

The nucleic acid-binding zinc finger protein of potato virus M is translated by internal initiation as well as by ribosomal frameshifting involving a shifty stop codon and a novel mechanism of P-site slippage

Andrea Gramstat*, Dirk Prüfer and Wolfgang Rohde

Max-Planck-Institut für Züchtungsforschung, Carl-von-Linné-Weg 10, D-50829 Köln, Germany

Received June 16, 1994; Revised and Accepted August 23, 1994

ABSTRACT

The genes for the capsid protein CP and the nucleic acid-binding 12K protein (pr12) of potato virus M (PVM) constitute the 3' terminal gene cluster of the PVM RNA genome. Both proteins are presumably translated from a single subgenomic RNA. We have identified two translational strategies operating in pr12 gene expression. Internal initiation at the first and the second AUG codon of the pr12 coding sequence results in the synthesis of the 12K protein. In addition the protein is produced as a CP/12K transframe protein by ribosomal frameshifting. For these studies parts of the CP and pr12 coding sequences including the putative frameshift region were introduced into an internal position of the β -glucuronidase gene. Mutational analyses in conjunction with *in vitro* translation experiments identified a homopolymeric string of four adenosine nucleotides which together with a 3' flanking UGA stop codon were required for efficient frameshifting. The signal AAAAUGA is the first frameshift signal with a shifty stop codon to be analyzed in the eukaryotic system. Substitution of the four consecutive adenosine nucleotides by UUUU increased the efficiency of frameshifting, while substitution by GGGG or CCCC dramatically reduced the synthesis of the transframe protein. Also, UAA and UAG could replace the opal stop codon without effect on the frameshifting event, but mutation of UGA to the sense codon UGG inhibited transframe protein formation. These findings suggest that the mechanism of ribosomal frameshifting at the PVM signal is different from the one described by the 'simultaneous slippage' model in that only the string of four adenosine nucleotides represents the slippery sequence involved in a -1 P-site slippage.

INTRODUCTION

The genome of the carlavirus PVM (potato virus M) consists of a single-stranded RNA molecule of 8534 nucleotides (1) and

contains six major open reading frames (ORFs; Fig. 1A). These are divided into three clusters separated by small intergenic regions in which the first block from the 5' end is represented by a continuous ORF for a 223 kDa protein with sequence motifs characteristic for viral replicases. The second block coding for three proteins of molecular weights 25, 12 and 7 kDa, respectively, shows homology to the 'triple gene block' present also in the hordei-, furo-, carla- and potexvirus genomes (2) as well as in *Nicotiana velutina* mosaic virus (3). Its gene products are apparently involved in opening plasmodesmata between neighbouring cells (4) and promoting cell-to-cell movement of virus or viral RNA/protein complexes and spread of the virus from the initial site of infection as recently demonstrated for the hordeiviruses (5). The third gene cluster is separated by an intergenic region of 21 nucleotides from the triple gene block and encodes the 34 kDa capsid protein CP and a 12 kDa protein (pr12) of yet unknown function. This protein contains a zinc finger motif and displays properties of a nucleic acid-binding protein (6).

Expression strategies for the carlavirus genome have only recently attracted attention. Host plants infected with potato carlavirus S (PVS) contain a 1.3 kb subgenomic (sg) RNA that is colinear with the 3' end of PVS genomic RNA (7). It directs the synthesis of the 33 kDa PVS coat protein as well as a minor 44 kDa polypeptide which is immunoprecipitable by CP antisera and could possibly represent a CP/pr11 transframe protein. Although a PVM sgRNA of similar size has not yet been described, we decided to study the translational mechanism(s) operating in the synthesis of pr12 from a mRNA in which the 12K ORF was located at an internal position comparable to that in the putative PVM sgRNA.

Here we present evidence by *in vitro* transcription/translation experiments that two different translational mechanisms, namely internal initiation at pr12 AUG translational start codons and ribosomal frameshifting, operate in the expression of the 12K ORF to form the 12K protein and a CP/12K transframe protein, respectively. In contrast to the generally accepted heptanucleotide signal mediating ribosomal frameshifting, the slippery sequence on which the PVM CP/12K frameshift occurs consists of only

*To whom correspondence should be addressed

four adenosine nucleotides which cooperate with a stop codon. This is the first time that a detailed analysis of a frameshift event in eukaryotes has been investigated that involves a nucleotide string and a shifty stop codon.

MATERIALS AND METHODS

Synthesis of chimeric PVM/GUS constructs

For the replacement of the internal *EcoRV* fragment of the β -glucuronidase (GUS) gene (8) by PVM cDNA sequences, the GUS gene was isolated after restriction of pRT104GUS DNA (9) by *XhoI* and *EcoRI* and subcloned in sense orientation into the *SmaI* site of the transcription vector pSP65 (10) to yield pSP65GUS. Digestion with *EcoRV* removed an internal 230 bp fragment (coordinates 584 to 814 of the wildtype GUS gene), which was replaced by the 530 bp *BamHI/SalI* fragment of PVM cDNA clone pvmB (6). This fragment was treated with Klenow DNA polymerase to produce flush ends and ligated into *EcoRV*-restricted pSP65GUS to produce pFS-WT in which CP and the 12K coding sequences were in frame with the flanking GUS sequences. The integrity of this and all other constructs was verified by sequence analysis.

Site-directed mutagenesis

A 1324 bp PCR fragment (6) corresponding to the sequence from position 7210 to 8534 of the PVM RNA genome (1) and containing the complete CP and 12K ORFs was cloned between CaMV 35S promoter and terminator of the vector pRT100 (9). The promoter/CP + 12K/terminator fragment isolated from this construct after *HindIII* digestion was inserted into the *HindIII* site of the vector pSELECT (Promega) and transformed into the *E. coli* strain JM109 (11). After infection with the helper phage RK408, single-stranded template DNA for mutagenesis was isolated and used for site-directed mutagenesis according to the Promega protocol. Mutations were verified by sequence analysis, appropriate plasmid DNAs were digested with *BamHI/SalI* (except for mutant pFS-mut1 for which *BamHI* and *XhoI* were used), and the isolated DNA fragments were cloned into the *EcoRV*-restricted vector pSP65GUS as described above.

