Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

J. Mol. Biol. (1992) **227**, 463–479

Mutational Analysis of the "Slippery-sequence" Component of a Coronavirus Ribosomal Frameshifting Signal

Ian Brierley†, Alison J. Jenner

Division of Virology, Department of Pathology, University of Cambridge Tennis Court Road, Cambridge, CB2 1QP, U.K.

and Stephen C. Inglis

Cantab Pharmaceuticals Research Limited, 184 Cambridge Science Park, Milton Road Cambridge, CB4 4GN, U.K.

(Received 18 February 1992; accepted 29 May 1992)

The ribosomal frameshift signal in the genomic RNA of the coronavirus IBV is composed of two elements, a heptanucleotide "slippery-sequence" and a downstream RNA pseudoknot. We have investigated the kinds of slippery sequence that can function at the IBV frameshift site by analysing the frameshifting properties of a series of slippery-sequence mutants. We firstly confirmed that the site of frameshifting in IBV was at the heptanucleotide stretch UUUAAAC, and then used our knowledge of the pseudoknot structure and a suitable reporter gene to prepare an expression construct that allowed both the magnitude and direction of ribosomal frameshifting to be determined for candidate slippery sequences. Our results show that in almost all of the sequences tested, frameshifting is strictly into the −1 reading frame. Monotonous runs of nucleotides, however, gave detectable levels of a −2/+1 frameshift product, and U stretches in particular gave significant levels (2% to 21%). Preliminary evidence suggests that the RNA pseudoknot may play a role in influencing frameshift direction. The spectrum of slip-sequences tested in this analysis included all those known or suspected to be utilized *in vivo*. Our results indicate that triplets of A, C, G and U are functional when decoded in the ribosomal P-site following slippage (XXXYYYN) although C triplets were the least effective. In the A-site (XXYYYYN), triplets of C and G were non-functional. The identity of the nucleotide at position 7 of the slippery sequence (XXXYYYN) was found to be a critical determinant of frameshift efficiency and we show that a hierarchy of frameshifting exists for A-site codons. These observations lead us to suggest that ribosomal frameshifting at a particular site is determined, at least in part, by the strength of the interaction of normal cellular tRNAs with the A-site codon and does not necessarily involve specialized "shifty" tRNAs.

Keywords: Ribosomal frameshifting; slippery sequence; transfer RNA; frameshift direction; RNA pseudoknot

1. Introduction

The coronavirus infectious bronchitis virus (IBV‡) contains a signal in the genomic RNA which directs an efficient −1 (5′-wards) ribosomal frameshift. The 5′ end of the viral RNA contains two briefly overlapping open reading frames (ORFs) 1a and 1b, with 1b in the −1 frame with respect to 1a (Boursnell *et al.*, 1987). The frameshift signal is located at the overlap region and its consequence is that a proportion of ribosomes reading the 1a frame shift into the 1b frame just before encountering the 1a termination codon and continue to translate, producing a 1a-1b fusion protein (Brierley *et al.*, 1987). The frameshift is highly efficient; when synthetic mRNAs containing the 1a/1b overlap region within a reporter gene are translated in rabbit reti-

† Author to whom all correspondence should be addressed.

‡ Abbreviations used: IBV, infectious bronchitis virus; ORF, open reading frame; RSV, Rous sarcoma virus; MMTV, mouse mammary tumour virus; HIV, human immunodeficiency virus; bp, base-pair(s); MuLV, Moloney murine leukemia virus; RCNMV, red clover necrotic mosaic virus; EAV, equine arteritis virus; kb, 10^3 base-pairs; nt, nucleotide(s).

0022-2836/92/180463-17 $08.00/0

culocyte lysates or *Xenopus* oocytes, some 30% of the ribosomes are induced to change reading frame (Brierley *et al.*, 1987, 1989). Mutagenesis experiments have revealed that the necessary information for efficient frameshifting is contained within an 86 nucleotide stretch located at the overlap region (Brierley *et al.*, 1989). The frameshift signal is composed of two elements, a heptanucleotide sequence UUUAAAC, the site where the ribosome is thought to change reading frame (Jacks *et al.*, 1988*b*) positioned at a defined distance upstream from an RNA pseudoknot structure in the mRNA (see Pleij & Bosch, 1989). Both elements are essential for efficient frameshifting. However, in the case of the pseudoknot structure, no primary sequence determinant appears to be involved in the frameshift process; as long as the overall structure is maintained, frameshifting is highly efficient (Brierley *et al.*, 1991). The IBV frameshift signal belongs to a class of efficient −1 frameshift sites first described for the vertebrate retrovirus Rous sarcoma virus (RSV: Jacks & Varmus, 1985). In recent years, examples of this kind of frameshift expression strategy have been predicted or documented to occur in many other systems (for a review see Atkins *et al.*, 1990). In RSV, frameshifting links the overlapping *gag* and *pol* coding regions enabling production of the *gag-pol* polyprotein from which reverse transcriptase is derived. An alignment of the *gag-pol*, *gag-pro* or *pro/pol* overlap sequences of many retroviruses or related eukaryotic elements known or suspected to utilize frameshifting highlighted a conserved heptanucleotide motif containing two homopolymeric triplets of the order XXXYYYN (where X = A, G or U, Y = A or U and N is A, C or U: Jacks *et al.*, 1988*b*). Amino acid sequencing of the transframe protein has revealed that this sequence motif (termed the "slippery" sequence) is the actual site of the ribosomal frameshift in mouse mammary tumour virus (MMTV, *gag-pro* junction; Hizi *et al.*, 1987), human immunodeficiency virus type 1 (HIV-1; Jacks *et al.*, 1988*a*), RSV (Jacks *et al.*, 1988*b*), and is likely to be the slip-site in all the other systems including IBV. Sequence analysis has suggested that RNA pseudoknots are present downstream from the slip-sites in a large number of cases (Brierley *et al.*, 1989; ten Dam *et al.*, 1990). Thus slippery sequences and RNA pseudoknots (or other secondary structures) appear to be common elements of this class of −1 frameshift event. To account for the conserved slip-site motif in retroviruses, Jacks *et al.* (1988*b*) proposed a simultaneous slippage model of frameshifting (see Fig. 1) in which two adjacent ribosome-bound tRNAs decoding the slip sequence in the zero reading frame (i.e. X XXY YYZ) slip back simultaneously by one nucleotide such that both tRNAs are in the −1 phase (XXX YYY) and are base-paired to the mRNA in two out of three anticodon positions. Following translocation, translation proceeds in the −1 frame. Support for this model comes from studies of the slippery-sequences of RSV (Jacks *et al.*, 1988*b*), IBV (Brierley *et al.*, 1989) and

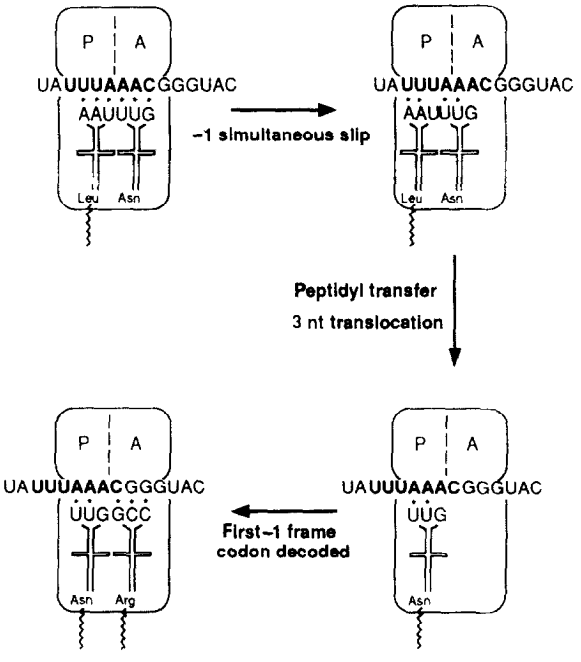

Figure 1. The simultaneous slip model of frameshifting, as proposed by Jacks *et al.* (1988*b*). In this example, $tRNA^{Leu}$ carrying the nascent peptide (jagged line) and $tRNA^{Asn}$ are shown bound to the IBV 1a frame codons UUA and AAC in the ribosomal P(peptidyl)- and A(aminoacyl)-sites. In about 30% of cases, both tRNAs slip back by 1 nucleotide in the 5′ direction, such that they now pair with the 1b frame codons UUU and AAA and form base pairs (dots) only with the first 2 nucleotides of each codon. Following peptidyl transfer and translocation, the next 1b frame codon, CGG, is brought into the A-site, where it is decoded by $tRNA^{Arg}$. In this example, which is based on that presented by Jacks *et al.* (1988*b*), the tRNA anticodon sequences are shown as standard Watson–Crick base-pairs. The actual anticodon sequences are described and discussed in the text.

the yeast double-stranded RNA virus, L-A (Dinman *et al.*, 1991), where it has been found that mutations that reduce the homopolymeric nature of the slip-sequence greatly reduce frameshifting, presumably by reducing the post-slippage mRNA–tRNA complementarity. In those systems in which the amino acid sequence of the transframe protein has been determined, it is known that the frameshift occurs within the slippery sequence at the second codon of the tandem slippery pair, i.e. at that codon decoded in the ribosomal A-site (XXXYYYN; zero frame, A-site codon underlined). Amongst natural retroviral slip sequences (Jacks *et al.*, 1988*b*) and in probably all cases of this class of frameshift signal found to date in viruses (ten Dam *et al.*, 1990), only three A-site codons (AAC, UUA and UUU) are represented, and it has been suggested that only certain "shifty" tRNAs are able to slip at the A-site (Jacks *et al.*, 1988*b*). The potential involvement of specialized tRNAs in the frameshift process was further implicated by the observation that the presence of a G nucleotide at position 7 of the slip-sequence (XXXYYYG) greatly reduces frame-

shifting in eukaryotic systems (Inglis *et al.*, 1990; Dinman *et al.*, 1991) but not in prokaryotic systems (Weiss *et al.*, 1989; Tsuchihashi, 1991).

Here, we describe a detailed analysis of the kind of slippery sequences that can function at the IBV frameshift site. Our approach was to replace the IBV slip-sequence with candidate sequences and then to test the ability of synthetic mRNAs containing the variant sites to direct frameshifting in a cell-free translation system. We hoped to be able to determine the frameshift efficiency of a large number of sites, particularly position 7 variants, since this would allow the slipperiness of a range of A-site-decoding tRNAs to be investigated. We firstly confirmed that the site of frameshifting in IBV was at the UUUAAAC sequence, and then used our knowledge of the pseudoknot structure (Brierley *et al.*, 1991) and a suitable reporter gene to design an expression construct in which all the termination codons within the pseudoknot were removed such that frameshifting into any reading frame could be monitored simultaneously. Thus, for each candidate slip sequence, both the magnitude and, to some extent, the direction of frameshifting could be determined. Our results show that in almost all of the sequences tested, frameshifting was strictly into the −1 reading frame. However, detectable levels of −2/+1 frameshifting (0·5%) were seen with monotonous runs of nucleotides (A_6N, C_6N, G_6N) or when a U triplet was present in the A-site. Significant levels of −2/+1 frameshifting were observed with monotonous U tracks (U_6N; 2% to 5%) and eight consecutive U bases gave very high levels (21%). These observations support the widely held view that runs of U bases in mRNAs are particularly slippy. With respect to the kind of sequences that could functionally replace the IBV slip sequence, we found that triplets of A, C, G and U were functional when placed in that region of the slip-sequence decoded in the ribosomal P-site following slippage (XXXYYYZ; −1 frame P-site codon underlined), but C-triplets were the least slippy. Only triplets of A and U were functional in the ribosomal A-site (XXXYYYZ; −1 frame A-site codon underlined), in agreement with related studies of RSV (Jacks *et al.*, 1988*b*) and L-A frameshifting (Dinman *et al.*, 1991). Surprisingly, a number of mutations created in the first slippery codon (XXXYYYN, decoded in the P-site) were found to display high levels of frameshifting, despite the apparent reduction in the homopolymeric nature of the slip sequence. This suggests that in certain cases only minimal post-slippage mRNA–tRNA pairing is required for efficient frameshifting. The identity of the nucleotide at position 7 of the slip-sequence was found to be a critical determinant of frameshift efficiency and we show here that a hierarchy of frameshifting is seen for the various pre-slippage A-site codons. For the series, XXXAAAN, the hierarchy for N was C > A ~ U ≫ G; for XXXUUUN, the hierarchy was U > A ~ C > G. These observations lead us to suggest that frameshifting at a particular site is determined, at least in part, by the strength of interaction of normal tRNAs with the A-site codon and does not necessarily involve specialized "shifty" tRNAs.

