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.

Article No. jmbi.1999.2689 available online at http://www.idealibrary.com on IDEAL J. Mol. Biol. (1999) 288, 321–335

JMB

Evidence for an RNA Pseudoknot Loop-Helix Interaction Essential for Efficient −1 Ribosomal Frameshifting

Jan Liphardt, Sawsan Napthine, Harry Kontos and Ian Brierley*

Division of Virology
Department of Pathology
University of Cambridge
Tennis Court Road, Cambridge
CB2 1QP, UK

*Corresponding author

RNA pseudoknots are structural elements that participate in a variety of biological processes. At −1 ribosomal frameshifting sites, several types of pseudoknot have been identified which differ in their organisation and functionality. The pseudoknot found in infectious bronchitis virus (IBV) is typical of those that possess a long stem 1 of 11-12 bp and a long loop 2 (30-164 nt). A second group of pseudoknots are distinguishable that contain stems of only 5 to 7 bp and shorter loops. The NMR structure of one such pseudoknot, that of mouse mammary tumor virus (MMTV), has revealed that it is kinked at the stem 1-stem 2 junction, and that this kinked conformation is essential for efficient frameshifting. We recently investigated the effect on frameshifting of modulating stem 1 length and stability in IBV-based pseudoknots, and found that a stem 1 with at least 11 bp was needed for efficient frameshifting. Here, we describe the sequence manipulations that are necessary to bypass the requirement for an 11 bp stem 1 and to convert a short non-functional IBV-derived pseudoknot into a highly efficient, kinked frameshifter pseudoknot. Simple insertion of an adenine residue at the stem 1-stem 2 junction (an essential feature of a kinked pseudoknot) was not sufficient to create a functional pseudoknot. An additional change was needed: efficient frameshifting was recovered only when the last nucleotide of loop 2 was changed from a G to an A. The requirement for an A at the end of loop 2 is consistent with a loop-helix contact similar to those described in other RNA tertiary structures. A mutational analysis of both partners of the proposed interaction, the loop 2 terminal adenine residue and two G·C pairs near the top of stem 1, revealed that the interaction was essential for efficient frameshifting. The specific requirement for a 3′-terminal A residue was lost when loop 2 was increased from 8 to 14 nt, suggesting that the loop-helix contact may be required only in those pseudoknots with a short loop 2.

Keywords: RNA pseudoknot; ribosomal frameshifting; loop-helix interaction; RNA structure probing; retrovirus

Introduction

RNA pseudoknots are structural elements that participate in a variety of biological processes. They are formed when residues in a hairpin-loop base-pair with nucleotides outside the loop, yielding two stems that are connected by single-stranded loops (Pleij *et al.*, 1985). Pseudoknots were first recognised experimentally from studies of the folding of the 3′-end of turnip yellow mosaic virus RNA (TYMV; Rietveld *et al.*, 1982) and since this time have been found in virtually all classes of RNA, including ribosomal RNAs, catalytic and self-splicing RNAs and messenger RNAs (reviewed by ten Dam *et al.*, 1992). Several pseudoknots have been described that play a role in protein biosynthesis. They are an essential component of the internal ribosome entry sites of some picornaviruses (Wang *et al.*, 1995; Rijnbrand *et al.*, 1997), are targets for the translational repression of certain ribosomal (Tang & Draper, 1989; Phillipe *et al.*,

Abbreviations used: 3D, three-dimensional; MD, molecular dynamics.

E-mail address of the corresponding author: ib103@mole.bio.cam.ac.uk

0022-2836/99/180321–15 $30.00/0

1990; Gluik *et al*., 1997; Benard *et al*., 1998) and bacteriophage mRNAs (Shamoo *et al*., 1993; Du & Hoffman, 1997), and play a pivotal role in the processes of termination codon suppression and ribosomal frameshifting in many viral and some cellular genes (reviewed by Farabaugh, 1996).

Programmed −1 ribosomal frameshifting is a mode of gene expression used primarily by retroviruses to produce structural and enzymatic proteins at a defined ratio from a single polycistronic messenger RNA. The signals that specify the frameshift event are composed of two essential elements, a heptanucleotide slippery sequence, located immediately upstream of a region of RNA secondary structure, most commonly a pseudoknot (Brierley *et al*., 1989; ten Dam *et al*., 1990). This stimulatory RNA structure is thought to function by modulation of the ribosomal elongation cycle (Dinman *et al*., 1997), perhaps during a ribosomal pause (Tu *et al*., 1992; Somogyi *et al*., 1993).

The role of pseudoknots in the frameshift process was first investigated using a signal derived from the genome of the avian coronavirus infectious bronchitis virus (IBV; Brierley *et al*., 1989, 1991). The IBV pseudoknot is of the hairpin-type (ten Dam *et al*., 1992) and is thought to possess coaxially stacked stems of 11 bp (stem 1, S1) and 6 bp (stem 2, S2) connected by single-stranded loops of 2 nt (loop 1, L1) and 32 nt (loop 2, L2). The structure of the IBV pseudoknot is representative of those pseudoknots that are present at the frameshift sites of the corona-, toro- and arteriviruses which possess a long S1 of 11-12 bp and usually a long L2 (30-164 nt). A second group of pseudoknots contain stems of 5 to 7 bp and shorter loops (ten Dam *et al*. 1990; Brierley, 1995). These pseudoknots are typified by those present at the frameshift sites of simian retrovirus 1 *gag-pro* (SRV-1; ten Dam *et al*., 1994) and the retrovirus mouse mammary tumor virus (MMTV) *gag-pro* (Chen *et al*., 1995). The three-dimensional (3D) structure of a functional variant of the MMTV pseudoknot (VPK) has been determined by NMR (Shen & Tinoco, 1995). This pseudoknot is bent at the S1-S2 junction with an intervening unpaired adenosine stacking between the two stems. The bent conformation is essential for efficient frameshifting (Shen & Tinoco, 1995; Chen *et al*., 1996; Kang & Tinoco, 1997) and may reflect a requirement for a structure-dependent interaction between a component of the translation apparatus and the pseudoknot.

At present, our understanding of the molecular basis of the frameshift process is rudimentary. The presence of two apparently distinct classes of RNA pseudoknot, however, provides the opportunity to gain mechanistic insights from a functional comparison of the two. On this basis, we recently investigated the effect on frameshifting of modulating S1 length and stability using a series of IBV-based pseudoknots as model systems (Napthine *et al*., 1999). We found that efficient frameshifting depended upon the presence of a minimum of 11 bp; pseudoknots with a shorter S1 (from 4 to 10 bp) were either non-functional or had a reduced frameshift efficiency. The inability of these shorter pseudoknots to stimulate frameshifting is not fully understood, but it seemed likely that those with an S1 length of 5 or 6 bp could be functionally restored by inclusion of an intercalating A residue between S1 and S2, creating a kinked pseudoknot similar to those discussed above.

Here, we describe the sequence manipulations that are required to convert such a short non-functional pseudoknot (pKA5, with an S1 of 6 bp) into a highly efficient kinked frameshifter pseudoknot. Surprisingly, introduction of an unpaired A residue at the junction between the two pseudoknot stems of pKA5 to create an MMTV-like structure was insufficient to produce an efficient frameshift signal. This was true even when the type and orientation of adjacent base-pairs was varied and the slippery sequence pseudoknot spacing distance was increased to the same length as that seen at the MMTV frameshift site (seven nucleotides). Based on a possible structural analogy with the TYMV pseudoknot (Kolk *et al*., 1998), we decided to change the last base of L2 from a G to an A residue. The TYMV structure, which is not associated with a frameshift signal, is stabilised by a loop to helix triple interaction between an adenine residue at the 3′-end of L2 and the minor groove of S1, and it was possible therefore that a similar interaction would be important in some frameshifter pseudoknots. To test this, a variant of pKA5 with an intercalated A residue was modified by changing the 3′-terminal base of L2 (the putative L2 donor base). The results were striking in that the presence of an A residue at the end of L2 generated a highly efficient frameshifter pseudoknot (KA-A, 31 %). This observation was consistent with a requirement for an interaction between the L2 terminal A and the minor groove of S1. Further support for this model was obtained from the translational properties of mutants with alterations in S1, the putative loop acceptor region. Reversing the orientation of two G·C base-pair near the top of S1, a change predicted to interfere with a loop-helix contact, decreased frameshifting to about 9 %. The specific requirement for an A at the end of L2 was not seen in pseudoknots with a longer L2, however, suggesting that the loop-helix contact may be important only in those pseudoknots with a short L2.