In vitro transcription/translation experiments

Single-stranded RNAs were obtained by *in vitro* transcription of *HindIII*-linearized pSP65GUS constructs by SP6 RNA polymerase in the presence of the cap analogue m⁷GpppG according to established protocols (10). RNAs were translated in the rabbit reticulocyte system (Amersham, Braunschweig) in the presence of ³⁵S-labelled methionine and translational products were analysed on SDS-containing 10% polyacrylamide gels as described previously (12).

Establishment and analysis of transgenic potato lines

Transgenic potato lines were obtained by *Agrobacterium tumefaciens*-mediated leaf disc transformation according to established protocols (13). Expression of the PVM transgenes was verified by Western analysis using monospecific antisera to bacterially expressed CP and pr12 fusion proteins (6, 13).

RESULTS

Identification of translational mechanisms for pr12 synthesis

For a detailed analysis of translational mechanisms a 530 bp fragment was released from the PVM cDNA clone pvmB (6)

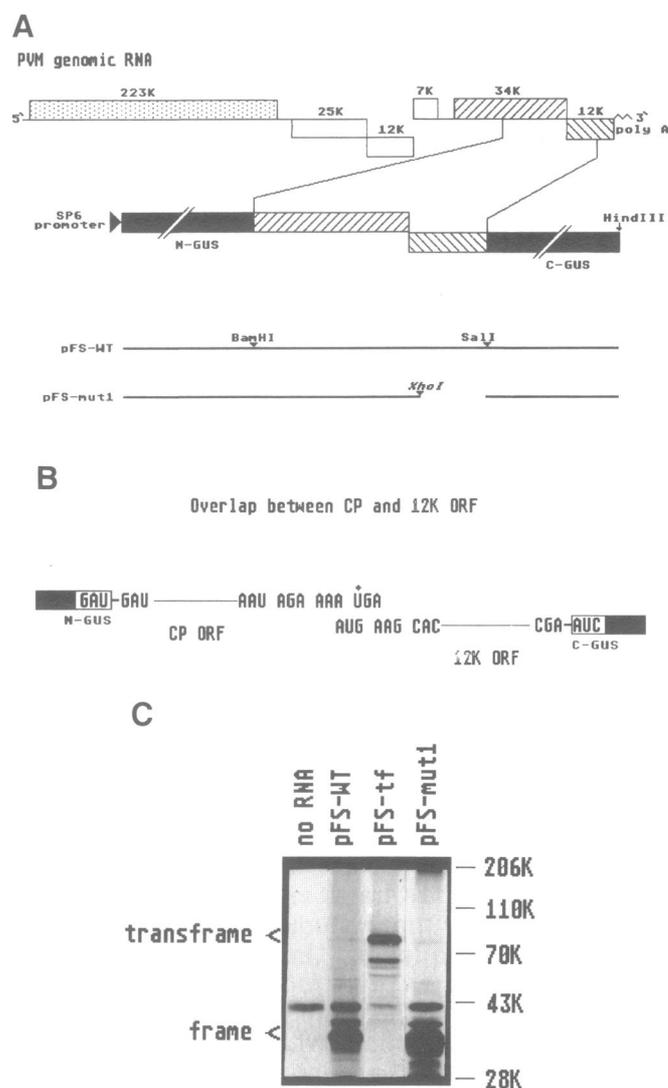


Figure 1. Analysis of the CP/12K frameshift. **A.** Construction of the wildtype template for *in vitro* transcription/translation experiments. The indicated *BamHI/SalI* cDNA fragment corresponding to coordinates 7716 to 8245 of the PVM RNA genome (1) was translationally fused at its N- and C-termini to the GUS gene (8) in the internal *EcoRV* site to yield pFS-WT in the transcription vector pSP65. The deletion mutant pFS-mut1 was obtained by site-directed mutagenesis and *XhoI* restriction as described in the Materials and Methods section. Solid lines represent sequences retained in the constructs. **B.** Overlap region of CP and 12K ORFs. The CP UGA termination codon is indicated by an asterisk. GUS coding sequences (N-GUS and C-GUS) are represented by black areas. The boxed sequences GUA and AUC represent the 5' and 3' termini of the internal GUS *EcoRV* sites. **C.** PAGE analysis of products obtained by *in vitro* translation of RNAs transcribed from *HindIII*-linearized pFS-WT, pFS-tf and pFS-mut1 template DNAs by SP6 RNA polymerase. RNAs were translated in the rabbit reticulocyte system in the presence of [³⁵S]methionine, the products were separated on a 10% SDS/polyacrylamide gel and detected by fluorography. Frame: the main translational product GUS-N/CP-C terminating at the CP ORF stop codon; transframe: the GUS-N/CP-C/12K-N/GUS-C transframe protein. The 42K protein represents the product of endogenous mRNA activity in the translation system (lane no RNA). Lane pFS-tf: translational products of a mutant RNA in which by insertion of a single nucleotide between CP ORF and 12K ORF a continuous GUS-N/CP-C/12K-N/GUS-C ORF was created; lane pFS-mut1: translational products of a mutant RNA in which all sequences 3' of the CP UGA stop codon were deleted except for the flanking 13 nucleotides.

by *BamHI/SalI* restriction. This fragment contains 423 bp corresponding to the CP C-terminus (CP-C) and the N-terminal 108 bp of the 12K ORF (12K-N). It was cloned into the GUS