2. Materials and Methods

(a) *Site-specific mutagenesis*

Site-directed mutagenesis was carried out by a procedure based on that described by Kunkel (1985; Brierley *et al.*, 1989). All the plasmids used in this study contain the intergenic region of the filamentous bacteriophage f1 (Dotto *et al.*, 1981) enabling single-stranded plasmid DNA to be generated following infection of plasmid-carrying bacteria with bacteriophage R408 (Russel *et al.*, 1986). Uracil-containing, single-stranded DNA substrates for mutagenesis were prepared by R408-superinfection of plasmid-carrying *Escherichia coli* RZ1032 cells (dut⁻ ung⁻; Kunkel, 1985). Oligonucleotides for mutagenesis were synthesized using an Applied Biosystems 381A DNA synthesizer and the mutagenesis reactions performed as before (Brierley *et al.*, 1989). Mutants were identified by dideoxy sequencing (Sanger *et al.*, 1977) of single-stranded DNA templates rescued from *E. coli* JM101 (Yanisch-Perron *et al.*, 1985).

(b) *Construction of plasmids*

Plasmid pEMBL8+ (Dente *et al.*, 1983) was digested with *Rsa*I and a 519 bp fragment containing the bacteriophage f1 intergenic region isolated. Plasmid pST2+ (Digard *et al.*, 1989), which contains the influenza A/PR8/34 PB2 gene (Young *et al.*, 1983) inserted into the *Bgl*II site of pSP64-T (Krieg & Melton, 1984), was digested with *Pvu*II (within the vector sequence) and ligated with the 519 bp *Rsa*I f1 fragment, to produce plasmid pING12. The orientation of the f1 fragment in pING12 is such that the non-coding strand of the plasmid is packaged into R408 bacteriophage particles during single-stranded DNA production (see section (a), above). Plasmid pFScass5 was created as follows. Firstly, 120 pmol of oligonucleotide 1 (sequence, 5′TAGATCTCAAAGGCTCGCTTTGCAAGGGGTACCAGCCGAGCCTGACTGATACCCCGTATCAGTTTAAAGCCCTA TAGTGAGTCGTATTAA 3′; 90mer) was annealed with 400 pmol of oligonucleotide 2 (sequence, 5′ TAGATCTTAATACGACTCACTATAG 3′; 25mer) by heating at 70 °C for 5 min in 7 μl of water, then incubating at 23 °C for 1 h after addition of 1 μl of a 10× concentrate of DNA polymerase large fragment (Klenow) reaction buffer (10× = 500 mM-Tris·HCl (pH 7·5 at 20 °C), 100 mM-$MgCl_2$, 10 mM-dithiothreitol, 0·5 mg bovine serum albumin/ml). After this time, 1 μl of a mixture of the 4 deoxynucleotide triphosphates (dNTPs) was added (final concn 2·5 mM each dNTP) and the single-stranded regions of the annealed primers copied by addition of 1 μl of DNA polymerase Klenow fragment (5 units; Boehringer Mannheim) and incubation at 23 °C for 1 h. After this time, a further 5 units of Klenow fragment were added and incubation continued for 2 h. The resulting double-stranded DNA was isolated by precipitation with ethanol following extraction with phenol/chloroform (1 : 1, v/v), and phosphorylated using bacteriophage T4 polynucleotide kinase as described (Brierley *et al.*, 1989). The kinased template was self-ligated using phage T4 DNA ligase and standard reaction conditions (13 °C, 16 h, Maniatis *et al.*, 1982) and then digested with *Bgl*II. The resulting double-stranded, *Bgl*II-flanked DNA fragment was ligated into the influenza PB2 gene of pING12 at the *Bgl*II site (position 456, Young *et al.*, 1983).

Recombinants were screened for insertion of a single copy of the fragment in the correct orientation by dideoxy sequencing. The resulting plasmid was subjected to 2 rounds of site-directed mutagenesis. Firstly, a nucleotide in the PB2 gene was changed from A to C at position 508 in the coding region (Young *et al.*, 1983) to change a nonsense codon (TAA) to a serine codon (TCA). Secondly, the sequence 5′ TGAGATCT 3′ present at the 3′ end of the DNA inserted into the *Bgl*II site of pING12 was converted to 5′ TCAGCTGT 3′. This removed a termination codon (TGA) and introduced a *Pvu*II restriction site (CAGCTG). The resulting construct was termed pFScass5, and the rationale behind its construction is discussed in the text. All plasmids were maintained in *E. coli* JM101.

(c) *Preparation of plasmid DNA for* in vitro *transcription*

Plasmids for transcription were prepared by the alkaline lysis mini-preparation method (Birnboim & Doly, 1979) and linearized by digestion with either *Sma*I (pFS8 derivatives) or BamH1 (pFScass5 derivatives); extracted once with phenol/chloroform (1 : 1, v/v) and the aqueous phase passed through a Sephadex G-50 spin column (Maniatis *et al.*, 1982). Linearized template was concentrated by precipitation with ethanol and transcribed using bacteriophage T7 or SP6 RNA polymerase as described (Brierley *et al.*, 1987, 1991).

(d) *Translation of synthetic mRNAs* in vitro

Serial dilutions of purified mRNAs were translated in rabbit reticulocyte lysates as described (Brierley *et al.* 1987) and translation products analysed on SDS/10% or 17·5% (w/v) polyacrylamide gels according to standard procedures (Hames, 1981). The relative abundance of non-frameshifted or frameshifted products on the gels was estimated by scanning densitometry of direct autoradiographs and adjusted to take into account the differential methionine content of the products. Scans were performed on exposures that were in the range where film response to excitation was linear. Frameshift efficiencies were calculated from those dilutions of RNA where translation was highly processive (RNA concentrations of 10 μg to 25 μg RNA/ml of reticulocyte lysate).

3. Results

(a) *The site of frameshifting is the UUUAAAC sequence*

We began by confirming that IBV frameshifting does indeed occur at the sequence UUUAAAC. To do this, we employed the frameshift expression plasmid pFS8 (see Fig. 2(a) and Brierley *et al.*, 1991), which contains the essential components of the IBV frameshift signal cloned within a reporter gene (influenza A/PR8/34 PB1; Young *et al.*, 1983) under the control of the bacteriophage T7 RNA polymerase promoter. Transcription of the plasmid with T7 RNA polymerase and subsequent translation of the synthetic RNA in rabbit reticulocyte lysates produces two major species, a 45 kDa protein corresponding to ribosomes that terminate at the 1a stop-codon within the 1a/1b overlap region, and a 95 kDa frameshift product. Approximately 30% of ribosomes change frame within the overlap region. To confirm that ribosomes frameshift at the UUUAAAC sequence, two mutant derivatives of pFS8 were prepared. In pFS8.5, the 1a frame (zero frame) glycine codon immediately downstream from the UUUAAAC sequence was converted by muta-

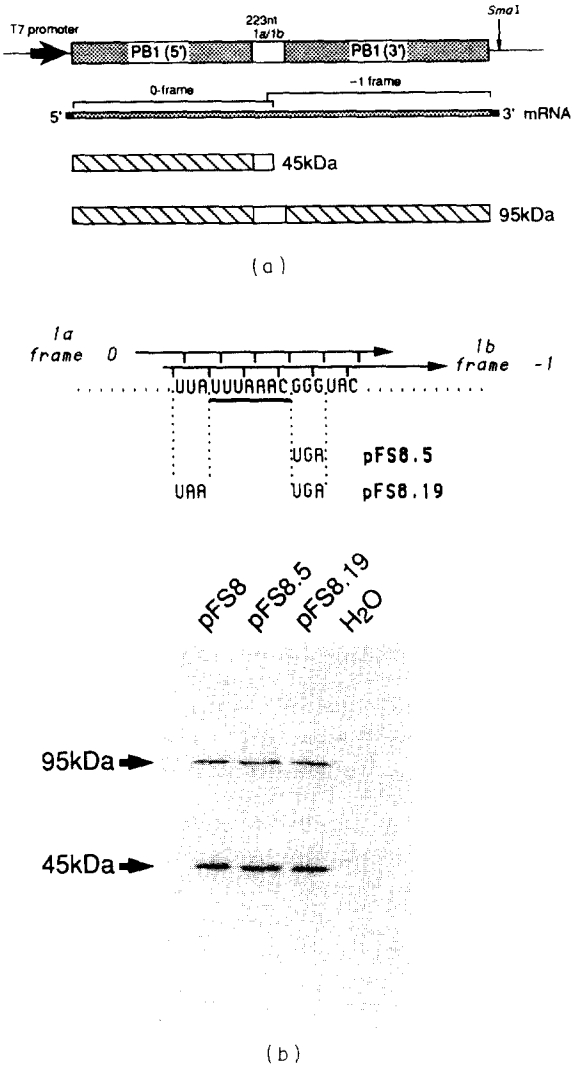

Figure 2. Definition of the IBV frameshift site. (a) The diagram of plasmid pFS8 shows the IBV 1a/1b ORF overlap region containing the frameshift signal (in white) flanked by the influenza PB1 reporter gene (shaded). Linearization of the plasmid with *Sma*I and *in vitro* transcription using T7 RNA polymerase yields an mRNA (2·8 kb) that, when translated in rabbit reticulocyte lysates, produces a 45 kDa product corresponding to ribosomes that terminate at the 1a termination codon within the 1a/1b overlap region, and a 95 kDa (−1) frameshift product corresponding to a PB1 (5′)-1a-1b-PB1 (3′) fusion protein. (b) The nucleotide sequence around the predicted frameshift site, UUUAAAC, is shown. Two mutant derivatives of pFS8 were prepared in which termination codons were introduced downstream (pFS8.5) or both upstream and downstream (pFS8.19) from the UUUAAAC sequence. The gel shows the reticulocyte lysate translation products synthesized in response to mRNAs derived from *Sma*I-digested pFS8 or mutant templates. Polypeptides were labelled with [^{35}S]methionine, separated on a SDS/10% polyacrylamide gel, and detected by autoradiography.

genesis to a stop codon (UGA). In pFS8.19, the same mutation was introduced but in addition the leucine codon just upstream from the *UUUAAAC* stretch (in the −1 or 1b frame) was changed to a termination codon (UAA). In pFS8.19, the upstream and downstream ORFs of the reporter gene overlap solely by the seven nucleotides that constitute the putative slip-site UUUAAAC. For ribosomal frameshifting to be observed in this mutant, ribosomes would have to slip within the heptanucleotide sequence. As can be seen in Figure 2(b), both pFS8.5 and pFS8.19 show frameshifting at wild-type levels, demonstrating that frameshifting does indeed occur at this position.

