Structural Elements Contributing to Efficient -1 Ribosomal Frameshifting in BWYV Pseudoknot

So Jung Park, Hee Jung Park, Sangho Lee,† and Yang-Gyun Kim*

*Departments of Chemistry and †Biological Science, School of Natural Sciences, Sungkyunkwan University, Suwon 440-746, Korea. *E-mail: ygkimmit@skku.edu*

Received January 12, 2009, Accepted January 28, 2009

Key Words: Ribosomal frameshifting, RNA pseudoknot, BWYV, SRV-1

The programmed -1 ribosomal frameshifting (-1RFS), one of the programmed alternative reading mechanisms in translation, moves tRNAs inside ribosome toward -1 direction during translation at a predetermined sequence on an mRNA with a certain frequency.[1,2] Occurrence of -1RFS enables the production of two gene products from a single mRNA. This unusual mechanism is wide-spread among retroviruses, plant viruses, retrotransposons, bacteria and yeast. Two *cis*-acting elements in mRNA are generally required to stimulate efficient -1RFS.[1,2] One is the slippery sequence consisting of a hepta-nucleotide-motif 'XXXYYYN' (X or Y represents the same base, and N is any) where -1 frameshifting of tRNAs actually occurs. The other element is an RNA structure, mostly RNA pseudoknots, locating 4-8 nucleotides downstream of the slippery sequence.[3]

Many studies have correlated the stability of an RNA structure, notably RNA pseudoknots, to the high efficiency of -1RFS.[4-8] Frameshifting-stimulating pseudoknots, referred to as pseudoknots unless specified otherwise, exhibit classic H-type pseudoknot-folds with two stacked stems (S1 and S2) and two loops (L1 and L2) (Fig. 1A). Structural elements responsible for the stability and the integrity of the pseudoknots include quasi-continuous base-stacking in S1 and S2 and an extensive S1-L2 triplex interaction.

Structural diversity among pseudoknots appears to be one of major culprits for various viral systems to achieve different -1RFS efficiencies optimized to their own needs. The pseudoknots in BWYV (beet western yellow virus) and SRV-1 (simian retrovirus-1) provide an interesting instance for the structural diversity. BWYV and other luteoviral pseudoknots, which mediate P1-P2 production through -1RFS, are relatively small and tightly folded with a characteristically very short S2 (3 base pairs) (Fig. 1A). These luteoviral pseudoknots produce considerable degrees of -1RFS and show a strong sequence preference in both L1 and L2 for efficient -1RFS.[5] From the structure-function analysis of the luteoviral pseudoknots,[5,9-14] it is well established that tertiary interactions are crucial: in particular, C8 in the L1 of the BWYV pseudoknot engages with G12, A25 and C26 to form the complex quadruple-base interaction. In contrast, SRV-1 pseudoknot, located at gag-pro junction, contains a relatively long S2 (5 base pairs without counting the A-U base pair in the stem junction) (Fig. 1A). Unlike the loops in the BWYV pseudoknot, the length and the base identity of loops in the SRV-1 pseudoknot are rather insignificant. It has been shown that even large deletion of L2 can be tolerable for efficient -1RFS.[8,15]

Structural comparison suggests that the stems of the pseudoknots of BWYV and SRV-1 employ different mechanisms to ensure the stability of the pseudoknots. The SRV-1 pseudoknot takes advantage of intrinsically high stem stability provided by the long S2 stem for efficient -1RFS. In contrast, the BWYV and other luteoviral pseudoknots possess tightly folded structures through the quadruple-base interaction between S2-L1 as well as the triplex interactions between S1-L2, which are enough to compensate for low stem stability caused by the short S2 stem.[12,16] Thus, it is of great interest to transform the BWYV pseudoknot into SRV-1-like pseudoknot