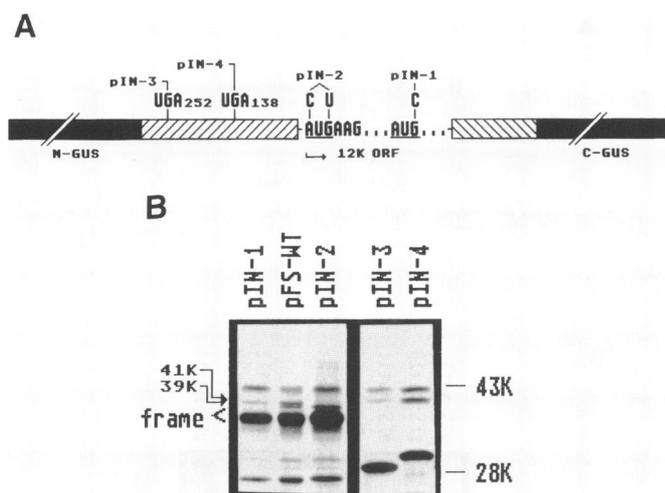


Figure 2. Expression of the 12K ORF by internal initiation. **A.** Internal initiation occurs at the first and the second AUG codon and is independent of upstream stop codons. The first and second AUG codon of the 12K ORF were mutated into CUU (pIN-2) and AUC (pIN-1), respectively. Furthermore, additional UGA stop codons were introduced into the CP ORF 252 (pIN-3) or 138 (pIN-4) nucleotides upstream of the original CP stop codon. **B.** Analysis of *in vitro* translations of RNAs from the mutants described in **A.** PAGE was performed as described in the legend to Figure 1C. Only enlargements of the lower portions of the gels are shown. Lanes pIN-3 and pIN-4 show truncated GUS-N/CP-C proteins as expected.

gene replacing an internal *EcoRV* GUS fragment as described previously (12) to yield construct pFS-WT in the transcription vector pSP65 (Fig. 1A). In this construct fusion of the N-proximal half of the GUS gene (GUS-N) to CP-C resulted in a continuous GUS-N/CP-C ORF for a 37 kDa protein, while the 12K N-terminus was in frame with the C-proximal half of the GUS gene (GUS-C) and formed a 41 kDa 12K-N/GUS-C protein. Ribosomal frameshifting from PVM CP to PVM 12K coding sequences that could possibly occur in the CP/12K overlap region (Fig. 1B) would result in a GUS-N/CP-C/12K-N/GUS-C transframe protein with an approximate molecular weight of 78 kDa.

An analysis by polyacrylamide gel electrophoresis (PAGE) of *in vitro* translation products from RNA transcribed from linearized pFS-WT plasmid DNA by SP6 RNA polymerase is shown in Fig. 1C. The most prominent translational product was the 37K chimeric GUS-N/CP-C protein (Fig. 1C, frame, lane pFS-WT). In the upper part of the gel a minor protein band showed an apparent molecular weight of 78 kDa as expected for the putative transframe protein (Fig. 1C, transframe, lane pFS-WT). This protein band is the major translational product (Fig. 1C, transframe, lane pFS-tf), when a single adenosine nucleotide is inserted by site-directed mutagenesis immediately upstream of the CP UGA stop codon (Fig. 1B) into the frameshift signal (see below) to produce a continuous ORF GUS-N/CP-C/12K-N/GUS-C for the 78K protein. Additional translation products besides the 43 kDa protein which is due to endogenous activity of the reticulocyte lysate (Fig. 1C, lane no RNA) appeared in the 50 to 70 kDa region and could have been due to premature termination or internal initiation.

Further studies were directed towards the origin of a doublet band with an apparent molecular weight of approximately 40 kDa (Fig. 1C, lanes pFS-WT, pFS-tf; see also Fig. 2B), as the

appearance of a protein of similar size was expected, if internal translation initiation would operate on pFS-WT RNA to produce the chimeric 12K-N/GUS-C protein. For a detailed analysis, the first as well as the second pr12 AUG codon located 42 nucleotides downstream of the pr12 N-terminus were removed by site-directed mutagenesis of pFS-WT to yield mutants pIN-2 and pIN-1, respectively (Fig. 2A). Translation of the corresponding RNAs demonstrated that the 39K and 41K proteins observed for pFS-WT RNA (Fig. 1C; Fig. 2B, lane pFS-WT) were in fact synthesized by internal initiation on the first as well as the second pr12 AUG codon. PAGE analysis of translational products from pIN-1 RNA showed only the 41K band as expected for translational start at the first AUG codon, while this protein disappeared and was replaced by the 39K protein, when pIN-2 RNA was translated *in vitro* (Fig. 2B, lanes pIN-1, pIN-2).

The mechanism of internal initiation at these pr12 AUG codons was not due to termination/reinitiation, i.e. by ribosome disassembly at the CP UGA stop codon and immediate initiation complex reformation at pr12 AUGs. A first indication for other mechanisms like leaky scanning, internal entry of ribosomes or a combination of both (shunt mechanism) came from experiments with the transframe mutant pFS-tf in which the original CP stop codon was placed out of frame by a single nucleotide insertion to form a continuous GUS-N/CP-C/12K-N/GUS-C ORF and which still allowed formation of the 39K and 41K proteins (Fig. 1C, lane pFS-tf). To further substantiate this finding UGA stop codons were introduced into the CP coding frame 138 (pIN-4) or 252 (pIN-3) nucleotides upstream of the first pr12 AUG (Fig. 2A). The results of *in vitro* translation from the RNAs of these constructs as depicted in Fig. 2B argued for the presence of internal ribosome entry sites (IRES) or a shunt mechanism by which scanning ribosomes would bypass the artificially introduced stop codons.

