

Structural requirements for efficient translational frameshifting in the synthesis of the putative viral RNA-dependent RNA polymerase of potato leafroll virus

Alicja B.Kujawa^{1,2*}, Gabrièle Drugeon¹, Danuta Hulanicka² and Anne-Lise Haenni¹

¹Institut Jacques Monod, 2 Place Jussieu, Tour 43, 75251 Paris Cedex 05, France and ²Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Rakowiecka 36, 02-532 Warsaw, Poland

Received December 24, 1992; Revised and Accepted April 2, 1993

ABSTRACT

The putative RNA-dependent RNA polymerase of potato leafroll luteovirus (PLRV) is expressed by -1 ribosomal frameshifting in the region where the open reading frames (ORF) of proteins 2a and 2b overlap. The signal responsible for efficient frameshift is composed of the slippery site *UUUAAAU* followed by a sequence that has the potential to adopt two alternative folding patterns, either a structure involving a pseudoknot, or a simple stem-loop structure. To investigate the structure requirements for efficient frameshifting, mutants in the stem-loop or in the potential pseudoknot regions of a Polish isolate of PLRV (PLRV-P) have been analyzed. Mutations that are located in the second stem (S2) of the potential pseudoknot structure, but are located in unpaired regions of the alternative stem-loop structure, reduce frameshift efficiency. Deletion of the 3' end sequence of the alternative stem-loop structure does not reduce frameshift efficiency. Our results confirm that -1 frameshift in the overlap region depends on the slippery site and on the downstream positioned sequence, and propose that in PLRV-P a pseudoknot is required for efficient frameshifting. These results are in agreement with those recently published for the closely related beet western yellows luteovirus (BWYV).

INTRODUCTION

Ribosomal frameshifting is a strategy frequently employed by various organisms to produce more than one protein from overlapping reading frames. It may occur in either direction. A shift in the 3' direction ($+1$ frameshift) has been described in the yeast retrotransposon TY (1), the copia-like element of *Drosophila* (2) and the *Escherichia coli* release factor 2 (ref. 3), whereas a shift in the 5' direction (-1 frameshift) has been demonstrated for retroviruses (4–10), infectious bronchitis coronavirus (IBV; 11, 12), luteoviruses (13–15), red clover necrotic mosaic dianthovirus (16), the L-A double-stranded RNA virus of yeast (17), *dnaX* of *E.coli*, (18–20) and the bacterial transposon *IS1* (21,22).

The site at which -1 frameshifting occurs consists of a 7 nucleotide-long sequence, the frameshift or slippery site. The composition of the slippery sites established to date is *X.XXY.YYN* where *X* is A, U or G, *Y* is A or U, and *N* is any nucleotide (18; discussed in 23). The simultaneous slippage model of Jacks *et al.* (4,5) proposes that the tRNAs bound respectively in the ribosomal P site to *XXY* and in the A site to *YYN* simultaneously slip back by one nucleotide on the RNA to pair with *XXX* and *YYY* respectively. In addition to the slippery site, in all the -1 ribosomal frameshift events for which such investigations have been performed, efficient frameshift also requires a downstream-located structure, either a simple stem-loop or a pseudoknot (5,24–26).

Among plant RNA viruses, the RNA-dependent RNA polymerase (replicase) gene is expressed via -1 frameshifting in luteoviruses (13–15) and in a dianthovirus (16), and has been suggested for pea enation mosaic virus (27,28).

The genome of luteoviruses consists of a single-stranded (ss) positive sense RNA with a 5'-linked VPg and no 3' poly(A) sequence. Members of this group include barley yellow dwarf virus, BWYV and PLRV (reviewed in 29). Recently the complete nucleotide sequence about 5800 nucleotides of the genomic RNA of four PLRV isolates has been determined (30–32). The genome organization of PLRV reveals six ORFs. ORF2a overlaps ORF2b by 582 nucleotides. ORF2b possesses the conserved GDD-containing motif present in the replicase of all plant RNA viruses sequenced to date (33). Since ORF2b lacks an AUG initiation codon, it has been suggested (30) that ORF2b might be expressed by -1 frameshifting. A potential shifty heptanucleotide stretch within the overlap region in PLRV was proposed by ten Dam *et al.* (23). The composition of this heptanucleotide sequence in the genome of all the PLRV isolates is *UUUAAAU*, except for one of two Scottish isolates in which the seventh nucleotide in the slippery site is *C* instead of *U* (30).

Prüfer *et al.* (13) using a German isolate of PLRV (PLRV-G) have demonstrated that -1 frameshifting takes place within the *UUUAAAU* sequence. The same authors examined the nucleotide sequence downstream of the slippery site and observed that a stem-loop structure located 5 nucleotides downstream of the slippery site is required for efficient frameshifting. In on-going

* To whom correspondence should be addressed at Institut Jacques Monod, 2 Place Jussieu, Tour 43, 75251 Paris Cedex 05, France

studies on a highly virulent Polish isolate of PLRV (PLRV-P), we have found that there are a few nucleotide differences within the frameshift region between PLRV-P and PLRV-G that provide limited but important changes in the putative structure of the RNA involved in the frameshift event.

We report here that in PLRV-P the slippery site is the same as in PLRV-G and that -1 ribosomal frameshifting depends on this sequence. Furthermore in the PLRV-P isolate, frameshift depends on a different RNA structure than the one proposed for PLRV-G: a pseudoknot structure appears to be an essential element for the expression of the putative replicase of PLRV-P.