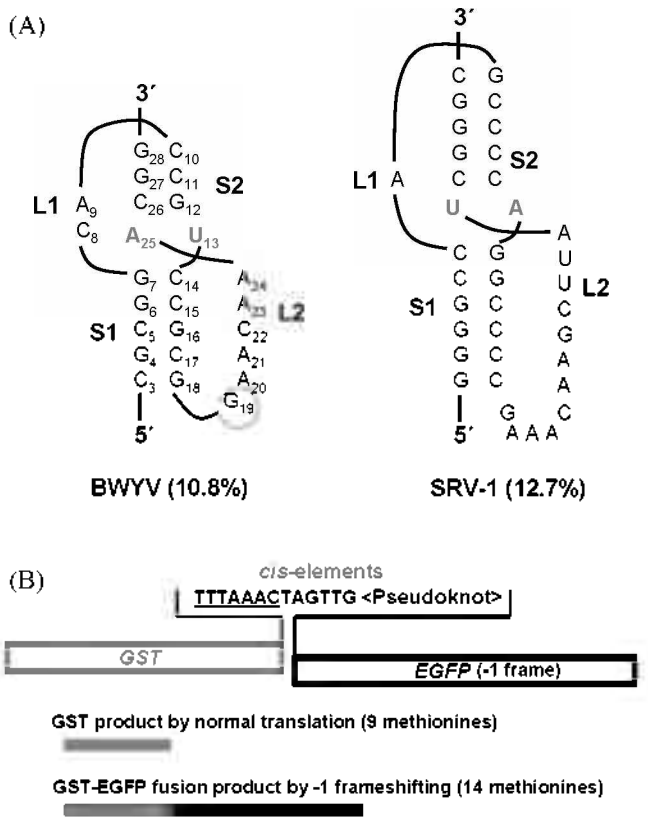

Figure 1. (A) Secondary-structure diagrams of the BWYV and SRV-1 pseudoknots. Numbers in parentheses indicate -1RFS efficiency determined using *in vitro* system. (B) Schematic representation of the reporter construct and resulting protein products used for *in vitro* -1RFS assay. Boxed regions indicate ORFs for *GST* (Grey) and *EGFP* (Black).

Table 1. Sequence and -1RFS efficiency of the wild-type and SRV-1-like mutants of the BWYV pseudoknot

Name	Pseudoknot Sequence								Efficiency (%)
	S1	L1	S2	J	S1	L2	J	S2	
WT[a]	CGCGG	CA	CCG	**U**	CCGCG	GAACAA	**A**	CGG	10.8
mS-B	CGCGG	CA	GCCCG	**U**	CCGCG	GAACAA	**A**	CGGGC	8.4
mS-B:C8A	CGCGG	AA	GCCCG	**U**	CCGCG	GAACAA	**A**	CGGGC	5.0
mS-B:A25G	CGCGG	CA	GCCCG	**U**	CCGCG	GAACAA	**G**	CGGGC	0.8
mSL-B	CGCGG	A	GCCCG	**U**	CCGCG	GAACAA	**A**	CGGGC	4.4
mSL-B:A9C	CGCGG	C	GCCCG	**U**	CCGCG	GAACAA	**A**	CGGGC	7.4
mSL-B:A9G	CGCGG	G	GCCCG	**U**	CCGCG	GAACAA	**A**	CGGGC	2.3
mSL-B:A9U	CGCGG	U	GCCCG	**U**	CCGCG	GAACAA	**A**	CGGGC	4.4
mSLJ-B	CGCGG	A	GCCCG	**A**	CCGCG	GAACAA	**U**	CGGGC	6.1
mS-B:UC19	CGCGG	CA	GCCCG	**U**	CCGCG	UCAACAA	**A**	CGGGC	26.8
mS-B:GC19	CGCGG	CA	GCCCG	**U**	CCGCG	GCAACAA	**A**	CGGGC	10.2
mSL-B:UC19	CGCGG	A	GCCCG	**U**	CCGCG	UCAACAA	**A**	CGGGC	12.1
mSL-B:GC19	CGCGG	A	GCCCG	**U**	CCGCG	GCAACAA	**A**	CGGGC	6.8
WT:UC19	CGCGG	CA	CCG	**U**	CCGCG	UCAACAA	**A**	CGG	30.5[b]
WT:GC19	CGCGG	CA	CCG	**U**	CCGCG	GCAACAA	**A**	CGG	13.3[b]

