

## Efficient Ribosomal Frameshifting Can Occur at the Beginning of the Translation

Yang-Gyun Kim,<sup>\*</sup> Jeong Suk Chi,<sup>†</sup> Seonmin Lee, Nyoun Soo Kwon,<sup>†</sup> Kwang Jin Baek,<sup>†</sup> and Hye-Young Yun<sup>†,\*</sup>

Department of Chemistry, School of Natural Sciences, Sungkyunkwan University, Suwon 440-746, Korea

<sup>\*</sup>E-mail: ygkimmit@skku.edu<sup>†</sup>Department of Biochemistry, College of Medicine, Chung-Ang University, Seoul 156-756, Korea. <sup>\*</sup>E-mail: hyyun@cau.ac.kr

Received June 13, 2011, Accepted July 19, 2011

**Key Words :** Translation, Programmed-1 ribosomal frameshifting, RNA pseudoknot, Gene expression, Bidirectional promoter

Programmed -1 ribosomal frameshifting (-1RFS) is one of well-known alternative decoding mechanisms found in nature. -1RFS mechanism has been first described as a mechanism that controls the relative expression levels of two proteins in metazoan viruses.<sup>1</sup> Furthermore, among many retroviruses, plant viruses, coronaviruses and certain bacterial and protozoan genes, this has been identified as a mechanism that modulates the translation of two proteins encoded by overlapping open reading frames present in one mRNA.<sup>2,3</sup> Their prevalence across such an evolutionarily diverse distribution manifests that such sites have evolved several times.

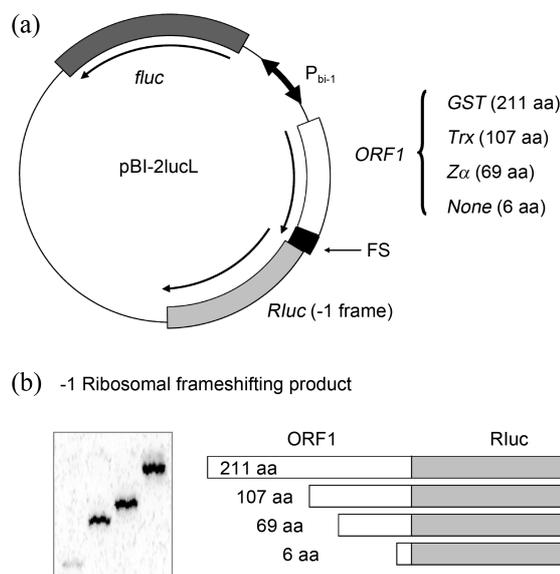
Sequence comparison and molecular genetic analysis of many of -1RFS sites show a canonical structure for these frameshift sites. It is now clear that two *cis*-elements in mRNA alone are enough to create this unusual alternative decoding event, although there are some possible influences of other *trans*-factors on frameshifting efficiency. Two *cis*-elements are involved in this alternative reading process of genetic codes during translation.<sup>4</sup> One is the slippery sequence, a heptanucleotide motif XXXYYYN, where -1 frameshifting occurs. The other is a downstream RNA secondary structure, usually a pseudoknot. The -1RFS system containing these two *cis*-elements can induce -1 frameshifting at a slippery site (slippery sequence) with 1 to over 30% of efficiency.

Recent studies have worked on not only the detailed mechanism and functional importance of -1RFS,<sup>5-7</sup> but also possible utilization of -1RFS for regulation of gene expression. For example, Dinman's group demonstrated that -1RFS signal from L-A virus successfully functions as a *cis*-acting mRNA destabilizing element.<sup>8</sup> In addition, small ligands targeting the embedded biotin-binding RNA sequence on mRNA that forms a pseudoknot-like structure when it bound with a biotin is sufficient for stimulating -1RFS when they replace biotin to bind the pseudoknot.

Here, we carried out the experiments to place -1RFS signal into genes of interest. This study was aimed to answer two questions. Firstly, can efficient -1RFS occur after very preceding translation, e.g. a less than 10 amino acid residues? Secondly, what is the effect of the preceding translation length on the efficiency of -1RFS? We used the following experimental strategy. As a reliable assay system for the

functional analysis of -1RFS *in vivo*, we chose the dual-luciferase assay system that measures the luminescence activities of *Renilla* luciferase (*Rluc*) and firefly luciferase (*fluc*) by a luminometer. In this assay system, activities of two luciferase-gene products are subsequently measured in the same tube. Thus, relative amounts of two luciferase proteins that reflect efficiency of -1RFS events can be determined by the activities of two luciferases.

