrate that ORF I is indeed expressed despite the lack of an ATG start codon, and the great variability of expression efficiency compared with those of ATG-containing constructs suggests a particular importance of cellular parameters for ORF I expression.

RTBV ORF I is not translated by stop codon suppression. It has been suggested that the RTBV ORF I might be expressed by suppression of the stop codon between sORF 11 and ORF I (Fig. 1B) (57). To test this suppression hypothesis, a frameshift mutation was introduced into sORF 11 (and into the overlapping sORF 12). In the resulting plasmid pCIC-21A.ATGBsfi (Fig. 3), sORF 12 terminates at the stop codon in frame with ORF I. We have previously shown that a reporter ORF linked to sORF 11 is translated much more efficiently (although still at low levels) than one linked to sORF 12 (11). If ORF I was really expressed by stop codon suppression, a significant reduction in the level of expression should have resulted. However, both mutant plasmids expressed CAT with the same efficiency as that of the parent plasmids (Fig. 3). To verify that the RTBV part of the fusion ORF was translated and that translation did not initiate further downstream in the authentic CAT coding region, a frameshift was introduced at the connection between the RTBV-derived sequence and the CAT ORF (pCIC-21A.Ecfi). No CAT expression was obtained from this construct (Fig. 3). As expected, the same manipulation of the ATG-containing ORF had no significant effect on CAT expression, since the introduced ATG codon lies downstream of the frameshift site (pCIC-21A.ATGEcfi [Fig. 3]). These data exclude suppression of the stop codon between sORF 11 and ORF I as a translation mechanism for the RTBV ORF I. They also strongly suggest that the RTBV ORF I sequence of the CAT fusion ORF is indeed translated. To further delimit the ORF I region that is important for translation initiation, the ATG-less CAT ORF was fused to the 12th codon (pCIC-12A [Fig. 1B]). Expression from pCIC-12 was clearly detectable (Fig. 3), suggesting that the sequences between ORF I codons 12 and 21 are not absolutely required for ORF I expression, although the lower level of CAT activity obtained from pCIC-12A (compared with that for pCIC-21A) could be indicative of a stimulating effect of these sequences.

Involvement of an ATT codon in ORF I translation. Since in CIC-12A the CAT ORF is translated in the absence of an ATG codon, it must be assumed that initiation occurs at a non-ATG start codon (a codon that deviates from ATG at one position), much as in a variety of other cases (6, 9, 10, 24, 27, 29). In the reading phase of ORF I, two candidate non-ATG codons (ATT at positions +686 and +698 corresponding to codons 7 and 11 of ORF I) are located near the 5′ end of ORF I (Fig. 1B). The sequence context of the ATT at +698 matches the consensus for efficiently used start codons (6, 27, 43, 50). In pCIC-12A, a *Bam*HI site was introduced upstream of the first ATT codon by a point mutation (CIC-12A.M1) and the CAT ORF was fused to a site closely downstream of the second ATT via an *Xho*I linker (Fig. 5A). The short *Bam*HI-*Xho*I fragment containing the ATT codons was replaced by suitable synthetic oligodeoxynucleotides to introduce mutations in the ATT codons (plasmid series A in Fig. 5A). Mutation of the ATT codon at +686 to TTT had no effect on CAT expression (pCIC-12A.M2), while a similar mutation of the +698 ATT codon (pCIC-12A.M3) or deletion of this codon (pCIC-12A.M4) caused a complete loss of CAT expression. Mutation of the +698 ATT codon to ATG resulted in a 10- to 20-fold increase in CAT expression (pCIC-12A.M5). Surprisingly, a double mutation of the +686 ATT codon to ATG and of the +698 ATT codon to TTT (pCIC-12A.M6) did not result in detectable CAT expression in the presence of the leader.

The ATT codon serves as an initiation codon. The most obvious explanation for the results described above would be a direct involvement of the +698 ATT codon in translation initiation. Alternatively, the ATT codon could be part of a signal that promotes association of ribosomes or other scanning complexes with the region downstream of the leader, similarly to an ATG codon being required for internal ribosome entry on picornaviral RNA (for reviews, see references 1, 39, and 42). To discriminate between these two possibilities, a second series of plasmids (series C, Fig. 5C) was constructed in which ORF I was separated from the CAT ORF by the introduction of an artificial sequence. In this series, initiation at an ORF I codon opens only a short ORF (ORF I′ [16 codons when starting at position +698]), which is separated from the CAT ORF by an intercistronic region of 70 nucleotides (Fig. 1B). The intercistronic region does not contain any features supporting unusual translation mechanisms (20). The CAT ORF now constitutes a separate translation unit which is identical in all constructs. Mutations in the ORF I parts of the constructs which interfere with ribosome access are expected to have similar effects in series A and C. Mutations that influence initiation at an ORF I codon directly should result in opposite effects for the two series, because the sORF I′ in series C is expected to inhibit scanning-related translation of a downstream ORF in a manner proportional to its own translation efficiency. For instance, a related ATG-initiated sORF inhibited scanning-dependent downstream translation on a model RNA by a factor of two to three (21). A similar level of inhibition would have been expected for ATG-initiated ORFs I′, and a much lower one would have been expected for ATT-initiated sORFs I′. The results show that mutation of both ATT codons to TTT abol-