The *in vitro* efficiencies of internal initiation as well as ribosomal frameshifting were calculated from radioactivities present in the 39K/41K doublet and the 78K transframe protein, corrected for background activities and different methionine contents of the proteins as described previously for PLRV frameshifting efficiency (12), and put in relation to the in-frame 37K GUS-N/CP-C protein. As calculated from four different determinations each, internal initiation occurred at a frequency of 3.8% and ribosomal frameshifting at a frequency of 0.3%. With respect to the low frequency of ribosomal frameshifting it has to be considered that transframe protein synthesis is determined in relation to the synthesis of the PVM capsid protein which is the most abundant viral protein in infected plants (see also Fig. 5).

Analysis of the frameshift site

Within the overlapping region between CP ORF and 12K ORF (Fig. 1B), the putative frameshift region, a heptanucleotide sequence reminiscent of the shifty signals described for animal (14) or plant viruses (12) was not present. The only related homopolymeric heptamer signal UAGAAAA contains an amber stop codon in the -1 frame. We, therefore, first investigated the possibility that the string of four adenosine nucleotides was required for ribosomal frameshifting.

As a first approach this string was mutated by replacing individual adenines by cytosine (Fig. 3A, mutants pFS-mut2 to pFS-mut5). Every single point mutation abolished the frameshift event (Fig. 3B, left panel) as also evident from overexposed autoradiograms (data not shown). Simultaneous

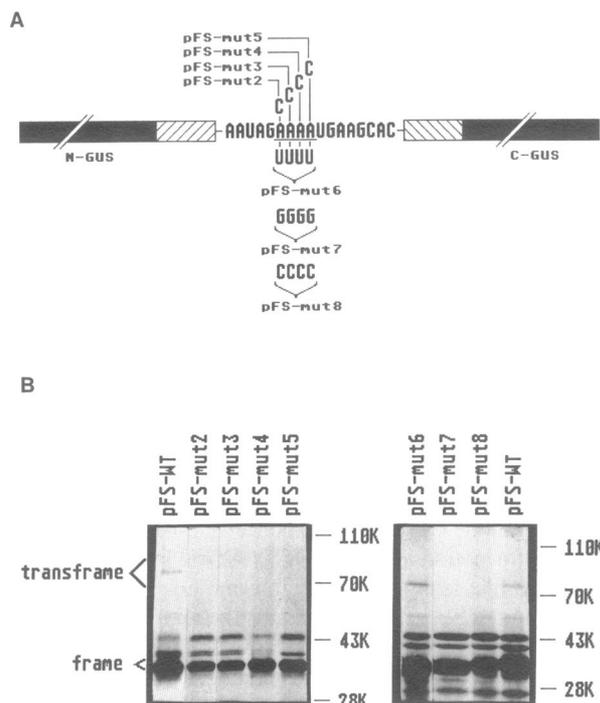


Figure 3. Analysis of the adenosine string. **A.** Individual adenosine nucleotides of the A string (coordinates 8135–8138 of the PVM RNA genome) were converted into cytidines by site-directed mutagenesis to yield mutants pFS-mut2, pFS-mut3, pFS-mut4, and pFS-mut5. By simultaneous substitution of all four adenosines by strings of uridines, guanines or cytidines the mutants pFS-mut6, pFS-mut7 and pFS-mut8 were obtained. **B.** Analysis of translation products was done as described in the legend to Figure 1C. Five independent determinations were done for the calculation of translation efficiencies for pFS-mut6 RNA. The standard deviation for the average value of 0.4% was ± 0.05 .

substitution of the four adenosine residues, however, by a uridine string as in mutant pFS-mut6 (Fig. 3A) not only maintained transframe protein synthesis, but led to an increase in frameshifting efficiency (from 0.3% to 0.4%, Fig. 3B). A drastic reduction was observed, when G (pFS-mut7) or C (pFS-mut8) strings replaced the A string (Fig. 3A, 3B). Even on overexposed autoradiograms the transframe protein was only visible as a minor signal (data not shown).

If the A string was part of a heptanucleotide frameshift signal, either the UAG stop codon in the -1 frame or the UGA stop codon in the zero frame would have to participate in this event (Fig. 1C). In this sequence (U AGA AAA UGA; the A string is underlined) the tRNAs involved in zero frame (CP) translation would base-pair with their respective anticodons 3' UCU 5' (tRNA^{Arg}) and 3' UUU 5' (tRNA^{Lys}). In case of a -1 frameshift to the 12K coding sequence (UAG AAA AUG), the amber stop codon UAG would occupy the peptidyl site of the ribosome (see also Fig. 6A). In order to study the putative involvement of this amber stop codon in the frameshift event, it was mutated to the opal stop codon UGA (pFS-mut9), the ochre stop codon UAA (pFS-mut10) and the sense codon UCG (pFS-mut11, Fig. 4A). All three mutations did not interfere with the synthesis of the transframe protein (Fig. 4B) suggesting that the UAG triplet was not involved in the frameshift event.

Similar experiments were conducted to study the effect of the zero (CP) frame UGA stop codon following the four adenosines

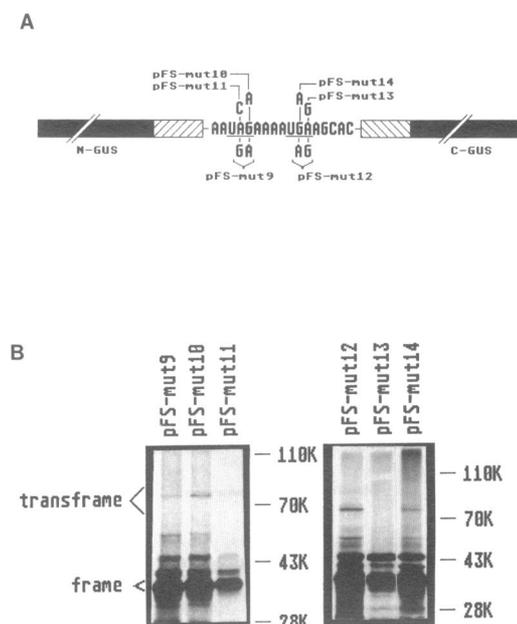


Figure 4. Analysis of stop codons flanking the adenosine string. **A.** The UAG amber stop codon in the -1 frame was mutated into the UGA (pFS-mut9) or UAA stop codon (pFS-mut10) or into the sense codon UCG (pFS-mut11). The CP ORF termination codon UGA was mutated into the amber stop codon UAG (pFS-mut12), the sense codon UGG (pFS-mut13) and ochre stop codon UAA (pFS-mut14). **B.** PAGE analysis of translation products was done as described in the legend to Figure 1C.