The reporter vectors were designed to have both *Rluc* and *fluc* genes in two separate open reading frames in the opposite direction using pBI-L (BD Biosciences) vector as a



**Figure 1.** Schematic map of the pBI-2lucL reporter vector and the reporter gene expression by -1RFS. (a) Map of the pBI-2lucL reporter vector. The reporter vector, pBI-2lucL, has two luciferase genes, firefly luciferase (*fluc*) and *Renilla* luciferase (*Rluc*), that are oriented in the opposite directions but both under the control of the same bidirectional promoter,  $P_{bi-1}$ . The -1RFS signal (FS) is placed between an *ORF1* and the 5' region of the *Rluc* gene that fused in -1 frame. Three genes with different lengths were fused to the *Rluc* gene via the FS; genes for the  $Z\alpha$  domain from human ADAR1 ( $Z\alpha$ ), thioredoxin protein (Trx) and glutathione S-transferase protein (GST). (b) Fusion protein products from various *ORF1*-reporter gene constructs by -1RFS. *ORF1*-*Rluc* fusion proteins are produced only when -1RFS occurs at the FS. Fusion proteins produced from -1RFS event were detected by an anti-*Renilla* luciferase monoclonal antibody (left panel).

**Table 1.** Sequences of -1RFS signals used in this study

FS	Slippery sequence-Spacer-Pseudoknot sequence
WT	<u>UUUAAAC</u> AUGCAU <b>CGCGGCACCGUCCGCCGAACAACGG</b>
U19C19a	<u>UUUAAAC</u> AUGCAU <b>CGCGGCACCGUCCGCCUCAACAACGG</b>

Two -1RFS signals, WT and U19C19a, were used in this study, which differ in RNA pseudoknot sequence (bold). WT has the same sequence as the wild-type frameshifting pseudoknot from BWYV, while U19C19a contains mutations in the loop 2 region that increases -1RFS efficiency three times higher than that of the wild-type pseudoknot.<sup>12</sup>

backbone (Fig. 1(a)). Both luciferases in these reporter vectors can be expressed simultaneously by the  $P_{bi-1}$  promoter, which is a bidirectional promoter containing two minimal CMV promoters.<sup>9</sup> The transcriptional activation of the  $P_{bi-1}$  promoter for luciferase genes can be induced by doxycycline.<sup>10</sup>

The -1RFS signal that consisted of either wild-type (WT) or a mutant (U19C19a) frameshifting pseudoknot from Beet western yellow virus (BWYV)<sup>11,12</sup> and the UUUAAAC slippery sequence was inserted into the upstream region of the *Rluc* gene (Fig. 1 and Table 1). In these constructs, Rluc protein would be translated if -1RFS occurs at the slippery site of a -1RFS signal placed in the upstream of the *Rluc* gene. Using an in-frame construct as a positive control defining 100% frameshifting efficiency and an out-of-frame construct as a negative control defining 0% frameshifting efficiency, -1RFS efficiency was calculated "indirectly" from ratios of two luciferase activities. We assumed that both luciferase genes are transcribed simultaneously by the bidirectional promoter,  $P_{bi-1}$ . However, it does not guarantee that amounts of transcriptions from both directions are the same although their transcription ratio would be steady. Thus, it should be noted that frameshifting efficiencies from this assay system do not reflect accurate efficiencies of -1 frameshifting event occurred in the inserted -1 frameshifting *cis*-elements, but rather it is a relative amount of two luciferase proteins produced from the same reporter vector.