MATERIALS AND METHODS

Enzymes, reagents and program

All chemicals were of the highest grade available. The T4 ligation system, [^{14}C]-labeled protein markers, [^{35}S]methionine (37 TBq/mmol) and [^{35}S]cysteine (48 TBq/mmol) were from Amersham. *Pfu* DNA polymerase was from Stratagene, T4 DNA polymerase from Boehringer Mannheim, T7 RNA polymerase from Pharmacia, and rRNasin inhibitor and RQTM DNase from Promega. The 'T7-Gen *in vitro* Mutagenesis Kit' was from United States Biochemicals, and Geneclean from BIO 101. Econofluor was from NEN. The oligonucleotides were synthesized on a Pharmacia LKB Assembler Plus. Polymerase chain reactions (PCR) were performed using a Perkin Elmer Cetus Thermal Cycler. The Zuker program (34) served to analyze possible foldings of the RNAs.

Plasmid constructions

In all constructs, designations of plasmids and transcripts are preceded by the letter 'p' and 't' respectively. The transcription vector used for all constructs was p204 (ref. 35). It includes the T7 promoter and the chloramphenicol acetyltransferase (CAT) gene.

A 919 bp-long *HindIII-HindIII* fragment from the PLRV-P cDNA clone pUP9 (A.K. *et al.*, unpublished results) corresponding to nucleotides 1583–2502 (numbering is according to Mayo *et al.* [30]) of the PLRV genome and containing the ORF2a/ORF2b overlap region was used for the construction of the frameshift detection plasmids. The sequence of this fragment was determined on both strands. For cloning purposes, new restriction sites were introduced at the ends of the cDNA fragment by PCR following the conditions provided by the supplier, and two primers. Primer P1 (5'-GGGTCGAC-CATGGCCTCAACCCATCCTC-3') contains, starting from the 5' end, a sequence to create *SaII* and *NcoI* sites, followed by a sequence complementary to nucleotides 2477–2491 of the PLRV genome. Primer P2 (5'-GGGGATCCATGGCTCATG-ATTATGAC-3') contains from the 5' end, a sequence to create *BamHI* and *NcoI* sites, followed by nucleotides corresponding to positions 1621–1636 of the PLRV genome.

The cDNA fragment resulting from amplification corresponds to nucleotides 1621–2491 of the PLRV genome. After digestion by *NcoI*, this fragment was subcloned into p204 such that the truncated ORF2b is in frame with the 141-terminal nucleotides of the 3' region of the CAT (designated 3'CAT) gene. On the other hand, the presence of the *BamHI* and *SaII* sites in the amplified cDNA made it possible to clone the cDNA fragment into the replicative form of the M13mp18 phage. The corresponding ssDNA of the resulting plasmid was used for all *in vitro* mutageneses. The flanking *NcoI* sites then served to

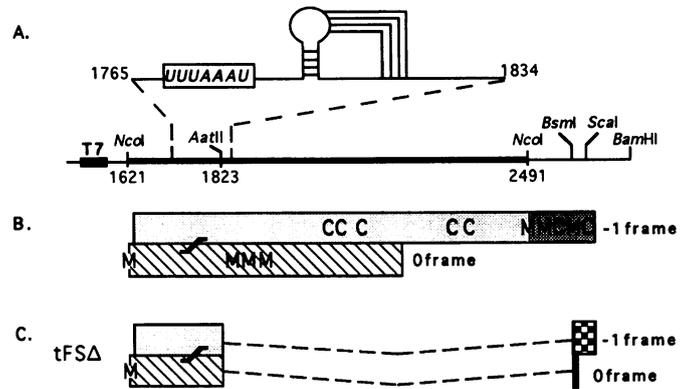


Figure 1. Schematic representation of the frameshift detection vector and characteristics of the insert-derived ORFs. A. The *NcoI*₁₆₂₁-*NcoI*₂₄₉₁ fragment (thick black line) containing the PLRV-P overlap region was inserted into p204 similarly cleaved, such that the 3' part of the CAT gene is in the same reading frame as the -1 frame protein. T7 = T7 promoter. The thin black line refers to p204-derived sequences. Appropriate restriction sites are indicated. Above is represented the PLRV-P frameshift signal; the heptanucleotide sequence is boxed and nucleotides are in italics. B. The 0 frame protein (▨) and the -1 frame protein (▤) are viral- and CAT-derived amino acid stretches respectively of the wild-type construct and of all the mutated derivatives are schematized. The frameshift site is indicated by \curvearrowright . The positions of all methionine (M) and cysteine (C) residues are indicated. C. The ORFs resulting from the *AatII-BsmI* deletion (hatched line) in tFS Δ are represented; ▨ and ▤ indicate the new extensions of the 0 frame and -1 frame proteins respectively. Other indications are as in B.

subclone the cDNA fragment into p204 (Fig. 1A). The presence of the 5'-proximal *NcoI* site had the added advantage of introducing an AUG initiation codon into the resulting *in vitro* transcript.