residues (Fig. 1B). In the sequence A AAA UGA the opal stop codon UGA was converted into the amber (UAG, pFS-mut12) or ochre stop codon (UAA, pFS-mut14) as well as into the UGG sense codon (pFS-mut13; Fig. 4A). In this latter case translation in the zero frame was terminated by a stop codon 6 nucleotides downstream of the original CP UGA stop codon. Synthesis of the transframe protein was observed with either stop codon-containing RNA (pFS-mut12, pFS-mut14, Fig. 4B, right panel), but translation of pFS-mut13 RNA in which the sense codon UGG had replaced the stop codon, did not produce the transframe protein. These results indicated that a stop codon following the A string is necessary for ribosomal frameshifting.

The two stop codons located 6 and 27 nucleotides, respectively, downstream of the CP stop codon in frame with the CP coding sequence were not required for frameshifting, as their individual conversion into sense codons had no effect on transframe protein synthesis (data not shown). Deletion mutants were produced to investigate the influence of other sequences located downstream of the AAAAUGA signal on frameshifting efficiency. In the mutant pFS-mut1 (Fig. 1A) an *Xho*I restriction site was created by site-directed mutagenesis 13 nucleotides downstream of the last coding triplet AAA of the CP ORF and used for the deletion of all 3' located pr12 sequences (Fig. 1A, pFS-mut1). As obvious from an analysis of the respective *in vitro* translation experiment (Fig. 1C), the deletion does not decrease or abolish the synthesis of the transframe protein.

Analysis of a putative *in vivo* CP/12K transframe protein

Attempts were made to identify a 46 kDa CP/12K transframe protein *in vivo*. Firstly, it was established that the CP/12K

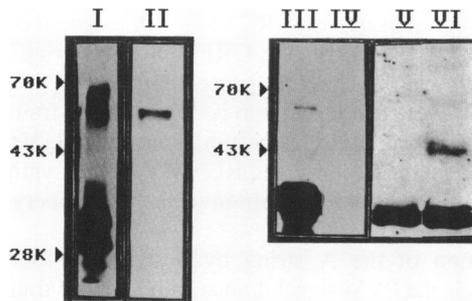


Figure 5. Western blot analysis of PVM proteins *in vivo*. Proteins from purified virus particles or protein extracts from PVM-infected or transgenic potato plants were separated by PAGE in SDS-containing gels, blotted onto nitrocellulose membranes and duplicates were incubated with CP- and 12K-specific antisera. Bound antibodies were reacted with goat anti-rabbit antisera/peroxidase conjugate and made visible with the ECL detection system. Lane I: Purified virus particles reacted with CP antiserum; II: virus particles reacted with 12K antiserum; III: extracts from PVM-infected potato plants reacted with CP antiserum; IV: extracts from non-infected potato plants reacted with CP antiserum; V: extracts from non-transformed potato plants reacted with 12K antiserum; VI: extracts from a transgenic potato plant expressing the inframe CP/12K sequence and reacted with 12K antiserum.

transframe protein was stable in the plant cell, as this 46 kDa protein was readily detected by 12K-specific antiserum in transgenic potato lines expressing the inframe CP/12K sequence (Fig. 5, lane VI). However, in transgenic potato lines expressing the wild-type CP/12K sequence, the amounts of 46K transframe protein were below the detection level (data not shown). This protein is also not detectable in protein extracts from PVM-infected plants or purified virus. Instead, both CP- and 12K-specific antisera reacted with a protein with an apparent molecular weight of 60–65 kDa which is present in extracts of PVM-infected plants as well as on PVM particles (Fig. 5, lanes I–III).

DISCUSSION

The nucleic acid-binding 12K protein is synthesized by internal initiation and ribosomal frameshifting

We have demonstrated here by *in vitro* experiments that the potato virus M 3'-terminal gene which encodes a 12K polypeptide with characteristics of a regulatory protein (6) is expressed by two different translational mechanisms. When the results obtained for the experimental system studied (insertion of CP/12K coding sequences into an internal position of the GUS gene) are applied to the actual situation in the PVM-infected potato plant, then the 12K protein would be synthesized in two different forms. While internal initiation at the first and second pr12 AUG would produce the original 12K protein, ribosomal frameshifting would direct the formation of a 46 kDa CP/12K transframe protein.

