

A RIBOSOMAL FRAMESHIFT SIGNAL IN THE POLYMERASE-ENCODING REGION OF THE IBV GENOME

S.C. Inglis, N. Rolley and I. Brierley

Division of Virology, Department of Pathology
University of Cambridge, Tennis Court Road, Cambridge, U K

INTRODUCTION

Nucleotide sequencing studies on the genomic RNA of infectious bronchitis virus (IBV) has revealed the presence, at its 5' end of two very large open reading frames (ORFs) termed F1 and F2 (Bourneil *et al.*, 1987). The 5' proximal ORF, F1 has the capacity to encode a polypeptide of approximately 440K, while the distal ORF, F2, which overlaps F1 by about 40 nucleotides (ntds) in a different (-1) reading frame, can encode a 300K polypeptide. Our previous work has indicated that both these open reading frames can be expressed as a single polypeptide from genomic mRNA by ribosomal frame-shifting within the overlap region (Brierley *et al.*, 1987, 1989). This conclusion was based on experiments which showed that sequences from the F1-F2 junction region, when introduced into a heterologous mRNA could direct efficient ribosomal frame-shifting during translation *in vitro* and *in vivo*. Thus the primary translation products of IBV genomic RNA in the infected cell consists of a 440K product, corresponding to the F1 ORF, and an additional 740K 'read-through' product representing fusion of the F1 and F2 ORFs (Figure 1).

Fine mapping of the frame-shift signal

Recently we have begun to map sequences around the F1/F2 junction in IBV RNA which contain the signal for this ribosomal frameshift event by site-directed mutagenesis (Brierley *et al.*, 1989). The essence of this work is that a minimum sequence of approximately 86 ntds

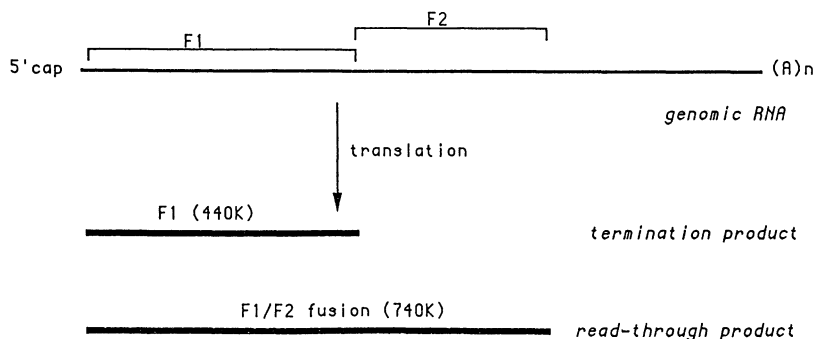


Fig. 1. Expression of the polymerase-encoding region of the IBV genome.

Construct	Sequence
pFS7.13	UAC
	slip site UAUUUAACGGGUACGGGGUAGCAGUGAGGCUCGCGCUGAUACCCCUUGCUAGUGGAGUGAGUCCUGAGUUGUARAAGCGAGCCUUU
pFS7.1	C
pFS7.12	
pFS7.14	U
pFS7.15	UCGGA
pFS7.16	UCGGA
pFS7.17	C
pFS7.18	CCCCAU
pFS7.19	AUGGGG
pFS7.20	CCCCAU AUGGGG
pFS7.22	
pFS8.5	UGA

Construct	Frame-shift efficiency	Construct	Frame-shift efficiency
pFS7.1	-	pFS7.18	-
pFS7.12	-	pFS7.19	-
pFS7.13	-	pFS7.20	++
pFS7.14	-	pFS8.5	++
pFS7.15	-	pFS7.14	-
pFS7.16	++	pFS7.22	-

Fig. 2. Effect on frameshifting activity of specific nucleotide substitutions within the F1/F2 frameshift signal. Sequences from the F1/F2 junction were cloned into a reporter gene, altered by site-directed mutagenesis, and assayed for frameshifting by *in vitro* transcription and translation as described (Brierley *et al.*, 1989).

