Published online 8 September 2011 *Nucleic Acids Research, 2012, Vol. 40, No. 1* **303–313**
doi:10.1093/nar/gkr686

mRNA pseudoknot structures can act as ribosomal roadblocks

Jesper Tholstrup[1], Lene B. Oddershede[2] and Michael A. Sørensen[1,*]

[1]Department of Biology, Ole Maaløes vej 5, University of Copenhagen, DK-2200 Copenhagen and [2]Niels Bohr Institute, Blegdamsvej 17, University of Copenhagen, DK-2100 Copenhagen, Denmark

Received March 24, 2011; Revised August 5, 2011; Accepted August 7, 2011

ABSTRACT

Several viruses utilize programmed ribosomal frameshifting mediated by mRNA pseudoknots in combination with a slippery sequence to produce a well defined stochiometric ratio of the upstream encoded to the downstream-encoded protein. A correlation between the mechanical strength of mRNA pseudoknots and frameshifting efficiency has previously been found; however, the physical mechanism behind frameshifting still remains to be fully understood. In this study, we utilized synthetic sequences predicted to form mRNA pseudoknot-like structures. Surprisingly, the structures predicted to be strongest lead only to limited frameshifting. Two-dimensional gel electrophoresis of pulse labelled proteins revealed that a significant fraction of the ribosomes were frameshifted but unable to pass the pseudoknot-like structures. Hence, pseudoknots can act as ribosomal roadblocks, prohibiting a significant fraction of the frameshifted ribosomes from reaching the downstream stop codon. The stronger the pseudoknot the larger the frameshifting efficiency and the larger its roadblocking effect. The maximal amount of full-length frameshifted product is produced from a structure where those two effects are balanced. Taking ribosomal roadblocking into account is a prerequisite for formulating correct frameshifting hypotheses.

INTRODUCTION

The reading frame of the vast majority of mRNAs is determined by the start codon after which the downstream cistron is translated in the same frame. Maintenance of the reading frame occurs without further signals to the ribosome. However, examples of genes containing information for programmed frameshifts can be found in most organisms, or in some of their IS sequences, transposable elements, retroelement-derived sequences or viruses. The sequence-information needed for programmed ribosomal frameshift varies and both +1 and −1 frameshifts can be induced (1–3).

Here, we focus on the frameshifting signal found in several viruses (1), including infectious bronchitis virus (IBV) and SARS-CoV. The signal leads to programmed ribosomal −1 frameshift, whereby multiple proteins are produced from a single polycistronic messenger RNA (mRNA) (4,5). The frameshift efficiency, i.e. the fraction of ribosomes, which change reading frame, is important to ensure a correct stoichiometric relationship between the different products of translation. It has been shown that altered frameshift efficiency has detrimental effects on the proliferation of HIV-I and the yeast L-A viruses (6,7). In order to induce −1 frameshift, these viruses rely on three physical features on the mRNA: a heptanucleotide sequence, a spacer and a downstream structure (8). The heptanucleotide sequence, called the slippery sequence, is where the −1 frameshift occurs and typically has the following sequence: X XXY YYZ, where X, Y and Z denote nucleotide species and spaces indicate initial reading frame. The spacer is a stretch of 6–9 nt positioning the ribosome correctly at the slippery site when encountering the downstream structure. The downstream structure is most often found to be a pseudoknot. The pseudoknot structure probably functions as a physical barrier deforming upon approach of the translating ribosome (9), thereby assisting the frameshifting process; however, geometry and surface charge of the structure may also play a role for the frameshifting (10).

In bacteria and yeast, programmed frameshift signals can have rather different elements, as, e.g. the upstream Shine–Dalgarno binding element in the autoregulatory RF2 gene frameshift site first described in *Escherichia coli* (11) or the different pattern of the +1 frameshift stimulating heptanucleotide sequences present in *Saccharomyces* Ty elements (2). However, many frameshift signals deviate little from those described for the virus-derived system used here and many signals are of

*To whom correspondence should be addressed. Tel: +45 3532 3711; Fax +45 3532 2128; Email: mas@bio.ku.dk

such general character that ribosomes from different kingdoms of life will respond to them by shifting frame (12). This happens not always with the same efficiency as in the original organism (12,13) and there are even examples found where a frameshift element can direct the ribosomes into −2 or +1 frameshift depending on the test organism (14). Here, we challenged *E. coli* ribosomes by constructing artificial frameshifting signals containing pseudoknot-like structures with strong stems. Using a refined frameshift assay, involving two-dimensional (2D) gel electrophoresis of pulse labelled proteins, we show that a significant amount of frameshifted ribosomes permanently stall within the strongest pseudoknots which therefore efficiently act as roadblocks.

The small ribosomal subunits have been shown to be sensitive towards mRNA secondary structure in the process of translation initiation and mRNA structures can exclude initiation both in eukaryotes during the scanning process (15) and in prokaryotes for binding between the mRNA and the 3′-end of 16S RNA (16). The fully assembled and translating 70S or 80S ribosomes seem to be more robust. It is, however, broadly accepted that mRNA secondary structures can function as obstacles to translating ribosomes (17,18) although examples exists of large secondary structures in mRNA that are translated without any ribosomal delay (19). Nevertheless, there is compelling evidence from *in vitro* experiments showing that ribosomes may pause upstream to such structures, most pronounced if the structures form pseudoknots (20–22). Possibly the lack of rotational freedom in the helix of stem 1, due to the pairing in stem 2, makes pseudoknot structures harder to 'unzip' by the ribosome than simple stem–loop structures (23). This may explain why pseudoknots can pause ribosomes. Examples from nature show the existence of diverse peptide sequences, often present in regulatory circuits, which will stall ribosomes (24), but to our knowledge, a permanent halt of ribosomes caused by mRNA structures has not been shown previously.

Recent single molecule investigations suggest that the mechanical strength of pseudoknots correlate with the ability of the pseudoknot to stimulate frameshift (25–27), at least in a certain interval. However, the calculated Gibbs free energy does not always correlate with frameshift efficiency. Not only the strength of the stems, but also the interaction between the loop and the stems might be of importance for the ability to induce frameshift and for the overall mechanical strength and brittleness of the structure. If the pseudoknot becomes too strong the ribosome, frameshifted or not, might not be able to open it and continue translation, whereby the pseudoknot acts as a roadblock. Often in literature (25–31) frameshifting assays were performed on constructs exhibiting the common feature that the stop codon for the normal reading frame was located at the entrance of the pseudoknot (or inside the pseudoknot) and the stop codon for the successful −1 frameshift was located downstream of the pseudoknot. In most frameshifting assays, the amount of frameshifting is determined by quantifying the amount of full-length frameshifted versus non-frameshifted products. However, for this to be a correct measure, the frameshifted ribosome must continue translation through the pseudoknot and beyond to the −1 frameshifted stop codon. If the −1 frameshifted ribosome permanently stalls inside the pseudoknot, it would falsely be interpreted as if the ribosome did not frameshift. Therefore, there is a serious pitfall in the classical methods which renders the amount of frameshifted ribosomes to be non-correctly determined, i.e. be underestimated, potentially leading to false hypotheses regarding the physical mechanism of frameshifting.

The observation that strong pseudoknot-like structures can stop translation lead to the hypothesis that the largest amount of frameshifted product will be produced if the pseudoknot is mechanically strong but without a significant roadblocking effect. Most likely, this is exactly the balance exhibited by naturally occurring viral pseudoknots.

MATERIALS AND METHODS

Bacterial growth

Escherichia coli strain MAS90 [*E. coli K-12, recA1 Δ(pro-lac) thi ara F′: lacI^{q1} lacZ::Tn5 proAB^{+}*]. Liquid cultures were grown in minimal MOPS media (32) using glycerol as carbon source. Cultures were incubated with shaking at 37°C for at least 10 generations in the log phase prior to being used in frameshift assays.