Stems (S1, S2), loops (L1, L2) and junctions (J) between stems and loops are indicated above pseudoknot sequences. The -1RFS efficiency of the wild-type BWYV pseudoknot, 10.8%, was used to normalize -1RFS efficiency of each pseudoknot. [a]Wild-type BWYV pseudoknot. [b]These results were published previously in the article by Kim *et al.*[5]

to investigate the contributions of the stem stability and the quadruple-base interactions to the efficient -1RFS. Here we examined the effects of the structural components of the BWYV pseudoknot on the -1RFS efficiency by altering the length and sequences of the S2 and the L1 to closely resemble the SRV-1 pseudoknot.

To study the effect of the length change in S2 on the -1RFS efficiency of the BWYV pseudoknot, we added two base pairs at the top side to the S2 of the BWYV pseudoknot, which resulted in a mutant pseudoknot containing five base-pair S2, referred to as mS-B (Fig. 2A, left panel). This extended S2 in mS-B is expected to provide more stability to the pseudoknot. However, it is likely to cause perturbation of the quadruple-base interactions between L1 and S2 as well. Various mutations on mS-B were introduced (Table 1) and their -1RFS efficiencies determined by *in vitro* translation assay. Plasmids for the *in vitro* assay have two genes in tandem, gluthathione *S*-transferase (*GST*) followed by enhanced green fluorescent protein (*EGFP*). These two genes are connected by -1 frame. Accordingly, the downstream *EGFP* gene would be translated as a fusion protein product (GST-EGFP) with the upstream *GST* gene only when -1RFS occurs during the translation (Fig. 1B). All pseudoknot-forming sequences were inserted after the slippery sequence with the 6-base spacer.

Results from the *in vitro* translational assay showed that -1RFS efficiency of mS-B was reduced to *ca.* 80% level of that of the BWYV wild type (see Fig. 1A, left panel and Fig. 2A, left panel). The two mutations at nucleotides participating the quadruple-base interaction, mS-B:C8A and mS-B:A25G, decreased -1RFS efficiencies even further (Fig. 2A, left panel). In particular, -1RFS was almost abolished by the A25G mutation in mS-B as similarly seen in the same mutation in the wild-type BWYV pseudoknot.[5] These observations suggest that the quadruple-base interaction may still play a role in the stabilization of mS-B.

Next we modified mS-B by reducing L1 to one nucleotide

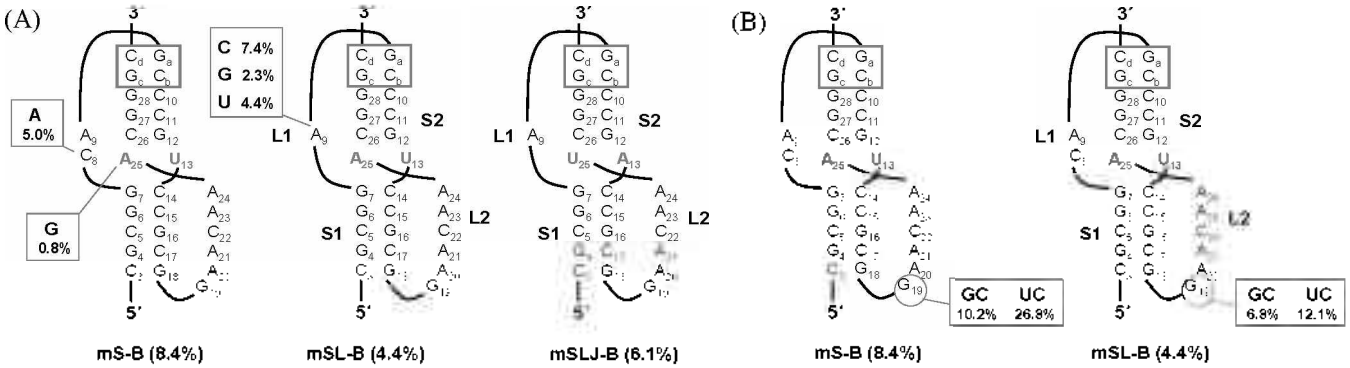

Figure 2. Constructs of the BWYV pseudoknot mimicking the SRV-1 pseudoknot and their effects on -1RFS efficiency. The -1RFS efficiency of the wild-type BWYV pseudoknot in this study is 10.8%. The numbers near the mutations correspond to the -1RFS efficiency. (A) SRV-1-like pseudoknot constructs (mS-B, mSL-B and mSLJ-B) and their mutants. (B) Mutations at the S1-L2 linker region of mS-B and mSL-B.