(b) *A frameshift cassette to monitor all reading frames*

We wished to design a construct in which frameshifting into any reading frame could be monitored simultaneously, since this was not possible with pFS8. As the wild-type IBV pseudoknot contains a number of stop codons, we firstly designed an artificial "minimal" pseudoknot based on our knowledge of this RNA structure (Brierley *et al.*, 1991), in which loop 2 was reduced from 32 to only 8 nucleotides in length, and all the termination codons were changed to sense codons (see Fig. 3). The changes were carefully chosen so as to form a pseudoknot whose structure and predicted stability bore close resemblance to the wild-type pseudoknot. As can be seen in Figure 3, the minimal pseudoknot differs from the wild-type in a number of ways; loop 2 is 24 nucleotides shorter, the G-A mismatched pair in stem 1 of the wild-type pseudoknot is replaced by a G·U base-pair, the G nucleotide of loop 1 of the wild-type structure is replaced by a C nucleotide and, finally, the minimal pseudoknot has no stop codons. There is no difference in the lengths of the two stems in each structure. Previous work (Brierley *et al.*, 1991) has shown that mutations that change the nucleotide sequence of the pseudoknot without greatly influencing the overall structure do not inhibit the frameshift process. Thus, reducing the length of loop 2 to eight nucleotides and changing the sequence of loop 1 ought not to influence frameshifting. However, by replacing the mismatched G-A pair in stem 1 by G·U, one may expect a moderate increase in the stability of stem 1 and hence a small increase in the frameshift efficiency (Brierley *et al.*, 1991; see later). Having designed the minimal pseudoknot, we then searched available reporter genes for regions in which significant lengths of ORF were present in all three frames. A suitable candidate found was the influenza PB2 gene that contained such a region just downstream from a unique *Bgl*II site beginning at position 456 (Young *et al.*, 1983). Our strategy was to clone the minimal pseudoknot into the *Bgl*II site such that the upstream portion of the PB2 gene was in-frame with the IBV 1a coding region. As the downstream PB2 information was open (to a greater or lesser extent) in all three reading frames, we

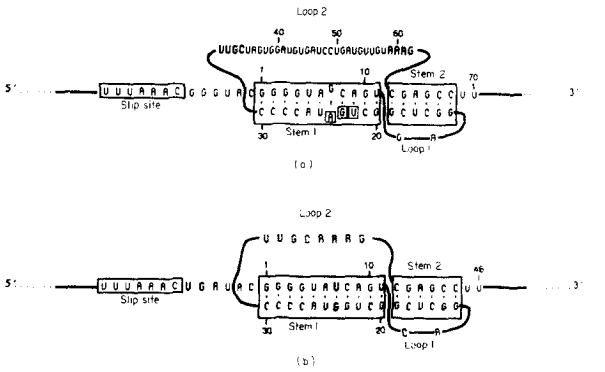

Figure 3. A comparison of (a) the wild-type IBV RNA pseudoknot structure and (b) the minimal pseudoknot used in pFScass5 and derivatives. In the wild-type structure, the 2 stacked stems are drawn horizontally, and form a quasi-continuous double helix of 16 bp (shaded) with a mismatched pair (G7-A24) in stem 1. The single-stranded connecting loops contain 2 (loop 1) and 32 (loop 2) nucleotides, respectively. The termination codon of the 1a ORF (UGA) is boxed, as is the slippery sequence (UUUAAAC). The nucleotides in loop 2 that were retained in the minimal pseudoknot are in a larger, bold font. In the minimal pseudoknot, which contains no termination codons, the stems are of the same length and base-pair composition as the wild-type pseudoknot, except that the G7-A24, mismatched pair is replaced by a U7·G24 pair. Three other differences are present; the loop 1 G nucleotide is changed to a C; loop 2 contains only 8 of the 32 nucleotides present in the wild-type structure, and the glycine codon immediately downstream from the slippery sequence is replaced by a termination codon (in the 1a reading frame).

could monitor frameshifting into any frame simultaneously by examining the size of the frameshift products on SDS/polyacrylamide gels. The strategy for the construction of this plasmid is shown in Figure 4 and detailed in Materials and Methods. An oligonucleotide (90mer) was synthesized that contained sequences complementary to the minimal pseudoknot, the slip-sequence and a bacteriophage T7 RNA polymerase promoter. A shorter oligonucleotide was then annealed that contained the reverse complement of the T7 promoter sequence and the single-stranded regions filled by primer extension using the large fragment of DNA polymerase I. The blunt-ended molecules were self-ligated and then digested with *Bgl*II. The resulting *Bgl*II pseudoknot "cassette" was cloned into the *Bgl*II site of plasmid pING12 at the aforementioned site in the PB2 gene. The construct was then modified in two ways. Firstly, an inconvenient termination codon (UAA) within one of the PB2 frames was changed to serine (UCA), such that the particular frame in question (zero frame) was extended by some 60 amino acid codons in order to ease assignment of translation products to reading frame. Secondly, a *Pvu*II restriction site was introduced at the 3′ end of the inserted cassette. This was carried out to remove a termination codon that was inadvertently missed during design of the long oligonucleotide, and to provide a suitable restriction site

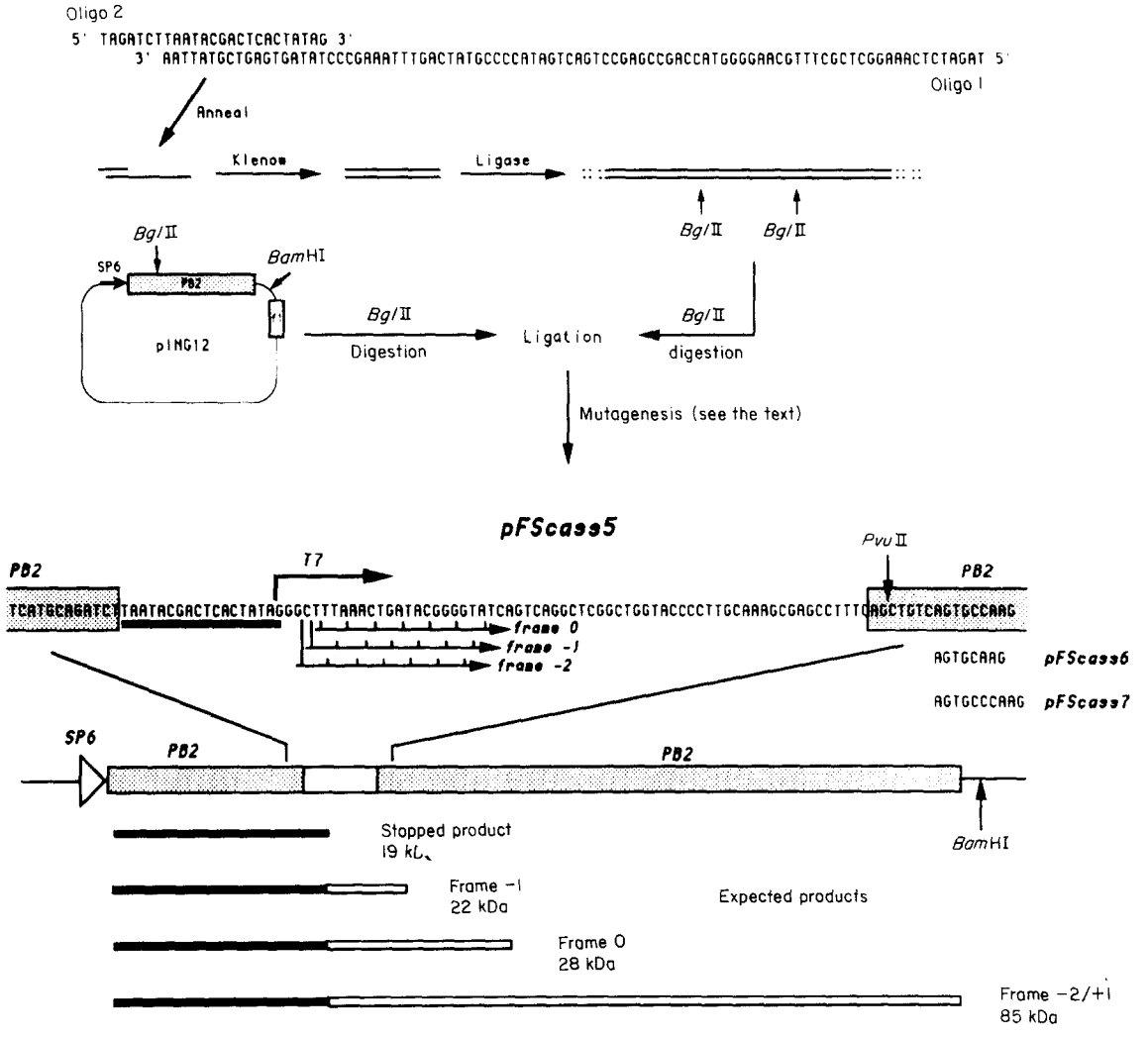

Figure 4. The construction of plasmid pFScass5 has been described in detail in the text. Briefly, an oligonucleotide containing sequences complementary to a T7 promoter and the IBV minimal pseudoknot (oligo1) was annealed to an oligonucleotide containing the T7 promoter sequence (oligo2) and the single-stranded regions filled with the Klenow fragment of DNA polymerase I. Following self-ligation, a *Bgl*II frameshift-signal-containing DNA fragment was isolated and ligated into a reporter gene, the influenza PB2 gene, to create pFScass5. The diagram of pFScass5 shows the minimal frameshift signal (in white) flanked by the influenza PB2 reporter gene (shaded). Linearization of the plasmid with *Bam*H1 and *in vitro* transcription using SP6 RNA polymerase yields an mRNA (2·4 kb) that, when translated in rabbit reticulocyte lysates, is predicted to produce a 19 kDa product corresponding to ribosomes that terminate at the new 1a termination codon (UGA) located immediately downstream from the UUUAAAC sequence (stippled), and a 22 kDa (−1) frameshift product. Any ribosomes that enter the −2/+1 reading frame would produce an 85 kDa product, and any zero frame ribosomes, a 28 kDa product. Plasmids pFScass6 and 7 differ from pFScass5 only by the deletion (pFScass6) or insertion (pFScass7) of a single C nucleotide just downstream from the inserted frameshift cassette.

for the generation of short, pseudoknot-containing T7 transcripts for structural studies (not addressed here). The resulting construct, pFScass5 is shown in Figure 4. As the natural 1a termination codon was within the pseudoknot structure, we introduced a substitute termination codon immediately downstream from the UUUAAAC sequence (in the original oligonucleotide) in order to terminate zero frame ribosomes. The PB2 reporter gene into which the cassette was cloned is under the control of a bacteriophage SP6 RNA polymerase promoter. Thus, synthetic transcripts derived by SP6-transcription of *Bam*HI linearized pFScass5 were expected to produce upon translation in the rabbit reticulocyte lysate system, a 19 kDa product corresponding to ribosomes that terminate without frameshifting and a 22 kDa −1 frameshift product. In addition, any ribosomes that enter the −2 (or +1) phase would give an 85 kDa product, and a 28 kDa product would be seen for any ribosomes that do not frameshift yet in some way suppress the termination codon and continue translation in the zero reading frame.