In the parental wild-type tFS (Fig. 1B), ORF2a (subsequently designated 0 frame or stopped protein) and the chimeric frameshift protein ORF2a-ORF2b-3'CAT (subsequently referred to as the -1 frame or frameshift protein), encode proteins containing 202 and 339 amino acids, that is 21.3K and 38K proteins, respectively.

pFS was digested with *AatII* present in the PLRV-P cDNA at the level of nucleotides 1823–1828, and with *BsmI* located in the 3' region of the CAT gene. It was made blunt end by T4 DNA polymerase and ligated using the T4 ligation system. The resulting plasmid pFS Δ contains a 762 bp-long deletion; its insert contains a 203 bp-long fragment from the PLRV genome that is linked upstream of the 3' region of the CAT sequence but is now no longer read in the CAT reading frame (Fig. 1C). In addition, nucleotide sequence verification of pFS Δ revealed that in three independent clones, the T residue which should have been maintained after *BsmI* digestion and fill-in was inadvertently lost. As a consequence, the resulting -1 frame protein (102 amino acids) is 33 amino acids longer than the 0 frame protein (69 amino acids).

Mutagenesis *in vitro*

The 'T7-Gen *in vitro* Mutagenesis Kit' was used as indicated by the supplier together with primers containing the appropriate mismatches.

A mutation in the slippery site was introduced using a primer complementary to nucleotides 1762–1782 of the PLRV genome. This oligonucleotide contains the elements for the creation of a unique *BstI* site at the level of the slippery site. The resulting

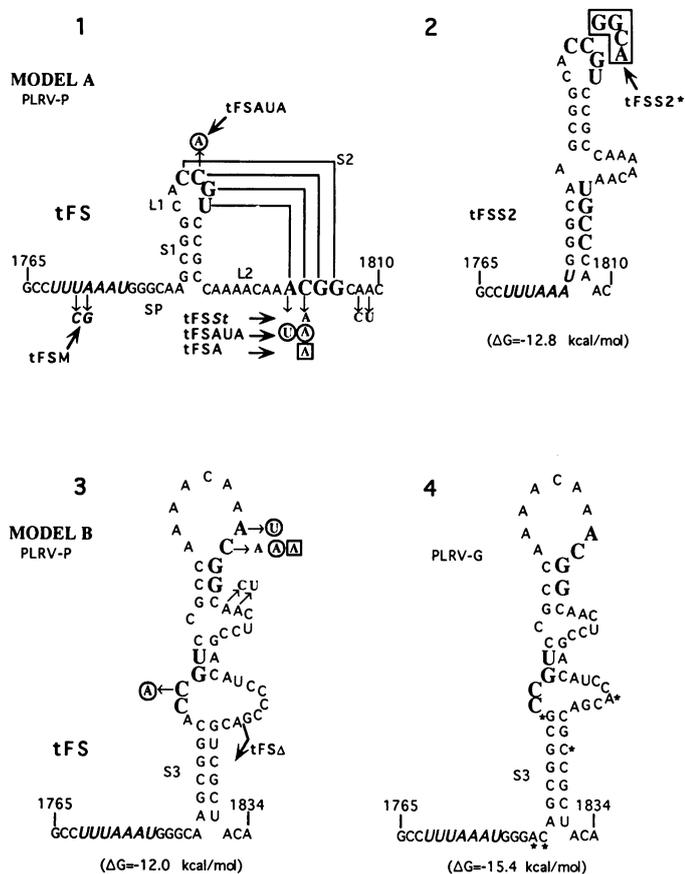


Figure 2. Possible folding patterns of the frameshift region of the PLRV genome. 1: Model A. In tFS, $\Delta G = -4.0$ kcal/mol, not taking S2 into account. The RNA pseudoknot structure is located six nucleotides downstream of the slippery site UUUAAAU. The positions of the mutations in tFSSr, tFSA and tFSAUA are indicated by the corresponding nucleotide changes. 2. In tFSS2 the sequence ACGG has been replaced by UGCC resulting in the formation of a new stem-loop structure ($\Delta G = -12.8$ kcal/mol). In mutant tFSS2* both parts of S2 have been exchanged; the resulting construct has the same structure and the same energy as tFSS2. 3: Model B. The stem-loop structure ($\Delta G = -12.0$ kcal/mol) is located five nucleotides downstream of the slippery site. The arrow indicates the 5' position of the deletion created to yield tFSA. The energy of the stem-loop structure in tFSSr, tFSA and tFSAUA is the same as in tFS. 4. Stem-loop structure in PLRV-G ($\Delta G = -15.4$ kcal/mol). The nucleotide differences compared to the sequence of the corresponding region of PLRV-P are indicated by asterisks. The numbering of the 5' and 3' nucleotides within the PLRV sequences considered, the shifty heptanucleotide (in bold italics), the spacer (SP), the stems (S1, S2 and S3), and the connecting loops (L1 and L2) are indicated. The large bold letters correspond to the 5' and 3' parts of S2 and also indicate the position these nucleotides occupy in the various constructs.

tFSM contains the sequence UUCGAAU instead of UUUAAAU (Fig. 2-1).

A unique *StuI* site was introduced using a primer complementary to nucleotides 1796–1816 of the PLRV genome, resulting in pFSSr. Its transcript contains the following alterations, C₁₈₀₄→A, A₁₈₀₈→C and A₁₈₀₉→U (Fig. 2-1). The ssDNA version of pFSSr was used for new rounds of mutations, leading to pFSA, pFSAUA, pFSS2 and pFSS2*, schematized in Fig. 2-1 and 2-2. Screening of these mutants was based on the absence of the *StuI* site, since all the oligonucleotide primers used were such that when the intended mutations were introduced, the *StuI* site was abolished.