Frameshifting efficiencies of -1RFS signals inserted at the upstream region of the *Rluc* gene of pBI-2lucL were determined in HEK293 Tet-On cells. Expression of luciferases from reporter vectors was induced by doxycycline to the  $P_{bi-1}$  promoter. Using the dual-luciferase assay system, the -1RFS efficiency was measured at various time points between 12 and 96 hr after transfection. The results showed that -1 frameshifting efficiencies have not been changed significantly between 18 and 48 hr, while there were significant increases after 72 hr (data not shown). In addition, there were too low luciferase activities to determine reliable measurements at 12 hr due to low protein expressions. Thus, we decided to use the 48 hr induction time for following experiments, where the -1RFS efficiency is stable and amounts of luciferase activities are reliable.

-1RFS efficiencies from pBI-2lucL-WT and -U19C19a reporter vectors were 0.539% and 3.667%, respectively, after 48 hr. These numbers were significantly lower than efficiencies from the previous report *in vivo*.<sup>12</sup> The difference in spacer sequences between our constructs and the constructs used in the previous study<sup>12</sup> may contribute to differ-

ences in -1 frameshifting efficiency. It is also possible that the efficiency numbers obtained from different assay systems may be not the same. In addition, the assay system used in this study may not reflect the actual efficiency of the -1RFS signals embedded in the *Rluc* gene. However, an alternative explanation is possible, which a very short preceding translation (in fact, only six amino acids) before reaching the -1RFS signal in the upstream of the *Rluc* gene in pBI-2lucL reporter constructs may be responsible for causing low frameshifting efficiencies. As far as we know, there is no report regarding the effect of the short preceding translation on -1RFS efficiency. To investigate this possibility, we examined the effect of the preceding translation length on -1RFS.

In viral systems, the -1RFS signal is positioned between two genes (*ORF1* and *ORF2*) so that a frameshifting product results in a fusion product of ORF1-ORF2.<sup>2</sup> However, there is a little known about the effect of the preceding translation length (i.e. ORF1) on -1RFS efficiency. Almost all experiments conducted in other studies used viral genes or other protein genes as *ORF1*s, lengths of which are usually well over 200 amino acid residues. In the initial result of this study, efficiencies of -1RFS were still substantial with the preceding translation of just 6 amino acid residues (Table 2). To examine the effect of the preceding translation length, we constructed new reporters that contain an *ORF1* gene with various lengths at the upstream region of the *Rluc* gene in pBI-2lucL. Three different genes were used as an *ORF1* in the -1RFS system. The  $Z\alpha$  domain gene (*Za*), the first Z-DNA binding domain from human ADAR1 protein, encodes 69 amino acid residues and it is well expressed and very soluble in *E. coli*.<sup>13</sup> Thioredoxin (*Trx*) gene and glutathione S-transferase (*GST*) gene encode 107 and 211 amino acid residues, respectively. These two proteins are well-known

**Table 2.** Efficiency of -1 RFS\* dependent on the length of the preceding translation and pseudoknot

	FS	WT	U19C19a
ORF1			
6 aa.		0.539 ± 0.053%	3.667 ± 0.270%
<i>Zα</i> (69 aa.)		0.687 ± 0.106%	2.401 ± 0.357%
<i>Trx</i> (107 aa.)		0.619 ± 0.125%	2.605 ± 0.644%
<i>GST</i> (211 aa.)		0.477 ± 0.077%	1.370 ± 0.117%

Efficiency of two -1RFS signals (FS), WT and U19C19a, with various lengths of preceding translations were measured in mammalian cells. \*Efficiency of -1RFS is defined as  $[(A_R/A_F \text{ of a reporter vector with -1RFS})/(A_R/A_F \text{ of an in-frame control vector})]$ , where  $A_R$  is activity of *Renilla* luciferase and  $A_F$  is activity of firefly luciferase.

soluble proteins that widely used as fusion partners of other proteins for enhancing expression and solubility.<sup>14,15</sup> We assumed that the N-terminal region of the -1RFS product would not affect the folding and activity of the *Renilla* luciferase located at its C terminus. Expression of full-length fusion protein products containing the C-terminal *Renilla* luciferase from these reporter constructs was confirmed by immunoblotting analysis using an anti-*Renilla* luciferase monoclonal antibody (Fig. 1(b)). It seems that the -1RFS products were expressed at different levels. In this -1RFS assay system, however, amounts of fusion proteins are not directly correlated with the -1RFS efficiency but rather varying levels of transfection. As described earlier, the use of the bi-directional promoter, P<sub>bi-1</sub>, for expression of two reporter genes in a single vector can ensure more accurate reflection of -1RFS efficiency. Thus, we conclude that fusion proteins resulted from -1RFS event in these reporter constructs are readily translated and expressed intact.