(c) *The IBV frameshift is strictly into the −1 phase*

Shown in Figure 5 are the translation products from pFScass5 and two related constructs, pFScass6 and 7, which differ from pFScass5 only by

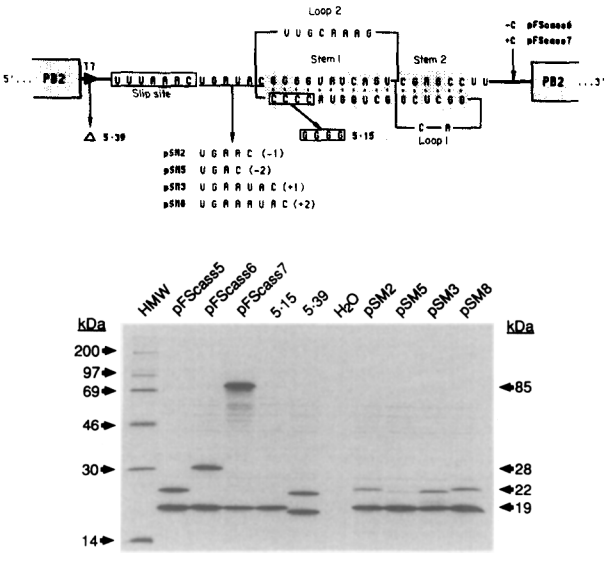

Figure 5. Characterization of the pFScass5 construct. This Figure shows the nucleotide sequence of the frameshift region in pFScass5. The slip-site UUUAAAC is boxed, and the nucleotides that separate the slip-site and pseudoknot (the spacer) are underlined. The 2 stacked stems of the pseudoknot are drawn horizontally and are shaded. A number of mutants created in pFScass5 are depicted. In 5.39, the T7 promoter was deleted but the 1a frame was maintained. A number of spacing mutants (pSM) are shown in which 1 or 2 nucleotides were deleted or inserted from the spacer sequence. These mutants were prepared in the background of either pFScass6 or 7 such that frameshifted ribosomes produce a 22 kDa product in each case (see the text). A mutation at the base of stem 1, and known to destabilize the wild-type RNA pseudoknot (Brierley *et al.*, 1991) was also prepared (5.15, CCCC to GGGG). The gel shows the reticulocyte lysate translation products synthesized in response to mRNAs derived from *Bam*HI-digested pFScass5 or mutant templates. Polypeptides were labelled with [^{35}S]methionine, separated on an SDS/17·5% polyacrylamide gel and detected by autoradiography.

the deletion (pFScass6) or addition (pFScass7) of a C nucleotide downstream from the cassette. Insertion or deletion of the single nucleotide ought to reassign the translation products to a different class of frameshift. Thus in pFScass6, the −1 frame is 28 kDa; the −2(+1), 22 kDa and the zero frame 85 kDa. In pFScass7, the −1 frame is 85 kDa, the −2(+1) frame 28 kDa and the zero frame, 22 kDa. As can be seen from the translations the constructs produced the expected −1 frameshift products; 22 kDa in pFScass5, 28 kDa in pFScass6 and 85 kDa in pFScass7. Thus, all the predicted open reading frames were present and correct. Estimates of the frameshift efficiency of the constructs indicate that some 40% of ribosomes enter the −1 phase at the frameshift site. This is some 10% higher than that seen in the wild-type pseudoknot construct (pFS8) and probably reflects, at least in part, the increased stability of the stem 1 helix in the pFScass constructs. (Using the base-stacking rules of Turner *et al.* (1988) it can be estimated that introducing the G·U base-pair in stem 1 increases the helix stability by 3·8 kcal mol^{-1} (1 cal = 4·184 J), from −12·4 kcal mol^{-1} (stem 1 of pFS8) to −16·2 kcal mol^{-1} (stem 1 of pFScass5).) In the pFScass5 translation, the absence of significant quantities of 85 kDa product suggests strongly that frameshifting is strictly into the −1 reading frame. Although a +2 shift would place ribosomes in the same phase, it is very unlikely that this occurs, since the ribosome-bound tRNAs could form only one P-site mRNA–tRNA base-pair and no A-site pairs following a +2 slip. In addition, we have shown previously that the primary sequence of the triplet downstream from the slip sequence is unimportant in the frameshift process (Brierley *et al.*, 1989). Furthermore, a +2 slip is not compatible with the results of amino acid sequencing of retroviral *trans*-frame proteins (see Introduction).

In order to check that the frameshift observed in pFScass5 was dependent upon the downstream pseudoknot, we made a destabilizing change at the base of stem 1 in pFScass5.15 (see Fig. 5). This mutation completely abolished frameshifting, as expected, highlighting the requirement for a pseudoknot at this frameshift site. In the pFScass5 translation, in addition to the intense 19 kDa and 22 kDa species, a large number of minor products were seen, with mobilities ranging, in the main, from about 20 kDa to 65 kDa. This size distribution raised the possibility that they may have been translated from a minor transcript generated as a result of aberrant recognition of the pFScass5 T7 promoter by the SP6 RNA polymerase during the *in vitro* transcriptions. However, when the T7 promoter was deleted in construct pFScass 5.39 (Fig. 5), the minor products were still present. It is therefore more likely that these proteins arise as a result of low-level aberrant translation events in the *in vitro* system. Although it is likely that some minor products may co-migrate with the authentic stopped and frameshift products on the gels, the comparative abundance of the minor bands is low, and hence can introduce only a small error into the calculated frameshift efficiencies.

Previous work with the IBV frameshift signal demonstrated that the precise distance between the slip-sequence and the pseudoknot (6 nucleotides) had to be maintained; increasing or decreasing this distance by three nucleotides greatly reduced or abolished frameshifting (Brierley *et al.*, 1989). We wished, therefore, to confirm that the six nucleotide spacing between the slip sequence and the minimal pseudoknot structure was optimal. Shown in Figure 5 are the translations of four spacing mutants in which this distance was decreased or increased by one or two nucleotides. In these mutants, the size of the −1 frameshift product was 22 kD in each case, since the mutations were introduced into the background of either pFScass6 (−2 deletion, pSM5, +1 insertion pSM3) or pFScass7 (−1 deletion, pSM2; +2 insertion, pSM8) in order to assign the −1 frameshift to this size class and ease visualization and quantification of the frameshift efficiencies. The

results show that the optimal spacing distance is indeed six nucleotides; shortening or lengthening the spacing by one or two nucleotides steadily decreased the frameshift efficiency. Spacing distances of five and seven nucleotides gave reasonable frameshifts (20·7% and 21·7% respectively), but with spacing distances of four or eight nucleotides, the efficiency was considerably reduced (6% and 13·2% respectively).

RNA pseudoknots have been proposed to play a role in the suppression of termination codons in some animal virus systems (ten Dam *et al.*, 1990) and are known to be important in translational readthrough of the termination codon that separates the *gag* and *pol* coding regions of the retrovirus Moloney murine leukaemia virus (MuLV: Honigman *et al.*, 1991; Wills *et al.*, 1991). We were interested to see if the IBV pseudoknot could also promote termination codon suppression. The pFScass6 construct allows such an event to be monitored, since any ribosomes that fail to frameshift and to terminate at the UGA codon downstream from the slip-sequence continue in the zero reading frame and produce an 85 kDa protein. However, although a distinct 85 kDa product was produced upon translation of mRNA derived from pFScass6 (see Fig. 5), the abundance of this product was very low (0·5% or less). Thus, it seems that at least at the stop codon–pseudoknot spacing distance in pFScass6 (3 nucleotides), the IBV pseudoknot cannot suppress significantly a UGA termination codon.

(d) *Slip-site requirements for efficient frameshifting*

Having characterized the pFScass5 construct, we set out to examine the kinds of slippery sequence that could specify efficient frameshifting in the context of the IBV pseudoknot. We concentrated our efforts on the analysis of homopolymeric runs, either fully monotonous (e.g. A_6N) or matched pairs of triplets (e.g. A_3U_3N), since it was known that point mutations that reduce the homopolymeric nature of the slip site greatly diminish or abolish the frameshift (Jacks *et al.*, 1988*a*,*b*; Brierley *et al.*, 1989; Dinman *et al.*, 1991). Shown in Figure 6 are the results of translations of a large number of slip site variants. These were chosen to include all those considered to be utilized *in vivo* (ten Dam *et al.*, 1990). Each particular slip site is represented by four constructs, since the identity of the seventh nucleotide of the candidate sequence was varied in each case. The frameshift efficiencies measured for each construct are detailed in Table 1. Although a large number of constructs were studied, several common features emerge from the analysis.

(i) *The pseudoknot can act as a frameshift enhancer*

Previous work has shown that slippery sequences in isolation can direct only low levels of frameshifting; in most cases, this level is 1% or less (A_6C (Jacks *et al.*, 1987); A_3U_3A (Jacks *et al.*, 1988*b*); U_3A_3C (Brierley *et al.*, 1989); G_3U_3A (Dinman *et al.*, 1991); G_3A_3C (ten Dam, Brierley, Inglis & Pleij, unpublished results)). An exception to this is the slip sequence of HIV-1, which comprises a homopolymeric U stretch (U_6A) and promotes higher levels of frameshifting (5%; Jacks *et al.*, 1988*a*; Wilson *et al.*, 1988). Examination of the level of frameshifting seen when these particular slip sequences are placed upstream from the IBV pseudoknot (Table 1) reveals that frameshifting is substantially increased. For the sites that in isolation show very low levels of frameshifting, the pseudoknot amplifies the slip enormously, resulting in efficiencies between 16% and 40%. Even the "naturally slippery" U_6N sequence is stimulated some five- to tenfold. These observations indicate that the pseudoknot can act as enhancer of frameshifting, and this is consistent with models for frameshifting in which the function of the pseudoknot is to impede the progress of ribosomes (Jacks *et al.*, 1988*a*,*b*). It is possible that the pseudoknot can arrest ribosomes in the act of decoding the slip site; perhaps increasing the decoding time such that ribosome-bound tRNAs can re-align on the mRNA with improved efficiency. As can be seen in Figure 6, however, a considerable

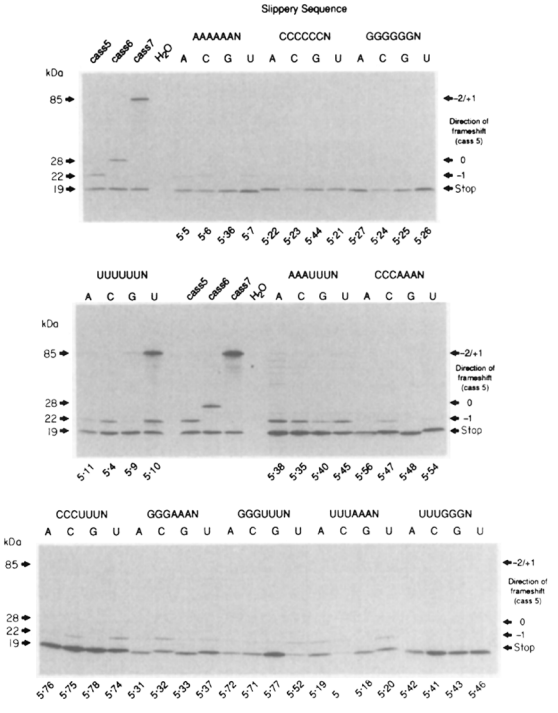

Figure 6. Slip-sequence requirements for frameshifting. In this Figure, the frameshift properties of a large number of slip-sequence mutants are shown. In most of these, the wild-type IBV slip-sequence UUUAAAC was replaced by a candidate sequence, for example AAAAAAA. For each class of sequence (e.g. AAAAAAN), N represents either A, C, G or U at the 7th position of the slip-sequence. The gels show the reticulocyte lysate translation products synthesized in response to mRNAs derived from *Bam*HI-digested pFScass5 or mutant derivatives. Polypeptides were labelled and analysed as described in the legend to Fig. 5. The bottom of each track is labelled with a number to identify each mutant.

number of sites tested do not show high levels of frameshifting. This suggests that the pseudoknot can amplify the signal only when the slip sequence contains inherent potential for slippage.

(ii) *Functional and non-functional slip sequences*

In the simultaneous slippage model for −1 retroviral frameshifting (Jacks *et al.*, 1988), the two ribosome-bound tRNAs paired in the zero frame slip back simultaneously by one nucleotide and each retain two (or more) out of three mRNA–tRNA base-pair contacts (see Fig. 1). However, as is clear from the results shown in Figure 6 and experiments with the slip sites of RSV (Jacks *et al.*, 1988*b*) and yeast L-A (Dinman *et al.*, 1991), simply having the potential for the formation of four out of six post-slippage pairs is not necessarily commensurate with an efficient slippery sequence. If one considers only the sites that conform to the XXXYYYN motif, there are, from our analysis (see Table 1), two main classes of slip sequence that are non-functional. In the first class, a C or G triplet is present in the part of the slip-sequence that is decoded in the ribosomal A-site following slippage (XXXCCCN or XXXGGGN; −1 frame A-site codon underlined). In the second class, the seventh nucleotide of the motif is a G (XXXYYYG). Thus the only A-site codons that can function in frameshifting in the reticulocyte lysate system are AAA, AAC, AAU and UUA, UUC, UUU. The requirements for specific nucleotides in that part of the slip sequence decoded in the ribosomal P-site were less stringent, with all four nucleotide triplets functioning to some extent (XXXYYYN; −1 frame P-site codon underlined). However, C-triplets gave, in general, very low frameshift levels, perhaps reflecting the fact that no P-site C-triplets have been described as yet in naturally occurring slip sequences (ten Dam *et al.*, 1990).