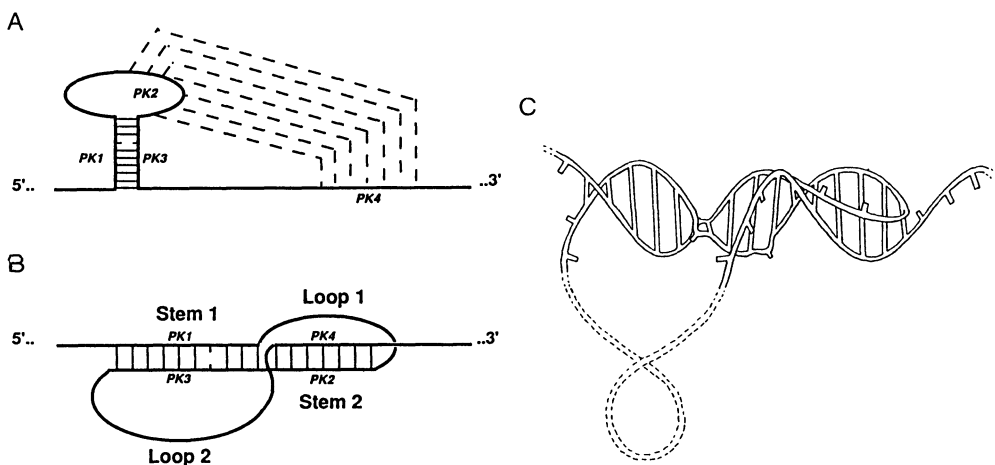


Fig. 3. Base-pairing between nucleotides in the loop of the stem-loop structure and a region downstream (A) results in the formation of an extended double helix, shown schematically (B), consisting of two helical regions (S1 and S2) connected by two single stranded loops (L1 and L2). An artist's impression of this structure is shown (C).

from the junction region is in itself sufficient to direct high efficiency frame-shifting in a heterologous context (Figure 2). Positioned at the 5' end of this sequence is a 7 nucleotide 'slippery sequence', UUUAAAC, similar to those described in retrovirus RNA frame-shifting signals (Jacks *et al.*, 1988a). By analogy with the situation in retroviruses (Jacks and Varmus, 1988a,b) this is likely to be the point at which frame-shifting actually occurs by a simultaneous "slippage" mechanism in which tRNAs decoding the UUA codon (occupying the ribosomal P site) and the AAC codon (occupying the A site) slip back on the mRNA to form mismatched (2 out of 3) pairings with UUU and AAA triplets. This model is strongly supported by the observation that introduction of a termination codon in the F1 reading frame immediately after the slippery site does not affect frame-shifting (Figure 2, pFS8.5). Thus frame-shifting must occur before this triplet is read, and since nucleotides upstream of the predicted slip site are dispensable, it seems highly likely that the UUUAAAC is indeed the site of frame-shifting. Also consistent with this is the observation that single nucleotide substitutions in the slippery sequence that would disrupt pairing in the slipped position are inhibitory to frame-shifting (Figure 2, pFS7.1, 7.14, 7.17).

In addition to the slippery site however, downstream sequences are required for efficient frame-shifting, and our mutational analysis (Brierley *et al.*, 1989) has shown that these downstream sequences fold to form an RNA tertiary structure (Figure 3), which has been termed a pseudoknot (Pleij *et al.*, 1981; Puglisi *et al.*, 1988) and which is an essential element of the signal. This structure consists essentially of a stem loop in which certain of the nucleotides in the loop are in fact paired with nucleotides provided by a downstream sequence, creating 2 stems (S1 & S2) which are stacked on top of each other and linked by two connecting loops (L1 & L2). Evidence in favour of this kind of model comes from the creation of complementary base mutations in each strand of both proposed stems (Figure 2). These changes, which would destabilise the proposed pseudoknotted structure, are highly inhibitory to frame-shifting (e.g. mutants pFS7.18, 7.19 for stem 1 and pFS 7.12, 7.15 for stem 2). However, when double mutations are made to restore potential base-pairing, creating complementary changes in both strands of each proposed stem, frame-shifting is restored to wild-type efficiency (pFS 7.20 and pFS 7.16).

The results strongly suggest that the proposed base-pairing does occur and that a pseudoknot is an essential element of the ribosomal frame-shifting signal. Although we cannot be sure, we believe the most likely structure for the pseudoknot to be that shown in Figure 4. In this structure, stem 1 consists of a 11 base pair helix, with 9 Watson-Crick pairs plus one G:U and one 'wobble' A:G pair. Stem 2 consists of a 6 base pair helix continuous with, and coaxially stacked on top of stem 1. The shorter of the connecting loops, L1, bridges the major groove of the stem 2 helix and is only 2 nucleotides long, the minimum number required to span the distance (about 11Å) between the top of stem 1 and the top of stem 2. The longer, L2, is 32 nucleotides in length, which is more than sufficient to bridge the distance (about 42Å) between the base of stem 1 and the base of stem 2.