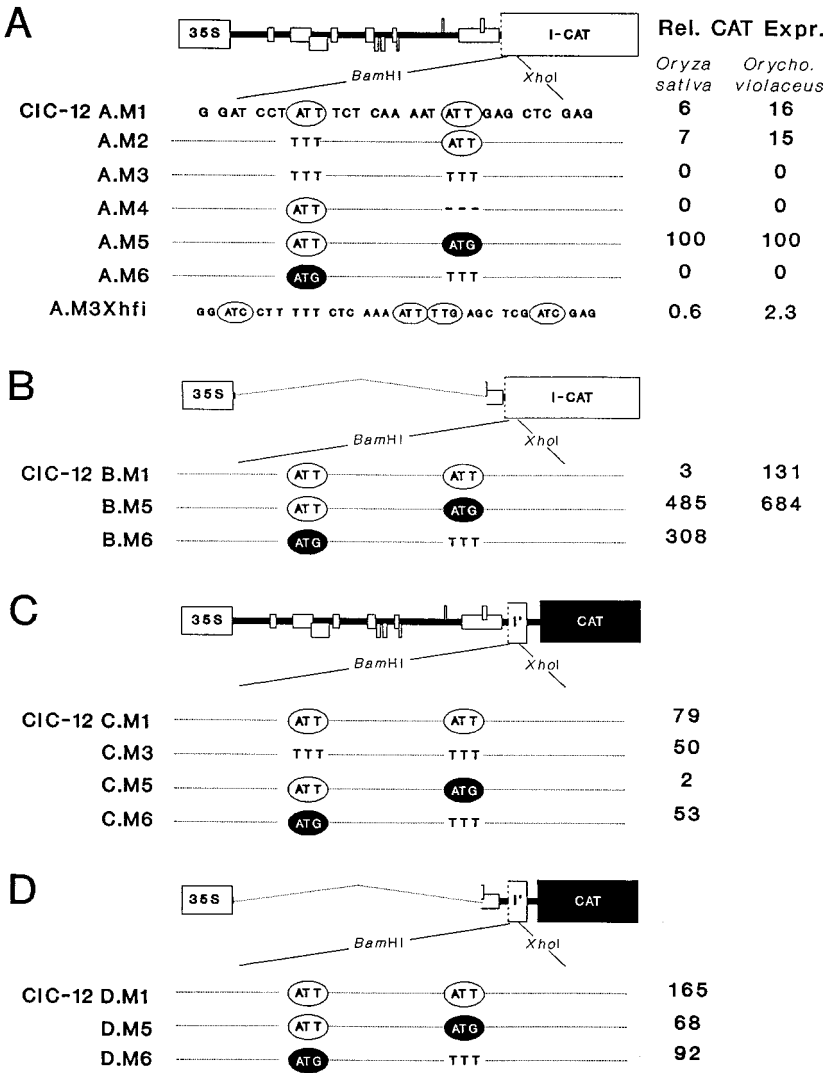

FIG. 5. Effects of start codon mutations on ORF I expression. The effects of start codon mutations in different derivatives of plasmid CIC-12 on CAT expression were assayed in as many as four different plasmids of series A to D and are shown as described in the legend to Fig. 2 for rice protoplasts (series A to D) and *O. violaceus* protoplasts (series A and B). Rel. expr., relative expression levels. (A) ORF organization of plasmid CIC-12, as described in the legend to Fig. 1. The sequences between the *Bam*HI site and the *Xho*I site of pCIC-12 and mutant derivatives are presented, and the reading phases of the CAT ORFs are indicated. Potential start codons are marked by filled (ATG) or empty (non-ATG) ellipses, and mutations of these codons are indicated. Nucleotides not printed are identical to those in CIC-12AM1. Dashes indicate a three-nucleotide deletion in A.M4. In mutant CIC-12A.M3Xhfi, the reading phases of ORF I with respect to the CAT ORF were altered as indicated by filling in of the *Xho*I site. (B) Plasmids with a deletion of the leader sequence, indicated as described for panel A. (C and D) Plasmids in which the CAT ORF is opened by an ATG start codon and is separated from the ORF I translation initiation site by a stop codon, which truncates the RTBV ORF I to an sORF (I′), and 70 nucleotides of intercistronic region (see Fig. 1B). Plasmids are indicated as described for panel A.

ishes expression of the directly fused CAT ORF but leads only to a slight reduction of unlinked CAT expression (compare in series C mutants M1 and M3 [Fig. 5C]), indicating that the mutation has only a small effect on the number of ribosomes reaching this region of the mRNA but rather directly influences translation initiation at ORF I. The mutation of the +698 ATT codon to ATG, which caused a strong increase of ORF I-linked CAT activity, resulted in a very low level of unlinked CAT activity (compare mutants M5 and M1 or M3 in Fig. 5C), suggesting that most of the ribosomes initiated at the mutant ATG codon and were not able to reinitiate after translation of the sORF. The mutation of the +686 ATT codon to ATG, which had no detectable influence on expression of the linked CAT ORF, also had no effect on expression of the unlinked CAT ORF (compare mutants M6 and M3 in Fig. 5C). The ATG codon at position +686 apparently is not recognized as a translation initiation codon in the presence of the complete leader, although it is recognized in a leader deletion mutant, in which it is the first ATG from the RNA 5′ end (compare mutants M6 in the plasmid series containing [Fig. 5A and C] or lacking [Fig. 5B and D] the leader).

Taken together, these findings suggest that ribosomes reach the region downstream of the leader at a well-defined position located downstream of the +686 ATT codon and initiate translation at the +698 ATT codon. The efficiency of this initiation is 5 to 10% compared with the efficiency of initiation at an ATG codon at this position. It has been shown that non-ATG initiation is particularly dependent on an optimal sequence context (6, 27, 45). This is in agreement with our observations, since the functional ATT codon has an optimal context with a

purine (A) at position −3, a G at position +4, and an A at position +5. With construct pCIC-12A.M3Xhfi, which was produced from pCIC-12A.M3 by filling in of the *Xho*I site between ORF I and CAT, only very low expression levels were obtained (Fig. 5A), although in this construct four possible non-ATG start codons were in frame with the CAT coding region, one of them even possessing the G residue at the +4 position.

Effect of the leader sequence on ORF I expression. In the case of a normal scanning mechanism, the particular structure of the leader sequence of the RTBV pregenomic RNA should cause inhibition of downstream translation. The influence of the leader was analyzed by deleting nucleotides 1 to 620 from plasmid series A (resulting in series B, [Fig. 5]) and from series C (resulting in series D [Fig. 5]). Removal of the leader sequence indeed resulted for most constructs in an increase in CAT expression. For instance, in protoplasts from the dicot plant *O. violaceus*, this increase was 5- to 10-fold for constructs with an ATT or an ATG start codon (Fig. 5A and B; compare A.M1 and B.M1 and A.M5 and B.M5). Expression of the ATG-less CAT ORF in the absence of the leader shows that no spliced-in ATG codon is required for ORF I translation and again substantiates that initiation occurs directly at the ATT codon. In protoplasts prepared from a rice (*O. sativa* cv. Oc) cell suspension culture, the effects of the leader deletion depended on the particular construct. Surprisingly, in the construct M1 with the natural ATT codon, deletion of the leader resulted in a reduction of expression (A.M1 and B.M1 in Fig. 5A and B). This reduction is not due to a reduced number of ribosomes associated with the truncated RNA, since with the ATG-containing construct A.M5, the leader deletion leads to a fivefold increase in CAT activity (A.M5 and B.M5 in Fig. 5A and B). In addition, with constructs C.M1 and D.M1, it is apparent that the leader deletion results in a general increase in ATG-initiated translation in the ORF I region, although in this case the effect is only twofold (Fig. 5C and D). These data indicate that the leader directly affects the capacity of ribosomes to initiate at the ATT codon, causing a higher relative efficiency of initiation (5 to 10% compared with that of ATG initiation; A.M1 versus A.M5) than in the absence of the leader (<1%; B.M1 versus B.M5). One component of this increased efficiency is a direct effect of the ATT codon on the number of ribosomes that overcome the inhibition by the leader, as indicated by the only twofold inhibitory effect in constructs with an ATT codon (C.M1 and D.M1) versus the fivefold inhibitory effect in constructs with an ATG codon (A.M5 and B.M5). A direct involvement of the ATT codon in ribosome access could also be deduced from the slight reduction observed for ATT-to-TTT mutations in the constructs C.M3 and C.M6 compared with C.M1 (Fig. 5C). A corresponding effect of the ATT-to-ATG mutation in C.M5 cannot be directly measured, since this mutation caused a strong inhibitory effect on initiation downstream of the ORF I start codon (Fig. 5C). It is evident that the inhibitory effect of the truncated ORF I on downstream CAT expression is modulated by the leader. In the absence of the leader, the ATG-initiated, truncated ORF I has only a moderate inhibitory effect, and this is independent of the precise location of the start codon (compare mutant D.M5 or D.M6 with D.M1 [Fig. 5D]). In the presence of the leader, the same ORF has a strong negative effect when the start codon is at position +698 (C.M5) but has no effect when the ATG start codon is at position +686 (C.M6 [Fig. 5C]).