Results

Conversion of a non-functional short pseudoknot to a functional "kinked" pseudoknot

In our previous investigation of the influence of IBV pseudoknot S1 length on frameshift efficiency, we made a series of constructs (the pKA series) encoding pseudoknots with different S1 lengths

containing only G·C (or C·G) pairs (Napthine *et al.*, 1999). Two of the shorter constructs, pKA3 (5 bp S1) and pKA5 (6 bp S1), were the starting point for the present analysis and are shown in Figures 1 and 3. The pKA3 construct (Figure 1) is a frameshift reporter construct containing a short IBV-based pseudoknot cloned into the influenza A/PR8/34 gene (Young *et al.*, 1983) at a unique *Bgl*II site. The position of the pseudoknot in this construct is just upstream of a region of the PB2 gene where significant lengths of open reading frame are present in all three frames. A frameshift into any one of these frames generates a product with a characteristic size (22, 28 or 85 kDa). This property was important for the analysis described here, since we wished to introduce single or pairs of nucleotides into the pseudoknot, which would change the exit phase of −1 frameshifted ribosomes. For pKA3-derived plasmids, ribosomes which terminate without frameshifting synthesise a 19 kDa non-frameshifted species. Ribosomes which undergo −1 frameshifting at the slippery sequence continue translation to produce a −1 frameshift product whose size is 22 kDa, 28 kDa or 85 kDa depending upon the exit reading frame, which is determined by the number of nucleotides introduced into S1.

The inability of constructs like pKA3 and pKA5 to stimulate frameshifting was puzzling, since some of the highly efficient pseudoknots present at other viral frameshift sites are predicted to have similar S1 lengths (5 or 6 bp). Studies on the MMTV pseudoknot (Chen *et al.*, 1995) have highlighted the functional importance of an intercalated, unpaired A nucleotide located between the two stems. The most likely explanation for the functional deficit of pKA3 and pKA5 was thus their lack of such an unpaired A at the stem-stem junction, and we suspected that they could be rendered functional by providing this base. As indicated in Table 1, however, inclusion of an A residue between S1 and S2 of pKA3 or pKA5 (to create pKA35 and pKA36) did not stimulate frameshifting and additional changes were therefore necessary.

On the basis of published pseudoknot sequences, mutational analysis of the MMTV frameshift signal (Chen *et al.*, 1995) and our own studies of the IBV pseudoknot stacking region (Brierley *et al.*, 1991), we suspected that the bases flanking the intercalated A residue were inappropriate, and would not allow correct folding of the pseudoknot. For this reason, we tested a variety of paired and unpaired G and C residues at the top of S1 or bottom of S2 within the context of a pseudoknot containing either a 5 bp (constructs pKA35, 42, 44, 45, 53) or 6 bp S1 (constructs pKA36, 43, 46, 47, 48, 51). Once again we were unable to prepare a structure which promoted efficient frameshifting. Variants of pKA35 and 36 with a 5 bp S2 were also inactive (pKA37, 38).

We continued the search, concentrating on spacer length and loop nucleotides. Increasing the spacer distance from 6 to 7 nt (the same length of spacer as present at the MMTV *gag/pro* frameshift signal) led to a small increase in frameshifting in an RNA with a 6 bp S1 (pKA-G, 5%). Alteration of the L1 nucleotides (5′ CA 3′) to either 5′ AA 3′ (pKA58, 60, 61) or a single A residue (pKA59, 62), within the context of a 6 or 7 nt spacer had no

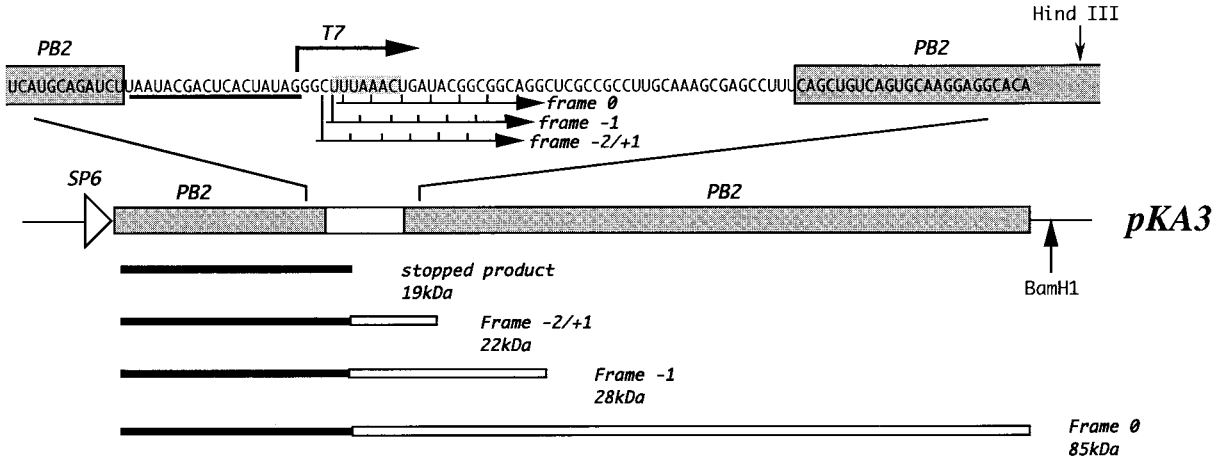


Figure 1. The frameshift reporter construct pKA3. Plasmid pKA3 (Napthine *et al.*, 1999; see Figure 3) contains a short, IBV-derived pseudoknot (white box) cloned into a reporter gene, the influenza PB2 gene (shaded boxes). Linearisation of the plasmid with *Bam*H1 and *in vitro* transcription using SP6 RNA polymerase yields an mRNA (2.4 kb) that, when translated in RRL, produces a 19 kDa non-frameshift product corresponding to ribosomes that terminate at the UGA termination codon (located immediately downstream of the slippery sequence UUUAAAC, shaded), and an 28 kDa −1 frameshift product. The 0-frame and −2/ + 1 frames are also open (to some extent) in this construct. Ribosomes which enter these frames produce 85 kDa and 22 kDa products, respectively. A bacteriophage T7 promoter is present just upstream of the frameshift region; this promoter is employed to generate short, pseudoknot-containing transcripts from *Hin*dIII-digested templates for secondary structure analysis.

Table 1. Conversion of a non-functional short pseudoknot to a functional kinked pseudoknot

Clone name	Sequence									FS (%)
	SS	Spacer	Stem1	L1	Stem2	INT	Stem1′	Loop2	Stem2′	
pKA3	UUUAAAC	UGAUAC	GGCGG	CA	GGCUCG		CCGCC	UUGCAAAG	CGAGCC	<2
pKA35						**a**				<2
pKA37					**d**GCUCG	**a**				<2
pKA42					GGCUC**c**	**a**	**g**CGCC			<2
pKA44			GGCG**c**			**a**	**g**CGCC			<2
pKA45					GGCUC**c**	**a**	**g**CGCC		**g**GAGCC	<2
pKA53			GGCG**c**		GGCUC**c**	**a**	**g**CGCC		**g**GAGCC	<2
pKA54		UGAUAC**a**	G**g**CG**c**		GGCUC**c**	**a**	**g**CGGC	**g**UUGCAAAG	**g**GAGCC	<2
pKA5			GGC**c**GG				CC**g**GCC			2.5
pKA36			GGC**c**GG			**a**	CC**g**GCC			<2
pKA38			GGC**c**GG		**d**GCUCG	**a**	CC**g**GCC			<2
pKA43			GGC**c**GG		GGCUC**c**	**a**	**g**C**g**GCC			<2
pKA46			GGC**c**GG		GGCUC**c**	**a**	**g**C**g**GCC		**g**GAGCC	<2
pKA47			GGC**c**G**c**		GGCUC**c**	**a**	**g**C**g**GCC			<2
pKA48			GGC**c**G**c**			**a**	**g**C**g**GCC			3
pKA51			GGC**c**G**c**		GGCUC**c**	**a**	**g**C**g**GCC		**g**GAGCC	3
pKA60			GGC**c**G**c**	**aA**	**d**GCUC**c**	**a**	**g**C**g**GCC		**g**GAGCC	3
pKA61			GGC**c**G**c**	**aA**		**a**	**g**C**g**GCC		**g**GAGCC	<2
pKA-G		UGAUAC**u**	GGC**c**G**c**			**a**	**g**C**g**GCC		**g**GAGCC	5
pKA58		UGAUAC**u**	GGC**c**G**c**	**aA**		**a**	**g**C**g**GCC		**g**GAGCC	3
pKA59		UGAUAC**u**	GGC**c**G**c**	**ad**		**a**	**g**C**g**GCC		**g**GAGCC	3
pKA62			GGC**c**G**c**	**ad**		**a**	**g**C**g**GCC		**g**GAGCC	<2
pKA-A		UGAUAC**u**	GGC**c**G**c**			**a**	**g**C**g**GCC	UUGCAAA**a**	**g**GAGCC	31

