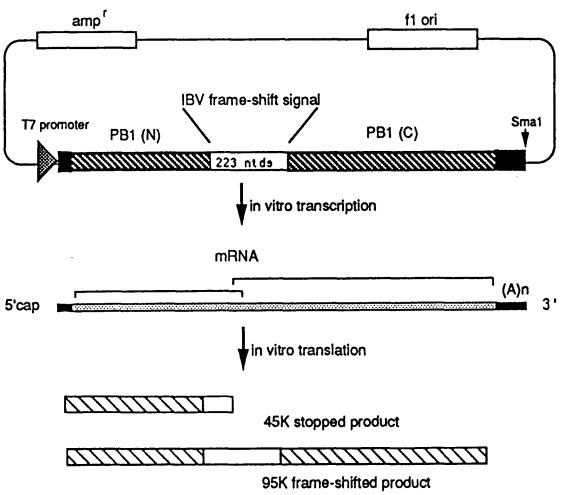
THE RIBOSOMAL FRAME-SHIFT SIGNAL OF INFECTIOUS BRONCHITIS VIRUS

S.C. Inglis, N. Rolley and I. Brierley
Division of Virology,
Department of Pathology,
University of Cambridge,
Tennis Court Road,
Cambridge, U.K.

Introduction

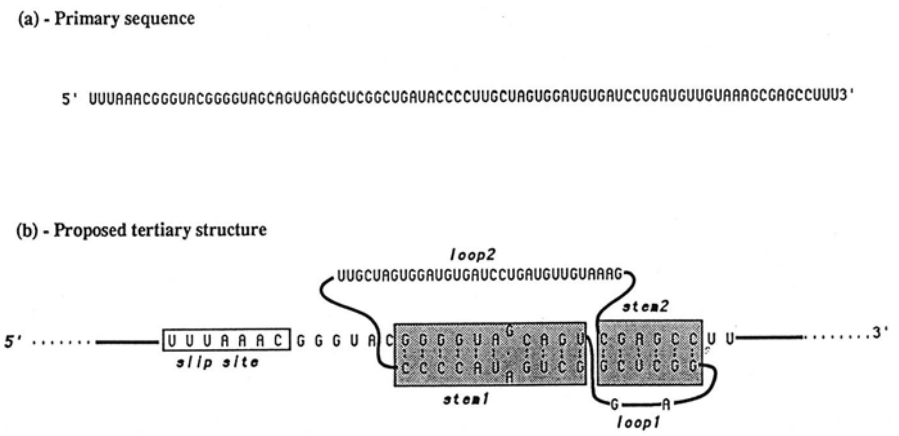
We recently described the first non-retroviral example of ribosomal frame-shifting in higher eukaryotes (Brierley et al., 1987). The shift occurs during translation of the genomic RNA of the coronavirus infectious bronchitis virus (IBV), just at the end of the F1 open reading frame (located at the 5' end of the genome); its consequence is that a proportion of ribosomes reading the F1 frame fail to terminate at the F1 stop codon, and instead begin reading the F2 open reading frame (ORF), which overlaps the end of F1, leading to the production of an F1-F2 fusion protein. This "-1" frame-shift is highly efficient (about 30%) and can be reproduced *in vitro* by cloning a short sequence from the junction of the F1/F2 ORFs into a suitable reporter gene (Figure 1). The recombinant gene may then be transcribed using the phage T7 RNA polymerase, and the resulting mRNA translated in the rabbit reticulocyte lysate cell-free system.

Figure 1 - Organisation of plasmid pFS8



NATO ASI Series, Vol. H 49
Post-Transcriptional Control of Gene Expression
Edited by J. E. G. McCarthy and M. F. Tuite

Figure 2 - Structure of the IBV frame-shift signal



Frame-shifting can therefore be monitored readily through production of a 'read-through' product .