We have described above that constructs with ATG-less CAT fusions to codon 21 resulted in CAT activity levels higher than those of fusions to codon 12 (Fig. 3; CIC-21A versus CIC-12A). A similar increase is also observed for ATG-containing constructs (Fig. 6A; CIC-21A.ATG versus CIC-12A.M5). However, CAT activity may not reflect the actual protein levels, because the different CAT proteins have different amino termini and may therefore have different activities or stabilities. However, series C derivatives with an ORF I-independent CAT ORF show a very similar difference in expression (CIC-12C versus CIC-21C [Fig. 6C]), especially for the ATG-containing constructs (CIC-12C.M5 versus CIC-21C.ATG). In this latter case, it must be considered that the upstream ORFs start at different ATG codons and could thus have a differential inhibitory effect. A positive effect is observed with CAT ORFs starting upstream (CIC-12A.M1 versus CIC-21A) and downstream (series C) of the apparently enhancing sequence between the ORF I codons 12 and 21, suggesting that it is largely due to an increased number of ribosomes available for initiation rather than to a better initiation efficiency at the respective start codons. An enhancing effect is obtained mainly in the presence of the leader. In its absence, only a slight positive effect on ATT-directed expression is observed (CIC-12B.M1 versus CIC-21B [Fig. 6B]), while the sequence exhibits a rather strong negative effect in constructs that allow the analysis of translation initiation downstream of codon 21 (CIC-12D.M1 versus CIC-21D; CIC-12B.M5 or D.M5 versus CIC-21B.ATG or D.ATG, respectively [Fig. 6]).

Analysis of ribosome migration. The results described above indicate that the position of translation initiation at ORF I is chosen with much precision. Initiation-competent translation complexes completely ignore an ATG codon at position +686 (pCIC-12A.M6, Fig. 5A) and recognize one at position +698 very efficiently, as shown by the almost complete absence of translation downstream of the sORF I′ in pCIC-12C.M5 (Fig. 5C). To reach ORF I, scanning complexes either have to pass through the leader, bypass the leader in a ribosome shunt, or bind directly to an internal RNA region. Alternatively, an RNA species may exist from which the leader is removed by splicing. We found no evidence for such spliced RNAs in the direct sequencing of RTBV RNA from infected plants or by PCR amplification of reverse transcripts of RNA from transfected protoplasts (not shown). Furthermore, the splice donor sequence in the beginning of the leader which normally is spliced to an acceptor upstream of RTBV ORF 4 (23) could be mutated without a negative effect on ORF I expression (CIC-12A.sd [Fig. 7]). Therefore, we regard involvement of RNA splicing in ORF I expression as highly unlikely. To discriminate between the other possibilities, a stem-loop structure that had been shown to inhibit scanning in an in vitro system (46) and in transfected plant protoplasts (20, 22) was inserted into various positions of the RTBV leader. Insertion at position +89, i.e., upstream of sORF1, resulted in complete inhibition of ORF I expression (CIC-12A.St1). Since sequences downstream of the stem insertion site are sufficient for normal ORF I expression in a leader deletion mutant (CIC12A.Δ89 [Fig. 7] and CI-CATΔCI 83 in reference 11), it is unlikely that the insertion destroyed a specifically required sequence element. It rather appears that this region is scanned. In contrast, insertion of the stem structure in the center of the leader (position +470 for CIC-12A.St2 [Fig. 7]) had no negative effect on CAT expression; it even slightly stimulated expression. Similar results were obtained with analogous stem insertion derivatives of CIC-12C (not shown). These data indicate that the leader region spanning from at least position +470 (stem insertion) to position +686 (nonfunctional ATG codon in construct pCIC-12A.M6 [Fig. 5A]) is not scanned by ribosomes or scanning complexes that become involved in ORF I translation.

The stem structure was also inserted into the derivative

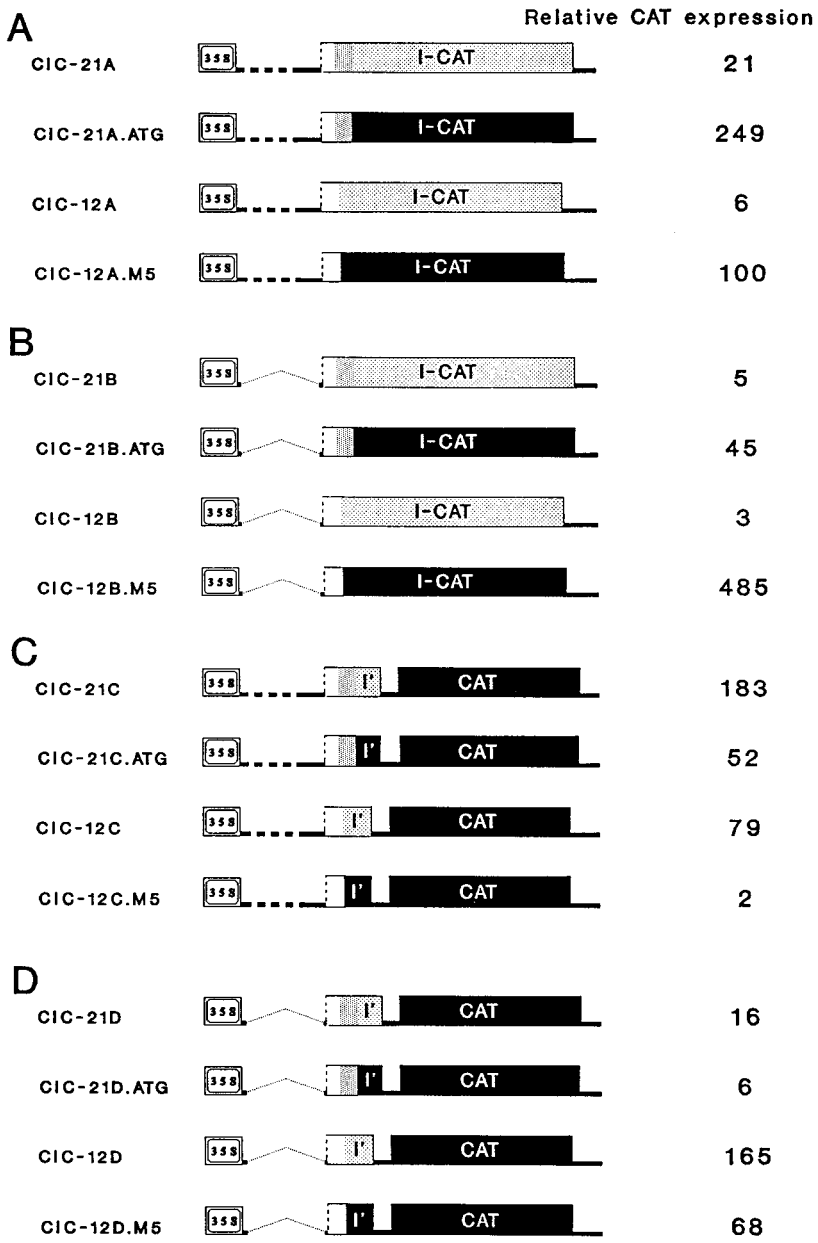

FIG. 6. Effect of sequences downstream of the start codon on expression. Derivatives of plasmids CIC-12 and CIC-21 are presented. ATG-containing and ATG-less ORFs are indicated as described in the legend to Fig. 3. Strong, broken lines (in panels A and C) indicate the complete leader sequence which is deleted in panels B and D. (A and B) CAT ORF initiated in the RTBV ORF I; (C and D) CAT ORF initiated at its own ATG start codon and separated from the ORF I by a stop codon and a short intercistronic region. The expression data are from four or more independent transfections of rice protoplasts.