This Table summarises the primary RNA sequence and frameshifting efficiencies of the transcripts of a number of the clones used in this study. Messenger RNAs derived from *Bam*HI-digested plasmids were translated in the rabbit reticulocyte lysate (RRL) *in vitro* translation system. Products were labelled with [^{35}S]methionine, separated on a SDS-15% polyacrylamide gel and detected by autoradiography. Sequence changes are given with respect to pKA3 and the base changes are shown in the context of their domain as lowercase bold letters (d, deletion). The slippery sequence, loop 1 and intercalating base are indicated by SS, L1 and INT respectively. The 3′ arms of the two stems are primed (′).

specific effect on frameshifting. When the spacer distance was 6 nt, frameshifting was very inefficient (pKA61, 62; <2%); constructs with a 7 nt spacer were slightly more active (pKA58, 59; 3%). Having failed to restore functionality with the above approaches, we turned our attention to changes in L2, which proved to be more fruitful.

Potential for a loop-helix interaction

The recent determination of the TYMV pseudoknot structure by NMR (Kolk *et al.*, 1998) alerted us to the role of adenine-related interactions in the formation of higher order RNA structures. In this RNA, L2 crosses the S1 minor groove and interacts with the opposing helix, in particular through hydrogen bonds with a loop terminal adenosine (A35). This residue's base moiety is tilted to an angle of about 90° with respect to the plane of the opposing base-pairs (Figure 2), which allows it to hydrogen bond with the minor groove faces of both G30 and G31 (Table 2). Significantly, structural constraints imposed by the main-chain of L2 seem to limit the base that can provide the required hydrogen bond donor/acceptor profile to an adenine (Kolk *et al.*, 1998). Inspection of the 3D structure of the VPK frameshifter pseudoknot determined by NMR (Shen & Tinoco, 1995) revealed an L2 adenine residue, A27, that faces the S2 minor groove (Figure 2, Table 2). This base, like A35 of TYMV, is tilted with respect to the plane of opposing base-pairs, except that it is positioned over the middle of the minor groove rather than at one side, allowing it to contact bases on both strands of the helix. Since both interactions seem to require a functionality that can only be provided by an adenine residue, we decided to change G29 of pKA-G to an A (construct pKA-A). As can be seen in Figure 3, frameshifting increased sixfold to 31% in response to this modification. However, changing G29 to U or C did not stimulate frameshifting markedly (6%, pKA-C; 10%, pKA-U), highlighting the specific requirement for an adenine residue at this position, consistent with the loop-helix interaction proposed above.

Having established the parameters needed for efficient frameshifting, we wished to confirm the importance of the intercalating residue A15.

Deletion of this base or replacing it with G, U or C caused a significant reduction in frameshift efficiency (from 31 % to 8-13 %, Figure 3, pKA71, 72, 73, 74) in good agreement with related studies on the MMTV pseudoknot (Chen *et al.*, 1995). Surprisingly, the A15 deletion mutant (pKA71) was relatively efficient (8 %) in comparison to pKA5 (2.5 %). This proved to be a consequence of the identity of the last L2 nucleotide; changing the equivalent residue in pKA5 to A (pKA90) increased frameshifting to 8 %. The U or C at this position showed slightly lower frameshift efficiency (pKA91 U, 6 % ; pKA92 C, 4 %). Three other parameters were investigated. Altering the base-pair at the bottom of stem 2 from G·C to C·G caused a modest reduction in frameshift efficiency (pKA103, 22 %), as did shortening the length of L1 to a single A residue (pKA80, 19 %). More dramatic was the effect of altering S1 length; increasing (pKA78) or decreasing (pKA79) S1 by a single base-pair caused a four- to sixfold reduction in frameshift efficiency, perhaps by preventing the loop-helix interaction (see Discussion).

RNA structure mapping of constructs KA-G and KA-A

The requirement for an A at the end of L2 for efficient frameshifting was clearly compatible with a loop-helix interaction. Nonetheless, did the deficit of function of the other RNAs highlight a specific contribution of the modified adenine residue, or was it a relatively uninformative consequence of odd interactions among the remaining components? Even minor base changes have the potential to modify pseudoknot conformation substantially (Kang & Tinoco, 1997), and it was therefore important to establish whether major conformational differences existed between the two RNAs. We determined the secondary structures of KA-A and KA-G by chemical and enzymatic probing using an end-labelling procedure (van Belkum *et al.*, 1988; Wyatt *et al.*, 1990; Polson & Bass, 1994). The two plasmids have an internal T7 promoter just upstream of the frameshift signal, allowing run-off transcripts to be prepared following linearisation with *Hin*dIII (Figure 1). Transcripts of 99 nt were prepared from each plasmid, end-labelled with [γ-^{33}P]ATP, gel purified and subjected to limited chemical and enzymatic digestion prior to analysis on denaturing 10 % or 15 % polyacrylamide gels. In these experiments, the Mg^{2+} level was kept at 2 mM (except for imidazole probing, where 10 mM Mg^{2+} was used) which is the approximate concentration of this ion in RRL (Jackson & Hunt, 1983) and sites of cleavage were scored only from those reactions where 80-90 % of the full-length RNA remained intact. A representative selection of the individual biochemical analyses and a diagrammatic summary are shown in Figure 4. The bases are numbered from the first base of the pseudoknot (G) in each case. The single strand-specific enzymatic probes employed were RNase T_1, which cleaves 3′ of unpaired G residues,