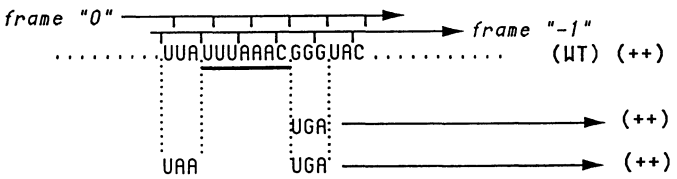
We have previously analysed in detail the sequence requirements for the IBV frameshift signal (Brierley et al., 1989) by deletion analysis and site-directed mutagenesis on the plasmid pFS8 (Figure1). These studies indicated that the signal can be narrowed down to a stretch of 86 nucleotides (ntds), which is in itself sufficient to direct efficient ribosomal frame-shifting in a heterologous genetic context (Figure 2). This sequence appeared to consist of two separate essential elements: a "slippery sequence", located at the 5'end at which the ribosome was believed actually to change frame, followed at a critical distance by a kind of tertiary RNA structure which has been called an RNA pseudoknot (Studnicka et al., 1978; Pleij et al., 1985). This is composed of two RNA helices (stems 1 and 2), which are co-axially-stacked to form a quasi-continuous double helix, with two connecting single-stranded loops (loops 1 and 2). The location of a slippery site at the 5' end of the frame-shift signal was suggested by comparison with known slip-sites from other viruses, and by deletion analysis and site-directed mutagenesis within this region. The requirement for downstream RNA tertiary structure was suggested by the observation that nucleotide changes within the two helices 1 and 2 were highly inhibitory to frame-shifting, but that these could be compensated fully by complementary changes on the opposite strand. These studies strongly suggested not only that such a structure was likely to be part of the frame-shift signal, but also that it was absolutely required for ribosomal slippage to occur with high efficiency. However as yet the mechanism by which ribosomal slippage occurs, and the precise structure and contribution of the pseudoknot remains unclear.

Here we describe further mutagenic studies on the frame-shift signal which seek to elucidate more precisely its structure and the mechanism by which it direct efficient ribosomal slippage.

Analysis of the "Slippery Site"

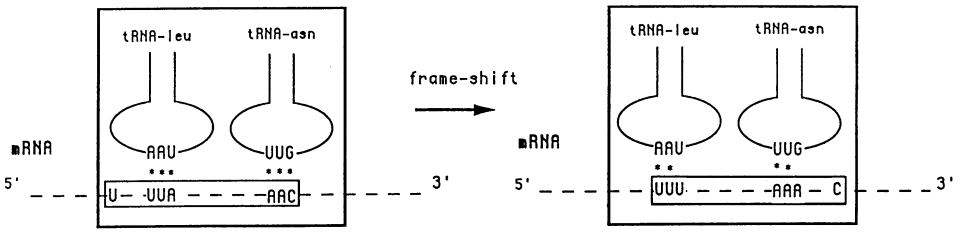
Based on the work of Jacks and colleagues (1985, 1988a, 1988b), who identified the precise location at which ribosomal frame-shifting occurred during translation of the gag-pol region of Rous sarcoma virus (RSV) and Human immunodeficiency virus (HIV) we previously suggested that the heptanucleotide sequence UUUAAAC was likely to be the point at which ribosomes change frame within the IBV sequence (Figure 2). This possibility was supported, but not formally proven by the deletion analysis described above (Brierley et al., 1989). We therefore introduced two new termination codons around the putative slip site (Figure 3) by site-directed mutagenesis, as described before (Brierley et al., 1989). The stop codon introduced downstream of the slip site is in the F1 reading frame, and therefore serves as a terminator for ribosomes which do not change frame. The stop codon introduced upstream of the slip site is in the "-1" frame (i.e. the F2 reading frame) and consequently the upstream and downstream open reading frames now overlap solely by the seven nucleotides which constitute the putative slip site UUUAAAC. For ribosomal frame-shifting to be observed in this mutant, ribosomes would have to slip within the heptanucleotide sequence. Translation of mRNA bearing this altered signal indicated that frame-shifting occurs with wild type efficiency (Figure 3)*, demonstrating that ribosomal slippage does indeed occur at this position.

Figure 3 - Definition of the IBV frame-shift site



*In this, and all subsequent figures, the symbol (++) represents the aproximate level of wild type frame-shifting (about 30%), (+) denotes frame-shifting between 10% and 20%, (+/-) - 5-10%, and (-) - less than 2%.

Figure 4 - Double-slippage model for -1 frame-shifting



Jacks et al. (1988) proposed a model for "-1" frame-shifting in retroviruses which involved simultaneous slippage on the messenger RNA of tRNAs occupying both the A and P site of the the translating ribosome (figure 4). The model suggests that such frame-shifts would only be allowed if 2 out of 3 base pair contacts are retained between the mRNA and each tRNA after slippage (i.e. 4 out of 6 contacts in total). Thus seven nucleotide sequences which begin with two repeated nucleotide triplets (XXXYYY) would be potentially slippery, and indeed almost all the sequences known, or thought, to direct -1 frame-shifting in retroviruses conform to this pattern, as does the IBV frame-shift site (UUUAAAC). One might expect therefore that alterations made to the slippery site which decreased potential pairing after slippage would be inhibitory to frame-shifting, and this is indeed the case. Mutation of the UUUAAAC sequence to either UUUAUAC or UUUACAC completely abolished frame-shifting, and a change to CUUAAAC was highly inhibitory (Brierley et al., 1989). If however the potential for formation of post-slippage pairs is all that is required, it might also be expected