Table 1

The −1 frameshift efficiencies in the analysis shown in Figure 6

Construct	Slip-sequence	Frameshift efficiency (%)
5·5	A AAA AAA	12·1
5·6	A AAA AAC	19·0
5·36	A AAA AAG	<1
5·7	A AAA AAU	10·1
5·22	C CCC CCA	4·4
5·23	C CCC CCC	4·8
5·44	C CCC CCG	<1
5·21	C CCC CCU	2·5
5·27	G GGG GGA	4·4
5·24	G GGG GGC	5·3
5·25	G GGG GGG	5·4
5·26	G GGG GGU	2·5
5·11	U UUU UUA	24·7
5·4	U UUU UUC	27·0
5·9	U UUU UUG	8·0
5·10	U UUU UUU	39·1
5·38	A AAU UUA	20·5
5·35	A AAU UUC	17.1
5·40	A AAU UUG	6·8
5·45	A AAU UUU	24·6
5·56	C CCA AAA	4·1
5·47	C CCA AAC	6·8
5·48	C CCA AAG	<1
5·54	C CCA AAU	3·2
5·76	C CCU UUA	3·3
5·75	C CCU UUC	4·7
5·78	C CCU UUG	<1
5·74	C CCU UUU	17·1
5·31	G GGA AAA	11·2
5·32	G GGA AAC	32·5
5·33	G GGA AAG	2·1
5·37	G GGA AAU	10·8
5·72	G GGU UUA	16·1
5·71	G GGU UUC	15·8
5·77	G GGU UUG	2·5
5·52	G GGU UUU	38·3
5·19	U UUA AAA	25·2
5	U UUA AAC	41·7
5·18	U UUA AAG	<1
5·20	U UUA AAU	22·0
5·42	U UUG GGA	1·0
5·41	U UUG GGC	1·8
5·43	U UUG GGG	2·1
5·46	U UUG GGU	<1

Each value is the average frameshift efficiency measured from the translation of 2 or more dilutions of the relevant RNA (see Materials and Methods). The variation in the efficiency measured for each dilution was 5% or less, i.e. a value of 12·1 represents a frameshift efficiency of between approximately 11·5 and 12·7%.

Examination of the candidate slip sites shown in Figure 6 reveals that within each group the identity of the seventh nucleotide greatly influences the frameshift. If this nucleotide is a G (A-site codon AAG or UUG), frameshifting is greatly reduced, as was seen in the yeast L-A system (Dinman *et al.*, 1991). As the presence of the IBV pseudoknot downstream from the slip site amplifies the frameshift level of inherently slippery sequences to easily quantifiable levels, it has been possible to determine a hierarchy of frameshifting with respect to the identity of the seventh nucleotide. When the A-site codon is AAN, this hierarchy is C > A ~ U ≫ G; when the codon is UUN, the hierarchy is U > A ~ C > G. The likeliest explanation for these observations is that the tRNAs that decode the A-site codons vary in their ability to frameshift. It is not known whether this is a result of the presence of unusual "shifty" tRNAs in the reticulocyte lysate, or whether the effect is simply due to differing strengths of interaction of the anticodons of normal tRNAs with the A-site codons. This is considered in the Discussion.

(iii) *Unusual P-site combinations*

Although most naturally occurring slip-sequences contain two homopolymeric triplets, there are three examples (to date) that do not conform to this organization: the slip sequence at the MMTV *pro-pol* overlap (GGAUUUA; Moore *et al.*, 1987; Jacks *et al.*, 1987), the related sequence of the red clover necrotic mosaic virus (RCNMV) RNA-1 p27-*pol* overlap (GGAUUUU; Xiong & Lommel, 1989) and that of the coronavirus-like equine arteritis virus (EAV) *1a-1b* overlap (GUUAAAC; den Boon *et al.*, 1991). It is possible to accommodate the MMTV *pro-pol* and

Table 2

The −1 frameshift efficiencies in the analysis shown in Figure 7

Construct	Slip-sequence	P-site codon	Frameshift efficiency (%)†	Post-slippage contacts‡
A. *P-site variants*				
5·72	G GGU UUA	Gly	16·1	GGG[1] ·· CCA[2]
5·49	G GAU UUA	Asp	15·5 (96)	GGA ·· CUA
5·32	G GGA AAC	Gly	32·5	GGG ··· CCU
5·51	G GAA AAC	Glu	22·3 (69)	GGA ··· CUU
5	U UUA AAC	Leu	41·7	UUU ·· AAU
5·29	G UUA AAC	Leu	35·1 (84)	GUU · AAU
5·55	G UCA AAC	Ser	13·8 (−)	GUC · AGU
B. *Position 1 variants*				
5·6	A AAA AAC	Lys	19·0 (100)	AAA ··· UUU
5·65	C AAA AAC	Lys	1·4 (7)	CAA ·· UUU
5·59	G AAA AAC	Lys	5·8 (31)	GAA ·· UUU
5·62	U AAA AAC	Lys	<1 (<5)	UAA ·· UUU
5·35	A AAU UUC	Asn	17·1 (100)	AAA ·· UUA
5·64	C AAU UUC	Asn	4·4 (26)	CAA · UUA
5·63	G AAU UUC	Asn	10·2 (60)	GAA ·· UUA
5·60	U AAU UUC	Asn	5·5 (32)	UAA · UUA
5.58	A GGA AAC	Gly	1·3 (4)	AGG ·· CCU
5·61	C GGA AAC	Gly	<1 (<5)	CGG ·· CCU
5·32	G GGA AAC	Gly	32·5 (100)	GGG ··· CCU
5·57	U GGA AAC	Gly	2·1 (6)	UGG ·· CCU
5·30	A UUA AAC	Leu	14·3 (34)	AUU · AAU
5·53	C UUA AAC	Leu	5·6 (13)	CUU · AAU
5·29	G UUA AAC	Leu	35·1 (84)	GUU · AAU
5	U UUA AAC	Leu	41·7 (100)	UUU ·· AAU

Table 2 *continued*

Construct	Slip-sequence	P-site codon	Frameshift efficiency (%)†	Post-slippage contacts‡
5·69	A UUU UUC	Phe	8·6 (32)	AUU · · · AAA
5·68	C UUU UUC	Phe	4·3 (16)	CUU · · · AAA
5·70	G UUU UUC	Phe	13·7 (51)	GUU · · · AAA
5·4	U UUU UUC	Phe	27 (100)	UUU · · · AAA

Each value is the average frameshift efficiency measured from the translation of 2 or more dilutions of the relevant RNA (see Materials and Methods). The variation in the efficiency measured for each dilution was 5% or less, i.e. a value of 12·1 represents a frameshift efficiency of between approximately 11·5 and 12·7%.

†Figures in parentheses represent the relative efficiency of frameshifting expressed as a proportion of the efficiency of the parental construct.

‡The predicted base-pair contacts formed between the mRNA (1) and the tRNA anticodon (2) are shown for each slip-sequence tested. The anti-codon sequences shown are based on standard Watson–Crick base-pairing; modified nucleosides in the anticodon loops are not shown. The minus sign adjacent to the frameshift efficiency of construct 5·55 (part A) indicates that we cannot express this efficiency as a proportion of the parental efficiency, since 5·55 is a double mutant and cannot be assigned to a particular parent.

RCNMV p*27-pol* slip-sequences within the simultaneous-slip model by invoking the formation of a post-slippage G·U mRNA–tRNA base-pair at the non-wobble position 2 in the P-site following a simultaneous-slip (see Table 2A). Indeed, when we changed the IBV slip-sequence to GGAUUUA (MMTV) or the related sequence GGAAAAC, some 15% and 22% of the ribosomes frameshifted at those sites respectively, supporting the idea that a G·U pair can be tolerated at this position (see Fig. 7 and Table 2). The EAV slip sequence (GUUAAAC), however, is more problematical, in that the simultaneous slip model would predict the formation of an A·U pair at position 2 and a non-Watson–Crick G-A mismatch at the non-wobble position 1 in the P-site. Nevertheless, this slip sequence has been shown to specify a 20% frameshift in the context of an RNA pseudoknot *in vivo* (den Boon *et al.*, 1991), suggesting that either the G·A mismatched pair can contribute to post-slippage tRNA–mRNA stability or that, under certain circumstances, stable post-slippage pairing is not required. When we tested this slip sequence in the context of the IBV pseudoknot, frameshifting occurred at high efficiency (35%: Table 2A, Fig. 7). We further investigated which nucleotide could be tolerated at position 1 of the slip sequence by completing a series of mutants in which the IBV sequence was modified at this position (NUUAAAC) and then prepared a series of other slip sequences in which the first position nucleotide was varied against a background of different first slippery codons (NAAAAAC, NAAUUUC, NGGAAAC, NUUUUUC; zero frame, P-site codon underlined). The sites were chosen such that the

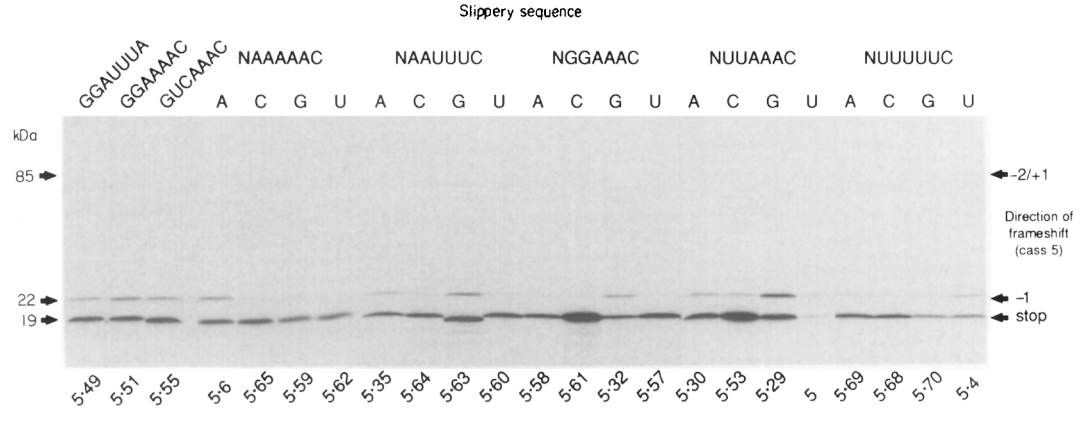

Figure 7. Unusual P-site combinations. In this Figure, the frameshift properties of a number of slip-sequence mutants are shown. In most of these, the wild-type IBV slip-sequence was replaced by a candidate sequence, e.g. GGAUUUA. For each class of sequence, N represents either A, C, G or U at the 1st position of the slip-sequence. The gel shows the reticulocyte lysate translation products synthesized in response to mRNAs derived from *Bam*HI-digested pFScass5 or mutant derivatives. Polypeptides were labelled and analysed as described in the legend to Fig. 5.

class of tRNA decoding the slip-sequence at the P-site varied in each case, since the observation of highly efficient frameshifting at the EAV slip sequence raised the possibility that the ability to slip efficiently in the P-site, yet mispair at position 1 may be related to the type of tRNA in the P-site.