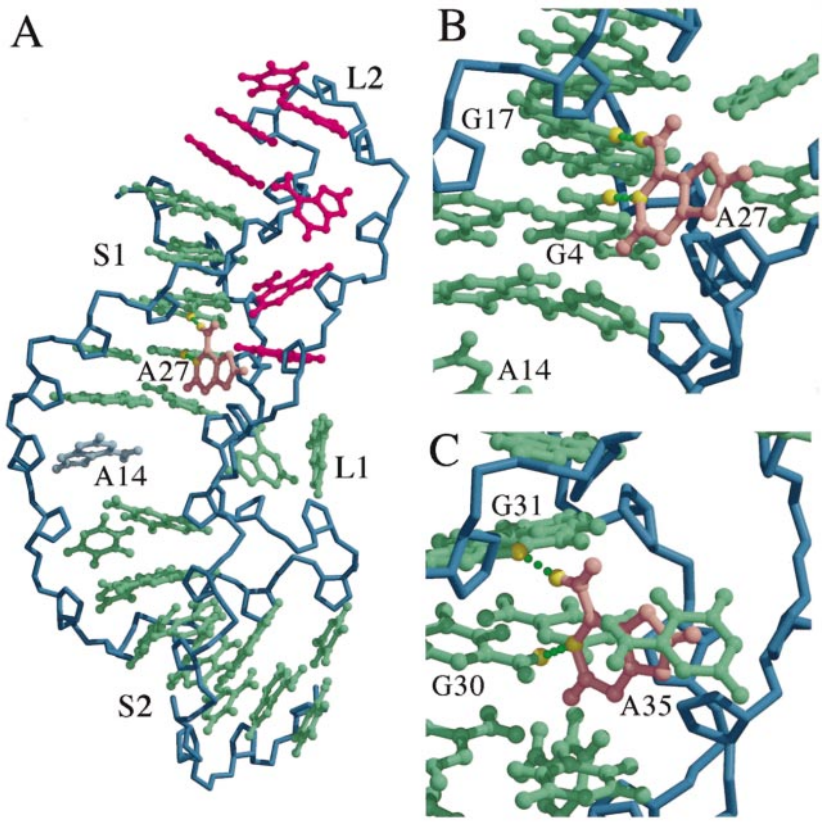


Figure 2. Loop-helix interactions in RNA pseudoknots. (a) The three-dimensional structure of the frameshifter pseudoknot VPK (Shen & Tinoco, 1995). The RNA backbone is in blue and the bases are colour-coded by domain or motif; stem residues in green, L2 in pink and A27 in peach. The pseudoknot is bent at the S1-S2 junction; the intercalating A14 (light blue) is stacked between the two stems. L2 crosses the minor groove of S1 with A27 potentially contacting G17 and G4 (coordinates as reported by Chen *et al.*, 1996). (b) Close up of the VPK structure showing the potential loop-helix contact involving A27, G4 and G17. (c) Close up of the equivalent region of the TYMV pseudoknot showing the hydrogen bonds between L2 and S1 (coordinates as reported by Kolk *et al.*, 1998). The orientation of the adenine moiety with regard to the plane of opposing G·C pairs is similar to that seen in VPK. Figures were generated using MOLSCRIPT (Kraulis, 1991) in combination with Raster3D (Merritt & Bacon, 1997).

Table 2. Predicted or observed hydrogen bonds between L2 terminal bases and stem 1 of RNA pseudoknots

Residue	Atom	Residue	Atom	Distance (Å)	Structure/ coordinates	Reference/method
G4	2-NH2	A27	N1	1.58	VPK	Shen & Tinoco (1995); NMR
G17	N3	A27	6-NH2	1.73		
G30	2-NH2	A35	N1	2.07	TYMV (1A60)	Kolk *et al.* (1998); NMR
G31	N3	A35	6-NH2	1.85		
G4	2-NH2	A27	N3	2.44	VPK (VPK_m4)	Le *et al.* (1998); MD simulation of VPK NMR distance constraints with sodium ions and water molecules
G4	N3	A27	2′-OH	2.05		
G17	2-NH2	A27	N7	2.44		
G17	2′-OH	A27	N1	1.96		
C18	O4′	A27	6-NH2	2.01		
A26	N1	A27	6-NH2	2.02		
C3	2′-OH	A26	N7	2.06		
C3	O2	A26	6-NH2	1.96		
C18	2′-OH	A26	N1	2.10		
C18	O2	A26	6-NH2	1.95		

and RNase U_2, which cleaves 3′ of single-stranded A or G residues, with a preference for A. To probe double-stranded regions, RNase V_1 was employed which cleaves in helical regions. RNase V_1 is not base-specific but cleaves RNA that is in helical conformation, whether base-paired (a minimum of

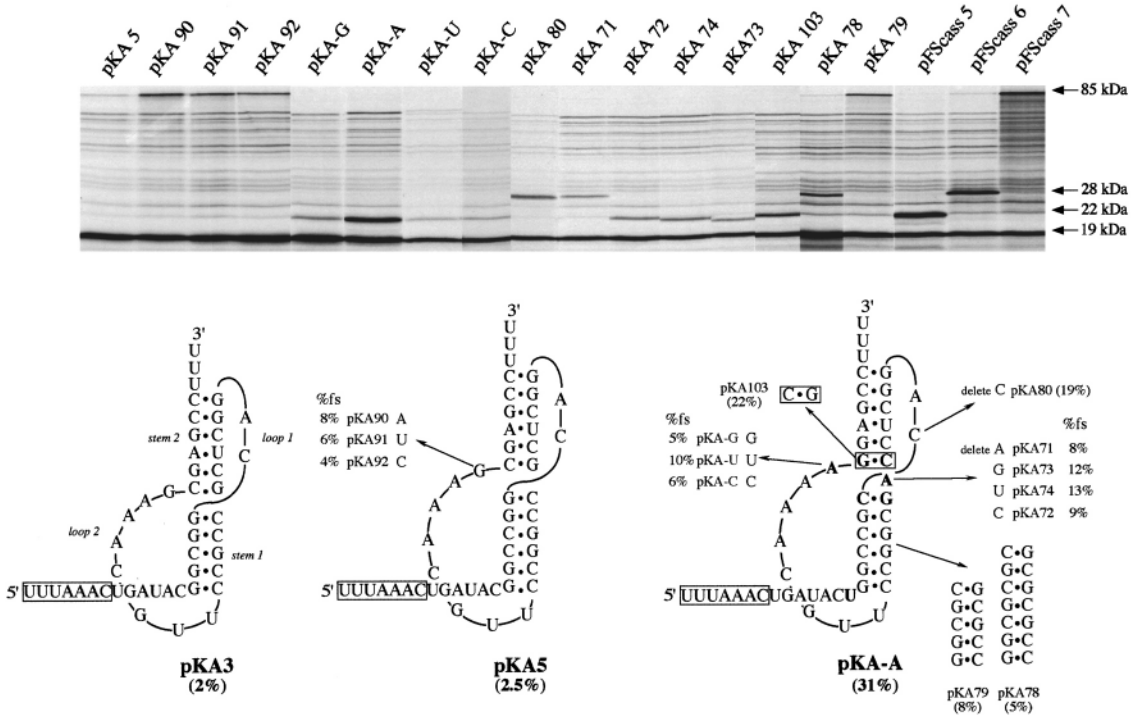


Figure 3. Analysis of the pKA-A frameshift signal by mutagenesis. This Figure shows the primary sequence and proposed secondary structure of the pKA3, pKA5 and pKA-A pseudoknots. The slippery sequence is boxed. Nucleotides of pKA-A that differ in orientation or presence from those of pKA5 are shown in bold. A series of mutations were created in pKA5 and pKA-A, and the effects on frameshifting measured by *in vitro* transcription and translation in the RRL. The upper portion of the Figure shows the RRL translation products synthesised in response to mRNAs derived from *Bam*HI-digested plasmids. Products were labelled with [^{35}S]methionine, separated on a SDS-15% polyacrylamide gel and detected by autoradiography. The frameshifted (22, 28 or 85 kDa) and non-frameshifted (19 kDa) species are marked with arrows. The size of the −1 frameshift product produced by the various mutant RNAs is determined by the number of nucleotides in the pseudoknot (see Results). The predicted sizes are 22 kDa (pKA-G, A, 69, 70, 72, 73, 74, 103 pFScass 5), 28 kDa (pKA71, 78, 80, pFScass 6) or 85 kDa (pKA5, 79, 90, 91, 92, pFScass 7). pFScass 5, 6 and 7 contain the minimal IBV frameshift signal (Brierley *et al.*, 1992) and were translated to mark the position of the 22, 28 and 85 kDa frameshift products. In translations of pKA-based constructs in RRL, a low level of background polypeptides are present which arise as a result of aberrant initiation events (see Brierley *et al.*, 1992). These can potentially introduce an inaccuracy of up to 2% in estimations of frameshift efficiency and we define a non-functional construct as one displaying a frameshift efficiency of $\leqslant$2%.

4 to 6 bp are required) or single-stranded and stacked. We also used the single strand-specific chemical probes imidazole (Vlassov *et al.*, 1995) and lead acetate (Krzyzosiak *et al.*, 1988; Kolchanov *et al.*, 1996).

From the structure probing data (Figure 4), it is clear that the two RNAs are pseudoknots and have a cleavage pattern that suggests they are very similar in structure, both to each other and to frameshifter pseudoknots of the IBV class described previously (Napthine *et al.*, 1999). Structure probing with either imidazole or lead (lanes 7-8, 13-14) generated a similar pattern of bands highlighting the major regions of single and double-stranded RNA. Both pseudoknot stems showed resistance to chemical cleavage, especially the G + C-rich S1. S2 was somewhat less resistant to imidazole, but was largely resistant to lead, consistent with the formation of this stem. RNase V_1 cuts were evident in S1 and were also seen in the 5′-arm of stem 2 (lanes 10-11). The single strand-specific enzymes RNase T_1 and U_2 (lanes 4-5, 16-17) reacted largely as expected, with those loop residues predicted to be single-stranded. Some cleavages were noted in stem regions. For both KA-G and KA-A, RNase T_1 cleavage was seen at G5, G16 and G19 within S1 and at G9, G10 and G33 (and weakly at G30 and G31) within S2. Similarly, an RNase U_2 cut was seen at A32 of S2 for both pseudoknots, although cleavage was slightly less in pKA-A. The intercalated residue A15 at the junction of the two stems was only weakly reactive with RNase U_2 in either RNA. The reactivity of residues G9 and G10 at the top of S2 with single-stranded enzymatic and chemical probes was seen in previous studies with IBV pseudoknot variants (pKA13 and pKA18; Napthine *et al.*, 1999) and may be indicative of a region of unusual conformation. Overall, however, KA-G and KA-A appear to possess a similar secondary structure given the restricted resolution of chemical and enzymatic probing.