Figure 5 - Mutational analysis of the IBV slip site

	P site	A site	post-slip contacts	Frame-shifting	%
WT	U UUA	AAC	4/6	++	30
	U UUA	AAG	4/6	+/-	1
	U UUA	AAU	4/6	++	30
	U UUU	UUC	5/6	++	30
	U UUG	GGC	4/6	+/-	1

that any sequence which can still allow potential 4/6 pairing (or greater) after slippage would be slippery. Consistent with this, the sequence UUUUUUC which should allow 5/6 pairs to form after slippage is at least as efficient as the WT IBV sequence (Figure 5). However the sequence UUUGGGC, though still active, was much less efficient, even though it also has the potential to form 4/6 post-slippage contacts. One possible explanation for this result is that heptanucleotide sequences involving several G:C pairs may be less slippery due to the extra energy required to break these contacts. However when we changed only the last nucleotide of the IBV slippery sequence, from a C to a G residue (which should alter neither the potential for pairing post-slippage, nor the overall energy of the tRNA:mRNA interactions) we found once again a marked reduction in frame-shift efficiency. This suggests that the "slipperiness" of a particular sequence is not simply due to the number and type of pairs which can form post-slippage, but in addition depends on the particular tRNAs with which it interacts. Thus the lysine-tRNA (which decodes AAG) appears less prone to slippage than the asparagine tRNA which decodes AAC. In this respect our results differ somewhat from those of Jacks et al., (1988b) who found that any nucleotide substitution could be tolerated equally well at the last position within the RSV frame-shift site. However in this case the sequence of the wild type slip site is quite different (AAAUUUA), and it could be that the tRNAs which decode the final UUA (leu) and third position variants thereof (UUU-phe, UUG-val and UUC-leu) may all be inherently slippery. In vitro translation studies in cell-free systems supplemented with particular kinds of tRNAs may help to resolve this question.

Analysis of pseudoknot structure

Our previous results (Brierley et al., 1989) indicated that formation of a pseudoknot downstream of the slippery site was required for efficient ribosomal slippage, and we proposed a likely model of the structure based on our initial mutagenesis data and on nucleotide sequence analysis. However the precise configuration of the pseudoknot remained uncertain, and so we set out to examine in detail the structure through site-directed mutagenesis, on the premise that nucleotide changes which destabilise the structure should be inhibitory to frame-shifting.

Our analysis of the nucleotides proposed to be part of stems 1 and 2 is summarised in Figure 6. In general the results bear out the model displayed in Figure 2, in that individual nucleotide changes within stems 1 and 2 are inhibitory to frame-shifting. However the results show that the inhibitory effect is much less for nucleotides located at the extremities of the helices than for those in the middle. This would be consistent with the idea that the overall stability of the structure is related to its ability to promote frame-shifting; thus the slight destabilisation of the structure produced by mismatching at the ends, would have a less dramatic

Figure 6 - Mutational analysis of the pseudoknot stems

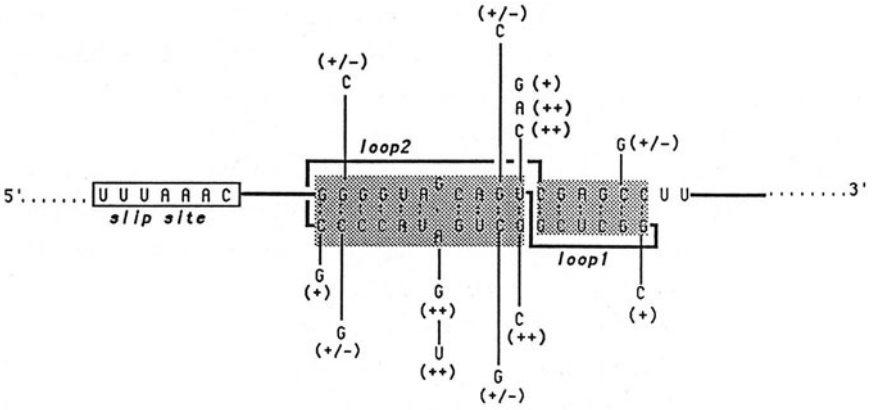
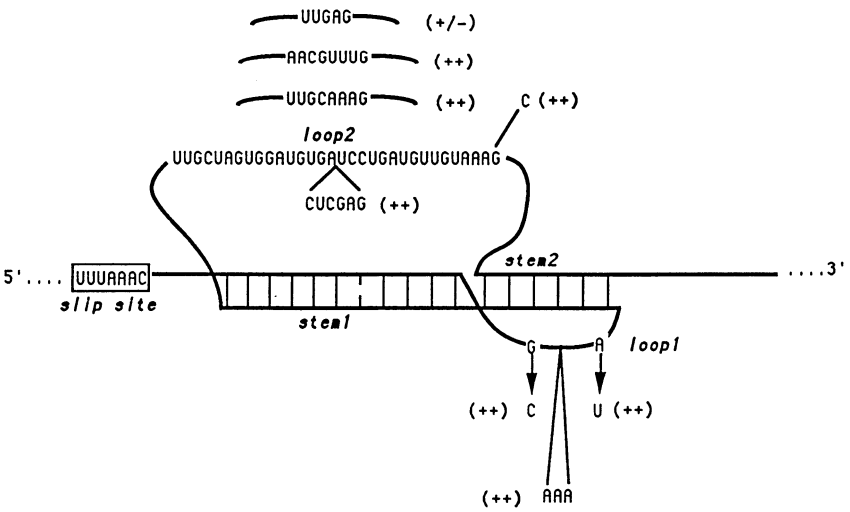