The *in vitro* translations of the position 1 mutants are shown in Figure 7. In Table 2B, the frameshift efficiency measured for each construct, the relative efficiency expressed as a proportion of the efficiency of the parental construct and the likely post-slippage mRNA–tRNA contacts formed are shown. It is clear from the results that, in most cases, a mutation in position 1 greatly reduces the frameshift efficiency and this supports the view that, in general, stable post-slippage tRNA–mRNA pairing is important in the frameshift process. However, a number of the position 1 mutants retained the ability to specify efficient frameshifting, namely GAAUUUC (60% of the parental level), AUUAAAC (34%), GUUAAAC (84%) and GUUUUUC (51%). The overall efficiency of frameshifting within each group did depend to some extent on the P-site tRNA; UUA and UUU-decoding tRNAs were, in general, more slippy than AAA and GGA-decoding tRNAs. However, it was conspicuous that within each group the most efficient mutant sites were those in which a G nucleotide was present at position 1; whether forming a G·U or a G·A pair post-slippage. The explanation for this observation is at present unclear. In a further experiment, we tested the sequence GUCAAAC, which ought to retain, in terms of standard Watson–Crick base-pairing, following slippage, a sole U·G base-pair at position 2 in the P-site. This construct (see Fig. 7 and Table 2A) displayed a frameshift efficiency of 14%, despite the apparent paucity of pairing in the P-site. Frameshifting can thus occur at high efficiency under certain circumstances, with minimal base-pairing in the P-site.

(iv) *Frameshift direction*

In the majority of slip-sequence variants tested in this study, frameshifting was strictly into the −1 reading frame. However, for a number of sites, an 85 kDa product was seen, and this corresponds to ribosomes shifting into the −2 reading frame (or the +1 frame; we cannot, at present, distinguish between the 2 possibilities). In all the cases in which this product appeared, the slip sequence contained a U triplet in the A site (i.e. N_3U_3N) or was fully monotonous (N_7). The abundance of the −2/+1 frameshift product was, in general, extremely low (efficiency less than 0·5%). However, when monotonous U stretches (U_6N) were tested, significant levels (2% to 5%) of −2/+1 frameshifting were observed. Indeed in the U_7 construct (pFScass5.10), which was a stretch of eight U bases (since the codon immediately downstream from the slip sequence was the zero-frame termination codon, UGA), frameshifting was remarkably efficient, with some 39% of ribosomes entering the −1 phase, and

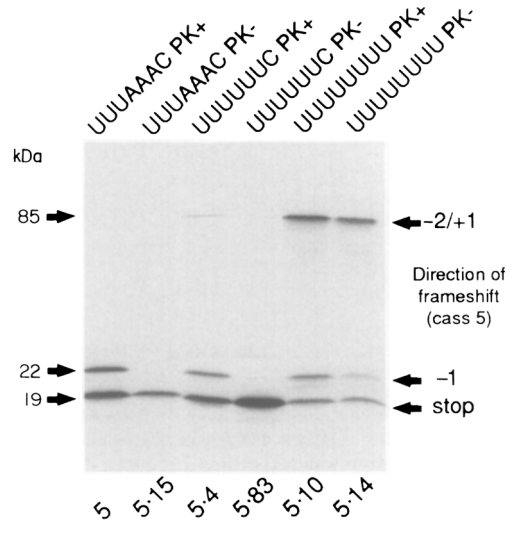

Figure 8. Effect of pseudoknot destabilization on frameshift direction. The gel shows the reticulocyte lysate translation products synthesized in response to mRNAs derived from *Bam*HI-digested pFScass5 or mutant derivatives. The slippery sequence present in each construct, and the presence (PK^+) or absence (PK^-) of a pseudoknot in these constructs is indicated above each track. In the PK^- constructs, the pseudoknot was destabilized by converting 4 C nucleotides at the base of stem 1 of the pseudoknot to G residues (as in construct 5·15, see Fig. 5). The polypeptides produced by each mutant were labelled and analysed as described in the legend to Fig. 5.

a further 21%, the −2/+1 frame. We wished to test whether −2/+1 frameshifting was sensitive to the presence of the RNA pseudoknot and so made two further constructs in which the pseudoknot structure downstream from the U_6C and U_8 constructs was destabilized by making a complementary change in stem 1 (to create pFScass5.83 (U_6C, PK^-) and pFScass5.14 (U_8, PK^-, see Fig. 8); the complementary change in stem 1 was identical with that made in pFScass5.15, see Fig. 5). Removal of the pseudoknot in the U_6C construct greatly reduced the level of both −1 (from 27% to 2%) and −2/+1 (from 2·2% to less than 0·5%) frameshifting; raising the possibility that both classes of frameshift are dependent upon the RNA pseudoknot. However, removal of the pseudoknot in the U_8 construct had a much less dramatic effect on frameshifting; with the −1 level reduced by about half (to 21·7%) and the −2/+1 level only marginally (from 21% to 20%). Frameshifting on the U_8 stretch therefore appears to be relatively independent of downstream RNA structure. Runs of U bases in coding sequences have been implicated in frameshift events in other systems, including the yeast mitochondrial *oxi*1 gene (Fox & Weiss-Brummer, 1980), bacteriophage T7 gene 10 (Dunn & Studier, 1983) and the trpE gene of *Salmonella typhimurium* (Atkins *et al.*, 1983). The high levels of frameshifting associated with the U_8 sequence suggests that when ribosomes encounter this stretch the ability to monitor reading frame is lost, since they are distributed into each of the three

reading frames at high efficiency (39%, −1; 21%, −2/+1; 40%, zero frame). Furthermore, even when the downstream pseudoknot was destabilized, reading frame maintenance was still aberrant (22%, −1; 20%, −2/+1; 58%, zero frame). Why the two sequences respond in different ways to the removal of the pseudoknot is not clear, but possible explanations are given in the Discussion. Frameshifting at the *gag-pol* overlap in HIV-1 occurs at a U_6A sequence and evidence has accumulated that, unlike most frameshift sites, no downstream RNA structures appear to be involved in frameshifting in this system (Wilson *et al.*, 1988; Madhani *et al.*, 1988). The possibility that a −2/+1 frameshift product could be produced from this sequence *in vivo* has been raised (Kingsman *et al.*, 1990). Our analysis of the U_6C construct described here suggests that such a frameshift product would be of low abundance.

Discussion

(a) *The A-site slippery codon: tRNAs and ribosomal frameshifting*

In the reticulocyte lysate system, and in the context of the IBV pseudoknot, only six codons are functional at the A-site, namely AAN (where N = A, C or U) and UUN (where N = A, C or U). Triplets of G or C and codons ending in G gave very low levels of frameshifting at this position. It is clear from this study that the IBV pseudoknot can act as an enhancer for the frameshift process, and appears to be considerably more effective than the pseudoknots predicted to form downstream from the RSV (Jacks *et al.*, 1988*b*) and L-A (Dinman *et al.*, 1991) slip sites. When the slip-sequences of RSV (A_3U_3A) and L-A (G_3U_3A) were placed upstream from the IBV pseudoknot, frameshifting at these sites was at a four- and ninefold greater level, respectively, than that seen in their natural contexts. The enhanced frameshifting afforded by the IBV pseudoknot allowed the determination of a hierarchy of frameshift efficiency of A-site codons. In order to explain this hierarchy, we propose that slip sequences are decoded by the normal pool of cellular tRNAs and the ability to slip at a particular codon is determined by the strength of the codon–anticodon interaction in the ribosomal A-site before tRNA slippage. We suggest that if this interaction is relatively weak, then slippage is more likely to occur. The ability to form two or more post-slippage mRNA–tRNA contacts is taken as a prerequisite for frameshifting at the A-site. The strength of the interaction between mRNA and tRNA is likely to be influenced considerably by the kind of base-pair that forms between the 3′ base of the codon and the 5′ base of the anticodon (position 34) at the wobble position (Crick, 1966; for a review, see Björk *et al.*, 1987). On this basis, we have surveyed the nature of the anticodons of the tRNAs predicted to be decoding the second slippery codon. We limited the analysis to tRNAs of non-transformed mammalian cells, excluding mitochondrial tRNAs and form our arguments on the basis of only those tRNAs whose sequence and modification pattern have been determined. As a source of information, we used the tRNA sequence compilations assembled by Sprinzl *et al.* (1989) and the EMBL and NEWEMBL databases.

For the AAN hierarchy, AAA and AAG are decoded by $tRNA^{Lys}$; AAC and AAU by $tRNA^{Asn}$. In mammalian cells, two lysine-acceptor tRNAs are present, $tRNA^{Lys}$ U^9UU (anticodon 5′ U^9UU 3′; U^9 = 5-methoxycarbonylmethyl-2-thiouridine) and $tRNA^{Lys}$ CUU (Raba *et al.*, 1979). On first inspection, one may expect $tRNA^{Lys}$ U^9UU to recognize both the AAA and the AAG lysine codons, and $tRNA^{Lys}$ CUU to recognize only AAG. However, the U^9 modification influences the codon preference of $tRNA^{Lys}$ U^9UU such that it preferentially recognizes the AAA codon with little affinity for the AAG codon (Lustig *et al.*, 1981). Thus, it is likely that the AAA and AAG codons are decoded by a separate tRNA. As the AAG·CUU codon–anticodon pairing is likely to be more stable than the AAA·UUU pairing, then, according to our hypothesis, AAG should frameshift poorly, since it forms strong pre-slip contacts with the tRNA. This is supported by the lack of frameshifting seen in slip-sequences ending with AAG. The AAA codon, however, would form a weaker pre-slip contact and, as is seen, may be expected to frameshift at a higher level. In the case of $tRNA^{Asn}$, which decodes AAC and AAU, only one tRNA species has been detected to date; $tRNA^{Asn}$ (anticodon 5′ QUU 3′; Q = queuosine: Chen & Roe, 1978, 1980; Roe *et al.*, 1979). The Q modification appears to influence codon choice in such a way that $tRNA^{Asn}$ QUU recognizes both AAC and AAU, with a slight preference for AAU (Meier *et al.*, 1985). If the wobble base-pairs between Q-C and Q-U are less strong than a G·C pair, then one may expect both AAC and AAU to be slippery codons. As $tRNA^{Asn}$ QUU has a slight preference for AAU, one would expect AAC to be more slippery, and this correlates with the experimental observation. Thus, for the AAN series, there is some support for the observed hierarchy C > U ~ A ≫ G, based on the strength of the pre-slip codon–anticodon pairing.

For the UUN series, UUA and UUG are decoded by $tRNA^{Leu}$; UUC and UUU by $tRNA^{Phe}$. Of the mammalian leucine tRNAs described to date, there are two species that are likely to decode UUA and UUG; namely, $tRNA^{Leu}$ (IAG; I = inosine: Pirtle *et al.*, 1980) and $tRNA^{Leu}$ (CAG: Sprinzl & Gauss, 1983). The presence of inosine at the wobble position allows the formation of I·A, I·C and I·U but not I·G base-pairs. Thus, $tRNA^{Leu}$ IAG recognizes the UUA but not the UUG codon. The isoacceptor $tRNA^{Leu}$ CAG will recognize only the UUG codon and may be expected to form a strong C·G wobble pair and not to frameshift efficiently. The I·A wobble pair formed upon decoding UUA by $tRNA^{Leu}$ IAG would be weaker than the C·G pair and thus one would expect UUA to frameshift at a higher efficiency than UUG, as is seen. In the case of $tRNA^{Phe}$, which decodes UUC and UUU, only one

tRNA species has been detected to date, $tRNA^{Phe}$ (G^3AA; G^3 = 2′-*o*-methylguanosine: Roe *et al.*, 1975; Keith & Dirheimer, 1978; Lin *et al.*, 1980). Unfortunately, the effects on codon choice of this modification (if any) have not been studied, so it is not possible to comment on the likely affinity of this tRNA for the UUC and UUU codons. If the frameshift hypothesis is correct, however, one would expect this tRNA to recognize both UUC and UUU with a slight preference for UUC. However, the G^3·C and G^3·U pairs would be predicted to be less strong than an unmodified G·C pair, since both UUC and UUU codons allow efficient frameshifting.