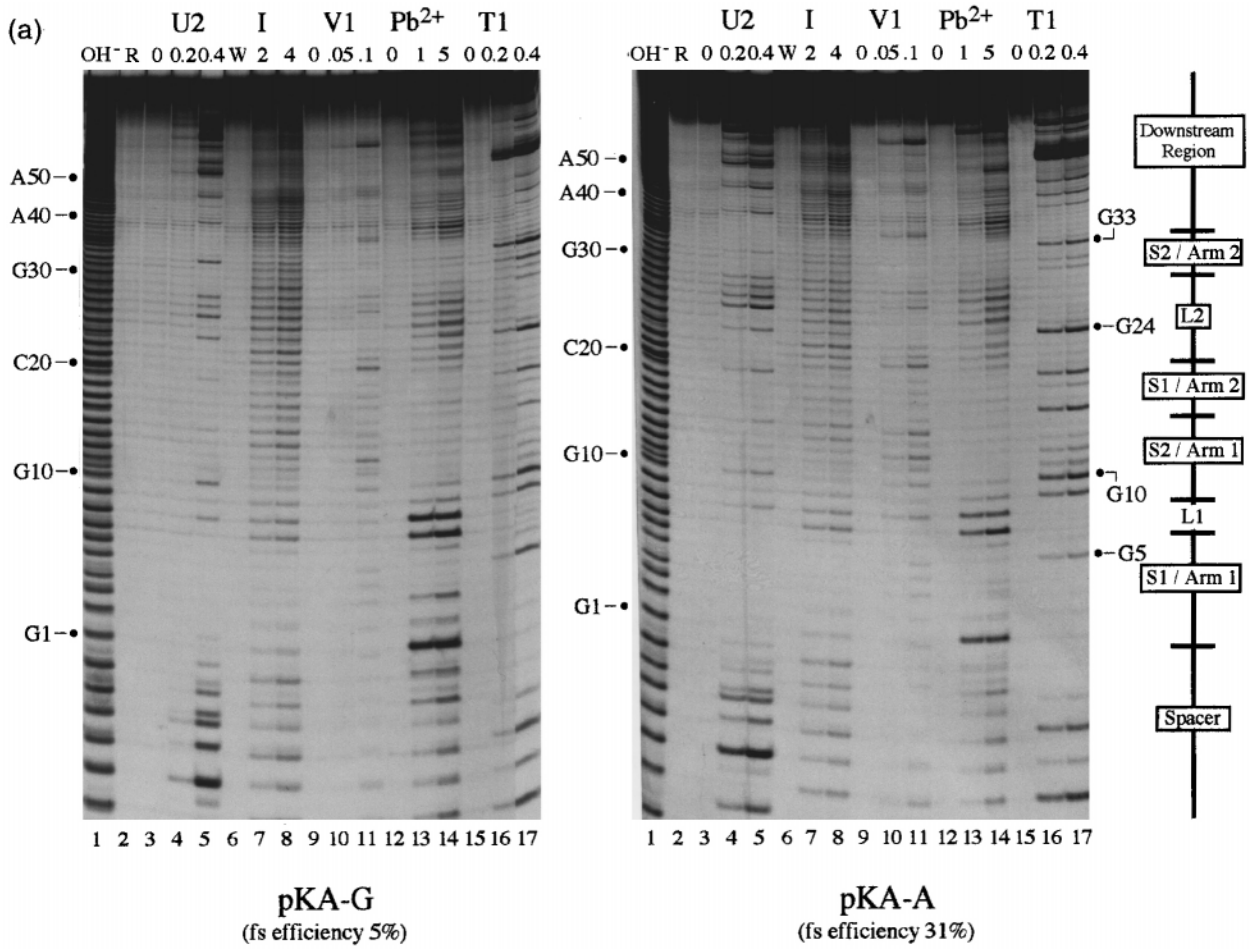
Modification of the putative loop acceptor region of S1 affects only RNAs with an adenine residue at position 29

To identify bases in S_1 that could potentially interact with L2 of pKA-A, we compared the predicted or observed hydrogen bonds between L2 terminal bases and S1 pairs that had been documented in previous studies (see Table 2). From these, we judged that the two G·C pairs near the top of S1 (C4·G18 and G5·C17, Figure 3) were the most likely candidate for a loop base acceptor. We modelled several changes in that region using NAMOT (Carter & Tung, 1996) to identify mutations that would affect base-pairing with A29 while unlikely, on their own, to significantly change the structure or thermodynamic stability of the KA pseudoknot. A simple base-flip to give G4·C18 and C5·G17 was theoretically ideal due to the absence of obvious confounding effects. This mutation would modify the hydrogen bond profile of the putative loop acceptor. In the flipped case, the N1 of A29 can no longer form an optimum linear hydrogen bond with the hydrogen of the guanine exocyclic amine. Further, the G18 N3 (which has one lone pair) is replaced by a pyrimidine O2 with its two lone pairs (see Kielkopf *et al.*, 1998a,b). The flip does not substantially alter the predicted stability of S1 (12.1 *versus* 12.4 kcal/mol; as the energetics of pseudoknot formation is not fully understood, the S1 stability predictions, calculated according to the rules reported by Turner *et al.* (1988), were performed on the basis that the stem formed independently and had a loop length of 8 nt).

Accordingly, variants of pKA-A were prepared in which the orientation of both C4-G18 and G5-C17 was flipped in the background of different L2 3′-terminal bases (pKA-FL-A, G, U and C, Figure 5). As can be seen in Figure 5, the frameshift efficiency of the flipped construct with an A at position 29 (pKA-FL-A) dropped approximately threefold to 9 %, a value similar to that seen in the non-flipped context of pKA-G. No significant effect of the flip was seen in any of the other structures, which promoted frameshifting to levels very close to those seen in the non-flipped context (pKA-FL-G, 6 %; pKA-FL-U, 14 %; pKA-FL-C, 8 %). That a reduction in frameshifting was seen only in KA-FL-A argues strongly against a general effect on frameshifting of the base changes created in S1.

Increasing L2 length to 14 nt changes KA's response to variation of the X29 base

In a recent mutational analysis of the SRV-1 pseudoknot (Sung & Kang, 1998), it was found that replacing the 3′-terminal L2 base, a uracil, with other bases had no effect on frameshifting. This was inconsistent with our finding that the frameshift efficiency of pKA-A depended crucially on the A29 residue because the pseudoknots are thought to be similar both in structure and function, with both RNAs adopting a bent conformation (see, for example, Sung & Kang, 1998, but note the conflicting NMR data by Du *et al.*, 1997). Therefore, we would have expected SRV-1 pseudoknot function to be sensitive to modifications of the L2 3′-terminal base, which was not the case. The recent study by Le *et al.* (1998), however, who modelled the structure of VPK in the presence of water and sodium ions, provides a potential explanation for the differing response of the SRV-1 and KA-A pseudoknots to such changes. They predicted a compact structure in which the first four L2 nucleotides (ACUC) of VPK make a sharp turn, and the remaining four A residues of L2 cross the minor groove with the last two A residues contacting the top of S1 directly. The compact packing of L2 onto S1 is facilitated by the shortness of the loop and by the presence of the A quartet. The SRV-1 pseudoknot differs from the MMTV, VPK and KA-A pseudoknots in L2, which is five bases longer in SRV-1 (13 *versus* 8 nt) and contains a U, rather than an A-rich stretch at the 3′-end of L2.