and inverting the junctional base pair (from A25:U13 to U13:A25). These constructs, mSL-B and mSLJ-B, respectively, mimic the SRV-1 pseudoknot by possessing the extended S2 (5 base pairs) and a shortened L1 (one nucleotide) (Fig. 2A, center and right panels). mSL-B exhibited reduced -1RFS efficiency (4.4%) to an half or less than mS-B (8.4%) and the BWYV wild type (10.8%) while mSLJ-B showed higher -1RFS efficiency (6.1%) than mSL-B (4.4%) (Fig. 2A). The results indicate that the disruption of the quadruple-base interaction by shortening L1 can be compensated to some extent by the inverted junctional U25:A13 pair.

Unlike BWYV pseudoknot as reported previously,[5] there is no marked sequence preference exhibited in the L1 of SRV-1 pseudoknot.[8] Accordingly, we investigated the sequence preference of the shortened L1 in mSL-B and compared with that in the wild-type BWYV pseudoknot. We examined the -1RFS efficiencies of all four nucleotides in the L1 of mSL-B (Fig. 2A). mSL-B with the cytosine in the L1 (mSL-B:A9C) showed the highest -1RFS efficiency (7.3%) among the four nucleotides tested, which is about 70% of that of the BWYV wild type (10.8%). mSL-B:A9U showed the same -1RFS efficiency (4.4%) as mSL-B while mSL-B:A9G exhibited the reduced -1RFS efficiency to 2.3%. The highest -1RFS efficiency by mSL-B:A9C implies that mSL-B may not completely discard the quadruple-base interaction in which the identity of cytosine in the L1 is critical for the efficient -1RFS. This cytosine preference suggests the possible formation of a quadruple-base interaction in mSL-B although its formation may not be as strong as in the wild-type BWYV pseudoknot.

Finally we performed mutational analysis of the S1-L2 linker region. Previously a drastic increase of -1RFS efficiency in mutations at G19 in the S1-L2 linker region was noted, seemingly irrespective of structure-stabilizing effect (Table 1).[5] It was postulated that G19 might contact a part of ribosome during the -1RFS event, which in turn may stimulate -1RFS through intermolecular interaction(s). To examine the effects of G19 on the -1RFS efficiencies in our SRV-1-like pseudoknots, we mutated G19 of the S1-L2 linker region to G19C19a and U19C19a by insertion and substitution in both mS-B and mSL-B (Fig. 2B). Although it does not appear that the changes of the S1-L2 linker region are major factors for integrity of BWYV pseudoknot, the increasing effect of these mutations on -1RFS efficiency was similarly reproduced in both mS-B and mSL-B (Fig. 2B) as seen in the wild-type BWYV pseudoknot (Table 1, WT:UC19 and WT:GC19). In addition, the identities of the bases at the S1-L2 linker region also affected the -1RFS efficiency greatly. Specifically G to U substitution was substantially more effective on promoting -1RFS and further enhancing in accompanying with C insertion.

Overall, transforming the BWYV pseudoknot into the SRV-1-like pseudoknot (mS-B) by increasing the length of S2 resulted in a bit weaker but still robust frameshifting pseudoknot. The extended S2 in mS-B (from 3 base-pairs to 5 base-pairs) might lead to almost complete loss or significant reduction of the quadruple-base interaction in the L1-S2 junction due to the base substitution and misalignment of the interacting bases. However, the enhanced stability of mS-B by having two extra base pairs appears to compensate for the weakened major interactions in the wild-type BWYV pseudoknot, at least considerably if not completely. Moreover, like BWYV pseudoknot, unique sequence preference of the S1-L2 linker region unrelated to the pseudoknot stability was still observed. Thus, our results support that the S1-L2 linker region could increase the -1RFS efficiency by unknown *trans*-interactions with other component(s) of translational machinery.