Plasmid construction

Pseudoknots were designed using custom-made software, which ensued that the codon usage was appropriate for expression in *E. coli* and that the sequences were likely to fold into the correct structure as determined by pknotsRG (33). Hence, the resulting sequences are artificial pseudoknot-like structures and there is always a risk that the structure does not fold as anticipated. The selected sequences were synthesized by GeneScript and were subsequently inserted into plasmid OFX302 [containing slippery sequence, spacer and pseudoknot (25)] between HindIII and ApaI restriction sites.

Frameshift assay

The *in vivo* frameshift assays were performed as described previously (25). Briefly, 1 ml of an exponentially growing culture was induced with Isopropyl β-D-Thiogalactopyranoside (IPTG) to a final concentration of 1 mM at an optical density of 0.4–0.7 measured at 436 nm (OD_{436}). After induction for 15 min, the culture was pulse-labelled with ∼10 μCi L-[^{35}S]-methionine for 20 s and chased with 100 mg L-methionine for 2 min before being transferred to 25 μl of chloramphenicol (100 μg/ml) on ice. The cells were harvested by centrifugation and proteins were boiled in SDS buffer and separated by 9% SDS–PAGE. The gel was dried and placed on a phosphor imager screen (Molecular Dynamics) and left to expose for 1–3 days. Relative amount of protein of the

relevant polypeptides was quantified using ImageQuant software and the frameshift efficiency (e) was determined as follows:

$$e = \frac{V_{FS}/n_{met,FS}}{V_{FS}/n_{met,FS}+V_{STOP}/n_{met,STOP}}$$

where V_{FS} is the relative radioactivity in the frameshift product, $n_{met,FS}$ is the number of methionines in the frameshift product, V_{STOP} is the relative radioactivity in the in-frame stop product and $n_{met,STOP}$ is the number of methionines in the in-frame stop product.

Two-dimensional SDS–PAGE

Two-dimensional SDS–gels were performed as described (34) with a few modifications (35) using samples from the frameshift assay described above. The frameshift efficiency was determined as described for the frameshift assay above, although polygonal shapes were used to encircle the polypeptides of interest and quantify the relative amount of radioactivity in them.

Polypeptides originating from stalled ribosomes were found as radioactive polypeptides with appropriate isoelectric point and molecular weight appearing on gels when the translated transcript contained a pseudoknot. These polypeptides were absent when a transcript without a pseudoknot was translated. The weakest stalled protein spots were difficult to distinguish from spots originating from endogenous gene expression on these gels (compare to the 0 construct in Supplementary Figure S5) and their determination is connected with some uncertainty. The statistical analysis used to compare the stalling efficiency between pseudoknot 22/6a and 22/6b was an unpaired one-tailed Student's t-test with a significance level of 0.05.

Northern blots

Total RNA was extracted from 1.5 ml culture samples by the 'Hot–phenol' extraction method and separated according to size by electrophoresis on 1.2% agarose, 6% formaldehyde gels in recirculating 1xMOPS buffer. Capillary blots were performed onto Hybond-N$^+$ (Perkin Elmer) membranes, and the RNA was crosslinked to the membrane by 0.12 J/cm2 UV light in a Stratalinker 1800. Riboprobes covering mRNA sequences as described in Figure 4 were made by T7 RNA polymerase transcripts from the pMAS39 'downstream' template (19) or from templates made by PCR where one primer included 'hanging out' T7 promoter sequences (gene*10* and *lacZ* 5′ probes). The riboprobes were synthesized in the presence of 32-P-UTP and the final specific activity was about 40 Ci/mmol of nucleotide. Hybridization and stripping of membranes were performed following standard protocols (Amersham, Hybond-N+ booklet, 2006). The membranes were wrapped in Saran wrap and placed on a phosphor imager screen (Molecular Dynamics) and left to expose over night. Signals were visualized using ImageQuant software.

RESULTS

mRNA pseudoknot constructs to separate programmed stop from ribosome stalling

We created a series of plasmids containing different pseudoknots and where the in-frame stop codon was placed either immediately upstream ('Upstream stop') or ~150 nt downstream ('Downstream stop') from the pseudoknot (Figure 1A). The 'Upstream stop' constructs had an in-frame stop codon in the spacer between the slippery sequence and the pseudoknot. This caused non-frameshifted ribosomes to produce a 28 kDa polypeptide (gene*10* from phage T7) while ribosomes undergoing a −1 frameshift continued through the pseudoknot and into *lacZ* producing a 148 kDa fusion protein of the T7 gene*10* and *lacZ* sequences. In the 'Downstream Stop' constructs we replaced the UAA stop codon immediately upstream from the pseudoknot with a lysine codon (AAA). This change caused non-frameshifting ribosomes to continue through the pseudoknot and terminate at a downstream UGA codon producing a 37 kDa polypeptide. The pseudoknot constructs based on the plasmid OFX302 (25) are detailed in Figure 1B. We systematically increased the length of stem 1 and in pseudoknot 22/6a through 22/6c, we exchanged GC with UA base pairs, thus, gradually decreasing the stability of stem 1.

Often, the number of ribosomes which undergo −1 frameshift has been determined from constructs such as our 'Upstream stop' constructs, by separating radioactively labelled proteins by SDS–PAGE and quantifying the relative amount of protein in each of the two polypeptides (28 versus 148 kDa). Given the limited resolution of SDS–PAGE, it is, however, impossible to clearly differentiate between polypeptides produced by ribosomes that terminate at the in-frame UAA stop codon and ribosomes that undergo −1 frameshift but stall within the pseudoknot. In order to overcome this problem, we invoked 2D SDS–PAGE (34) whereby polypeptides were separated not only by molecular weight but also by their isoelectric point (pI).

While polypeptides originating from ribosomes stalled in the pseudoknot varied only slightly in molecular weight, they varied significantly in their pI. Based on the 'Downstream Stop' construct, we calculated a theoretical 2D SDS–PAGE assay of a growing polypeptide as consecutive codons are translated (shown in Figure 2A). At around 28 kDa, the trace splits into two, the triangles denote the non-frameshifted product and the circles denote the −1 frameshifted product. Red symbols denote codons inside the pseudoknot. Experimental data originating from the 'Downstream Stop' construct is shown in Figure 2B, the theoretically expected features are indeed present, e.g. both the non-frameshifted (DS-stop) and the −1 frameshifted (FS) products are visible. The heat shock proteins GroEL and DnaK serve as landmarks on the gel. Interestingly, a series of polypeptides originating from ribosomes stalled inside the pseudoknot appeared (inside dashed red line). For comparison, a standard SDS–PAGE of the same sample is shown in Figure 2C, here, the second level of information