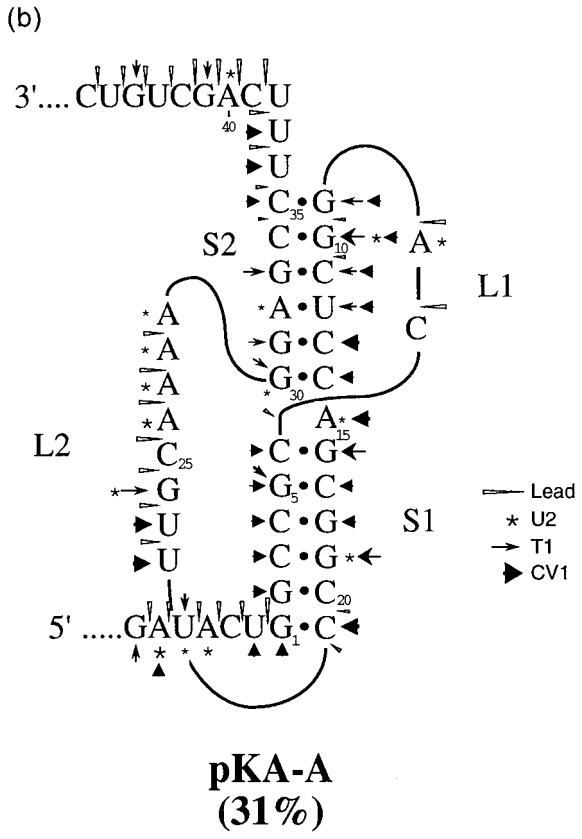


Figure 4. Structure probing of the KA-G and KA-A pseudoknots. RNA derived by T7 transcription of (a) pKA-G/*Hin*dIII or (b) pKA-A/*Hin*dIII was 5′ end-labelled with [γ-^{33}P]ATP and subjected to limited RNase or chemical cleavage using structure-specific probes. Sites of cleavage were identified by comparison with a ladder of bands created by limited alkaline hydrolysis of the RNA (OH^-) and the position of known RNase U_2 and T_1 cuts, determined empirically. Products were analysed on 15% acrylamide/7 M urea gels. Data was also collected from 10% gels (gels not shown). Enzymatic structure probing was with RNases U_2, V_1 and T_1. Uniquely cleaved nucleotides were identified by their absence in untreated control lanes (0). The number of units of enzyme added to each reaction is indicated. Chemical structure probing was with imidazole (I, hours) or lead acetate (Pb^{2+}; mM concentration in reaction). The water lane (W) in the imidazole panel represents RNA which was dissolved in water, incubated for 4 hours and processed in parallel to the imidazole-treated samples. R represents an aliquot of the purified RNA loaded directly onto the gel without incubation in a reaction buffer. (b) Summary of the KA-A probing results. The sensitivity of bases in the KA-A frameshift region to the various probes is shown. The size of the symbols is approximately proportional to the intensity of cleavage at that site. The imidazole probing data, which was similar to that seen with lead acetate, is omitted for clarity. The KA-G probing results are not summarised since they were essentially the same as those of KA-A.

Also, L1 of the SRV-1 pseudoknot comprises a single A residue, whereas the others have two bases in L1 (Figure 6). In SRV-1, the extended L2 length and altered sequence composition could well prevent a stimulatory contact with S1. If this is the case, then shortening L2 of SRV-1 should promote frameshifting. Indeed, ten Dam *et al.* (1995) demonstrated that removing two or three bases from L2 of SRV-1 increased frameshifting from 20 % to 34 %, although further base deletions to an L2 length of 9 and 7 nt had the opposite effect and decreased frameshifting to 18 and 10 %. On this basis, we would predict that the KA pseudoknots should be rendered relatively insensitive to L2 terminal base changes by increasing the L2 length to that of SRV-1.

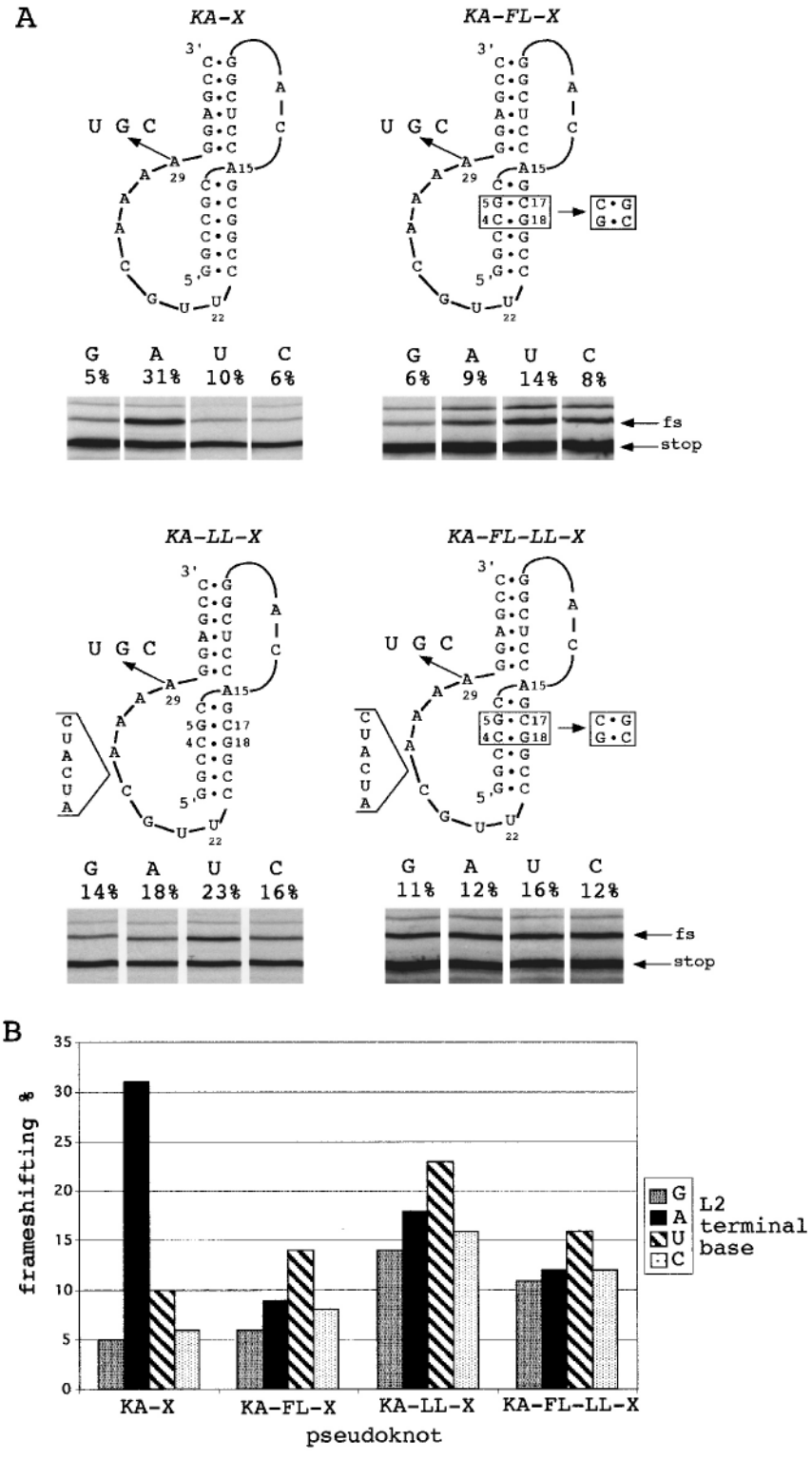


Figure 5. Evidence for a loop-helix interaction in pseudoknot KA-A. (a) The influence of the L2 terminal nucleotide (X29) on −1 ribosomal frameshifting was tested in the context of (KA-FL-X) a variant of pKA-A in which two S1 bases had been flipped (G5·C17; C4·G18; boxed); (KA-LL-X), a variant of pKA-A in which L2 had been increased to 14 nt by insertion of 6 nt (AUCAUC) and (KA-FL-LL-X), a variant of pKA-A in which both flip and L2 increase were present. The effect on frameshifting of varying base A29 in the context of the wild-type pseudoknot is shown for comparative purposes (KA-X). The RRL translation products synthesised in response to mRNAs derived from *Bam*HI-digested pKA-A and mutant derivatives are shown. Products were labelled with [^{35}S]methionine, separated on a SDS-15 % polyacrylamide gel and detected by autoradiography. The frameshifted (fs; 22 kDa) and non-frameshifted (stop; 19 kDa) species are marked with arrows. (b) A comparison of the frameshift efficiencies measured for each X29 variant in the four pseudoknot contexts; wild-type (KA-X), S1 bases flipped (KA-FL-X), L2 length increase (KA-LL-X), S1 base-pairs flipped and L2 length increase (KA-FL-LL-X).