Figure 7 - Mutational analysis of the pseudoknot loops



effect than more highly destabilising mismatches introduced within the stems. One exception to this general finding was observed however. Mutations affecting the G:U pair proposed to form at the top of stem1 were not found to affect frame-shifting efficiency greatly. This was a somewhat surprising result since one of the central features of the pseudoknot building principle (Pleij et al., 1985) is that the stems contributing to the tertiary structure should stack co-axially form at the top of stem1 were not found to affect frame-shifting efficiency greatly. This was a somewhat surprising result since one of the central features of the pseudoknot building principle (Pleij et al., 1985) is that the stems contributing to the tertiary structure should stack co-axially to form a quasi-continuous double helix, and that these stacking forces will help to stabilise the structure. Thus we expected that introduction of a mismatched nucleotide pair at the top of stem1 would destabilise the pseudoknot considerably and hence inhibit frame-shifting. That this did not happen suggests either that direct stacking is not required and that a bulge can be tolerated at this position, or perhaps that a mismatched nucleotide pair can be "held" in position by virtue of the stacking forces on either side. Direct structural analysis of the pseudoknot will be required to resolve this point.

Our model for the structure of the pseudoknot proposes that the two RNA helices are connected by two single-stranded loops of 2 and 32 nucleotides (Figure 2) which span the 6 base pairs of stem2 (loop1) and the 11 base pairs of stem1 (loop2) respectively. Theoretical considerations (Pleij et al., 1985) suggest that for the structure to form, loop1 would have to be a minimum of 2 nucleotides long, while loop2 would have to be 7 nucleotides or more. However, provided the loops are long enough, their precise nucleotide sequence shoud not be important. Thus if frame-shifting requires only the formation of the correct tertiary structure, and does not depend on the presence of particular nucleotides within the loops, then alteration of the sequence and length of the loops within the constraint of the required length should not have a dramatic effect. The results of this kind of analysis, summarised in Figure 7, show that this is indeed the case. We found that either of the nucleotides proposed to comprise loop1 could be changed without affecting frame-shifting, and that insertion of an extra 3 nucleotides in the loop could also be tolerated. Loop2 may be reduced in length to 8 nucleotides, and furthermore each of these 8 nucleotides can be substituted without affecting frame-shifting. However deletion of a further 3 nucleotides (leaving only 5 nucleotides) was highly inhibitory. Since we predicted that loop2 would have to be a minimum of 7 nucleotides in order for the pseudoknot structure to form, this result was not unexpected and increases our confidence that the proposed structure is largely correct. However direct structural analysis will be once again required to confirm the hypothesis.

Brierley I, Digard P, Inglis SC (1989) Characterisation of an efficient coronavirus ribosomal frameshifting signal: Requirement for an RNA pseudoknot. Cell 57: 537-547.

Jacks T, Varmus HE (1985) Expression of the Rous sarcoma virus *pol* gene by ribosomal frameshifting. Science 230: 1237-1242.

Jacks T, Madhani HD, Masiarz FR, Varmus HE (1988a) Signals for ribosomal frameshifting in the Rous sarcoma virus *gag-pol* region. Cell 55: 447-458.

Jacks T, Power MD, Masiarz FR, Luciw PA, Barr PJ, Varmus HE (1988b) Characterisation of ribosomal frameshifting in HIV-1 *gag-pol* expression. Nature 331: 280-283.

Pleij CWA, Rietveld K, Bosch L. (1985) A new principle of RNA folding based on pseudoknotting Nucl. Acids Res. 13: 1717-1731.

Studnicka GM, Rahn GM, Cummings IW, Salser WA (1978) Computer method for predicting the secondary structure of single-stranded RNA. Nucleic Acids Res 5: 3365-3387.

Tang CK, Draper DE (1989) Unusual mRNA pseudoknot structure is recognised by a protein translational repressor. Cell 57: 531-536.