The hypothesis outlined above is drawn from subject areas in which our understanding of the precise details is incomplete, particularly with respect to the energetics of base-pair formation between modified nucleosides in tRNA anticodons and mRNA codons. Furthermore, the effects of modification on codon choice described are based on the study of only a limited number of transfer RNAs. Nevertheless, there is some additional support for the hypothesis from studies of "retrovirus-like" frameshifting in *E. coli*. Weiss *et al.* (1989) demonstrated that the MMTV *gag-pro* frameshift signal (slip-sequence A_6C) worked poorly in *E. coli* (frameshift efficiency of 1·5%) but a point mutation that converted the slippery sequence to A_6G increased the efficiency to over 50%. Tsuchihashi (1991) has proposed an explanation for this observation based on the knowledge that only one $tRNA^{Lys}$ species has been reported in *E. coli*; $tRNA^{Lys}$ 5′ U^8UU 3′ (U^8 = 5-methylaminomethyl-2-thiouridine: Chakraburtty *et al.*, 1975). Tsuchihashi (1991) suggests that as the U^8 modification appears to weaken the interaction between this tRNA and the AAG codon (Yokoyama *et al.*, 1985), then the recognition of AAG may be poor and the $tRNA^{Lys}$ able to slip at high efficiency. In *E. coli*, therefore, AAG, in contrast to the situation in eukaryotes, is a very slippery codon, and this may well be a consequence of having only a single $tRNA^{Lys}$ species. In eukaryotes, the presence of the $tRNA^{Lys}$ CUU isoacceptor may well prevent frameshifting at the AAG codon in the context of "retrovirus-type" frameshift signal.

Hatfield and colleagues have proposed an alternative explanation for the limited number of codons known to be functional in frameshifting, in which frameshifting is mediated by a subset of hypomodified isoacceptor tRNAs (Hatfield *et al.*, 1990). The tRNAs that would decode the three A-site codons found in naturally occurring slip-sequences are $tRNA^{Asn}$ (codon AAC), $tRNA^{Phe}$ (UUU) and $tRNA^{Leu}$ (UUA). These tRNAs are characterized by the fact that $tRNA^{Asn}$ contains the highly modified Q base in the 5′ position of its anticodon, $tRNA^{Phe}$ the highly modified Wyebutosine (Y1) or Wyebutoxosine (Y2) base (Sprinzl *et al.* 1989) just to the 3′ side of the anticodon, and $tRNA^{Leu}$ lacks a highly modified base in its anticodon loop. The hypothesis suggests that hypomodified variants of $tRNA^{Asn}$ and $tRNA^{Phe}$ (and the already "hypo"-modified $tRNA^{Leu}$) would be able to slip at frameshift sites, since the lack of a highly modified base would create more space in and around the frameshift site, which may facilitate frameshifting by allowing greater flexibility of movement of the anticodon (Hatfield *et al.*, 1989; Hatfield & Oroszlan, 1990). Experimental support for the idea that hypomodified variants of $tRNA^{Asn}$ and $tRNA^{Phe}$ could be involved in the frameshift process comes from an analysis of the modification status of these tRNAs in cells infected with various retroviruses (Hatfield *et al.*, 1989). These studies have indicated that the relevant tRNA (Phe or Asn) is largely hypomodified in infected cells compared with uninfected cells. Whether this indicates an important role for hypomodified tRNAs in the frameshift process *per se* is uncertain, since it is known that ribosomal frameshift signals can work in a fairly wide range of uninfected cell types. If hypomodified tRNAs are involved in frameshifting, then one would predict that this would influence codon preference. For example, hypomodified $tRNA^{Asn}$ GUU (hypermodified anticodon is QUU) would be expected to prefer the AAC codon rather than AAU (Meier *et al.*, 1985). Similarly, unmodified $tRNA^{Lys}$ UUU (modified anticodon is U^9UU) would be expected to recognize the AAA codon in preference to AAG (Lustig *et al.*, 1981) but to still recognize both codons. A hierarchy of frameshifting similar to that demonstrated experimentally here could be predicted from these changed preferences, but only if frameshifting depends upon a strong pre-slippage tRNA–mRNA pair. The idea that strong pre-slip base-pairing is important, however, is not consistent with the high levels of frameshifting seen with the AAG codon in *E. coli*, since this would be decoded by $tRNA^{Lys}$ U^8UU (or UUU if unmodified), which would be expected to recognize the AAA codon strongly, and AAG weakly. Thus, we prefer the hypothesis that normal tRNAs are involved in decoding this class of frameshift site.

(b) *The P-site slippery codon*

Almost all of the mutants tested in the P-site analysis described here were concerned with position 1 of the slip-sequence, since we were intrigued by the observation that the slip-sequence of EAV (GUUAAAC) was functional *in vivo* (den Boon *et al.*, 1991). In our assay, this sequence stimulated frameshifting at a level of 35%, only slightly less efficient than the wild-type IBV sequence (UUUAAAC). This raised the possibility that the predicted post-slippage base-pairs, G-A at position 1 and U·A at position 2, may be sufficiently stable to allow frameshifting. An alternative possibility was that the leucine tRNA decoding UUA had some unique property that facilitated slippage without the formation of strong post-slip contacts with the mRNA. To test this, we prepared a number of position 1 variants of a range of slip-sequences and tested frameshifting. It was found that in most cases, and in agreement with the necessity for retention of the

homopolymeric nature of the slip-sequence, frameshifting was greatly reduced. In four of the mutants, however, significant frameshifting was observed (from 35% to 85% of the parental level). Three of these had G at position 1 (GUUAAAC, GUUUUUC, GAAUUUC) and it was clear that in all the mutants tested, the presence of G at the first position conferred the highest frameshift level within each class. The fourth functional position 1 mutant was AUUAAAC. We have been unable to arrive at a satisfactory explanation for these observations. One possibility is that during a frameshift, unusual base-pairs can form and are stable at position 1. Current models for the frameshift process suggest that the RNA structures downstream from the slip-site may function to stall ribosomes over the slippery sequence codons; this pause could directly influence the fidelity of mRNA decoding, and promote the realignment of ribosome-bound tRNAs into the −1 reading frame (Jacks *et al.*, 1988*a*,*b*). Thus, the post-slip G·G, G·A and A·A pairs that may form in the functional position 1 mutants (see Table 2) could be allowed under the circumstances of a paused ribosome. Surprisingly, the sequence GUCAAAC was also found to stimulate reasonable levels of frameshifting (14%), despite being expected to form only a G·A mismatch at position 1 and a U·G pair at position 2 post-slippage (see Table 2A). Thus, certain sequences do not appear to adhere strongly to the two-out-of-three post-slip contacts code of the simultaneous slip model. Clearly, more information is needed before the molecular basis of this P-site misreading is understood, particularly with respect to the functionality of G-nucleotides at position 1.

(c) *Frameshift direction*

In our analysis of the direction of frameshifting, we have found that at the majority of slip sequences, frameshifting was strictly into the −1 reading frame. Significant quantities of −2/+1 product, however, were seen when fully homopolymeric U runs (U_6N) were tested (2 to 5%) and a U_8 stretch gave high levels (21%). An important question concerning the frameshift process is whether RNA pseudoknots can influence frameshift direction. In the simplest case, RNA pseudoknots may function solely to pause ribosomes; such ribosomes can then frameshift in a direction determined by the nature of the slippery sequence. In the majority of frameshift sites, the likeliest slip is a −1 slip, since this produces the most stable post-slippage mRNA–tRNA base-pairing. In the case of homopolymeric runs, identical post-slip contacts can form following −1 or −2 slips (e.g. in the case of the U_6C sequence) or even −1, −2 or +1 slips (e.g. in the case of the U_8 sequence). If the pseudoknot plays a role in determining the direction of frameshifting, however, then in the presence of a pseudoknot one would expect a particular class of frameshift to predominate. We investigated this possibility by destabilizing the pseudoknot structure in constructs containing a U_6C or U_8 slip-sequence and tested for frameshifting (see Fig. 8). In the U_6C construct, this resulted in a reduction of both −1 and −2/+1 frameshifting. In the U_8 construct, however, removal of the pseudoknot had little effect on −2/+1 frameshifting, but reduced −1 frameshifting by about half. These apparently inconsistent effects can be explained if the pseudoknot is able to specifically enhance 5′-wards slippage (minus direction). In the U_6C construct, the 85 kDa −2/+1 product probably arises from a −2 slip rather than a +1 frameshift, since the phenylalanine tRNA decoding the A-site codon can form identical post-slip contacts with the mRNA following either a −1 or −2 slip, but not from a +1 slip. If the pseudoknot does enhance frameshifting in the 5′ direction, then the reduction in both −1 and −2/(+1) frameshifting observed in the U_6C PK^- construct is perhaps expected. In the case of the U_8 construct, destabilization of the pseudoknot once again caused a reduction in −1 frameshifting but, in this case, little effect on −2/+1 frameshifting was seen. This may be related to the likelihood that both forward and backward slips can occur at the U_8 sequence; perhaps in the absence of a pseudoknot, +1 frameshifts occur, whereas in the presence of the structure, frameshifting may occur mainly into the −2 (and −1) reading frame. As we are unable to distinguish between −2 and +1 shifts in this assay, these interpretations must be viewed as speculative, but the results lend some support to the idea that RNA pseudoknots can influence the direction of frameshifting.

In summary, our investigation into the magnitude and directionality of ribosomal frameshifting described here offers an explanation for the observation that many natural slip-sequences employ imperfect tandem slippery codons (e.g. IBV, UUUAAAC; RSV AAAUUUA; HIV-1; UUUUUUA). Firstly, for those sequences ending in AAN, a perfect second slippery codon (AAA) would frameshift less well; hence all natural sites of this class utilize the AAC codon (ten Dam *et al.*, 1990). Indeed, recent work with the MMTV frameshift signal has shown that AAC is the optimal codon (Chamorro *et al.*, 1992). Secondly, in the case of those sequences ending with UUN, although perfecting the triplet to UUU may increase the frameshift efficiency, it may also start to generate undesirable −2/+1 frameshifts.

We thank the Agricultural and Food Research Council and the Medical Research Council for supporting this investigation. We are grateful to Nicky Rolley for her assistance in the early part of this work. Many thanks to Dolph Hatfield for communicating information before publication.

References

Atkins, J. F., Nichols, B. P. & Thompson, S. (1983). The nucleotide sequence of the first externally suppressible −1 frameshift mutant and of some nearby leaky frameshift mutants. *EMBO J.* **2**, 1345–1350.

Atkins, J. F., Weiss, R. B. and Gesteland, R. F. (1990). Ribosome gymnastics–degree of difficulty 9·5, style 10·0. *Cell*, **62**, 413–423.

Birnboim, H. C. & Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucl. Acids Res.* **7**, 1513–1523.

Björk, G. R., Ericson, J. U., Gustafsson, C. E. D., Hagervall, T. G., Jönsson, Y. H. & Wikström, P. M. (1987). Transfer RNA modification. *Annu. Rev. Biochem.* **56**, 263–287.

Boursnell, M. E. G., Brown, T. D. K., Foulds, I. J., Green, P. F., Tomley, F. M. & Binns, M. M. (1987). Completion of the sequence of the genome of the coronavirus avian infectious bronchitis virus. *J. Gen. Virol.* **68**, 57–77.

Brierley, I., Boursnell, M. E. G., Binns, M. M., Bilimoria, B., Blok, V. C., Brown, T. D. K. & Inglis, S. C. (1987). An efficient ribosomal frame-shifting signal in the polymerase-encoding region of the coronavirus IBV. *EMBO J.* **6**, 3779–3785.

Brierley, I., Digard, P. & Inglis, S. C. (1989). Characterisation of an efficient coronavirus ribosomal frameshifting signal: requirement for an RNA pseudoknot. *Cell*, **57**, 537–547.

Brierley, I., Rolley, N. J., Jenner, A. J. & Inglis, S. C. (1991). Mutational analysis of the RNA pseudoknot component of a coronavirus ribosomal frameshifting signal. *J. Mol. Biol.* **220**, 889–902.

Chakraburtty, K., Steinschneider, A., Case, R. V. & Mehler, A. H. (1975). Primary structure of $tRNA^{Lys}$ of *E. coli* B. *Nucl. Acids Res.* **2**, 2069–2075.