The efficiency of -1RFS for the WT and the U19C19a constructs with long preceding ORF1s were examined through fusion constructs in reporter vectors derived from pBI-2lucL. The results showed that a short preceding translation as seen in the experiments with pBI-2lucL constructs could produce as similar efficiencies of frameshifting as from constructs with longer preceding translation, if not higher (Table 2). Although levels of frameshifting efficiencies were relatively low in the constructs with GST, they show clear -1RFS occurrences. In conclusion, the -1RFS signal (at least from BWYV -1RFS system) is independent on the location within mRNA for efficient -1RFS.

In the -1RFS event, fusion protein products are produced from the preceding translation (ORF1) and the following translation of the downstream ORF2 as a result of -1 frameshifting at the slippery sequence. Until now, there is no detailed study about effect of the very short preceding translations on -1RFS efficiency. In our study, we examined the effect of preceding translation length by placing various protein genes that have different sizes of ORFs before the -1RFS signal. As summarized in Table 2, our results showed that substantial levels of -1RFS occurs at -1RFS signal even when translation undergoes only six-amino acid long. The -1RFS efficiency after this short preceding translation was comparable as those of much longer preceding translations such as Za, Trx and GST. The mRNA translation machinery pauses at the RNA structural barrier such as a pseudoknot and increases the realignment of the aminoacyl tRNA and the peptidyl tRNA inside a ribosome in the 5' direction from the zero frame to the -1 frame.<sup>2,3</sup> Thus, the preceding translation may not directly associate with -1 frameshifting event (i.e., frequency of frameshifting) as two *cis*-elements of the -1RFS signal were exclusively responsible for generating -1RFS. Apparently, this explanation may be consistent with our result.

In conclusion, we did not find any considerable dependency of -1RFS efficiencies on length of preceding translation. Thereby, proteins can be efficiently produced by -1RFS from mRNA templates that contain only short extra

peptides at the N-terminus before the -1RFS signal. Two *cis*-elements on mRNAs are essential and sufficient for efficient -1RFS, which is an attractive feature for applying it to a regulator of protein expression. In addition, it has been known that a wide range of -1RFS efficiencies can be possible without drastic changes of *cis*-elements. Accordingly, -1RFS signals can be used as an embedded device to generate different levels of protein expression. Thus, this study is important since it provides the basis of a system that can control expression of proteins in the future.

## Experimental Procedures

**Construction of Reporter Vectors.** The *Renilla* luciferase gene (*Rluc*) in pRL-CMV vector (Promega) was altered to have a polycloning site by inserting a duplex DNA oligomer adjacent to its start codon. This modified *Rluc* gene was transferred into the polycloning site of the pBI-L (BD Biosciences) at *MluI/NotI* restriction sites to create the vector named as pBI-2lucL (Fig. 1(a)). Transcriptions of two luciferase genes in pBI-2lucL were simultaneously controlled by a bidirectional Tet-responsive promoter (P<sub>bi-1</sub>). Thus, expressions of two luciferase proteins from pBI-2lucL vector would be induced simultaneously when the P<sub>bi-1</sub> promoter was activated by doxycycline, a tetracycline analog.<sup>10</sup> Protein genes as an *ORF1* and a -1RFS signal were fused to the upstream of the *Rluc* gene in -1 frame (Fig. 1(a)). When -1 frameshifting occurs at the slippery sequence, fusion protein products between an *ORF1* protein and downstream *Renilla* luciferase were readily translated (Fig. 1(b)). To be used as an 100% -1RFS efficiency control, construction of in-frame fusion between each *ORF1* and *Rluc* genes was also carried out.