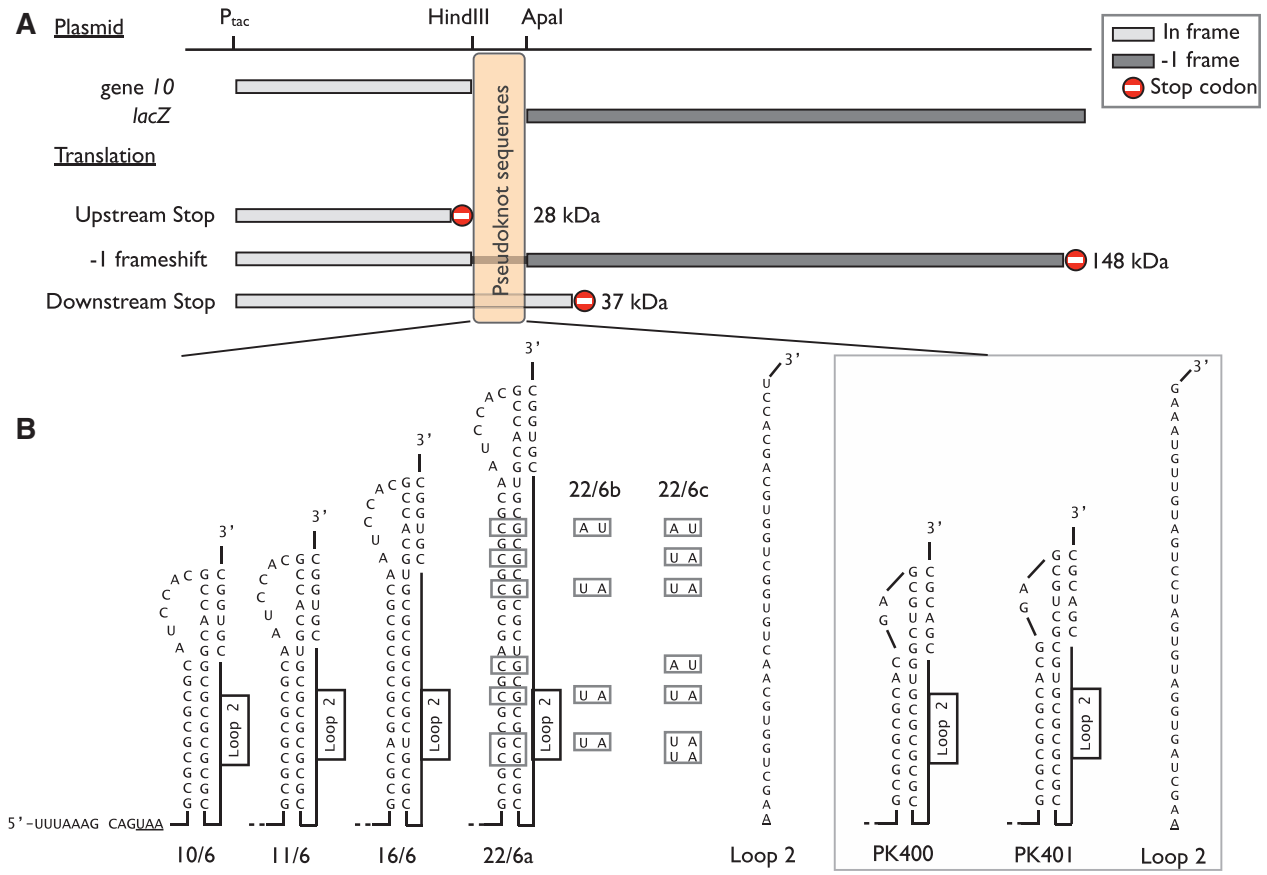

Figure 1. Frameshift assay and pseudoknot structures. (**A**) All plasmid constructs contain an IPTG inducible promoter in front of T7 gene*10* (light grey), a complete frameshift signal, and *lacZ* (dark grey). The frame shift stimulating pseudoknot-like structure is inserted downstream of gene*10*. Immediately, downstream from the pseudoknot *lacZ* is inserted in the −1 reading frame relative to gene*10*. In the 'Upstream Stop' construct the non-frameshifting ribosomes will translate gene*10* and terminate at a UAA stop codon in the spacer sequence and produce a 28 kDa polypeptide. Ribosomes undergoing −1 frameshift at the slippery sequence translate *lacZ* thus producing ~148 kDa polypeptide. In the 'Downstream Stop' construct the UAA stop codon is replaced by an AAA lysine codon thus resulting in ~37 kDa polypeptide being produced by non-frameshifting ribosomes which terminate at an UGA stop codon downstream from the pseudoknot. (**B**) Sequence and structure of the inserted pseudoknots, the slippery sequence and the spacer. In pseudoknot, 10/6, 22/6a, 22/6b and 22/6c the first base in loop 2 has been removed in order to maintain the downstream reading frame (underlined). The boxed insert in panel B shows the structure and sequence of previously described constructs (25).

(isoelectric point) is lost and the relative blurry bands are difficult to interpret.

Quantification of ribosome stalling and correlation with stem strength

The results shown in Figure 2 revealed that a 1D SDS–PAGE assay could not firmly identify polypeptides originating from a −1 frameshifted ribosome stalled in the pseudoknot from the non-frameshifted product in a 'Downstream Stop' construct. In order to quantify the amount of −1 frameshifted ribosomes stalled inside the pseudoknot, we performed a 2D SDS–PAGE separation of the radioactively labelled proteins originating from the 'Upstream Stop' construct (Supplementary Figures S4 and S5), which is the type of construct most commonly used throughout literature. The advantage of a 2D-gel analysis is that all the unfinished protein chains with different lengths concentrate in a common spot when they have the same pI. This made it possible to identify randomly stalled translation products inside the pseudoknot sequence and we quantified the amount of radioactivity in all identified additional spots. This produced a conservative estimate of the amount of stalled translations.

The result of quantifying the fraction of *in vivo* −1 frameshifted ribosomes, both those which made it all the way to the *lacZ* stop codon (gene*10*/*lacZ* fusion) and those which stalled inside the pseudoknot, is shown in Figure 3A. The hatched bars denote the −1 frameshift efficiency taking into account only the end product of −1 frameshift (148 kDa gene*10*/*lacZ* fusion). This frameshift efficiency was calculated as (intensity of FS product)/(intensity of non-FS product + intensity of FS product). The filled bars denote the −1 frameshift efficiency when both the end product (148 kDa gene*10*/*lacZ* fusion) and the products originating from stalled ribosomes are taken into account. This frameshift efficiency was calculated as (intensity of FS product + intensity of stalled product)/(intensity of non-FS product + intensity of FS product + intensity of stalled product).

In addition to the six artificial pseudoknot-like structures, we also analysed two earlier investigated pseudoknots PK400 and PK401 (25), with over-all