CIC-21C with unlinked CAT ORF (CIC-21C.St3) in the artificial sequence separating sORF I′ and the CAT ORF. The strong inhibition of CAT expression is indicative for ribosome scanning through this region downstream of the ATT start codon (Fig. 7).

Analysis of RNA levels. To exclude the possibility that the strong alterations of CAT expression levels for different mutants described above were caused by differences in transcription or RNA stability, RNA levels were determined by RNaseA/T_1 mapping relative to an internal standard. Data obtained from two typical experiments are shown in Fig. 8. As calculated from a number of independent experiments, RNA levels of constructs with leader modifications and with unmodified leaders (e.g., CIC-12A.M1 or C.M1) averaged at very similar levels. Between experiments, relative levels normally varied by approximately twofold but occasionally by more. Variations are even observed for the ratio of the two fragments generated from the internal standard which originate from the same primary transcript (e.g., lane CIC-21A.MS in Fig. 8). Therefore, we regard the determination of RNA levels as less precise and take the data only as evidence that the large variations observed for CAT expression with different constructs are not due to correspondingly different RNA levels. On average, constructs with deletion of the complete leader (series B) produced about twofold less RNA, and the construct with the mutated splice donor sequence (CIC-12A.sd [Fig. 7]) pro-

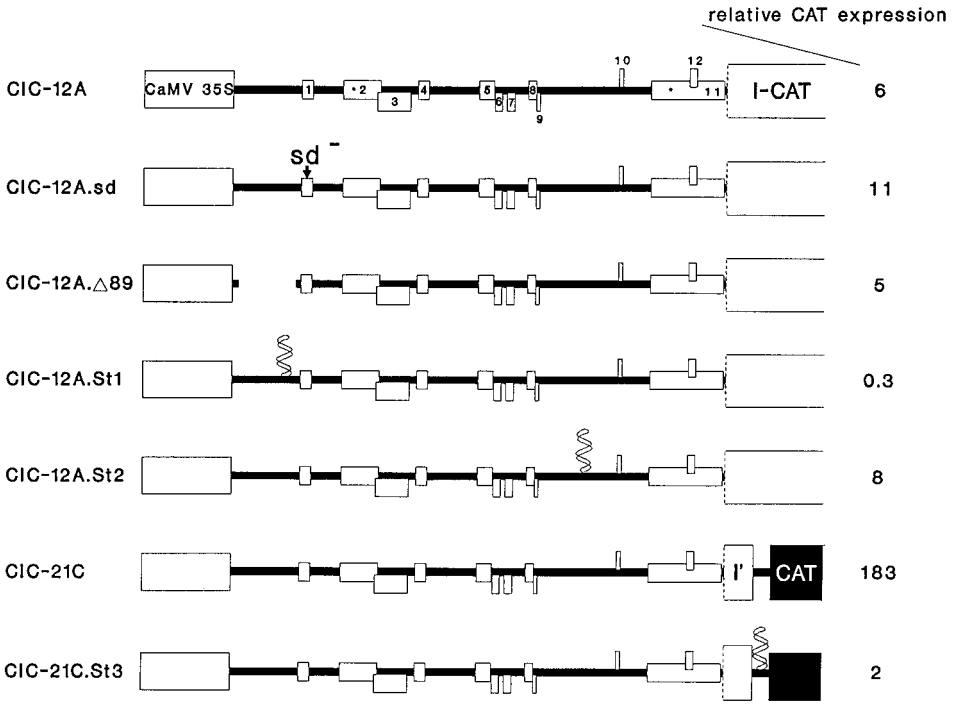

FIG. 7. Effect of leader modifications on ORF I expression. Expression constructs are indicated as described in the legends to Fig. 3 and 6. sd, double point mutation, changing the AGGT context of a splice donor site to AAAT. The gap indicates a deletion in the leader, and helices indicate insertion points of a stable stem structure. Expression values in rice protoplasts are indicated as described in the legend to Fig. 3.

duced about twofold more RNA than CIC-12A.M1. The slightly positive effect of the leader could be explained by the presence of a transcriptional enhancer near the beginning of the leader which however has only a weak effect on the CaMV 35S promoter (11). The positive effect of the splice donor mutation on RNA levels which corresponds to the effect on CAT activity (Fig. 7) could be explained by evading splicing to cryptic acceptor sites within or downstream of the CAT ORF or of abortive assembly of splice complexes.

These data indicate that the strong alterations in CAT activity levels observed with different constructs were caused mainly by posttranscriptional effects; however, it cannot be excluded that some of the minor changes are due to variations of the RNA levels.

Effect of mutations of the ATT codons on RTBV infection in plants. Two mutations were made on agroinoculation constructs, one (construct pRTRB-ATG686) changing the +686 ATT codon to ATG and the other (construct pRTRB-ATG 698) changing the +698 ATT codon to ATG. When agrobacteria containing these mutant constructs were injected into rice seedlings, symptoms were observed in both treatments. Symptoms caused by pRTRB-ATG698 were undistinguishable from those obtained upon agroinfection without viral sequences and, therefore, are caused by the inoculation procedure and not by virus multiplication. The symptoms caused by pRTRB-ATG686 were more severe. Western blots with antiserum to RTBV gene product P12 (31) revealed RTBV-specific products only in plants injected with pRTRB-ATG686 (Fig. 9). PCR of total DNA extracted from these plants by using primers spanning nucleotide positions 7872 to 373 gave the expected 493-bp product only from pRTRB-ATG686-injected plants (data not shown). Sequencing this PCR product showed that the ATG at +686 was retained in the virus population in the infected plant (data not shown).

DISCUSSION

The pregenomic RNA of all retroelements is used as a template for genome amplification and also as mRNA (for reviews, see references 34, 35, and 59). In almost all cases, this RNA is a polycistronic mRNA from which more than one protein is translated. This requires specific translation mechanisms, since eukaryotic translation normally initiates only once per mRNA-associated ribosome and usually only at 5′ proximal ORFs (44, 48). Retroviruses use frameshifts (see reference 12 and references therein) or stop codon suppression (30) to increase the number of proteins produced from one RNA, while caulimoviruses encode a transactivator that apparently induces reinitiation of translation in a relay race mechanism and thus allows translation of multiple, consecutive ORFs from one mRNA (7, 20, 26, 62). At first glance, the genome organization of plant badnaviruses closely resembles that of caulimoviruses. It has been shown by RNA mapping or can be deduced from sequence homology that the pregenomic RNA begins in all cases with a long, AUG-burdened leader sequence which is followed by three or four tightly arranged ORFs (8, 28, 32, 51, 57). In RTBV, the leader has a length of 698 nucleotides and contains 12 short ORFs and in total 14 AUG codons. According to the rules of conventional translation (44), such a structure should be very inhibitory.