Chamorro, M., Parkin, N. & Varmus, H. E. (1992). An RNA pseudoknot and an optimal heptameric shift site are required for highly efficient ribosomal frameshifting on a retroviral messenger RNA. *Proc. Nat. Acad. Sci., U.S.A.* **89**, 713–717.

Chen, E. Y & Roe, B. A. (1978). The nucleotide sequence of rat liver $tRNA^{Asn}$. *Biochem. Biophys. Res. Commun.* **82**, 235–246.

Chen, E. Y. & Roe, B. A. (1980). Structural comparison of human, bovine, rat and Walker 256 carcinosarcoma asparaginyl-tRNA. *Biochim. Biophys. Acta.* **610**, 272–284.

Crick, F. H. C. (1966). Codon-anticodon pairing. The wobble hypothesis. *J. Mol. Biol.* **19**, 548–555.

den Boon, J. A., Snijder, E. J., Chirnside, E. D., de Vries, A. A. F., Horzinek, M. C. & Spaan, W. J. M. (1991). Equine arteritis virus is not a togavirus but belongs to the coronaviruslike superfamily. *J. Virol.* **65**, 2910–2920.

Dente, L., Cesareni, G. & Cortese, R. (1983). pEMBL: a new family of single-stranded plasmids. *Nucl. Acids Res.* **11**, 1645–1655.

Digard, P., Blok, V. C. & Inglis, S. C. (1989). Complex formation between influenza virus polymerase proteins expressed in *Xenopus* oocytes. *Virology*, **171**, 162–169.

Dinman, J. D., Icho, T. and Wickner, R. B. (1991). A −1 ribosomal frameshift in a double stranded RNA virus of yeast forms a gag-pol fusion protein. *Proc. Nat. Acad. Sci., U.S.A.* **88**, 174–178.

Dotto, G. P., Enea, V. & Zinder, N.D. (1981). Functional analysis of bacteriophage f1 intergenic region. *Virology*, **114**, 463–473.

Dunn, J. J. & Studier, F. W. (1983). Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. *J. Mol. Biol.* **166**, 477–535.

Fox, T. D. & Weiss-Brummer, B. (1980). Leaky +1 and −1 frameshift mutations at the same site in a yeast mitochondrial gene. *Nature (London)*, **288**, 60–63.

Hames, B. D. (1981). An introduction to polyacrylamide gel electrophoresis. In *Gel Electrophoresis of Proteins—A Practical Approach* (Hames, B. D. & Rickwood D., eds), pp. 1–91, IRL Press, Oxford.

Hatfield, D. & Oroszlan, S. (1990). The *where, what* and *how* of ribosomal frameshifting in retroviral protein synthesis. *Trends Biochem. Sci.* **15**, 186–190.

Hatfield, D., Feng, Y-X., Lee, B. J., Rein, A., Levin, J. G. & Oroszlan, S. (1989). Chromatographic analysis of the aminoacyl-tRNAs which are required for translation of codons at and around the ribosomal frameshift sites of HIV, HTLV−1 and BLV. *Virology*, **173**, 736–742.

Hatfield, D., Smith, D. W. E., Lee, B. J., Worland, P. J. & Oroszlan, S. (1990). Structure and function of suppressor tRNAs in higher eukaryotes. *Crit. Rev. Biochem. Mol. Biol.* **25**, 71–96.

Hizi, A., Henderson, L. E., Copeland, T. D., Sowden, R. C., Hixson, C. V. and Oroszlan, S. (1987). Characterisation of mouse mammary tumour virus *gag-pol* gene products and the ribosomal frameshift by protein sequencing. *Proc. Nat. Acad. Sci., U.S.A.* **84**, 7041–7046.

Honigman, A., Wolf, D., Yaish, S., Falk, H. & Panet, A. (1991). *cis* acting RNA sequences control the gag-pol translation readthrough in murine leukaemia virus. *Virology*, **183**, 313–319.

Inglis, S. C., Rolley, N. & Brierley, I. (1990). The ribosomal frameshift signal of infectious bronchitis virus. In *Post-transcriptional Control of Gene Expression* (McCarthy, J. E. G. & Tuite, M. F., eds) pp. 603–610, Nato ASI series H: Cell Biology, vol. 49, Springer-Verlag, Berlin and Heidelberg.

Jacks, T. & Varmus, H. E. (1985). Expression of the Rous sarcoma virus *pol* gene by ribosomal frameshifting. *Science*, **230**, 1237–1242.

Jacks, T., Townsley, K., Varmus, H. E. & Majors, J. (1987). Two efficient ribosomal frameshifting events are required for synthesis of mouse mammary tumor virus *gag*-related polyproteins. *Proc. Nat. Acad. Sci., U.S.A.* **84**, 4298–4302.

Jacks, T., Power, M. D., Masiarz, F. R., Luciw, P. A., Barr, P. J. and Varmus, H.E. (1988*a*). Characterisation of ribosomal frameshifting in HIV−1 *gag-pol* expression. *Nature (London)*, **331**, 280–283.

Jacks, T., Madhani, H. D., Masiarz, F. R. & Varmus, H. E. (1988*b*). Signals for ribosomal frameshifting in the Rous sarcoma virus *gag-pol* region. *Cell*, **55**, 447–458.

Keith, G. & Dirheimer, G. (1978). The primary structure of rabbit, calf and bovine liver $tRNA^{Phe}$. *Biochim. Biophys. Acta*, **517**, 133–149.

Kingsman, A. J., Wilson, W. and Kingsman, S. M. (1990). HIV *pol* expression via a ribosomal frameshift. In *Post-transcriptional Control of Gene Expression*. (McCarthy, J. E. G. & Tuite, M. F, eds) pp. 623–633. Nato ASI series H: Cell Biology; vol. 49. Springer-Verlag. Berlin and Heidelberg.

Krieg, P. A. & Melton, D. A. (1984). Functional messenger RNAs are produced by SP6 *in vitro* transcription of cloned cDNAs. *Nucl. Acids Res.* **12**, 7057–7071.

Kunkel, T. A. (1985). Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Nat. Acad. Sci. U.S.A.* **82**, 488–492.

Lin, F.-K., Furr, T. D., Chang, S. H., Horwitz, J., Agris, P. F. & Ortwerth, B. J. (1980). The nucleotide sequence of two bovine lens phenylalanine tRNAs. *J. Biol. Chem.* **255**, 6020–6023.

Lustig, F., Elias, P., Axberg, T., Samuelsson, T., Tittawella, I. & Lagerkvist, U. (1981). Codon reading and translational error. Reading of the glutamine and lysine codons during protein synthesis *in vitro*. *J. Biol. Chem.* **256**, 2635–2643.

Madhani, H. D., Jacks, T. & Varmus, H. E. (1988). Signals for the expression of the HIV *pol* gene by ribosomal frameshifting. In *The Control of HIV Gene Expression* (Cullen, B. & Wong-Staal, F., eds), pp. 119–125, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Maniatis, T., Fritsch, E. F. and Sambrook, J. (1982). *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Meier, F., Suter, B., Grosjean, H., Keith, G. & Kubli, E. (1985). Queuosine modification of the wobble base in $tRNA^{His}$ influences *'in vivo'* decoding properties. *EMBO J.* **4**, 823–827.

Moore, R., Dixon, M., Smith, R., Peters, G. & Dickson, C. (1987). Complete nucleotide sequence of a milk-transmitted mouse mammary tumor virus: Two frameshift suppression events required for translation of *gag* and *pol*. *J. Virol.* **61**, 480–490.

Pirtle, R., Kashdan, M., Pirtle, B. & Dudock, B. (1980). The nucleotide sequence of a major species of tRNA from bovine liver. *Nucl. Acids Res.* **8**, 805–815.

Pleij, C. W. A. and Bosch, L. (1989). RNA pseudoknots: Structure, detection and prediction. *Methods Enzymol.* **180**, 289–303.

Raba, M., Limburg, K., Burghagen, M., Katze, J. R., Simsek, M., Heckman, J. E., Rajbhandary, U. L. & Gross, H. J. (1979). Nucleotide sequence of three isoaccepting lysine tRNAs from rabbit liver and SV40-transformed mouse fibroblasts. *Eur. J. Biochem.* **97**, 305–318.

Roe, B. A., Anandaraj, M. P. J. S., Chia, L. S. Y., Randerath, E., Gupta, R. C. & Randerath, K. (1975). Sequence studies on $tRNA^{Phe}$ from human placenta: Comparisons with known sequences of $tRNA^{Phe}$ from other normal mammalian tissues. *Biochem. Biophys. Res. Commun.* **66**, 1097–1105.

Roe, B. A., Stankiewicz, A. F., Rizi, H. L., Weisz, C., DiLauro, M. N., Pike, D., Chen, C. Y. & Chen, E. Y. (1979). Comparison of rat liver and Walker 256 carcinosarcoma tRNAs. *Nucl. Acids Res.* **6**, 673–688.

Russel, M., Kidd, S. & Kelley, M.R. (1986). An improved filamentous helper phage for generating single-stranded plasmid DNA. *Gene*, **45**, 333–338.

Sanger, F., Nicklen, S. & Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Nat. Acad. Sci., U.S.A.* **74**, 5463–5467.

Sprinzl, M. & Gauss, D. H. (1983). Compilation of transfer RNA sequences and modified nucleosides in transfer RNA. In *The Modified Nucleosides of Transfer RNA, II* (Agris, P. F. & Kopper, R. A., eds), pp. 129–226, Alan R. Liss Inc., New York.

Sprinzl, M., Hartmann, T., Weber, J., Blank, J. & Zeidler, R. (1989). Compilation of tRNA sequences and sequences of tRNA genes. *Nucl. Acids Res.* **17** (suppl.), 1–173.

ten Dam, E. B., Pleij, C. W. A. & Bosch, L. (1990). RNA pseudoknots: Translational frameshifting and read through on viral RNAs. *Virus Genes* **4**, 121–136.

Tsuchihashi, Z. (1991). Translational frameshifting in the *Escherichia coli* dna X gene *in vitro*. *Nucl. Acids Res.* **19**, 2457–2462.

Turner, D. H., Sugimoto, N. & Freier, S. M. (1988). RNA structure prediction. *Annu. Rev. Biophys. Biophys. Chem.* **17**, 167–192.

Weiss, R. B., Dunn, D. M., Shuh, M., Atkins, J. F. and Gesteland, R. F. (1989). *E. coli* ribosomes rephase on retroviral frameshift signals at rates ranging from 2 to 50 percent. *The New Biologist*, **1**, 159–170.

Wills, N. M., Gesteland, R. F. & Atkins, J. F. (1991). Evidence that a downstream pseudoknot is required for translational readthrough of the Moloney murine leukaemia virus *gag* stop codon. *Proc. Nat. Acad. Sci., U.S.A.* **88**, 6991–6995.

Wilson, W., Braddock, M., Adams, S. E., Rathjen, P. D., Kingsman, S. M. & Kingsman, A. J. (1988). HIV expression strategies: ribosomal frameshifting is directed by a short sequence in both mammalian and yeast systems. *Cell*, **55**, 1159–1169.

Xiong, Z. & Lommel, S. A. (1989). The complete nucleotide sequence and genome organisation of red clover necrotic mosaic virus RNA-1. *Virology*, **171**, 543–554.

Yanisch-Perron, C., Vieira, J. & Messing, J. (1985). Improved M13 phage cloning vectors and host strains: Nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene*, **33**, 103–119.

Yokoyama, S., Watanabe, T., Murao, K., Ishikura, H., Yamaizumi, Z., Nishimura, S. & Miyazawa, T. (1985). Molecular mechanism of codon recognition by tRNA species with modified uridine in the first position of the anticodon. *Proc. Nat. Acad. Sci., U.S.A.* **82**, 4905–4909.

Young, J. F., Desselberger, U., Graves, P., Palese, P., Shatsman, A. & Rosenberg, M. (1983). Cloning and expression of influenza virus genes. In *The Origin of Pandemic Influenza Viruses* (Laver, W. G., ed.), pp. 129–138, Elsevier Science, Amsterdam.

Edited by M. Yaniv