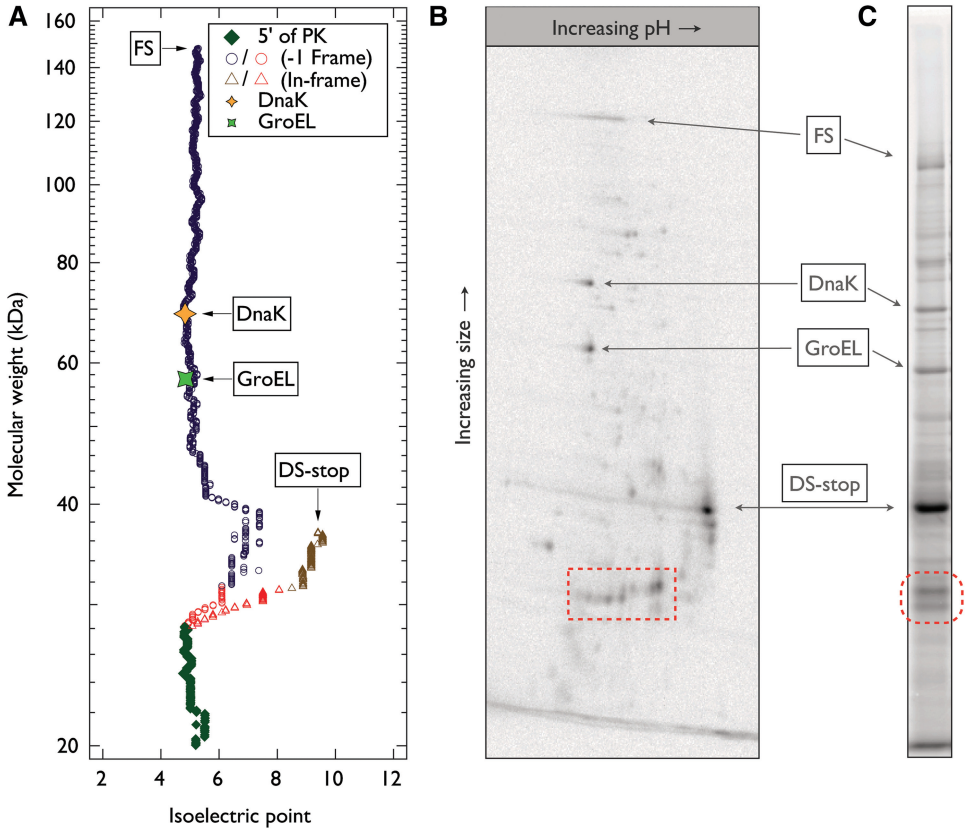

Figure 2. Stalled ribosomes. The pseudoknot used was 22/6a 'Downstrem stop'. Polypeptides produced by ribosomes stalled inside the pseudoknot are marked by red symbols. (**A**) Theoretically calculated size and isoelectric point of the growing polypeptide as consecutive codons are translated. Each symbol signifies the addition of a new amino acid and upstream from the pseudoknot this is indicated by green diamonds, inside the pseudoknot (in-frame) by red triangles, after the pseudoknot (in-frame) by brown triangles, inside the pseudoknot (−1 reading frame) by red circles, and after the pseudoknot (−1 reading frame) by purple circles. The expected size and isoelectric point of the completed peptides for both non-frameshifting ribosomes (DS-stop) and for −1 frameshifted ribosomes (FS) are indicated. The sizes and isoelectric points of DnaK (GenBank CAQ30531.1) and GroEL (GenBank CAQ34492.1) are indicated to provide landmarks. (**B**) Image of phosphor screen with L-[35S]-methionine-labelled proteins from a strain expressing 22/6a 'Downstream Stop' separated by 2D SDS–PAGE. (**C**) Same as **B** but only separated according to molecular weight by 1D SDS–PAGE.

structures more similar to naturally occurring pseudoknots (Figure 1 insert), inspired from structures in the infectious bronchitis virus (22,28,30). The pseudoknot structures in this type of virus are selected for their effects on vertebrate ribosomes, but the stem1 length variations were found to yield approximately the same relative stimulatory effect in *E. coli* (25) and suggest that stem1 strength is equally important for stimulating bacterial ribosomes to frameshift.

All pseudoknots investigated stalled some fraction of the frameshifted ribosomes, however, significantly more ribosomes stalled in the artificial pseudoknots than in those resembling naturally occurring pseudoknots (PK400 and PK401).

To quantify the amount of ribosomes stalling within a pseudoknot *in vivo* we calculated the ratio of (stalled + non-stalled frameshifted ribosomes) to (non-stalled frameshifted ribosomes), the result is shown in Figure 3B. For the IBV inspired pseudoknots, this ratio was close to 1 signifying that essentially no ribosomes stalled. However, the ratio was significantly larger than 1 for the more artificial pseudoknots which acted as roadblocks for a large amount of frameshifted ribosomes. The length of stem 1 did not significantly influence on the amount of frameshifted or stalled ribosomes. Interestingly, within pseudoknots with the same overall structure (22/6a–c) 22/6a stalls a significantly higher fraction of frameshifted ribosomes than 22/6b (verified by Student's *t*-test, $n = 4$, $\alpha = 0.05$, $P = 0.012$), which again stalls more than 22/6c. Hence, the ability to stall a ribosome correlated with the strength of the pseudoknot base pairs, the stronger the base pairs the more frameshifted ribosomes were stalled.

Messenger RNA pseudoknots are not a barrier to the RNA polymerase

Earlier studies have shown that the insertion of sequences able to form mRNA secondary structures into a gene may cause the RNA polymerase to stall or invoke a target for endonucleolytic attacks (19). Therefore, in our analysis of mRNA pseudoknot-stalled ribosomes, it was important to verify that there was no significant population of mRNAs that ended within the pseudoknot structure. If such truncated transcripts were abundant, it would be difficult

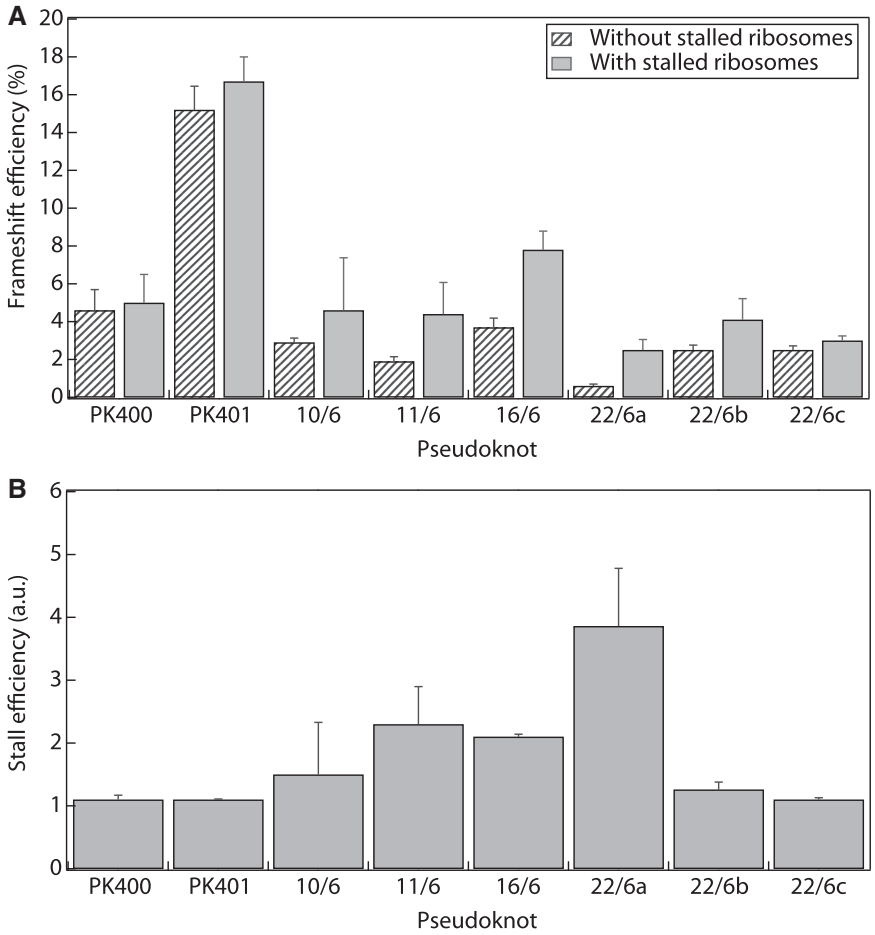

Figure 3. Frameshift and stall efficiency. (**A**) *In vivo* frameshift efficiency of different 'Upstream stop' constructs containing pseudoknots without taking peptides from stalled ribosomes into account (hatched bars) or when the peptides originating from stalled ribosomes are taken into account (filled bars). (**B**) Stall efficiency (i.e. ratio of all frameshifted ribosomes to non-stalled frameshifted ribosomes). Values are mean ± SEM, $n = 2$–4.

to distinguish between protein products from ribosomes stalled within the pseudoknot and protein products originating from ribosomes ending translation at 'non-stop' mRNAs having their 3′-ends within the pseudoknot sequence. In the latter case translation would be terminated by tmRNA trans-translation thus rendering the protein products unstable due to the tmRNA-encoded tag (36). In the following subsections 'Identification of transcripts from the T7gene10-PK-lacZ gene fusions', 'Messenger RNA stability' and 'Coupling between translation and transcription is required for full-length transcripts', we will show that the observed proteins did indeed originate from stalled ribosomes and that they were not caused by other effects.

Identification of transcripts from the T7gene10-PK-lacZ gene fusions. To identify the major class of transcripts from our pseudoknot containing constructs, we made a northern blot with RNA from all strains used to measure frameshift frequencies, which are those containing the upstream stop. We used three different probes hybridizing either upstream of the pseudoknot, immediately downstream of the pseudoknot or in the very end of the *lacZ* reading frame (Figure 4A).