In RTBV, translation of the first longer ORF is further complicated by the absence of a proper ATG start codon. We have shown here that an ATT codon downstream of the leader is used as an initiation codon for ORF I translation. Thus far, reports of naturally occurring deviations from ATG as the eukaryotic translation initiation signal consist only of CTG, ACG, and GTG. Initiation at ATT and other codons was observed only in artificial constructs (3, 13, 24, 45, 52, 55) and perhaps also in the Ω sequence of tobacco mosaic virus (dis-

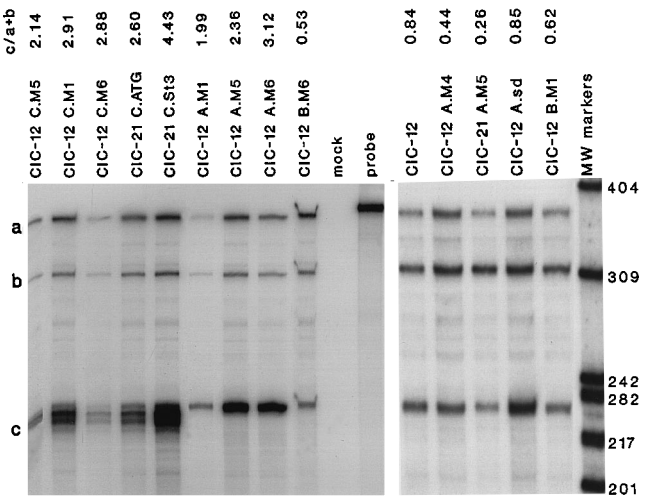

FIG. 8. RNase A/T_1 protection assay. RNA was isolated from rice protoplasts transfected with constructs as described in the legends to Fig. 5, 6, and 7. RNA levels were compared relative to a cotransfected standard construct, C4CΔBB, which contains a CAT ORF fused to the RTBV ORF IV (23). The labelled RNA antisense probe covers 222 nucleotides (band c) of the CAT ORF for the RTBV ORF I fusions and about 300 and 350 nucleotides (bands a and b) of spliced and unspliced RNAs from the standard construct. Gels from two different, representative experiments are shown. Because of differences at the fusion site between the CAT ORF and upstream sequences, RNA from the C series always gave rise to a double band, presumably because of a partial digestion at this junction point. No specific bands are observed in a mock-transfected control. Probe, undigested RNA; MW, size marker (*Hpa*II-digested pBR322). The relative expression levels of the different reporter plasmids (counts at position c divided by counts at positions a and b) as determined by analysis with a PhosphorImager are listed with the plasmid names.

cussed in reference 58), but it is unclear whether in the latter case a functional translation product results from this initiation event. In prokaryotes, ATT initiation has been described for only two genes (54, 60). In one of these, the *E. coli infC* gene (coding for initiation factor 3), ATT initiation is regulated in response to initiation factor concentration (49). It is unclear at present whether the ATT initiation in RTBV is also subject to a specific control mechanism. When located downstream of the leader, expression efficiencies of between 2 and 30% relative to corresponding ATG initiation were obtained with different protoplast types and batches (Fig. 4). This comparatively large variation may indicate some type of regulation by cellular parameters, as has been reported for non-ATG initiation of translation of the c-myc 1 protein in vivo (29) and for the general efficiency of non-ATG-prompted translation initiation in vitro (45).

The translation efficiency of ORF I (the first viral ORF downstream of the leader) is determined by at least two factors, both of which may vary in different cell types. One is the general efficiency of initiation at its unusual ATT start codon in the particular cellular environment, and the second is the efficiency of the mechanism that allows translation despite the presence of the long leader. In the absence of this leader, ATT initiation is less efficient in protoplasts of rice than of *O. violaceus*; in its presence, this difference is no longer apparent (Fig. 5), suggesting that the leader sequence positively influences ATT initiation in rice protoplasts. This is probably achieved through the mechanism that brings ribosomes to the respective initiation regions. From the effects of insertion of stem structures or an additional ATG codon into different parts of the leader, we conclude that this mechanism resembles the ribosome shunt described for the CaMV 35S RNA. The shunt process transfers ribosomes or other scanning complexes after initial scanning of only a small part of the 5′-terminal leader region directly to an acceptor region further downstream (22). In the case of RTBV, skipping of the central leader region by the shunt process is indicated by the tolerance of a stem structure insertion at position +470 (Fig. 7). Direct internal ribosome entry is unlikely because of the strong inhibitory effect of a stem structure insertion at position +89, within a region which is otherwise not required for ORF I expression.

The putative shunt acceptor site on RTBV RNA seems to be defined as precisely as the internal ribosome entry site on encephalomyocarditis virus RNA. For encephalomyocarditis virus, a clear discrimination between two ATG codons only 8 nucleotides apart is accomplished (40, 41), while for RTBV only the downstream one of two start codons positioned 12 nucleotides apart is recognized in the presence of the leader (Fig. 10). For RTBV, this is not due to the less favorable context of the first initiation codon, since ATGs at both positions are recognized with comparable efficiencies in leaderless constructs (Fig. 5D). The almost complete inhibition of translation downstream of the shortened ATG initiated ORF I in construct CIC-12C.M5 (Fig. 5C) additionally suggests that most of the ribosomes arriving downstream of the leader encounter the start codon at position +698. This indicates that almost all of the ribosomes either (re)enter the RNA between positions +686 and +698 or, as a less likely alternative, become initiation competent during scanning of this short region. It appears possible that the ORF I start codon is part of the actual shunt acceptor site. The difference between arrival of ribosomes at the start codon by scanning versus shunt (or internal entry) may be responsible for the altered initiation behavior in the presence of the leader. Shunted ribosomes might lack factors or interactions required for scanning, resulting in a prolonged pause at the potential initiation site and providing more time for the formation of an initiation complex. This would cause an increased initiation efficiency at otherwise inefficient start codons similar to ribosome pausing at suboptimal start codons due to suitably spaced downstream secondary structures (47). If the start codon in question is an ATG codon in good sequence context, the effect will not be apparent at the level of the translation product; however, even at such codons, the kinetics of initiation may be different. Such a difference may explain the strong inhibitory effect of an sORF starting at the proposed shunt acceptor site on further

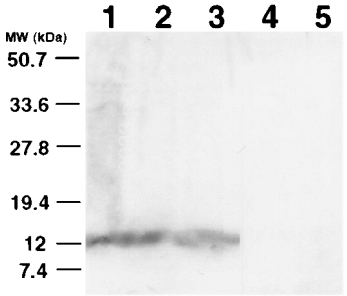

FIG. 9. Detection by Western blotting of gene product P12 from RTBV ORF II in extracts of rice inoculated with *A. tumefaciens* containing RTBV constructs. Lanes: 1, Philippines isolate (field infected); 2, construct pRTRB1162 (wild-type virus [14]); 3, construct pRTRB-ATG686; 4, construct pRTRB-ATG698; 5, noninfected rice. Molecular mass are indicated on the left (in kilodaltons). RTBV-specific gene products migrating with a molecular mass of 12 kDa are present in extracts of field-infected rice plants and of plants inoculated with agrobacteria containing constructs pRTRB1162 and pRTRB-ATG686 but not in extracts from uninfected rice or rice inoculated with agrobacteria containing construct pRTRB-ATG698.