We thus inserted a six base sequence (AUCAUC; Figure 5) into the middle of L2 of the pKA-A, G, U or C pseudoknots, increasing the loop length from eight to 14 bases without changing either the first four (UUGC) or final four bases (AAAX29). The *in vitro* frameshifting efficiency of the resultant constructs pKA-LL-A, C, G, and U are shown in Figure 5. The efficiency of the KA-LL-A pseudoknot was found to be 18 %, about twofold lower than the equivalent construct with a shorter L2. In contrast, frameshifting in the KA-LL mutants containing C, G or U at position X29 was approximately doubled (G29, 14 %; C29, 16 %; 23 %, U29) in comparison to the eight-base L2 constructs. Thus although frameshifting efficiency was still somewhat sensitive to the L2 terminal base, it was significantly less so than in the eight-base L2 structure and was decreased (KA-LL-A) or increased (KA-LL-C, -G, -U) towards the value reported for the SRV-1 structure (23 %).

Flipping the S1 C·G pairs in the context of a 14 nt L2 has no differential effect on X29 variants

Given the reduced influence of A29 on frameshifting in constructs with a longer L2, it seemed likely that alterations of the putative loop acceptor region would, similarly, have little effect on efficiency. Accordingly, base X29 was varied in the background of both the G·C flip and the 14 nt L2. In these mutants (KA-FL-LL-A, G, U, C; Figure 5) frameshifting was reduced a little in each case compared to the non-flipped KA-LL series; from 14-23 % to 11-16 % (Figure 5), but no significant differential effects were seen. The loop-helix interaction, therefore, seems to be important in KA-A-directed frameshifting only in the context of a short L2.

Discussion

Here, we were interested to determine the sequence changes required to convert an IBV-based non-functional pseudoknot, with a short S1, into a functional kinked pseudoknot. Although a variety of pseudoknots were constructed that contained an intercalating A residue at the junction between the two pseudoknot stems (an essential feature of a kinked pseudoknot), efficient frameshifting was not observed until the last nucleotide of L2 was changed from a G to an A residue (Table 1). This result was consistent with the possibility that frameshifting requires an interaction between L2 and residues in S1. Further mutational analysis was carried out to demonstrate that the loop-helix interaction, between the terminal L2 adenine residue and S1, is essential for efficient frameshifting. Interactions of this kind, involving adenosine residues, have been recognised increasingly in both pseudoknots and in other RNA structures. In the P4-P6 domain of the *Tetrahymena thermophila* intron, each adenosine in a GAAA tetraloop makes specific hydrogen bonds to the tetraloop receptor, comprising two adjacent C·G base-pairs, a 5 nt internal loop and a G·U base-pair (Cate *et al.*, 1996a,b). Similarly, the terminal adenosine A35 of L2 of the TYMV pseudoknot is buried deep within the minor groove of the opposing S1 helix (Kolk *et al.*, 1998).

Recently, potentially equivalent interactions have been proposed to occur in the pseudoknots associated with the ribosomal frameshift signals of an MMTV-variant, VPK (Le *et al.*, 1998) and the plant virus beet western yellows virus (BWYV; Su *et al.*, 1999). Although an L2-S1 contact was not proposed from the original VPK NMR data (Shen & Tinoco, 1995), the possibility of additional loop-helix interactions was noted following refinement by Le *et al.* (1998). Based on restrained molecular dynamics (MD) simulations with water molecules and sodium ions, they predicted a compact structure in which the first four nucleotides (A20 CUC) of L2 make a sharp turn and the other four bases (AAA A27) cross the minor groove with A27 and A26 directly contacting the top of S1 (Table 2). The 3D structure of the BWYV frameshifter pseudoknot has been solved by X-ray crystallography (Su *et al.*, 1999). In this RNA, the adenosine-rich L2 is hydrogen-bonded to the minor groove of S1 largely through contacts involving the 2′-hydroxyl group, forming a novel three-stranded structure with some similarity to the S1-L2 domain of the refined VPK pseudoknot.

We began our investigation of the potential loop-helix interaction in the KA-A pseudoknot by inspecting the predicted or observed hydrogen

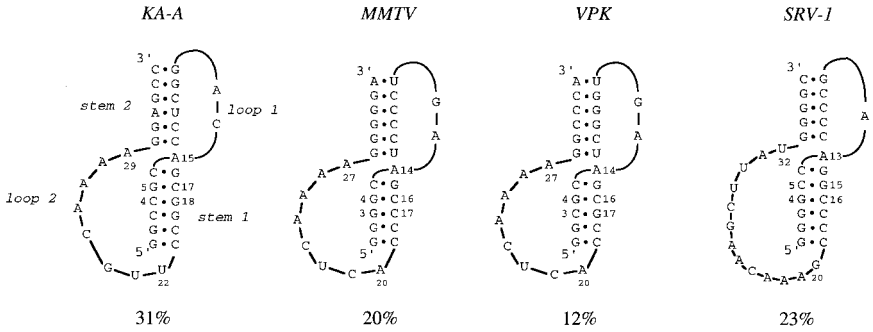


Figure 6. Comparison of the primary sequences and proposed secondary structures of the RNA pseudoknots pKA-A, MMTV, VPK, and SRV-1. The frameshift efficiencies of the various structures, as determined in RRL, are shown and were reported by Chamorro *et al.* (1992); MMTV), Chen *et al.* (1995; VPK) and ten Dam *et al.* (1994; SRV-1).

bonds between L2 terminal bases and S1 that had been noted in previous studies (see Table 2). From these, we judged that the region of pKA-A most likely to act as a loop base receptor would be the penultimate two base-pairs in S1 (C4·G18 and G5·C17). Since an interaction between the loop adenosine and the S1 helix is likely to require a particular placement of the hydrogen bond donor/acceptor capacities along the minor groove, the inhibitory effect found upon replacing A with any other base ought to be recapitulated by changes in the putative loop acceptor region. Flipping the two G·C base-pairs was predicted to change the hydrogen bond donor/acceptor profile in a way that was unlikely to distort the overall conformation of stem 1. Indeed, flipping the G·C base-pairs impaired frameshifting only in the construct with an adenosine residue at the end of L2 (pKA-FL-A), confirming the absence of any general effect on frameshifting of the S1 changes. Thus only the combination of an L2 terminal adenosine with a particular S1 G·C base-pair orientation was fully functional, and changes of either partner of the proposed interaction had identical inhibitory effects. This strongly supports the conclusion that A29 is contacting the S1 minor groove and that this contact is important for efficient frameshifting.

There are two caveats, however. The first is that we were unable to detect conformational differences between the KA-A pseudoknot and a pseudoknot with an impaired frameshift capacity (KA-G). We have interpreted the difference in frameshift efficiency seen with the two constructs to be attributable to an ability to form (KA-A) or to be unable to form (KA-G) the loop-helix interaction, which presumably exerts its effect on frameshifting by modifying the pseudoknot's structure. It is feasible, however, that the structure probing methods used were not sensitive enough to detect such, possibly very subtle, conformational changes. Clearly, structural analysis at higher resolution will be required to address this question. Secondly, the pKA-A construct was very sensitive to alterations in S1 length; increasing or decreasing S1 by a single base-pair reduced frameshifting dramatically, despite the fact that the putative S1 loop acceptor bases were still present. However, this may be explicable in terms of helix rotation; increasing or decreasing the length of S1 would presumably change the position of L2 with respect to the helix, which could well perturb the loop-helix interaction.