As seen in Figure 4B–D, there was an unspecific hybridization from all three probes to the 23S and 16S ribosomal RNAs. In *E. coli*, ribosomal RNA constitutes between 80% and 90% of total RNA depending on growth conditions and some cross-hybridization to these species is often seen in northern blots. Here, the uninduced culture in Figure 4B–D, lane '0 no IPTG', made it possible to estimate the unspecific probing to rRNA and the two bands were used as size markers on the blots.

Following induction with IPTG, all strains showed increased hybridization above the 23S RNA band compared to the uninduced control with all three probes. The so-called 0 construct was described in reference (25), and contains a slippery sequence and the UAA stop codon but no pseudoknot-like structure. In all strains, except the one with the 0 construct, there were a distinct band (Fl) representing the expected full-length transcript. The full-length transcript reached from transcription start to the stem–loop structure downstream of the 3′-end of the *lacZ* open reading frame ('hp' in Figure 4A). This mRNA stem–loop structure has been shown to stabilize the *lacZ* transcript by reducing 3′-end exonucleolytic attacks (37). The core plasmid contained no distinct transcription termination signal after the

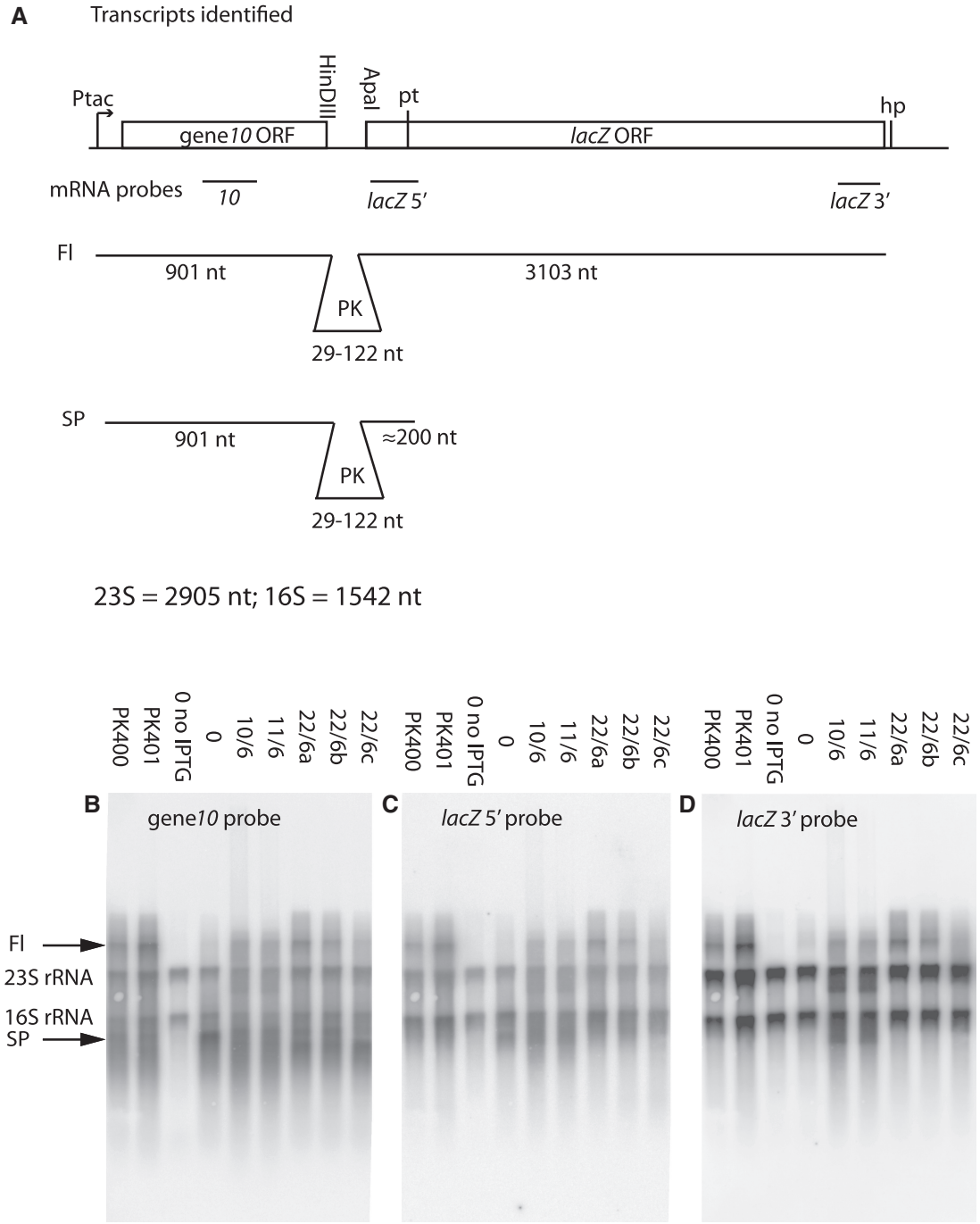

Figure 4. Transcripts from pseudoknot containing constructs. Northern blot showing transcripts from 'Upstream stop' constructs containing pseudoknot structures. Panel (**A**) Map of genes, transcripts and probes (not drawn to scale). PK: sequences of pseudoknot structures inserted between the HindIII–ApaI restriction sites; Fl, full-length transcript from +1 to the mRNA stabilizing hair-pin (hp) in the end of the *lacZ* open reading frame; SP, premature transcription termination fragment originating from transcription from +1 to the premature transcription termination site (pt) where RNA–polymerase terminates in cases where translating ribosomes are uncoupled from the transcribing polymerase; 23S and 16S rRNA: ribosomal RNA from the large (50S) and small (30S) ribosomal subunit, respectively. Panel (**B–D**) northern blots. The strains were induced by IPTG for 15 min before harvest for RNA extraction. Each lane contains 1 μg of RNA from a strain expressing the gene construct indicated above the lane. No IPTG: no induction of P_{tac} transcription. The blot was probed with three different Ribo-probes: (B) T7 gene*10* probe covering nucleotides +476 to +676 of the induced transcript; (C) *lacZ* 5′ probe covering 8–247 nt after the ApaI site (approximately nucleotides +1000 to +1300 of the transcript); (D) *lacZ* 3′ probe covering nucleotides +2769 to +3010 after the ApaI site (probe covering upstream to the last 50 nt of the open reading frame of *lacZ*).

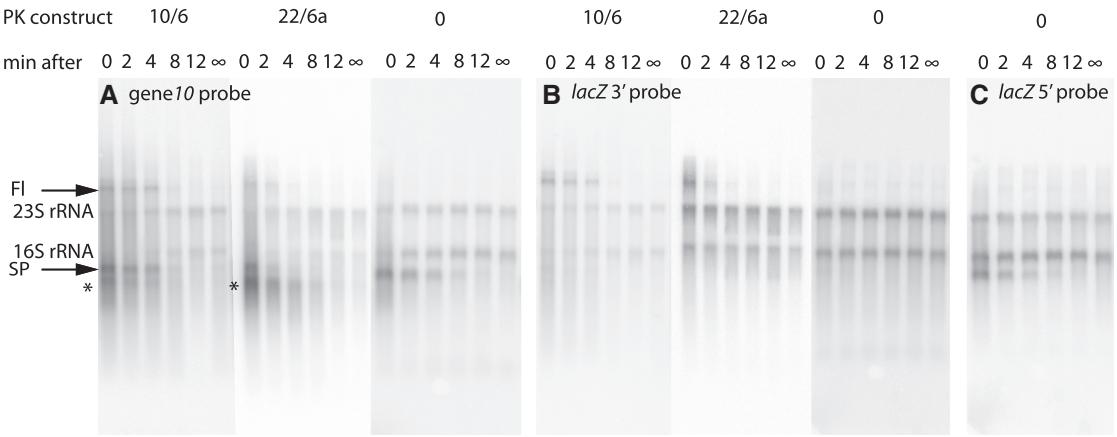

Figure 5. mRNA stability. Northern blots showing chemical stability of IPTG inducible transcripts from three PK containing plasmids all having the upstream stop UAA codon. The three strains were induced by IPTG for 15 min, then IPTG was removed by filtration at 37°C and aliquots were harvested for RNA extraction. Each lane contains 1 μg of RNA from the strain indicated above each blot (containing six lanes). The time of harvest after IPTG was removed is indicated above each lane; 0: sample harvested before filtration; ∞: sample harvested either 16 min after (10/6) or before induction (0 and 22/6a). * indicates an RNA band discussed in text. Other symbols: see caption of Figure 4.

lacZ gene, and accordingly we found transcripts that exceeded far beyond the full-length Fl band (Figure 4B–D).