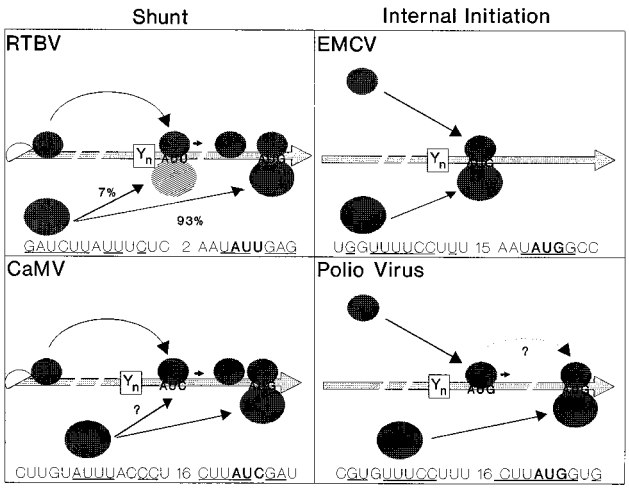

FIG. 10. Comparison of plant pararetroviral ribosome shunt and picornaviral internal ribosome entry. RNA is indicated as capped or uncapped. Small and large ribosomal subunits (associated with the respective translation factors) are indicated by small and large ellipses, respectively. Pyrimidine stretches are indicated by boxed Y_n, and the positions of AUG or potential non-AUG start codons are indicated. The sequences consisting of the pyrimidine stretches and the AUG or non-AUG codons are given, and the numbers of nucleotides separating these elements are indicated. Underlined sequences have the potential to base pair with the 3′ end of eukaryotic 18S rRNA, as suggested by Scheper et al. (61) for picornaviruses. Movement of ribosomal subunits to the site of 80S ribosome assembly is indicated by black arrows. In the literature on poliovirus and viruses with similar internal ribosome entry sites, it has usually been assumed that the distance between the two AUG codons is covered by scanning; however, recently published data (33) could also be interpreted in terms of a sort of shunt mechanism. EMCV, encephalomyocarditis virus.

downstream translation, while the same or a very similar sORF had only a moderate inhibitory effect when starting at sites that are reached by scanning (Fig. 5C and D).

The process that delivers ribosomes to the end of the RTBV RNA leader is active in protoplasts from at least two, quite unrelated plant species, albeit with different efficiencies. It remains to be determined which parts of the leader sequence are required for the RTBV ribosome shunt. A preliminary deletion analysis detected such sequences both in the 5′ and in the 3′ leader regions (11; unpublished observations). From the data presented here, it appears that there is a region downstream of the ATT start codon which is not absolutely required but which is involved in tuning the efficiency of ribosome delivery (Fig. 6). A positive influence of sequences downstream of the start codon on ribosome delivery and/or initiation has been described for a CTG initiation event on the *Drosophila* E74A mRNA (9) and has also been discussed for internal ribosome entry on picornaviral RNA (16, 42). For cardioviruses, the assumption could not be verified in in vitro translation experiments (38); however, like other features of picornaviral translation (1), such effects may depend on certain cellular parameters. Mutations in the region downstream of the internal ribosome entry site of poliovirus have a complex and cell type-dependent influence on downstream translation; however, in this case, the effect is not on ribosome binding but rather on transfer of ribosomes from the entry point to the initiation region by scanning (4, 41, 56) or, as could be concluded from recently published data (33), by a shunt process in some cell types.

We note that the supposed ribosome acceptor site of the RTBV pregenomic RNA resembles in part the internal ribosome entry site of picornaviral RNAs in containing a pyrimidine stretch and a potential start codon, although the spacing of 17 nucleotides between the first pyrimidine and the A of the ATT is smaller than in most picornaviruses (Fig. 10). For CaMV, the shunt acceptor near the end of the leader has been localized to a region between 50 and 65 nucleotides upstream of the ATG of the first long ORF (22). This region contains a potential non-ATG start codon in optimal downstream sequence context positioned downstream of a pyrimidine-rich sequence (Fig. 10). It remains to be established whether the structure of this element is indeed important for the CaMV ribosome shunt.

The particular translation mechanism for RTBV ORF I did not support a similarly efficient expression of an ORF starting with a variety of other non-ATG codons in the respective RNA region. It remains to be investigated whether this is due to the slightly different positions of these alternative start codons or to deviations from the particularly favorable sequence context of the ORF I ATT codon. This context for the ATT codon on the RTBV RNA and on all of our reporter constructs includes not only the classical consensus nucleotides A and G at positions −3 and +4, respectively (44), but also the A at position +5 which was recently shown to have a strong stimulatory effect on initiation at non-ATG start codons (6, 27).

The proposed shunt mechanism with the shunt acceptor between the +686 and +698 ATT codons would also explain the infectivity results with the mutants pRTRB-ATG686 (+686 ATT→ATG) and pRTRB-ATG698 (+698 ATT→ATG). The +686 ATG codon would be bypassed by the shunt, whereas the +698 ATG codon would efficiently recruit ribosomes and would strongly reduce leaky scanning beyond +698. A peculiarity of the RTBV sequence suggests that the absence of a proper start codon for ORF I, besides possibly allowing particular control mechanisms for ORF I expression, is also of importance for the expression of the RTBV ORFs II and III. It is notable that the 1000 nucleotides of sequence encompassing the RTBV ORFs I and II contain only a single ATG codon. This codon occurs in an unfavorable sequence context for translation initiation and opens ORF II. Initiation at that codon would not be much more efficient than at the ATT codon of ORF I. Such an arrangement is suitable for a leaky scanning mechanism for expression of the ORFs II and III. Parallel work (to be published elsewhere) has shown that the low initiation potential of the ORF I start codon is crucial for expression of ORFs II and III. Mutation of the ATT codon that we identified as the ORF I start codon to ATG not only increased the efficiency of ORF I translation but also reduced translation of ORFs II and III (16a). It remains to be seen whether the effect of the mutation of the +698 ATT codon to ATG on virus infectivity is caused by overexpression of ORF I or by interference with expression of the downstream ORFs.

ACKNOWLEDGMENTS

We are grateful to M. Müller for excellent supplies of protoplasts, to G. Chen for discussions and construct CIC-L^m, to K. Ullrich for discussions, and to H. Rothnie for comments on early versions of the manuscript and for help with the protection assays.

RTBV was held at the John Innes Centre under MAFF license no. PHF1419/473/53, and the genetic manipulations were carried out under MAFF license no. PHF1419/304/53.