**Mammalian Cell Culture and Dual-luciferase Assay.** HEK293 Tet-On cells (BD Biosciences) were cultured in Dulbecco's modified eagle's medium (DMEM) (Sigma) containing 1% Penicillin-Streptomycin (Invitrogen) and 100 µg/mL G418 (Duchefa) and supplemented with 10% Tet system approved fetal bovine serum (Hyclone). For the dual-luciferase reporter assay *in vivo*, cells were washed 3 times with DMEM and cultured in DMEM containing 10% Tet system approved fetal bovine serum and 1% Penicillin-Streptomycin with 1 µg/mL doxycycline. Cells were cultured up to 40% confluence in 24-well plates and reporter plasmids were then transfected by FuGene6 transfection reagent (Roche). The cells were assayed for transient expression of reporter genes at 48 hr after transfection, unless time is specified elsewhere. The dual-luciferase assay was performed using the Dual-luciferase™ reporter assay kit (Promega) according to the manufacturer's protocol. All individual *in vivo* assays to determine the average -1RFS efficiencies were repeated three times or more.

**Immunoblot Analysis.** Cells were harvested and lysed by adding lysis buffer (1% Triton X-100, 1 mM EDTA, 50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.1 µg/mL Aprotinin, 0.1 µg/mL Leupeptin, 0.1 mM AEBSF). Equal amounts of proteins were loaded in each lane and separated on a 10%

SDS polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to nitrocellulose membrane and then blocked for 45 min in 1 × TBST (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.1% Tween 20), 4% nonfat dry milk. The membrane was incubated for overnight in 10 mL of a 1:4000 dilution of a mouse anti-Renilla monoclonal antibody (US Biological) in 1 × TBST, 2% nonfat dry milk, and then washed for 1 hr with multiple changes in 1 × TBST. The membrane was incubated for 1 hr with a 1:5000 dilution of an HRP-conjugated anti-mouse secondary antibody (Amersham) and washed as above and then detected using a SuperSignal West Pico Chemiluminescence Substrate (Pierce).

**Acknowledgments.** This work was supported by the Korea Research Foundation Grant funded by the Korean Government (KRF-313-2008-2-C00540).

### References

1. Jacks, T.; Varmus, H. E. *Science* **1985**, *230*, 1237-1242.
  2. Farabaugh, P. J. *Microbiol. Rev.* **1996**, *60*, 103-134.
  3. Gesteland, R. F.; Atkins, J. F. *Annu. Rev. Biochem.* **1996**, *65*, 741-68.
  4. Jacks, T.; Madhani, H. D.; Masiarz, F. R.; Varmus, H. E. *Cell* **1988**, *55*, 447-458.
  5. Namy, O.; Moran, S. J.; Stuart, D. I.; Gilbert, R. J.; Brierley, I. *Nature* **2006**, *441*, 244-247.
  6. Liao, P. Y.; Choi, Y. S.; Dinman, J. D.; Lee, K. H. *Nucleic. Acids Res.* **2011**, *39*, 300-312.
  7. Plant, E. P.; Rakauskaitė, R.; Taylor, D. R.; Dinman, J. D. *J. Virol.* **2010**, *84*, 4330-4340.
  8. Plant, E. P.; Wang, P.; Jacobs, J. L.; Dinman, J. D. *Nucleic. Acids Res.* **2004**, *32*, 784-790.
  9. Baron, U.; Freundlieb, S.; Gossen, M.; Bujard, H. *Nucleic. Acids Res.* **1995**, *23*, 3605-3606.
  10. Gossen, M.; Freundlieb, S.; Bender, G.; Muller, G.; Hillen, W.; Bujard, H. *Science* **1995**, *268*, 1766-1769.
  11. Su, L.; Chen, L.; Egli, M.; Berger, J. M.; Rich, A. *Nat. Struct. Biol.* **1999**, *6*, 285-292.
  12. Kim, Y. G.; Su, L.; Maas, S.; O'Neill, A.; Rich, A. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 14234-14239.
  13. Schwartz, T.; Lowenhaupt, K.; Kim, Y. G.; Li, L.; Brown, B. A., 2nd; Herbert, A.; Rich, A. *J. Biol. Chem.* **1999**, *274*, 2899-2906.
  14. Smith, D. B.; Johnson, K. S. *Gene* **1988**, *67*, 31-40.
  15. LaVallie, E. R.; DiBlasio, E. A.; Kovacic, S.; Grant, K. L.; Schendel, P. F.; McCoy, J. M. *Biotechnology (N Y)* **1993**, *11*, 187-193.
-