In view of the low translational efficiencies calculated from the *in vitro* experiments for the synthesis of CP/12K, it is not surprising that this protein is only barely detected in extracts from PVM-infected plants (Fig. 5). In addition, in virus-infected plants and on purified virus, the transframe protein did not show the expected apparent molecular weight of 46 kDa as for example in transgenic plants. Rather, polyclonal antisera raised against the bacterially expressed CP or 12K proteins (6) reacted with

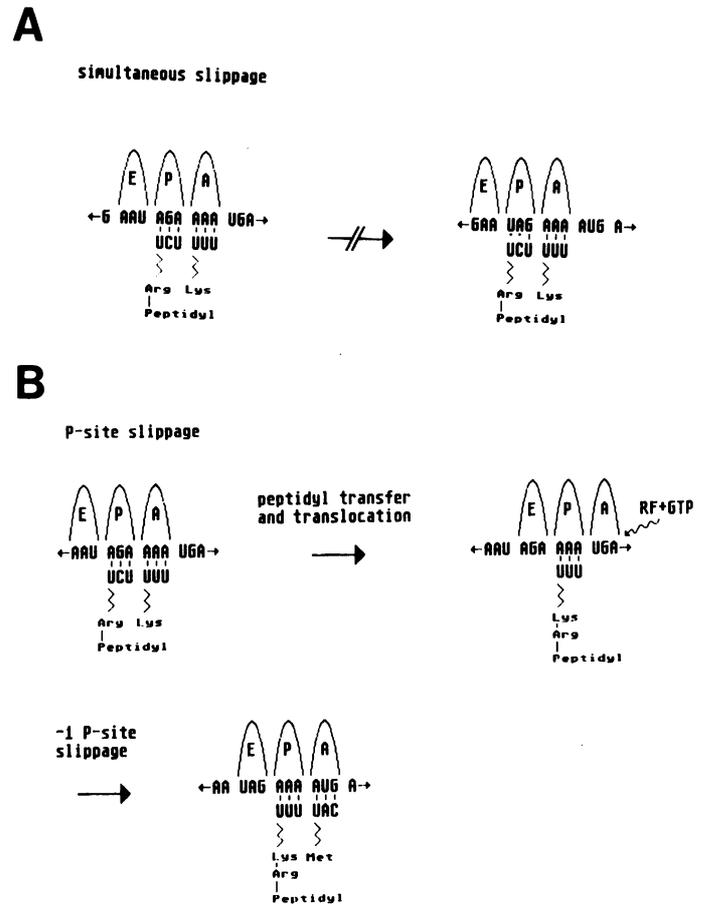


Figure 6. P-site slippage model proposed for CP/12K ribosomal frameshifting. **A.** Simultaneous slippage model according to Jacks *et al.* (28). Interaction of peptidyl-tRNA^{Arg} with the UAG stop codon in the -1 frame after a putative -1 ribosomal frameshift on the heptanucleotide sequence UAGAAA does not take place as shown in Fig. 4B. **B.** P-site slippage proposed for the CP/12K -1 ribosomal frameshift. Peptidyl-tRNA^{Lys} at the P-site awaits decoding of the UGA stop codon in the zero frame by GTP and release factors (RF). A -1 slippage at the P-site preserves optimal codon/anticodon interactions and leads to the synthesis of the CP/12K transframe protein.

a 60–65K protein. A putative 44K transframe protein has also been obtained for the carlavirus PVS with only minor amounts detectable by *in vitro* translation of PVS subgenomic RNA and immunoprecipitation with PVS CP-specific antiserum (7). It should be taken into account, however, that the synthesis of these proteins was determined in relation to the capsid protein CP, the most abundant protein in PVM- or PVS-infected cells. These preliminary results would indicate that ribosomal frameshifting also occurs *in vivo* and leads to the synthesis of a CP/12K transframe protein which accumulates on the virus particle and may be posttranslationally modified as a result of virus infection, as the inframe protein of 46 kDa is readily detected in protein extracts from appropriate transgenic plants (Fig. 5). As previously discussed (6), the resulting CP/12K transframe protein shows strong structural and functional homology (Cys–Cys zinc finger motif, nucleic acid-binding capacity) to the regulatory capsid proteins of tobacco streak and alfalfa mosaic virus (15). The capsid proteins of these viruses are necessary for infection by

forming specific protein/RNA interactions through a basic domain (16, 17, 18). By analogy the CP/12K transframe protein could possibly perform a similar role during initial stages of replication in the host plant or the virus-transmitting insect vector.

Internal initiation to produce the 12K protein occurs *in vitro* with comparable efficiency at the first and the second AUG of the 12K ORF. The mechanism that directs ribosomes to initiate at these internally located pr12 AUG(s) remains to be solved. Translation of CP/12K in-frame mutant RNA pFS-tf (Fig. 1C) indicated that the CP UGA stop codon is not involved, and excludes a termination/reinitiation mechanism for 12K ORF expression. Also the introduction of upstream stop codons into the CP coding sequence as in mutants pIN-3 and pIN-4 (Fig. 2A, 2B) had no effect on internal initiation and would argue against a simple leaky scanning mechanism by which ribosomes would have to bypass six internal AUG codons in the wildtype GUS-N/CP-C sequence before initiating at the pr12 AUG translational start codon(s). Rather, the internal initiation observed for pr12 translation may be controlled by internal ribosome entry sites (IRES) as for example in cowpea mosaic virus M RNA (19). Deletion analysis of the CP coding sequence in the artificial GUS-N/CP-C/12K-N/GUS-C construct (pFS-WT, Fig. 1A) will help to clarify this question.

A novel signal directs synthesis of the CP/12K transframe protein

Ribosomal frameshifting as a mechanism for the translation of a single protein from two or more overlapping ORFs has been described in detail for retroviral systems (14, 20, 21). The first examples for frameshifting in plants came from recent studies on the luteoviruses potato leafroll virus (PLRV, 12) and barley yellow dwarf virus (22). In all instances where a detailed analysis of the frameshift signal had been conducted, a homopolymeric heptamer described as the 'slippery sequence' was identified as the signal for -1 ribosomal frameshifting, and on the basis of consensus sequences four classes of signals were established. These include the signal sequences (the consensus sequences are underlined) AAAUUUA (Rous sarcoma virus, RSV; 23), AAAAAAC (mouse mammary tumor virus; 24), AAAUUUU (the 17.6 transposable element from *Drosophila*; 25), and UUUAAAU (PLRV; 12, 26). As a second type of information often required for a high efficiency of frameshifting (21), stem-loop structures downstream of the slippery sequence were identified as for example in RSV (23), infectious bronchitis coronavirus (27) or PLRV (12). Jacks and coworkers (23) suggested that the secondary structures slow down the ribosome thereby providing the opportunity for a -1 frameshift by 'simultaneous slippage' of peptidyl- and aminoacyl-tRNAs at the frameshift signal and decoding of a new set of codons (28).