REFERENCES

1. **Agol, V. I.** 1991. The 5′-untranslated region of picornaviral genomes. Adv. Virus Res. **40:**103–180.
2. **Bao, Y., and R. Hull.** 1993. Mapping the 5′-terminus of rice tungro bacilliform viral genomic RNA. Virology **197:**445–448.
3. **Beames, S., S. Braunagel, M. D. Summers, and R. E. Lanford.** 1991. Polyhedrin initiator codon altered to AUU yields unexpected fusion protein from a baculovirus vector. BioTechniques **11:**378–383.

4. **Belsham, G. J.** 1992. Dual initiation sites of protein synthesis on foot-and-mouth disease virus RNA are selected following internal entry and scanning of ribosomes in vivo. EMBO J. **11:**1105–1110.
5. **Bevan, M.** 1984. Binary *Agrobacterium* vectors for plant transformation. Nucleic Acids Res. **12:**8711–8721.
6. **Boeck, R., and D. Kolakofski.** 1994. Positions +5 and +6 can be major determinants of the efficiency of non-AUG initiation codons for protein synthesis. EMBO J. **13:**3608–3617.
7. **Bonneville, J.-M., H. Sanfaçon, J. Fütterer, and T. Hohn.** 1989. Posttranscriptional transactivation in cauliflower mosaic virus. Cell **59:**1135–1143.
8. **Bouhida, M., B. E. L. Lockhart, and N. E. Olszewski.** 1993. An analysis of the complete sequence of a sugarcane bacilliform virus genome infectious to banana and rice. J. Gen. Virol. **74:**15–22.
9. **Boyd, L., and C. S. Thummel.** 1993. Selection of CUG and AUG initiator codons for *Drosophila* E74A translation depends on downstream sequences. Proc. Natl. Acad. Sci. USA **90:**9164–9167.
10. **Carroll, R., and D. Derse.** 1993. Translation of equine infectious anemia virus bicistronic *tat-rev* mRNA requires leaky ribosome scanning of the *tat* CTG initiation codon. J. Virol. **67:**1433–1440.
11. **Chen, G., M. Müller, I. Potrykus, T. Hohn, and J. Fütterer.** 1994. Rice tungro bacilliform virus: transcription and translation in protoplasts. Virology **204:**91–100.
12. **Chen, X., M. Chamorro, S. I. Lee, L. X. Shen, J. V. Hines, I. Tinoco, Jr., and H. E. Varmus.** 1995. Structural and functional studies of retroviral RNA pseudoknots involved in ribosomal frameshifting: nucleotides at the junction of the two stems are important for efficient ribosomal frameshifting. EMBO J. **14:**842–852.
13. **Clements, J. M., T. M. Laz, and F. Sherman.** 1988. Efficiency of translation initiation by non-AUG codons in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **8:**4533–4536.
14. **Dasgupta, I., R. Hull, S. Eastop, C. Poggi-Pollini, M. Blakebrough, M. I. Boulton, and J. W. Davies.** 1991. Rice tungro bacilliform virus DNA independently infects rice after *Agrobacterium*-mediated transfer. J. Gen. Virol. **72:**1215–1221.
15. **Datta, S. K., A. Peterhans, K. Datta, and I. Potrykus.** 1990. Genetically engineered fertile Indica-rice recovered from protoplasts. Bio/Technology **8:** 736–740.
16. **Davies, M. V., and R. J. Kaufman.** 1992. The sequence context of the initiation codon in the encephalomyocarditis virus leader modulates efficiency of internal translation initiation. J. Virol. **66:**1924–1932.

16a. **Fütterer, J.** Unpublished observations.