Experimental Section

Plasmid construction. Plasmids used for *in vitro* translation assay were constructed by essentially the same methods described previously.[5] In brief, the glutathione *S*-transferase (*GST*) gene and the enhanced green fluorescent protein (*EGFP*) gene downstream of the *GST* gene were inserted into the pGEM-3Z vector (Promega), respectively. Various pseudoknot sequences were inserted downstream of the UUUAAAC slippery (Fig. 1B). The resulting constructs contain the two genes connected in out-of-frame. Thus, GST-EGFP fusion protein is translated only when -1RFS occurs at the slippery site.

***In vitro* frameshifting assay.** The T_NT T7-coupled transcription/translation system (Promega) was used according to the manufacturer's protocol as previously described.[5] To calculate -1RFS efficiency, the ^{35}S-Met labeled samples were run on 12% SDS polyacrylamide gels to separate translational products. PhosphorImager (Molecular Dynamics) was employed to quantify signals of frameshifting (14 methionines) and non-framshifting (9 methionines) products. As Frameshifting efficiencies were calculated with the formula (I[FS]/14)/[(I[FS]/14) + (I[NFS]/9)], where I[FS] is the signal intensity of the frameshifting product and I[NFS], the signal intensity of the nonframeshifting product. All individual *in vitro* assays were accompanied by the wild-type BWYV pseudoknot. Assays were repeated three times or more to determine the average -1RFS efficiencies for all constructs. The -1RFS efficiency of the wild-type BWYV pseudoknot, 10.8% as reported previously,[5] was used to normalize -1RFS efficiency of each pseudoknot.

Acknowledgments. This work was supported by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD) (KRF-2005-070-C00106).

References

1. Gesteland, R. F.; Atkins, J. F. *Annu. Rev. Biochem.* **1996**, *65*, 741.
2. Farabaugh, P. J. *Microbiol. Rev.* **1996**, *60*, 103.
3. ten Dam, E. B.; Pleij, C. W.; Bosch, L. *Virus Genes* **1990**, *4*, 121.
4. Kollmus, H.; Hentze, M. W.; Hauser, H. *RNA* **1996**, *2*, 316.
5. Kim, Y. G.; Su, L.; Maas, S.; O'Neill, A.; Rich, A. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 14234.
6. Kollmus, H.; Honigman, A.; Panet, A.; Hauser, H. *J. Virol.* **1994**, *68*, 6087.
7. Giedroc, D. P.; Theimer, C. A.; Nixon, P. L. *J. Mol. Biol.* **2000**, *298*, 167.
8. ten Dam, E. B.; Verlaan, P. W.; Pleij, C. W. *RNA* **1995**, *1*, 146.
9. Kim, Y. G.; Maas, S.; Wang, S. C.; Rich, A. *RNA* **2000**, *6*, 1157.
10. Cornish, P. V.; Hennig, M.; Giedroc, D. P. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 12694.

11. Cornish, P. V.; Stammler, S. N.; Giedroc, D. P. *RNA* **2006**, *12*, 1959.
12. Nixon, P. L.; Cornish, P. V.; Suram, S. V.; Giedroc, D. P. *Biochemistry* **2002**, *41*, 10665.
13. Nixon, P. L.; Giedroc, D. P. *J. Mol. Biol.* **2000**, *296*, 659.
14. Nixon, P. L.; Rangan, A.; Kim, Y. G.; Rich, A.; Hoffman, D. W.; Hennig, M.; Giedroc, D. P. *J. Mol. Biol.* **2002**, *322*, 621.
15. ten Dam, E.; Brierley, I.; Inglis, S.; Pleij, C. *Nucleic Acids Res.* **1994**, *22*, 2304.
16. Pallan, P. S.; Marshall, W. S.; Harp, J.; Jewett, F. C., 3rd; Wawrzak, Z.; Brown, B. A., 2nd; Rich, A.; Egli, M. *Biochemistry* **2005**, *44*, 11315.