From the model shown it is possible to make predictions about the effect of certain kinds of mutations on the pseudoknot structure and function, and to test these readily by mutagenesis (Figure 4). One obvious prediction is that much of loop 2 is redundant, and indeed 24 nucleotides can be removed from L2 without affecting frame-shifting (pFS 8.2). Similarly, insertion of an additional 6 nts has no effect (pFS 8.1). Another prediction is that nucleotides in L1 can be changed to their complements without effect, and once again this is the case (pFS 8.7). On the other hand, making complementary changes in the proposed stems should be inhibitory. For stem 1 this is demonstrated clearly by mutants pFS 8.4 and pFS 8.6 in which the top 3 nucleotides in each strand of stem 1 have been altered; in each case frame-shifting is dramatically reduced, but in the double mutant, pFS 8.12, once again it is restored to wild type efficiency. Changing only the top nucleotide of this trio (pFS 8.11) similarly has a drastic inhibitory effect. In stem 2 conversion of a central G residue to a C is highly inhibitory (pFS 7.23), and destabilisation of the nucleotide pair immediately above also reduced frame-shifting, although in this case the reduction was much less dramatic. This may indicate a lower contribution of this nucleotide pair to the overall stability of the structure, which would be consistent with its position at the very end of stem 2. Surprisingly however, changing a G residue at the base of stem 2 to a C (pFS 8.10) had no effect on frame-shifting, suggesting that the proposed G:U pairing may not occur. It is possible however that in this mutant, the U residue normally opposing the wild type G could be displaced from the structure by the adjacent

G on the end of loop 2, creating a new base pair with the mutated C. This would create an even more stable pseudoknot, and indeed this mutant displayed slightly more efficient frameshifting than did the wild type sequence. An alternative explanation however is that the G residue in the wild type structure is in reality unpaired and 'bulges' out of the helix without affecting its overall structure or stability.

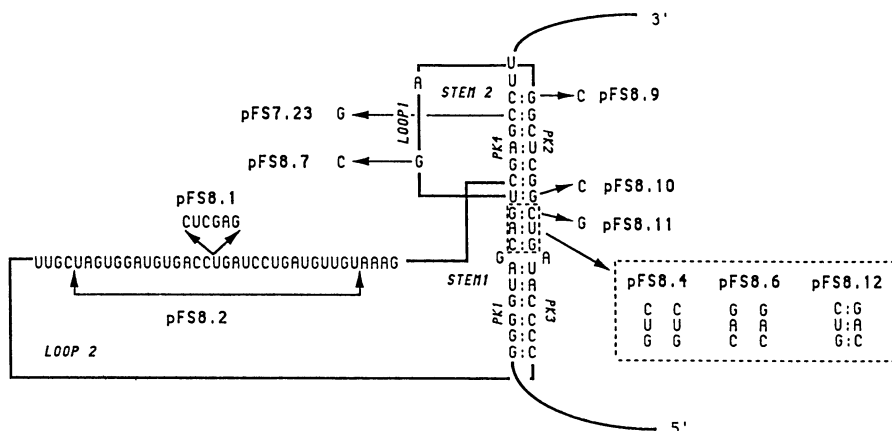
So far then our experimental observations support the proposed model, but clearly confirmation will require direct structural analysis using a range of biochemical and biophysical techniques.

Role of the pseudoknot in frame-shifting

It is not yet clear how the pseudoknot causes ribosomes to change reading frame during mRNA translation. The most obvious possibility is that the necessity to unwind the knot slows or stalls the ribosome as it passes through the frame-shift region, such that there is an increased chance of slippage at the slippery site. In such a situation presumably the position of the knot with respect to the slip site would be critical, and indeed insertion or deletion of 3 nucleotides between the two (Figure 2) severely inhibits frame-shifting. However it is equally possible that a protein factor recognises the knotted RNA, and that somehow this interaction could promote frame-shifting. Further experiments will be required to distinguish between these possibilities.

Significance of frame-shifting for IBV gene expression

Our data suggest that translation of IBV genomic RNA in infected cells leads to the production of a 440K F1 product and a 740K F1-F2 read-through product, in a ratio of about 3:1. As yet we know very little about the nature of these proteins other than that the primary translation products are almost certainly cleaved, proteolytically to produce a number of smaller polypeptides (Brierley *et al.* this volume). Since IBV genomic RNA alone is infectious



Construct	Frame-shifting efficiency	Construct	Frame-shifting efficiency
pFS8.1	++	pFS8.9	+
pFS8.2	++	pFS8.10	++
pFS8.4	-	pFS8.11	-
pFS8.6	-	pFS8.12	++
pFS8.7	++	pFS7.23	-

Fig. 4. Testing of predicted pseudoknot structure through site-directed mutagenesis of specific nucleotides. Mutated frameshift sites were constructed and assayed for frameshifting as before.