Mutant pFSA contains the nucleotide change C₁₈₀₄→A introduced using a primer complementary to nucleotides

1793–1816 of the PLRV genome. In the transcript derived from mutant pFSAUA, the following nucleotide changes were introduced, C₁₇₈₈→A, A₁₈₀₃→U and C₁₈₀₄→A, using a primer complementary to PLRV nucleotides 1776–1814. The transcript obtained from mutant pFSS2 contains the sequence UGCC instead of ACGG in positions 1803–1806. This mutation was introduced using a primer complementary to PLRV nucleotides 1791–1822. In the transcript produced by mutant pFSS2* (pseudo-wild-type) the sequences CCGU in positions 1787–1790 and ACGG in positions 1803–1806 were replaced by GGCA and UGCC respectively, using a primer complementary to PLRV nucleotides 1778–1820.

The presence of the desired mutations or deletions was confirmed by restriction analyses and by direct sequencing using the 'AutoRead Sequencing Kit' and an automated laser fluorescent A.L.F. DNA Sequencer from Pharmacia.

Transcription and translation

Transcripts of the corresponding plasmids linearized with *Bam*HI and also with *Sca*I in the case of pFSA were obtained as follows. DNA (1 μ g) was incubated for 1 h at 37°C with 10 μ l of 5× concentrated BRL buffer for T7 RNA polymerase, 10 mM DTT, 2.5 μ l of 1 mg/ml of BSA, 0.5 mM each of ATP, CTP and UTP, 0.2 mM GTP, 1 mM m⁷GpppG, 40 U of RNasin and 63 U of T7 RNA polymerase in a total volume of 50 μ l. The transcribed RNAs were treated with 1 U of RQ™ DNase for 15 min at 37°C, purified by Sephadex G-50 column chromatography, phenol extracted and ethanol precipitated. The integrity and size of the transcripts were estimated by native agarose gel electrophoresis, and the concentration was determined by A₂₆₀ measurements.

The transcripts (0.2 μ g) were translated in a reticulocyte lysate (36) in the presence of 370 kBq of [³⁵S]methionine or [³⁵S]cysteine per 10 μ l incubation. After translation, 2 μ l were precipitated by 5% hot trichloroacetic acid (TCA) to determine the total radioactivity incorporated; the remaining 8 μ l were analyzed (37) by 0.1% SDS-15% polyacrylamide gel electrophoresis. The relevant bands were cut from the dried gel and their radioactivity was counted in Econofluor. Two methods were used to estimate efficiency of frameshift. 1) The ratio of counts contained in the frameshift protein (band a) over the counts contained in band a + those contained in the stopped protein (band c) was determined. 2) The ratio of counts contained in band a over the average of the total TCA counts deposited onto each well was estimated. The values obtained with wild-type tFS are defined as 100% frameshift. For tFSA the radioactivity measured in band a was furthermore corrected for the difference in methionine content of this protein. In the case of tFSA, it was not possible to evaluate frameshift efficiency using method 1, because of the variability in removal of the only methionine residue in the stopped protein of this construct, i. e. the initiator methionine.

RESULTS

Mutation in the frameshift site

Based on their amino acid composition, the 0 frame and the -1 frame proteins can be distinguished by labeling with [³⁵S]cysteine. Indeed in all but one construct (tFSA, see below), the 0 frame protein lacks cysteine residues whereas 7 cysteines are present in the -1 frame protein downstream of the slippery

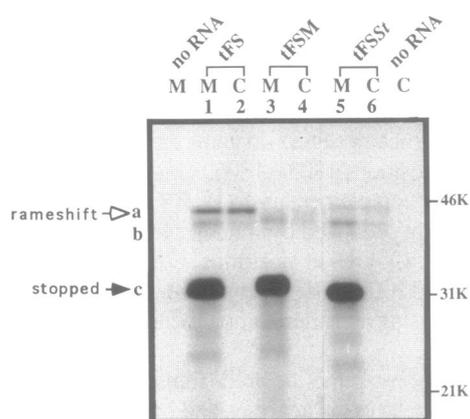


Figure 3. Translation *in vitro* of transcripts tFS, tFSM and tFSSr derived from the corresponding *Bam*HI-linearized DNA templates. Translation was in the presence of [³⁵S]methionine (M) or [³⁵S]cysteine (C). Transcripts used as templates are as designated above the lanes. To the left the positions of the frameshift protein (band a) and the stopped protein (band c) are indicated by an open \rightarrow and a closed \rightarrow arrow respectively; the position of band b is also indicated. To the right are shown the size and position of the [¹⁴C]-labeled protein markers run in a parallel well. All lanes are from the same gel.

Table 1. Frameshift efficiency of various transcripts determined using two methods.

Transcript	Frameshift efficiency %	
	a	a
	a+c	total cpm
tFS	100	100
tFSM	30	30
tFSSr	27	21
tFSΔ/ <i>Bam</i> HI	—	134
tFSΔ/ <i>Sca</i> I	—	130
tFSA	39	43
tFSAUA	45	44
tFSS2	37	46
tFSS2*	30	32

The results presented are averages of 3 to 4 experiments, except for the tFSΔ constructs. All transcripts are derived from templates linearized by *Bam*HI except tFSΔ that was linearized by *Bam*HI or *Sca*I. a = counts contained in the frameshift protein; c = counts contained in the stopped protein; total cpm = average of the total counts deposited onto the wells. The values obtained with tFS (about 2% of the total counts deposited) are defined as 100% frameshift.

site (Fig. 1B); each protein contains 4 methionines, of which one is the initiator methionine.