In the beginning of *lacZ*, ~200 nt into the open reading frame, there is a site, called 'pt' in Figure 4A, where the RNA polymerase is caused to terminate if there is inefficient translation initiation of the *lacZ* gene (38). In the 0 construct there is no pseudoknot to stimulate frameshift at the slippery site. Therefore, virtually no ribosomes were expected to follow the RNA polymerase from gene *10* into the *lacZ* part of our gene fusion. As expected, Figure 4B and C, lane '0' shows a prominent band ('SP' for stop polymerase) corresponding in size and probe-ability to this premature termination product. Also, corresponding low amounts of high molecular weight transcripts are detected for this construct. All the other constructs shown in Figure 4 contained frameshift stimulating pseudoknots and a inspection of the northern blot showed that the 'SP' bands probed with both gene*10* and *lacZ5′* sequences were present in sizes which correspond to the sizes of the pseudoknots inserted.

Messenger RNA stability. The wild type *lacZ* mRNA half-life is close to the average mRNA half-life in *E. coli* (120 s) and transcription takes close to 80 s due to the length of the *lacZ* gene (three times longer than the average gene). Therefore, a northern blot of wild type *lacZ* mRNA under steady state transcription will always include a lot of unfinished native transcripts, as well as mRNAs under degradation. Here, our gene*10-lacZ* fusion was even longer and transcription should take ~120 s. Accordingly, all induced strains included in Figure 4 show a distinct smear of mRNA fragments recognized by all three probes. In order to examine the half-life of our artificial transcripts, we made experiments where transcription from the P_{tac} promotor was stopped due to removal of the inducer (Figure 5). Two minutes after IPTG removal, any remaining smear should originate from mRNA degradation because most of the RNA polymerase should have reached the end of the gene fusion at this time.

From the experiment, shown in Figure 5, it is evident that both the 'Fl' and the 'SP' mRNA fragments had a half-life close to the average 2 min *E. coli* mRNA half-life. In addition, both the pseudoknot containing constructs (10/6 and 22/6a) revealed the existence of a short mRNA fragment that was recognized only by the gene*10* probe but not the *lacZ5′* and 3′ probes (indicated by 'asterisks' in Figure 5). This fragment includes the transcription start in the 5′-end and the pseudoknot in the 3′-end. We suggest that the pseudoknot acts as an exonuclease barrier like the natural stem–loop structure in the 3′-end of the wild type *lacZ* transcript (37) and thereby induces a degradation intermediate of a distinct length with increased half life compared to unstructured mRNA sequences like those from construct 0. Alternatively, but not mutually exclusive, a pseudoknot acts like a rho-independent termination signal to the RNA polymerase. However, the sequences were not followed by a row of uridine residues, which would be necessary to make a stem–loop structure into a functional transcription terminator.

Coupling between translation and transcription is required for full-length transcripts. The final test of our model for the transcription pattern in our artificial gene fusion was to establish translational coupling beyond the slippery sequence and into the polar termination site (SP) in *lacZ*. By changing the upstream stop codon between the slippery site and the pseudoknot region into a sense codon ribosomes should, frameshifted or not, follow the RNA polymerase into the beginning of the *lacZ* sequence.

The 22/6a and the 0 constructs are the two constructs with the lowest frequency of frameshifting. Therefore, they have the least ribosome traffic into the *lacZ* sequence. Alteration of the UAA stop codon into a lysine AAA codon in the spacer between the slippery sequence and the pseudoknot changed the pattern of transcripts immensely. These two downstream stop variants ('DS. stop' in Figure 6), which did not contain a stop codon upstream from the structure, expressed significantly more full-length ('Fl') transcript and only insignificant

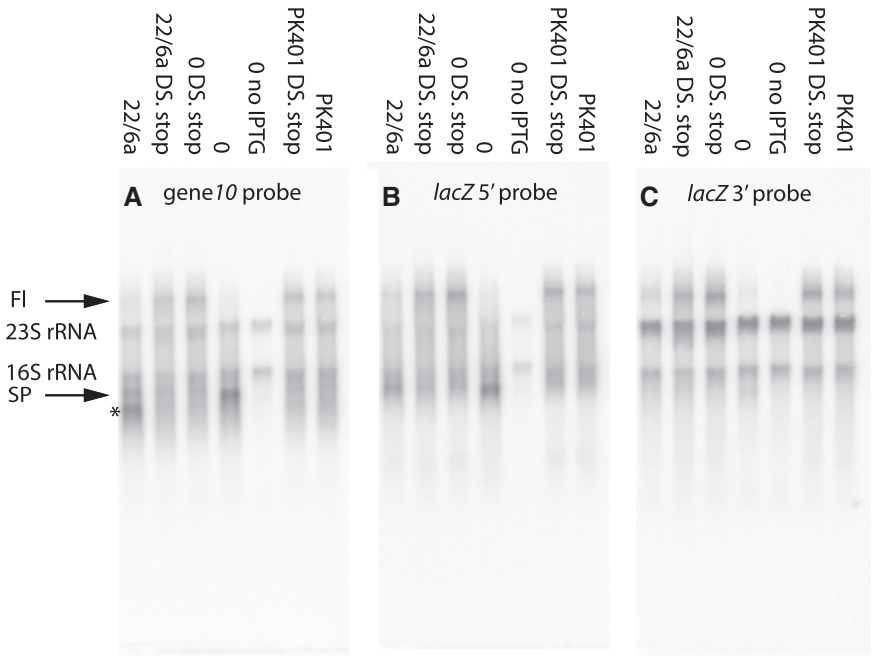

Figure 6. Transcriptional polarity in the *lacZ* gene. Northern blot comparing transcripts from 'Upstream stop' and 'Downstream stop' (DS. stop) constructs. The strains were induced by IPTG for 15 min before harvest for RNA extraction. Each lane contains 1 μg of RNA from the strain indicated above the lane. No IPTG: no induction of P_{tac} transcription. Other symbols: see caption of Figures 4 and 5. In PK401 DS. stop, the nucleotide T encoding the most 3′ U in Loop 2 (Fig. 1) was mutated into an A to avoid an internal stop codon.

amounts of premature transcription stop fragment ('SP') compared to their sister constructs containing the UAA stop codon upstream from the structure (Figure 6). Our control construct, PK401, which stimulated 14% frameshift, showed no premature transcription stop fragment ('SP') and therefore no change in transcription pattern was observed as a consequence of removing the upstream UAA stop codon (Figure 6) thus confirming that the major effect causing the 'SP' fragment is polarity in the *lacZ* gene and not transcription termination caused by the pseudoknot sequences.

Also, the very short band marked by 'asterisks' that appeared from the 22/6a construct was not present in the 'DS. stop' variant (Figure 6). This exclude this mRNA fragment to be causal for the appearance of stalled protein products, because 22/6a ('DS. stop') is the construct that caused the highest frequency of stalling (compare Figure 2 and Supplementary Figure S5).