(Schochetman *et al.*, 1977) one can assume also that among the F region translation products there is a virus RNA-dependent RNA polymerase, although as yet the activity has not been identified.

Why then is ribosomal frame-shifting required for expression of the F region? One possibility is that it provides a mechanism by which a defined ratio of products may be synthesised from the upstream F1 and downstream F2 ORF. This kind of explanation has been advanced for retrovirus *gag-pol* expression which occurs by a similar mechanism. However whereas for the retroviruses RSV and HIV, the ratio of *gag* to *gag-pol* product is about 20:1 (Jacks *et al.*, 1985, 1988a), for IBV the F1:F2 ratio would be only 3:1. It is difficult to imagine therefore why such a relatively crude control mechanism would be favoured by the virus. Another explanation is that the frame-shift signal simply provides a mechanism for creating two different protein products from the same genomic region. Thus successful virus replication might require a protein consisting of the terminal region of F1 alone, and another consisting of fused sequences from both F1 and F2. This end is achieved in other ways by other viruses. For example in the picornaviruses, incomplete proteolytic cleavage at particular sites can generate alternative protein products (Kitamura *et al.*, 1981), and in the alphavirus Sindbis, a similar result arises from the presence of a "leaky" stop codon (Strauss *et al.*, 1983, Lopez *et al.*, 1985) in the non-structural coding region of the genomic RNA. Ribosomal frame-shifting therefore may represent yet another variation on this theme.

Acknowledgements

This work was supported by an AFRC "Link" grant, LRG171, awarded to SCI. We are grateful to members of the Institute of Animal Health, Houghton Laboratory, and in particular Mike Bournsnel for useful discussion

References

- Bournsnel, M.E.G., Brown, T.D.K., Foulds, I.J. Green, P.F., Tomley, F.M. and Binns, M.M., 1987, Completion of the sequence of the genome of the Coronavirus Avian Infectious Bronchitis Virus, J. Gen. Virol., 68: 57.
- Brierley, I., Bournsnel, M.E.G., Binns, M.M., Bilimoria, B., Blok, V.C., Brown, T.D.K. and Inglis, S.C., 1987, An efficient ribosomal frame-shifting signal in the polymerase-encoding region of the coronavirus IBV, EMBO J., 6: 3779.
- Brierley, I., Digard, P. and Inglis, S.C., 1989, Characterisation of an efficient coronavirus ribosomal frameshifting signal: Requirement for an RNA pseudoknot, Cell, 57: 537.
- Jacks, T. and Varmus, H.E., 1985, Expression of the Rous sarcoma virus *pol* gene by ribosomal frameshifting, Science, 230: 1237.
- Jacks, T., Madhani H.D., Masiarz, F.R., and Varmus, H.E., 1988a, Signals for ribosomal frameshifting in the Rous sarcoma virus *gag-pol* region Cell, 55: 447.
- Jacks, T., Power, M.D., Masiarz, F.R., Luciw, P.A., Barr, P.J. and Varmus, H.E., 1988b, Characterisation of ribosomal frameshifting in HIV-1 *gag-pol* expression, Nature, 331: 280.
- Kitamura, N., Semler, B.L., Roth, P.G., Larsen, G.R., Adler, C.J., Dorner, A.J., Emini, E.A., Hanecak, R., Lee, J.J., van der Werf, S., Anderson, C.W. and Wimmer, E., 1981, Primary structure, gene organisation, and polypeptide expression of poliovirus RNA Nature, 291: 547.
- Lopez, S., Bell, J.R., Strauss, E.G. and Strauss, J.H., 1985, The nonstructural proteins of Sindbis as studied with an antibody specific for the C-terminus of the nonstructural read-through polyprotein, Virology, 141: 235.
- Pleij, C.W.A., Rietveld, K. and Bosch, L., 1985, A new principle of RNA folding based on pseudoknotting, Nucl. Acids Res., 13: 1717.
- Puglisi, J.D., Wyatt, J.R. and Tinoco, I., 1988, A pseudoknotted RNA oligonucleotide, Nature, 331: 283.
- Schochetman, G., Stevens, R.H. and Simpson, R.W., 1977, Presence of infectious polyadenylated RNA in the coronavirus avian infectious bronchitis virus, Virology, 77: 772.
- Strauss, E.G., Rice, C.M. & Strauss, J.H., 1983, Sequence coding for the alphavirus nonstructural proteins is interrupted by an opal termination codon, Proc. Natl. Acad. Sci. USA, 80: 5271.