Using tFS and [³⁵S]methionine (Fig. 3, lane 1), a band of about 45K (band a) and faint band (band b) of about 40K appeared, as well as an intense band (band c) migrating at about 30K. With [³⁵S]cysteine, band a predominated and band c was not visible (Fig. 3, lane 2). It can be concluded that band a is the frameshift protein, whereas band c is the stopped protein. The protein patterns obtained with all subsequent mutated transcripts have confirmed this conclusion. The origin of band b migrating slightly faster than the frameshift protein remains enigmatic. The reasons for which it most certainly does not correspond to the frameshift protein are presented in the Discussion.

The position of the proteins in bands a and c obtained with tFS as well as with all the mutated constructs is considerably different from what is expected based on the calculated size (38K

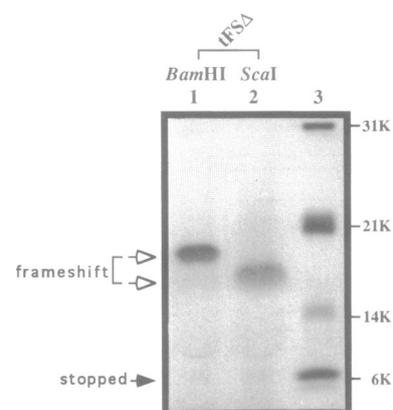


Figure 4. Effect of tFSΔ on efficiency of frameshifting as determined by *in vitro* translation. pFSΔ was linearized *Bam*HI or *Sca*I and the *in vitro*-derived transcripts were translated in the presence of [³⁵S]methionine (lanes 1 and 2 respectively). Lane 3 corresponds to the migration pattern of the [¹⁴C]-labeled protein markers whose sizes are indicated to the right. Other indications are as in Fig. 3.

and 21.3K) of the corresponding ORFs; this is particularly striking in the case of the stopped protein and may be explained by its high isoelectric point which is between 9.1 and 10.1 for the different transcripts. A similar observation has been made concerning the unusual behavior of the turnip yellow mosaic virus 69K protein whose isoelectric point is 11.5 (38,39).

To verify that the heptanucleotide *UUUAAA*U is indeed involved in frameshift, it was mutated to *UUCGAAU* in tFSM (Fig. 2–1). In the presence of [³⁵S]methionine, translation of the frameshift protein was considerably reduced and essentially the stopped protein was synthesized (Fig. 3, lane 3); in the presence of [³⁵S]cysteine the same reduced level of frameshift protein was produced and the stopped protein was not detected (Fig. 3, lane 4). Thus, the mutations introduced into the heptanucleotide sequence strongly reduced frameshift as already observed by Prüfer *et al.* (13). In tFSM the mutated heptanucleotide *UUCGAAU* codes for Ser.Asn (UCG.AAU) in the 0 frame and for Phe.Glu (UUC.GAA) in the –1 frame, as opposed to Leu.Asn (UUA.AAU) in the 0 frame and Phe.Lys (UUU.AAA) in the –1 frame in the parental *UUUAAA*U sequence. Frameshift is hindered with tFSM because tRNA^{Asn} very poorly shifts to the –1 frameshift position, in accordance with the model proposed by Jacks *et al.* (5).

These results together with those of Prüfer *et al.* (13) suggest that the integrity of the heptanucleotide sequence *UUUAAA*U is necessary for ribosomal frameshifting and expression of the frameshift protein. Frameshift efficiency with tFSM was 30% of wild-type (Table 1).

Structure requirements for efficient translational frameshifting

Based on theoretical grounds, the slippery site appears to be followed by a sequence that has the potential to adopt two different configurations depending on the length of the sequence considered: the configuration in Model A (Fig. 2–1) involves a possible pseudoknot as proposed by ten Dam *et al.* (23), whereas the configuration in Model B (Fig. 2–3) involves a stem-loop structure resembling the one proposed by Prüfer *et al.* (13) as schematized in Fig. 2–4. The sequence of Model B includes the sequence which in Model A can create the

pseudoknot; one can thus expect one or the other model to exist and to promote frameshift. To distinguish which model participates in frameshifting, a series of mutants were constructed.

In tFS Δ , Model A is not altered since the deletion begins beyond the region required to form the pseudoknot structure; however, Model B is affected because the region downstream of position 1823 has been deleted; this deletion disrupts base-pairing of stem S3 in Model B (Fig. 2–3). Possible folding of tFS Δ demonstrated that the vector sequence following the deletion does not replace the nucleotides removed from the 3' part of S3 to restore this stem. If the pseudoknot structure of Model A is sufficient to promote frameshift, translation of tFS Δ derived from a BamHI-linearized template (tFS Δ /BamHI) should result in the synthesis of a stopped and a frameshift protein of calculated size of 7.2K and 11.3K respectively. With tFS Δ , the distinction between 0 frame and –1 frame protein by differential labeling is not possible because the first 6 cysteine residues beyond the frameshift site have been removed by the deletion (Fig. 1C), and the 7th cysteine downstream of the BsmI site was lost; the frameshift protein contains 3 methionine residues (of which the initiator methionine), whereas the only methionine contained in the stopped protein is the initiator methionine (for details, see Materials and Methods).