Our conclusion is that the stable proteins observed from within the pseudoknot structures (Figure 2, Supplementary Figure S1, S2, S4 and S5) were products from stalled ribosomes. The stalling of the ribosomes was directly caused by the tertiary structure and not by some secondary effect, as, e.g. stop codon-less mRNA fragments ending within the structure sequences.

DISCUSSION

The structures analysed in this study are artificial and were designed to fold into pseudoknot-like structures with a gradually increasing mechanical strength. The mechanical strength was adjusted by changing the base pairs of the two stems, which seems to be a reasonable way of crudely varying the mechanical strength, as the energy involved in base pairing is higher than the energies involved in, e.g. the electrostatic interaction of the loop with the stems. It is, however, likely that the loop–stem interaction, surface charges or other players than just mechanical strength influence frameshift stimulating effect of mRNA structures. As there is a consensus in recent literature that pseudoknot mechanical strength correlates with frameshifting efficiency (23–25), it was intriguing that the amount of frameshifted product was reduced by the stronger pseudoknot 22/6a compared to the weaker 22/6b or c (Figure 3A). This proved to be caused by stalling of a significant amount of frameshifted ribosomes by the strong pseudoknots (Figure 3B). Future studies will show whether significant stalling can also be caused by naturally occurring pseudoknots.

Quantitative northern blot analysis was used to examine whether the observed translation products ending within the pseudoknot structure arose from fragments of mRNA produced either by low RNA–polymerase processivity or specific endonucleolytic attacks by RNases at the pseudoknot sequences. No evidence was found of a specific population of transcripts that could explain the amounts of protein products attributed to originate from pseudoknot-stalled ribosomes. Also, our protein-stability assay showed that the translational products from the stalled ribosomes were stable for at least 80 min (Supplementary Figure S1), thus indicating that the stalled ribosomes are not rescued by tmRNA and that the stalled proteins do not originate from truncated mRNA.

We also checked whether the protein products from within the pseudoknot structure could arise from very

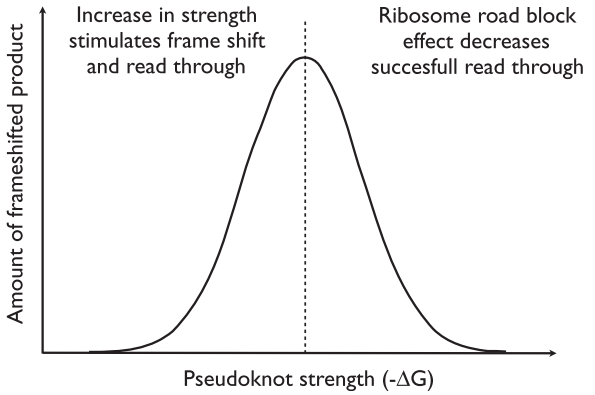

Figure 7. Model of frameshifting efficiency. Increasing the strength of a pseudoknot causes the pseudoknot to induce frameshifting at a higher frequency. However, the stronger the pseudoknot the larger the likelihood that it will act as a roadblock for the ribosome, reducing the amount of frameshifted product produced. The optimal frameshifting efficiency is achieved by balancing the two contributions.

slow rather than permanently stalled ribosomes. A pulse chase experiment (Supplementary Figure S2) revealed that within 16 min there was no sign of a redistribution of label between the stalled spots and the stop codon-terminated downstream stop product, thus proving the possibility of very slow ribosomes to be unlikely.

It is possible that the newly discovered ribosome rescue factor, ArfA (39) could be active at pseudoknot-stalled ribosomes and that nascent proteins would be more stable than if saved by tmRNA. However, as can be seen in Supplementary Figure S3, the growth of strains expressing pseudoknot 22/6a was severely affected by induction and showed a decrease in growth rate correlating to the amount of stall product observed. Because ribosomes are limiting in growing cells (40), the sequestration of ribosomes by engagement in induced overexpression of a gene from a plasmid will often cause a strain to grow slower than the uninduced counterpart. The enhanced reduction in growth rate upon induction of 22/6a compared to the 0 construct (Supplementary Figure S3) could indicate that stalled ribosomes were not rescued at a sufficiently high rate and we suggest that either the ribosomal rescue systems were titrated by the large amount of mRNA induced from the plasmid alleles, or alternatively, that no rescue is possible for pseudoknot-stalled ribosomes.

Our results are in agreement with the observation that the amount of protein produced from an mRNA can be reduced when a pseudoknot is located upstream (29). Also, they provide a possible explanation for the reduction in frameshift efficiency observed by, e.g. Napthine *et al.* (30) when increasing the thermodynamic stability of stem 1 above a certain threshold. This apparent reduction in frameshift efficiency (observed by 1D SDS–PAGE) could be caused by the fact that a significant fraction of the 'frameshifted' ribosomes permanently stalled within the pseudoknot.

We propose that pseudoknot induced frameshifting efficiency can be viewed as a balance between two competing effects (as visualized in Figure 7), the mechanically stronger the pseudoknot, the larger the frameshifting efficiency (25–27), however, the stronger the pseudoknot the larger the likelihood of stalling the frameshifted ribosome, thus preventing the translation of full-length frameshift product. Possibly, evolution optimized viral pseudoknots to balance these two effects. Hence, in measurements of frameshifting efficiency it is important to take into account the roadblocking effect of mRNA pseudoknots.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors thank M. Warrer for excellent technical support. J.T., L.B.O. and M.A.S. designed research, analysed data and wrote paper. J.T. and M.A.S. performed research.

FUNDING

Funding for open access charge: University of Copenhagen Excellence Program.

Conflict of interest statement. None declared.

REFERENCES

1. Brierley,I., Gilbert,R.J.C. and Pennell,S. (2010) Pseudoknot-Dependent Programmed-1 Ribosomal Frameshifting: Structures, Mechanisms and Models. In Atkins,J.F. and Gesteland,R.F. (eds), *Recoding: Expansion of Decoding Rules Enriches Gene Expression*, Vol. 24. Springer, New York/Dordrecht/Heidelberg/London, pp. 149–174.
2. Farabaugh,P.J. (2010) Programmed Frameshifting in Budding Yeast. In Atkins,J.F. and Gesteland,R.F. (eds), *Recoding: Expansion of Decoding Rules enriches Gene Expression*, Vol. 24. Springer, New York/Dordrecht/Heidelberg/London, pp. 221–247.
3. Fayet,O. and Prère,M.-F. (2010) Programmed Ribosomal-1 Frameshifting as a Tradition: The Bacterial Transposable Elements of the IS*3* Family. In Atkins,J.F. and Gesteland,R.F. (eds), *Recoding: Expansion of Decoding Rules Enriches Gene Expression*, Vol. 24. Springer, New York/Dordrecht/Heidelberg/London, pp. 259–280.
4. Brierley,I. and Dos Ramos,F.J. (2006) Programmed ribosomal frameshifting in HIV-1 and the SARS-CoV. *Virus Res.*, **119**, 29–42.
5. Firth,A.E. and Atkins,J.F. (2009) A conserved predicted pseudoknot in the NS2A-encoding sequence of West Nile and Japanese encephalitis flaviviruses suggests NS1′ may derive from ribosomal frameshifting. *Virol. J.*, **6**, 14.
6. Shehu-Xhilaga,M., Crowe,S.M. and Mak,J. (2001) Maintenance of the Gag/Gag-Pol ratio is important for human immunodeficiency virus type 1 RNA dimerization and viral infectivity. *J. Virol.*, **75**, 1834–1841.
7. Dinman,J.D. and Wickner,R.B. (1992) Ribosomal frameshifting efficiency and gag/gag-pol ratio are critical for yeast M1 double-stranded RNA virus propagation. *J. Virol.*, **66**, 3669–3676.
8. Giedroc,D.P. and Cornish,P.V. (2009) Frameshifting RNA pseudoknots: structure and mechanism. *Virus Res.*, **139**, 193–208.
9. Namy,O., Moran,S.J., Stuart,D.I., Gilbert,R.J. and Brierley,I. (2006) A mechanical explanation of RNA pseudoknot function in programmed ribosomal frameshifting. *Nature*, **441**, 244–247.