17. **Fütterer, J., K. Gordon, J. M. Bonneville, H. Sanfaçon, B. Pisan, J. Penswick, and T. Hohn.** 1988. The leading sequence of caulimovirus large RNA can be folded into a large stem-loop structure. Nucleic Acids Res. **16:** 8377–8390.
18. **Fütterer, J., K. Gordon, P. Pfeiffer, H. Sanfaçon, B. Pisan, J. M. Bonneville, and T. Hohn.** 1989. Differential inhibition of downstream gene expression by the cauliflower mosaic virus 35S RNA leader. Virus Genes **3:**45–55.
19. **Fütterer, J., K. Gordon, H. Sanfaçon, J. M. Bonneville, and T. Hohn.** 1990. Positive and negative control of translation by the leader sequence of cauliflower mosaic virus pregenomic 35S RNA. EMBO J. **9:**1697–1707.
20. **Fütterer, J., and T. Hohn.** 1991. Translation of a polycistronic mRNA in the presence of the cauliflower mosaic virus transactivator protein. EMBO J. **10:** 3887–3896.
21. **Fütterer, J., and T. Hohn.** 1992. Role of an upstream open reading frame in the translation of polycistronic mRNAs in plant cells. Nucleic Acids Res. **20:** 3851–3857.
22. **Fütterer, J., Z. Kiss-Làszlò, and T. Hohn.** 1993. Non-linear ribosome migration on cauliflower mosaic virus 35S RNA. Cell **73:**789–802.
23. **Fütterer, J., I. Potrykus, M. P. Valles Brau, I. Dasgupta, R. Hull, and T. Hohn.** 1994. Splicing in a plant pararetrovirus. Virology **198:**663–670.
24. **Gordon, K., J. Fütterer, and T. Hohn.** 1992. Efficient initiation of translation at non-AUG triplets in plant cells. Plant J. **2:**809–813.
25. **Gowda, S., H. B. Scholthof, F. C. Wu, and R. J. Shepherd.** 1991. Requirement of gene VII in *cis* for the expression of downstream genes on the major transcript of figwort mosaic virus. Virology **185:**867–871.
26. **Gowda, S., F. C. Wu, H. B. Scholthof, and R. J. Shepherd.** 1989. Gene VI of figwort mosaic virus (caulimovirus group) functions in posttranscriptional expression of genes on the full-length RNA transcript. Proc. Natl. Acad. Sci. USA **86:**9203–9207.
27. **Grünert, S., and R. J. Jackson.** 1994. The immediate downstream codon strongly influences the efficiency of utilization of eukaryotic translation initiation codons. EMBO J. **9:**3618–3630.
28. **Hagen, L. S., M. Jacquemont, A. Lepingle, H. Lot, and M. Tepfer.** 1993. Nucleotide sequence and genomic organization of cacao swollen shoot virus. Virology **196:**619–628.
29. **Hann, S. R., K. Sloan-Brown, and G. D. Spotts.** 1992. Translational activation of the non-AUG-initiated c-*myc* 1 protein at high cell densities due to methionine deprivation. Genes Dev. **6:**1229–1240.
30. **Hatfield, D. L., J. G. Levin, A. Reim, and S. Oroszlan.** 1992. Translational suppression in retroviral gene expression. Adv. Virus Res. **41:**193–239.
31. **Hay, J., F. Grieco, A. Druka, M. Pinner, S.-C. Lee, and R. Hull.** 1994. Detection of rice tungro bacilliform virus gene products *in vivo*. Virology **205:**430–437.
32. **Hay, J. M., M. C. Jones, M. L. Blackebrough, I. Dasgupta, J. W. Davies, and R. Hull.** 1991. An analysis of the sequence of an infectious clone of rice tungro bacilliform virus, a plant pararetrovirus. Nucleic Acids Res. **19:**2615–2621.
33. **Hellen, C. U. T., T. V. Pestova, and E. Wimmer.** 1994. Effect of mutations downstream of the internal ribosome entry site on initiation of poliovirus protein synthesis. J. Virol. **68:**6312–6322.
34. **Hohn, T., and J. Fütterer.** 1991. Pararetroviruses and retroviruses: a comparison of expression strategies. Semin. Virol. **2:**55–70.
35. **Hohn, T., and J. Fütterer.** 1992. Transcriptional and translational control of gene expression in cauliflower mosaic virus. Curr. Opin. Genet. Dev. **2:** 90–96.
36. **Hsiao, K.** 1991. A fast and simple procedure for sequencing double stranded DNA with Sequenase. Nucleic Acids Res. **19:**2787.
37. **Hull, R.** 1992. Genome organization of retroviruses and retroelements: evolutionary considerations and implications. Semin. Virol. **3:**373–382.
38. **Hunt, S. R., A. Kaminski, and R. J. Jackson.** 1993. The influence of viral coding sequences on the efficiency of internal initiation of translation of cardiovirus RNAs. Virology **197:**801–807.
39. **Jang, S. K., T. V. Pestova, C. U. T. Hellen, G. W. Witherell, and E. Wimmer.** 1990. Cap-independent translation of picornavirus RNAs: structure and function of the internal ribosome entry site. Enzyme **44:**292–309.
40. **Kaminski, A., G. J. Belsham, and R. J. Jackson.** 1994. Translation of encephalomyocarditis virus RNA: parameters influencing the selection of the internal initiation site. EMBO J. **13:**1673–1681.
41. **Kaminski, A., M. T. Howell, and R. J. Jackson.** 1990. Initiation of encephalomyocarditis virus RNA translation: the authentic start site is not selected by a scanning mechanism. EMBO J. **9:**3753–3759.
42. **Kaminski, A., S. L. Hunt, C. L. Gibbs, and R. J. Jackson.** 1994. Internal initiation of mRNA translation in eukaryotes. Genet. Eng. **16:**115–155.
43. **Kozak, M.** 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell **44:** 283–292.
44. **Kozak, M.** 1989. The scanning model for translation: an update. J. Cell. Biol. **108:**229–241.
45. **Kozak, M.** 1989. Context effects and inefficient initiation at non-AUG codons in eukaryotic cell free translation systems. Mol. Cell. Biol. **9:**5073–5080.
46. **Kozak, M.** 1989. Circumstances and mechanisms of inhibition of translation by secondary structure in eukaryotic mRNAs. Mol. Cell. Biol. **9:**5134–5142.
47. **Kozak, M.** 1990. Downstream secondary structure facilitates recognition of initiator codons by eukaryotic ribosomes. Proc. Natl. Acad. Sci. USA **87:** 8301–8305.
48. **Kozak, M.** 1992. Regulation of translation in eukaryotic systems. Annu. Rev. Cell Biol. **8:**197–225.
49. **LaTeana, A., C. L. Pon, and C. O. Gualerzi.** 1993. Translation of mRNAs with degenerate triplet AUU displays high initiation factor 2 dependence and is subject to initiation factor 3 repression. Proc. Natl. Acad. Sci. USA **90:** 4161–4165.
50. **Lütcke, H. A., K. C. Chow, F. S. Mickel, K. A. Moss, H. F. Kern, and G. A. Scheele.** 1987. Selection of AUG codons differs in plants and animals. EMBO J. **6:**43–48.
51. **Medberry, S. L., B. E. L. Lockhart, and N. Olszewski.** 1990. Properties of commelina yellow mottle virus's complete DNA sequence, genomic discontinuities and transcript suggest that it is a pararetrovirus. Nucleic Acids Res. **18:**5505–5513.
52. **Mehdi, H., E. Ono, and K. C. Gupta.** 1990. Initiation of translation at CUG, GUG and ACG codons in mammalian cells. Gene **91:**173–178.
53. **Nagel, R., A. Elliot, A. Masel, R. G. Birch, and J. M. Manners.** 1990. Electroporation of binary Ti plasmid vector into *Agrobacterium tumefaciens* and *Agrobacterium rhizogenes*. FEMS Microbiol. Lett. **67:**325–328.
54. **Nivinskas, R., R. Vaišknaite, and A. Raudonikiene.** 1992. An internal AUU codon initiates a smaller peptide encoded by bacteriophage T4 baseplate gene 26. Mol. Gen. Genet. **232:**257–261.
55. **Peabody, D. D.** 1989. Translation initiation at non-AUG triplets in mammalian cells. J. Biol. Chem. **264:**5031–5035.
56. **Pelletier, J., and N. Sonenberg.** 1988. Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. Nature (London) **334:**320–325.
57. **Qu, R., M. Bhattacharyya, G. S. Laco, A. de Kochko, S. Rao, M. B. Kaniewska, S. Elmer, D. E. Rochester, C. E. Smith, and R. N. Beachy.** 1991. Characterization of the genome of rice tungro bacilliform virus: comparison with commelina yellow mottle virus and caulimoviruses. Virology **185:**354–364.
58. **Rohde, W., A. Gramstat, J. Schmitz, E. Tacke, and D. Prüfer.** 1994. Plant viruses as model systems for the study of non-canonical translation mechanisms in higher plants. J. Gen. Virol. **75:**2142–2149.
59. **Rothnie, H. M., Y. Chapdelaine, and T. Hohn.** 1994. Pararetroviruses and retroviruses: a comparative review of viral structure and gene expression strategies. Adv. Virus Res. **44:**1–67.

60. **Sacerdot, C., G. Fayat, P. Dessen, M. Springer, J. A. Plumbridge, M. Grunberg-Manago, and S. Blanquet.** 1982. Sequence of a 1.26-kb DNA fragment containing the structural gene for *E. coli* initiation factor IF3: presence of an AUU initiator codon. EMBO J. **1:**311–315.
61. **Scheper, G. C., H. O. Voorma, and A. A. M. Thomas.** 1994. Basepairing with 18S ribosomal RNA in internal initiation of translation. FEBS Lett. **352:**271–275.
62. **Scholthof, H. B., S. Gowda, F. C. Wu, and R. J. Shepherd.** 1992. The full-length transcript of caulimovirus is a polycistronic mRNA whose genes are transactivated by the product of gene VI. J. Virol. **66:**3131–3139.
63. **Scholthof, H. B., F. C. Wu, S. Gowda, and R. J. Shepherd.** 1992. Regulation of caulimovirus gene expression and the involvement of cis-acting elements on both viral transcripts. Virology **190:**403–412.
64. **Yin, Y., and R. N. Beachy.** 1995. The regulatory regions of the rice tungro bacilliform virus promoter and interacting nuclear factors in rice (*Oryza sativa* L.). Plant J. **7:**969–980.
65. **Zijlstra, C., and T. Hohn.** 1992. Cauliflower mosaic virus gene VI controls translation from dicistronic expression units in transgenic Arabidopsis plants. Plant Cell **4:**1471–1484.