Translation of tFS Δ /BamHI in the presence of [³⁵S]methionine yielded a strong band migrating as an 18K protein, and a faint band migrating as a 7K protein (Fig. 4, lane 1). To verify that these bands correspond to the frameshift and the stopped protein respectively, pFS Δ was linearized with ScaI. Translation of the resulting tFS Δ /SacI should yield a frameshift protein shorter by 28 amino acids (*i.e.* 3.1K) than the corresponding protein produced with tFS Δ /BamHI, and it should contain 2 instead of 3 methionine residues; this should have no effect on the methionine content of the stopped protein. Translation demonstrated that the intense band now migrates as a protein of about 15K (Fig. 4, lane 2), whereas the position of the 7K protein remains unchanged. Consequently, the 18K and 15K proteins correspond to the frameshift product, whereas the 7K protein is the stopped protein. The intensity of the stopped protein band varied between experiments, probably depending on the extent to which the initiator methionine was removed during translation (40). The efficiency of frameshift with tFS Δ ranged from 130% to 134% as compared to tFS (Table 1). A similarly high level of frameshift was also obtained in a wheat germ extract (not shown). Higher than wild-type levels of frameshift have been described for IBV (12). Thus, the information located downstream of the slippery site and up to the AatII site of the viral genome is sufficient to promote frameshift. Our results also show that a structure different from Model B is probably involved in ribosome slippage.

The three mutations in tFSS t are located in unpaired regions in Model B (Fig. 2–3), and therefore do not change the structure of this model whose free energy remains the same as in tFS. However, the C₁₈₀₄→A mutation affects base-pairing in S2 of Model A (Fig. 2–1). When tFSS t was introduced into a reticulocyte lysate, the efficiency of frameshifting was reduced to 27%–21% as compared to tFS (Fig. 3, compare lanes 1 and 2 to lanes 5 and 6; Table 1), thereby favoring Model A. The slight shift in migration of the stopped protein produced by tFSS t and by other mutated transcripts with respect to the wild type transcript (see also Fig. 5), is very likely due to changes in amino acid composition that result from the mutations introduced. Indeed it has been reported that a single amino acid substitution in the

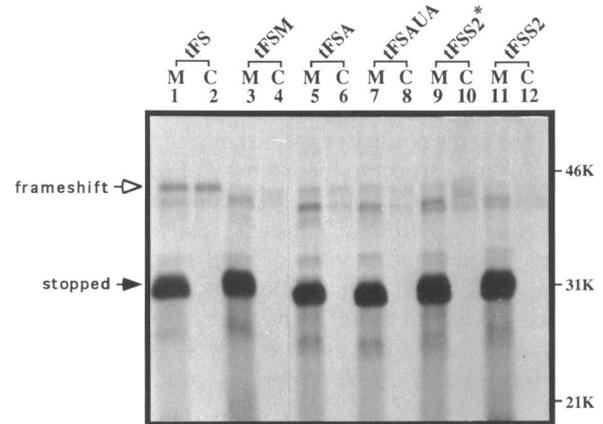


Figure 5. Translation *in vitro* of transcripts tFS, tFSA, tFSAUA, tFSS2, and tFSS2*, together with tFS and tFSM. The transcripts were obtained from cDNA templates linearized by BamHI and translated in the presence of [³⁵S]methionine (M) or [³⁵S]cysteine (C); they are as designated above the lanes. All lanes are from the same gel. Other indications are as in Fig. 3.

alfalfa mosaic virus capsid affects the electrophoretic mobility of this protein which now migrates as a 29K, rather than as a 27.5K protein in the wild-type (41,42). In any event, since 2 of the 3 mutations in tFSS t are located downstream of S2, it was conceivable that reduction in frameshift might have been due to these downstream mutations.

For this reason, tFSA and tFSAUA were produced (Fig. 2–1 and 2–3) that would destabilize stem S2 of Model A, but not the structure of Model B (as verified by computer search for Model B). When tFSA or tFSAUA were used as template for *in vitro* translation, a reduction in frameshift to about 41% and 44% respectively was observed (Fig. 5, lanes 5 to 8; Table 1).

If Model A is the more likely model, mutations destabilizing either S1 or S2 should lead to a decrease in frameshift efficiency. To investigate this possibility, in tFSS2 (Fig. 2–2) nucleotides ACGG on the 3' side of S2 were replaced by UGCC. In this case, frameshifting was strongly reduced (Fig. 5, lanes 11 and 12; Table 1). Restoration of a stable S2 structure was attempted in the double pseudo-wild-type mutant tFSS2* (Fig. 2–2) in which the base-paired strands of S2 were exchanged; however, contrary to expectations the efficiency of frameshifting remained far below that of the wild-type tFS (Fig. 5, lanes 9 and 10; Table 1). Likely interpretations of these data are presented below.

DISCUSSION

The results presented here confirm that the –1 ribosomal frameshifting site UUUAAAU is essential to express the putative replicase of PLRV-P using a reticulocyte lysate as also observed by Prüfer *et al.* (13). Modification of the slippery site to UUCG-AAU causes a reduction in the synthesis of the frameshift protein to 30% of the wild-type.