10. Pallan,P.S., Marshall,W.S., Harp,J., Jewett,F.C., Wawrzak,Z., Brown,B.A., Rich,A. and Egli,M. (2005) Crystal Structure of a Luteoviral RNA Pseudoknot and Model for a Minimal Ribosomal Frameshifting Motif. *Biochemistry*, **44**, 11315–11322.
11. Weiss,R.B., Dunn,D.M., Dahlberg,A.E., Atkins,J.F. and Gesteland,R.F. (1988) Reading frame switch caused by base-pair formation between the 3′ end of 16S rRNA and the mRNA during elongation of protein synthesis in Escherichia coli. *EMBO J.*, **7**, 1503–1507.
12. Weiss,R.B., Dunn,D.M., Shuh,M., Atkins,J.F. and Gesteland,R.F. (1989) E. coli ribosomes re-phase on retroviral frameshift signals at rates ranging from 2 to 50 percent. *New Biol.*, **1**, 159–169.
13. Garcia,A., van Duin,J. and Pleij,C.W.A. (1993) Differential response to frameshift signals in eukaryotic and prokaryotic translational systems. *Nucleic Acids Res.*, **21**, 401–406.
14. Ivanov,I.P., Gesteland,R.F., Matsufuji,S. and Atkins,J.F. (1998) Programmed frameshifting in the synthesis of mammalian antizyme is +1 in mammals, predominantly +1 in fission yeast, but -2 in budding yeast. *RNA*, **4**, 1230–1238.
15. Kozak,M. (1989) Circumstances and mechanisms of inhibition of translation by secondary structure in eucaryotic mRNAs. *Mol. Cell. Biol.*, **9**, 5134–5142.
16. Hall,M.N., Gabay,J., Debarbouille,M. and Schwartz,M. (1982) A role for mRNA secondary structure in the control of translation initiation. *Nature*, **295**, 616–618.
17. Von Heijne,G., Nilsson,L. and Blomberg,C. (1977) Translation and messenger RNA secondary structure. *J. Theor. Biol.*, **68**, 321–329.
18. Qu,X., Wen,J.-D., Lancaster,L., Noller,H.F., Bustamante,C. and Tinoco,I. (2011) The ribosome uses two active mechanisms to unwind messenger RNA during translation. *Nature*, **475**, 118–121.
19. Sørensen,M.A., Kurland,C.G. and Pedersen,S. (1989) Codon usage determines the translation rate in *Escherichia coli*. *J. Mol. Biol.*, **207**, 365–377.
20. Tu,C., Tzeng,T.H. and Bruenn,J.A. (1992) Ribosomal movement impeded at a pseudoknot required for frameshifting. *Proc. Natl Acad. Sci. USA*, **89**, 8636–8640.
21. Somogyi,P., Jenner,A.J., Brierley,I. and Inglis,S.C. (1993) Ribosomal pausing during translation of an RNA pseudoknot. *Mol. Cell. Biol.*, **13**, 6931–6940.
22. Kontos,H., Napthine,S. and Brierley,I. (2001) Ribosomal pausing at a frameshifter RNA pseudoknot is sensitive to reading phase but shows little correlation with frameshift efficiency. *Mol. Cell. Biol.*, **21**, 8657–8670.
23. Plant,E.P. and Dinman,J.D. (2005) Torsional restraint: a new twist on frameshifting pseudoknots. *Nucleic Acids Res.*, **33**, 1825–1833.
24. Ito,K., Chiba,S. and Pogliano,K. (2010) Divergent stalling sequences sense and control cellular physiology. *Biochem. Biophys. Res. Commun.*, **393**, 1–5.
25. Hansen,T.M., Reihani,S.N., Oddershede,L.B. and Sorensen,M.A. (2007) Correlation between mechanical strength of messenger RNA pseudoknots and ribosomal frameshifting. *Proc. Natl Acad. Sci. USA*, **104**, 5830–5835.
26. Chen,G., Chang,K.Y., Chou,M.Y., Bustamante,C. and Tinoco,I. Jr (2009) Triplex structures in an RNA pseudoknot enhance mechanical stability and increase efficiency of -1 ribosomal frameshifting. *Proc. Natl Acad. Sci. USA*, **106**, 12706–12711.
27. Green,L., Kim,C.-H., Bustamante,C. and Tinoco,I. Jr (2008) Characterization of the Mechanical Unfolding of RNA Pseudoknots. *J. Mol. Biol.*, **375**, 511–528.
28. Brierley,I., Digard,P. and Inglis,S.C. (1989) Characterization of an efficient coronavirus ribosomal frameshifting signal: requirement for an RNA pseudoknot. *Cell*, **57**, 537–547.
29. Plant,E.P., Rakauskaite,R., Taylor,D.R. and Dinman,J.D. (2010) Achieving a golden mean: mechanisms by which coronaviruses ensure synthesis of the correct stoichiometric ratios of viral proteins. *J. Virol.*, **84**, 4330–4340.
30. Napthine,S., Liphardt,J., Bloys,A., Routledge,S. and Brierley,I. (1999) The role of RNA pseudoknot stem 1 length in the promotion of efficient -1 ribosomal frameshifting. *J. Mol. Biol.*, **288**, 305–320.
31. Liphardt,J., Napthine,S., Kontos,H. and Brierley,I. (1999) Evidence for an RNA pseudoknot loop-helix interaction essential for efficient-1 ribosomal frameshifting. *J. Mol. Biol.*, **288**, 321–335.
32. Neidhardt,F.C., Bloch,P.L. and Smith,D.F. (1974) Culture medium for enterobacteria. *J. Bacteriol.*, **119**, 736–747.
33. Reeder,J., Steffen,P. and Giegerich,R. (2007) pknotsRG: RNA pseudoknot folding including near-optimal structures and sliding windows. *Nucleic Acids Res.*, **35**, W320–W324.
34. O'Farrell,P.H. (1975) High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.*, **250**, 4007–4021.
35. Sørensen,M.A. and Pedersen,S. (1997) Determination of the Peptide Elongation rate In Vivo. In Martin,R. (ed.), *Methods in Molecular Biology: Protein Synthesis: Methods and protocols*, Vol. 77. The Humana Press, Inc., Totowa, New Jersey, NJ, USA, pp. 129–142.
36. Farrell,C.M., Grossman,A.D. and Sauer,R.T. (2005) Cytoplasmic degradation of ssrA-tagged proteins. *Mol. Microbiol.*, **57**, 1750–1761.
37. Cannistraro,V.J., Subbarao,M.N. and Kennell,D. (1986) Specific endonucleolytic cleavage sites for decay of Escherichia coli mRNA. *J. Mol. Biol.*, **192**, 257–274.
38. Stanssens,P., Remaut,E. and Fiers,W. (1986) Inefficient translation initiation causes premature transcription termination in the lacZ gene. *Cell*, **44**, 711–718.
39. Chadani,Y., Ono,K., Ozawa,S., Takahashi,Y., Takai,K., Nanamiya,H., Tozawa,Y., Kutsukake,K. and Abo,T. (2010) Ribosome rescue by Escherichia coli ArfA (YhdL) in the absence of trans-translation system. *Mol. Microbiol.*, **78**, 796–808.
40. Vind,J., Sørensen,M.A., Rasmussen,M.D. and Pedersen,S. (1993) Synthesis of proteins in Escherichia coli is limited by the concentration of free ribosomes: expression from reporter genes does not always reflect functional mRNA levels. *J. Mol. Biol.*, **231**, 678–688.