This model for frameshifting cannot be the mechanism in the synthesis of the PVM CP/12K transframe protein, as a -1 frameshift on the A string will place the peptidyl-tRNA onto the UAG stop codon in the -1 frame (Fig. 6A), and mutations of this sequence did not interfere with frameshifting (Fig. 4A, 4B). Our data suggest that for the slippage to occur only the four adenosine nucleotides are responsible which form the frameshift window between the UAG stop codon in the -1 frame and the UGA stop codon in the zero frame. The A string could be replaced by a U string (Fig. 3A, 3B) and the UGA stop codon by UAG or UAA (Fig. 4A, 4B) without loss of frameshifting. Signals similar to the one analysed here have been observed in the genome of other viruses as well (14). The plant virus red

clover necrotic mosaic virus (RCNMV) for example contains the sequence UUUUUAG in the ORF1/ORF2 overlap region of RNA-1 (29) with a U string preceding the ORF1 UAG stop codon. However, this UAG stop codon 3' of the frameshift site does not participate in the frameshift. Rather, the U string is part of a shifty heptanucleotide, and RCNMV ORF2 synthesis as an ORF1/ORF2 transframe protein involves the slippery sequence G GAU UUU (30, 31).

Substitution of the A string by a string of four G or C nucleotides in the PVM signal dramatically reduced frameshifting, whereas the U string in the mutant pFS-mut6 even increased frameshifting (Fig. 3A, 3B). A comparable observation has been reported for the heptanucleotide signal G GGU UUA of the double-stranded L-A virus from yeast (32). Substitution of the G triplet by UUU gave the highest level of frameshifting and suggested that homopolymeric U sequences represent very 'slippery' sequences. Moreover, the second triplet UUU could only be substituted by AAA: efficient frameshifting was not observed with GGG and CCC.

It remains to be seen whether specific 'shifty' tRNAs are responsible for these different efficiencies. A remarkable example of how tRNAs can influence frameshifting is the synthesis of the γ -subunit of *E. coli* DNA polymerase III holoenzyme which is synthesized with 81% efficiency by -1 frameshifting at an A AAA AAG heptanucleotide signal. It is the absence in *E. coli* of a tRNA^{Lys} with a CUU anticodon necessary for optimal interaction with the AAG lysine codon that causes this high efficiency of frameshifting (33).

Involvement of a shifty stop codon

Evidence has been presented here for the involvement of a shifty stop codon in the ribosomal frameshifting event. All three nonsense codons were shown to be effective in directing frameshifting at the 3' end of the PVM CP coding sequence. Similar conclusions have been drawn by Weiss and coworkers (34) on the effectiveness of stop codons in *E. coli*. Their studies also confirmed that substitution of the stop codon by a sense codon dramatically decreased frameshifting, and a similar observation was made for the synthesis of the CP/12K transframe protein, when the UGA stop codon was replaced by the UGG sense codon in mutant pFS-mut13 (Fig. 4A, B). Also the two stop codons located 3' to the CP UGA stop did not exert any influence.

The exact mechanism by which ribosomes shift at the PVM A AAA UGA signal remains to be determined. A modified version of the simultaneous slippage model has been proposed (35) in which peptide bond formation takes place before slippage at the frameshift signal and where the ribosome provides a third site (exit site, E) besides the sites for peptidyl- (P) and aminoacyl-tRNAs (A; Fig. 6A). According to this model a downstream stem-loop structure may slow down translocation of the 3' codon within the frameshift signal from the A- to the P-site. Slippage then occurs with the growing polypeptide chain remaining stationary at the P-site, while the tRNAs transiently exist as hybrids in the E/P- and P/A-sites immediately after peptide bond formation.

Stem-loop structures that are located downstream of the frameshift signal and supposedly cause ribosomes to pause cannot play a role in the CP/12K frameshift, as the deletion of 3' sequences had no effect on frameshift efficiency. It has been suggested, however, that a stop codon may perform a function similar to that of stem-loop structures in that its decoding proceeds by a factor of 20 slower than that of a sense codon (34). Although

we cannot totally exclude at this stage ribosome hopping (36) during CP to 12K frameshifting, the involvement of every single nucleotide of the A string supports a -1 ribosomal frameshift mechanism. Therefore, with peptidyl-tRNA^{Lys} at the P-site, the stop codon in the PVM signal A AAA UGA would cause the ribosome to stall for the binding of release factors (RF) and GTP and release of the in-frame GUS-N/CP-C product (37). As in the -1 frame the A string is preceded by a UAG stop codon that does not participate in the frameshift, we propose that -1 slippage of the peptidyl-tRNA^{Lys} (awaiting decoding of the zero frame UGA stop codon at the A-site by release factors) occurs with low efficiency at the P-site as shown in Fig. 6B and allows the entry of met-tRNA^{Met} for the decoding of the AUG codon now present at the A-site. In fact the absence of frameshifting in pFS-mut13 (exchange of the UGA stop codon by the UGG sense codon) would support this kinetic interpretation of frameshifting versus termination regulation (38). UGG is the only sense codon for tryptophan, and four UGG codons are present in the PVM CP and 12K ORFs and a total of 20 UGG codons are found in the GUS-N/CP-C and 12K-N/GUS-C ORFs. Therefore, tRNA^{Trp} must be readily accessible *in vitro* and *in vivo*, and UGG decoding cannot be a rate-limiting step that would allow for frameshifting to occur. As a consequence, the mechanism of P-site slippage proposed to operate in the synthesis of the PVM CP/12K transframe protein involves the string of four A nucleotides in conjunction with an adjacent shifty stop codon.