In tFS as in most mutant constructs tested, (Fig. 3 and 5) a protein (band b) migrating slightly faster than the frameshift protein (band a) can also be detected. The reasons that lead us to propose that band a rather than band b corresponds to the frameshift protein are the following. 1) Although band b was clearly visible when translation was performed in a reticulocyte lysate, it was virtually absent when the wheat germ system served

for *in vitro* translation (not shown). 2) Using other mutants containing 15- or 33-nucleotide-long deletions downstream of nucleotides 1794 or 1788 respectively (not shown), the positions of band b remained unchanged, whereas the positions of band c were shifted to the positions expected of slightly shorter stopped proteins. 3) With respect to the intensity of the stopped protein (band c) and in the presence of [³⁵S]methionine the intensity of band b remained virtually constant whatever the mutation introduced, whereas the intensity of band a varied depending on the mutant. This was further verified by calculating the ratio of the counts present in band b over the total average counts deposited onto each well: again this ratio remained virtually constant (not shown). The origin of band b thus remains unknown.

For PLRV-G, it was shown (13) that the stem-loop structure located downstream of the slippery site (Fig. 2–4) is required for efficient frameshifting, but that it is not involved in a pseudoknot; however, the experiments performed in support of this conclusion are not shown, so that it is not possible to make direct comparisons with the results presented here.

The nucleotide sequence in PLRV-P differs slightly compared to the one in PLRV-G within the frameshift region. The few base changes observed between these two PLRV isolates are reflected by differences in their putative secondary structures and in their free energy: stem S3 in Model B of PLRV-P is one base pair shorter than in PLRV-G. Furthermore, the overall structure of the potential pseudoknot in PLRV-P, PLRV-G as also in virtually all the other PLRV isolates whose sequence has been determined in the frameshift region is the same. In addition, a very similar pseudoknot structure has recently been reported in the frameshift region of the closely related BWYV (15). The strong sequence similarity that exists between PLRV and BWYV covers—but does not extend beyond—the pseudoknot structure. In PLRV-P, the pseudoknot structure begins 6 nucleotides downstream of the slippery site. The translation experiments performed with tFSA indicate that the slippery site and the region up to nucleotide 1823 of the viral genome are sufficient to promote frameshift; even though the sequence involved in S3 has been removed, the remaining sequence can still fold into a pseudoknot. These results are in strong support of Model A.

One can expect that mutations within the unpaired regions of Model B such as in tFSA and tFSAUA would be without effect on frameshift efficiency as is the case of IBV (12). Frameshift was reduced with these two mutants. Even the single substitution present in tFSA had a pronounced effect, probably by weakening S2. A similar reduction in frameshifting was observed (43) for the feline immunodeficiency virus (FIV) pseudoknot structure, when one nucleotide of S2 was changed to its complementary residue.

Frameshift was greatly decreased in mutant pFSS2 in which the 5' part of S2 in tFS was changed to its complementary sequence. The pseudo-wild-type mutant pFSS2* in which the 5' and 3' sequences of S2 were exchanged to reconstruct the pseudoknot revealed an equally low level of frameshifting. A similar result has been reported with a pseudo-wild-type mutant of IBV (12), whose frameshift efficiency was not restored to wild-type levels. A likely interpretation for these data was obtained when comparing the possible folding patterns of the frameshift region of tFSS2 and of tFSS2* by computer analysis. As seen in Fig. 2–2, the changes introduced in tFSS2 cause a new structure to appear, in which UGCC forms the 3' side of a new stem that includes part of the slippery site, and seems quite stable.

This new structure is maintained in the pseudo-wild-type tFSS2* (Fig. 2–2). Although it is not clear how ribosome movement is affected by different RNA folding patterns, this structure may be preferentially formed, and as a result could significantly reduce the level of frameshift. Other mutations within S2 that would not lead to new folding patterns, in particular those that would lead to the creation of artificial but functional pseudoknots could provide further information on the role of the structure located downstream of the frameshift site in ribosome slippage. Another possible interpretation of these data is that the spacer has now been abolished (Fig. 2–2). Correct length of the spacer is believed to be important in the pausing process as demonstrated for IBV (11) and FIV (43).

The results presented here and supported by results obtained in a wheat germ system (not shown) confirm that frameshifting requires two elements, a slippery site and a downstream positioned sequence. They further strongly suggest that in PLRV-P a pseudoknot rather than a simple stem-loop structure is the folding pattern adopted by the downstream sequence to provide efficient frameshifting.

ACKNOWLEDGEMENTS

We wish to express our gratitude to François Chapeville and Wlodek Zagorski for their interest and encouragements in this work, and to Dan Levin for constructive suggestions and fruitful discussions. We thank Andrzej Palucha for his help in cloning the PLRV genome, and Marek Zagulski for his help in sequencing the mutants. A.K. benefited from a grant from the Twin-Association between the Institut Jacques Monod and the Centre de Génétique Moléculaire (France), and the Institute of Biochemistry and Biophysics (Poland). The project was also partly financed by a grant from the Ligue Nationale Française Contre le Cancer. The Institut Jacques Monod is an 'Institut Mixte, CNRS—Université Paris VII'.

REFERENCES

1. Belcourt, M.F. and Farabaugh, P.J. (1990) *Cell*, 62, 339–352.
2. Saigo, K., Kugimiya, W., Matsuo, Y., Inouye, S., Yoshioka, K. and Yuki, S. (1984) *Nature*, 312, 659–661.
3. Weiss, R.B., Dunn, D.M., Dahlberg, A.E., Atkins, J.F. and Gesteland, R.F. (1988) *EMBO J.*, 7, 1503–1507.
4. Jacks, T., Power, M.D., Masiarz, F.R., Luciw, P.A., Barr, P.J. and Varmus, H.E. (1988) *Nature*, 331, 280–283.
5. Jacks, T., Madhani, H.D., Masiarz, F.R. and Varmus, H.E. (1988) *Cell*, 55, 447–458.
6. Wilson, W., Braddock, M., Adams, S.E., Rathjen, P.D., Kingsman, S.M. and Kingsman, A.J. (1988) *Cell*, 55, 1159–1169.
7. Hatfield, D. and Oroszlan, S. (1990) *Trends Biochem. Sci.*, 15, 186–190.
8. Chamorro, M., Parkin, N. and Varmus, H.E. (1992) *Proc. Natl. Acad. Sci. USA*, 89, 713–717.
9. Parkin, N.T., Chamorro, M. and Varmus, H.E. (1992) *J. Virol.*, 66, 5147–5151.
10. Vickers, T.A. and Ecker, D.J. (1992) *Nucleic Acids Res.*, 20, 3945–3953.
11. Brierley, I., Digard, P. and Inglis, S.C. (1989) *Cell*, 57, 537–547.
12. Brierley, I., Rolley, N.J., Jenner, A.J. and Inglis, S.C. (1991) *J. Mol. Biol.*, 220, 889–902.
13. Prüfer, D., Tacke, E., Schmitz, J., Kull, B., Kaufmann, A. and Rohde, W. (1992) *EMBO J.*, 11, 1111–1117.
14. Brault, V. and Miller, W.A. (1992) *Proc. Natl. Acad. Sci. USA*, 89, 2262–2266.
15. Garcia, A., Van Duin, J. and Pleij, C.W.A. (1993) *Nucl. Acids Res.*, 21, 401–406.
16. Xiong, Z., Kim, K.H., Kendall, T.L. and Lommel, S.A. (1993) *Virology*, 193, 213–221.

17. Dinman, J.D., Icho, T. and Wickner, R.B. (1991) *Proc. Natl. Acad. Sci. USA*, 88, 174–178.
18. Tsuchihashi, Z. and Kornberg, A. (1990) *Proc. Natl. Acad. Sci. USA*, 87, 2516–2520.
19. Blinkowa, A.L. and Walker, J.R. (1990) *Nucleic Acids Res.*, 18, 1725–1729.
20. Flower, A.M. and McHenry, C.S. (1990) *Proc. Natl. Acad. Sci. USA*, 87, 3713–3717.
21. Sekine, Y. and Ohtsubo, E. (1989) *Proc. Natl. Acad. Sci. USA*, 86, 4609–4613.
22. Escoubas, J.M., Prère, M.F., Fayet, O., Salvignol, I., Galas, D., Zerbib, D. and Chandler, M. (1991) *EMBO J.*, 10, 705–712.
23. ten Dam, E.B., Pleij, C.W.A. and Bosch, L. (1990) *Virus Genes*, 4, 121–136.
24. Pleij, C.W.A. (1990) *Trends Biochem. Sci.*, 15, 143–147.
25. Le, S.-Y., Shapiro, B.A., Chen, J.-H., Nussinow, R. and Maizel, J.V. (1991) *Genet. Anal. Tech. Appl.*, 8, 191–205.
26. ten Dam, E., Pleij, K. and Draper, D. (1992) *Biochemistry*, 31, 11665–11676.
27. Demler, S.A. and de Zoeten, G.A. (1991) *J. Gen. Virol.*, 72, 1819–1834.
28. Demler, S.A., Rucker, D.G. and de Zoeten, G.A. (1993) *J. Gen. Virol.* 74, 1–14.
29. Matthews, R.E.F. (1991) *Plant Virology*. Academic Press, New York.
30. Mayo, M.A., Robinson, D.J., Jolly, C.A. and Hyman, L. (1989) *J. Gen. Virol.*, 70, 1037–1051.
31. van der Wilk, F., Huisman, M.J., Cornelissen, B.J.C., Huttinga, H. and Goldbach, R. (1989) *FEBS Lett.*, 245, 51–56.
32. Keese, P., Martin, R.R., Kawchuk, L.M., Waterhouse, P.M. and Gerlach, W.L. (1990) *J. Gen. Virol.*, 71, 719–724.
33. Koonin, E.V. (1991) *J. Gen. Virol.*, 72, 2197–2206.
34. Zuker, M. and Stiegler, P. (1981) *Nucleic Acids Res.*, 9, 133–148.
35. Valle, R.P.C., Drugeon, G., Devignes-Morch, M.-D., Legocki, A.B. and Haenni, A.-L. (1992) *FEBS Lett.*, 306, 133–139.
36. Morch, M.-D., Drugeon, G., Szafranski, P. and Haenni, A.-L. (1989) *J. Virology*, 63, 5135–5158.
37. Laemmli, U.K. (1970) *Nature*, 227, 680–685.
38. Morch, M.-D., Boyer, J.-C. and Haenni, A.-L. (1988) *Nucleic Acids Res.*, 16, 6157–6173.
39. Weiland, J.J. and Dreher, T.W. (1989) *Nucleic Acids Res.*, 17, 4675–4687.
40. Smith, A.E. (1973) *Eur. J. Biochem.*, 33, 301–313.
41. Huisman, M.J., Cornelissen, B.J.C., Groenendijk, C.M., Bol, J.F. and Van Vloten-Doting, L. (1989) *Virology*, 171, 409–416.
42. Smit, C.H. (1981) Ph. D. Thesis, Leiden.
43. Morikawa, S. and Bishop, D.H.L. (1992) *Virology*, 186, 389–397.