ACKNOWLEDGEMENTS

The authors wish to thank Dr E. Tacke for useful discussions and A. Kaufmann for excellent technical assistance. This work was supported by grant Ro 330/6-2 of the Deutsche Forschungsgemeinschaft.

REFERENCES

- Zavriev, S.K., Kanyuka, K.V. and Levay, K.E. (1991) *J. Gen. Virol.* 72, 9–14.
- Morozov, S.Y., Dolja, V.V. and Atabekov, J.G. (1989) *J. Mol. Evol.* 9, 52–62.
- Randles, J. and Rohde, W. (1990) *J. Gen. Virol.* 71, 1019–1027.
- Beck, D.L., Guilford, P.J., Voot, D.M., Andersen, M.T. and Forster, R.L.S. (1991) *Virology* 183, 695–702.
- Petty, I.T.D., French, R., Jones, R.W. and Jackson, A.O. (1990) *EMBO J.* 9, 3453–3457.
- Gramstat, A., Courtzozanis, A. and Rohde, W. (1990) *FEBS Lett.* 276, 34–38.
- Foster, G.D. and Mills, P.R. (1990) *J. Gen. Virol.* 71, 1247–1249.
- Jefferson, R.A., Burgess, S.M. and Hirsch, D. (1986) *Proc. Natl. Acad. Sci. USA* 83, 8447–8451.
- Töpfer, R., Pröls, M., Schell, J. and Steinbiss, H.H. (1988) *Plant Cell Rep.* 7, 225–228.
- Melton, D.A., Krieg, P.A., Rebagliati, M.R., Maniatis, T., Zinn, K. and Green, M.R. (1984) *Nucleic Acids Res.* 12, 7035–7056.
- Hanahan, D. (1985) In Glover, D.M. (ed.), *DNA Cloning—A Practical Approach*. IRL Press, Oxford, Vol. I, pp. 109–135.
- Prüfer, D., Tacke, E., Schmitz, J., Kull, B., Kaufmann, A. and Rohde, W. (1992) *EMBO J.* 3, 1111–1117.
- Tacke, E., Schmitz, J., Prüfer, D. and Rohde, W. (1993) *Virology* 197, 274–282.
- Hatfield, D.L., Levin, J.G., Rein, A. and Oroszlan, S. (1992) *Adv. Virus Res.* 41, 193–239.
- Sehnke, P.C., Mason, A.M., Hood, S.J., Lister, R.M. and Johnson, J.E. (1989) *Virology* 168, 48–56.
- Bol, J.F., van Vloten-Doting, L. and Jaspars, E.M.J. (1971) *Virology* 46, 73–85.
- Van Vloten-Doting, L. (1975) *Virology* 65, 215–225.
- Jaspars, E.M.J. (1985) In Davies, J.W. (ed.), *Molecular Plant Virology*, CRC Press, Boca Raton, pp. 155–221.
- Verver, J., Le Gall, O., van Kammen, A. and Wellink, J. (1991) *J. Gen. Virol.* 72, 2339–2345.
- Hatfield, D.L., Smith, D.W.E., Lee, B.J. and Oroszlan, S. (1990) *Crit. Rev. Biochem. Mol. Biol.* 25, 71–96.
- Hatfield, D.L. and Oroszlan, S. (1990) *Trends Biochem. Sci.* 15, 186–190.
- Brault, V. and Miller, W.A. (1992) *Proc. Natl. Acad. Sci. USA* 89, 2262–2266.
- Jacks, T., Power, M.D., Masiarz, F.R., Luciw, P.A., Barr, P.J. and Varmus, H.E. (1988a) *Nature* 331, 280–283.
- Jacks, T., Townsley, K., Varmus, H.E. and Majors, J. (1987) *Proc. Natl. Acad. Sci. USA* 84, 4298–4302.
- Saigo, K., Kugimiya, W., Matsuo, Y., Inouye, S., Yoshioka, K. and Yuki, S. (1984) *Nature* 312, 659–661.
- Kujawa, A.B., Drugeon, G., Hulanicka, D. and Haenni, A.L. (1993) *Nucleic Acids Res.* 21, 2165–2171.
- Brierly, I., Digard, P. and Inglis, S.C. (1989) *Cell* 57, 537–547.
- Jacks, T., Madhani, H.D., Masiarz, F.R. and Varmus, H.E. (1988b) *Cell* 55, 447–458.
- Xiong, Z. and Lommel, S.A. (1989) *Virology* 171, 543–554.
- Xiong, Z., Kim, K.H., Kendall, T.L. and Lommel, S.A. (1993) *Virology* 193, 213–221.
- Kim, K.H. and Lommel, S.A. (1994) *Virology* 200, 574–582.
- Dinman, J.D., Icho, T. and Wickner, R.B. (1991) *Proc. Natl. Acad. Sci. USA* 88, 174–178.
- Tsuchihashi, Z. and Brown, P.O. (1992) *Genes & Dev.* 6, 511–519.
- Weiss, R.B., Dunn, D.M., Atkins, J.F. and Gesteland, R.F. (1990) *Progr. Nucl. Acids Res. Mol. Biol.* 39, 159–183.
- Weiss, R.B., Dunn, D.M., Shuh, M., Atkins, J.F. and Gesteland, R.F. (1989) *New Biologist* 1, 159–169.
- Atkins, J.F., Weiss, R.B. and Gesteland, R.F. (1990) *Cell* 62, 413–423.
- Caskey, C.T. (1980) *Trends Biochem. Sci.* 5, 234–237.
- Valle, R.P.C. and Morch, M.-D. (1988) *FEBS Lett.* 235, 1–15.