An alternative explanation can be suggested for the differences in frameshifting seen with variants of the terminal L2 base, which concerns potential stacking interactions in the kinked region. Studies of functional and non-functional MMTV-based pseudoknots (Chen *et al.*, 1995, 1996; Shen & Tinoco, 1995; Kang *et al.*, 1996; Kang & Tinoco, 1997) have shown that frameshifting requires the pseudoknot to be in a specific kinked conformation. It has been proposed that this conformation can only be achieved when a single intercalated residue (A15) is present between the two pseudoknot stems and that the stability of base-pairs adjacent to A15 may be influenced by how well they stack onto proximal loop residues. It is conceivable, therefore, that the functionality of pseudoknots with variations in residue 29 (pKA-A, G, U and C) is influenced by the kind of stacking interactions that occur between these bases and the adjacent residues in loop 2 (A28) or stem 2 (G30). Clearly, interactions that occur between bases at the junction of the two pseudoknot stems will play some role in determining the overall conformation of the structure. Indeed, changing the base-pair at the bottom of S2 from G30·C14 to C30·G14 (pKA 103, Figure 2) had a modest impact on frameshifting efficiency, lowering it to 22 %. However, although no specific experiments were carried out to test a contribution of stacking to the frameshifting phenotype of the X29 variants, the available evidence is more supportive of an effect on a helix-loop interaction. Firstly, the preference for an A at the end of loop 2 was maintained even without the intercalating A15 residue (pKA71 and pKA90), which presumably would modify the potential for base stacking of A29 onto G30. Secondly, a specific reduction of frameshifting was seen only in the S1 flip construct KA-FL-A, but not in the other X29 variants. These findings, in combination with the available structural information are most consistent with a loop-helix contact model.

The differential effects on frameshifting seen with X29 variants of the KA-A pseudoknot are not seen with the SRV-1 pseudoknot (Sung & Kang, 1998). We speculate that this is a consequence of the increased length and altered sequence composition of L2, which may preclude a loop-helix interaction of the kind seen with KA-A. Indeed, increasing the length of L2 in the KA-X29 series reduced the differential effect of the L2 terminal base, bringing the frameshift efficiencies closer to the value reported for the SRV-1 pseudoknot (23 %; ten Dam *et al.*, 1994). Thus the KA-A and SRV-1 pseudoknots respond differently to L2 mutations. In the KA-A pseudoknot with an 8 nt L2 (and very likely in the MMTV pseudoknot), the indicated loop-helix interaction seems to contribute significantly to frameshifting efficiency. In contrast, in KA-A variants with a longer L2, frameshifting efficiency was reduced (from 31 % to 18 %, pKA-LL-A) in response to the length increase, and the G·C flip in S1 (pKA-FL-LL-A) had only a small inhibitory effect. Other factors, and not a loop-helix interaction involving an L2 terminal adenosine residue, are thus likely to be responsible for the ability of the SRV-1 pseudoknot to direct frameshifting. One possibility is L1 length; NMR studies suggest that it possesses only a single A residue in L1 (Du *et al.*, 1997). A variant of pKA-A in which L1 was reduced to a single A residue (pKA80, Figure 2) showed reduced frameshifting, indicating that a single L1 nucleotide is not optimal in this class of pseudoknot, and further supports

the idea that the KA-A and SRV-1 pseudoknots are not structurally equivalent.

The pseudoknots described here are part of a larger set of RNA structures that stimulate frameshifting. In HIV-1, for instance, the stimulatory element is a simple hairpin structure rather than a pseudoknot (Kang, 1998). An intermediate structure is seen at the frameshift site of Rous sarcoma virus (RSV), where the stimulatory RNA structure is pseudoknotted but able to tolerate disruptions of S2 to some degree (Marczinke *et al.*, 1998). At the other extreme are pseudoknots of the IBV class that require both a long S1 and a stable S2 for function. How can the apparent diversity of frameshifter pseudoknots be reconciled in terms of the frameshift process itself? At present, models for pseudoknot stimulated frameshifting tend to invoke a common 3D structure as the determinant of function. In this model, pseudoknots stimulate frameshifting through the formation of an RNA-RNA or RNA-protein interaction, perhaps by mimicking a tRNA (Chen *et al.*, 1996). To date, however, there is no evidence that a structural feature exists that is shared by all stimulatory RNA elements, although this is not inconceivable. A competing model that has some theoretical appeal is that the stimulatory RNAs share mechanical properties, such as initial resistance to unwinding. Dinman (1995) has suggested that the role of S2 of a pseudoknot may be to prevent routine unwinding of S1. Stem 1 must be able to rotate around its helix axis to unwind and the association of the bases at the top of the helix with the downstream sequences (to form S2) might prevent such rotation. The present findings do not particularly favour either hypothesis. A loop-helix interaction will almost certainly modify the 3D structure of the pseudoknot, thus influencing factor recognition, but also the stability of S1 and the ease with which it is unwound. In order to discriminate between the models, a determination the 3D structure of the KA-A and KA-FL-A RNA pseudoknots would be advantageous, as would the development of methods to investigate the unwinding (as opposed to melting) behaviour of RNA pseudoknots.

Materials and Methods

Site-specific mutagenesis

Site-directed mutagenesis was carried out by a procedure based on that of Kunkel (1985) as described (Brierley *et al.*, 1989). Mutants were identified by dideoxy sequencing of single-stranded templates (Sanger *et al.*, 1977). Sequencing through G + C-rich regions was facilitated by including formamide (to 20% (v/v)) in the sequencing gel and replacing dGTP with deaza-GTP in the sequencing mixes.

Construction of plasmids

The plasmids used in this study were derived from plasmids pKA3 or pKA5 (Napthine *et al.*, 1999; see Figure 2) by site-directed mutagenesis. These plasmids are derivatives of pFScass 6 (Brierley *et al.*, 1992), a frameshift reporter construct (see the legend to Figure 1).

In vitro transcription and translation

In vitro transcription reactions employing the bacteriophage SP6 RNA polymerase were carried out essentially as described by Melton *et al.* (1984) as described (Napthine *et al.*, 1999). In ribosomal frameshift assays, serial dilutions of purified mRNAs were translated in RRL as described (Brierley *et al.*, 1989). Translation products were analysed on SDS-15% (w/v) polyacrylamide gels according to standard procedures (Hames, 1981). The relative abundance of non-frameshifted or frameshifted products on the gels was determined by direct measurement of [^{35}S]methionine incorporation using a Packard Instant Imager 2024 and adjusted to take into account the differential methionine content of the products. 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). The frameshift efficiencies quoted are the average of at least three independent measurements which varied by less than 10%, i.e. a measurement of 20% frameshift efficiency was between 18% and 22%. The calculations of frameshift efficiency take into account the differential methionine content of the various products (19 kDa, 11; 22 kDa and 28 kDa, 12; 85 kDa, 35).

RNA structure probing

RNAs for secondary structure probing were prepared by *in vitro* transcription of *Hin*dIII-cut plasmid templates using bacteriophage T7 RNA polymerase. Transcription reactions were on a 200 μl scale and contained 20 μg plasmid DNA, 2.5 mM of each rNTP and 500 units of T7 RNA polymerase (NEB) in a buffer containing 40 mM Tris (pH 8), 15 mM $MgCl_2$ and 5 mM DTT. After three hours at 37°C, 100 units of DNase I was added and the incubation continued for a further 30 minutes. Nucleic acids were harvested by extraction with phenol/chloroform (1:1) and precipited in ethanol. DNA fragments were removed by Sephadex G-50 chromatography and the RNA transcripts concentrated by precipitation in ethanol. The RNA was quantified by spectrophotometry and its integrity checked by electrophoresis on a 2% (w/v) agarose gel containing 0.1% (w/v) sodium dodecyl sulphate. Transcripts (10 μg) were 5′-end-labelled with [γ-^{33}P]ATP using a standard dephosphorylation-rephosphorylation strategy (ten Dam *et al.*, 1995), purified from 10% acrylamide-urea gels and dissolved in water.

The structure probing experiments followed the general principles outlined by others (van Belkum *et al.*, 1988; Wyatt *et al.*, 1990; Polson & Bass, 1994). All reactions contained 10-50,000 cpm of 5′ end-labelled RNA transcript and 10 μg baker's yeast carrier tRNA. RNAse probing reactions were carried out in 50 μl reaction volumes. Probing with RNase T_1 (Amersham) was on ice for 20 minutes in 50 mM sodium cacodylate (pH 7), 2 mM $MgCl_2$ and 0-1 units of T_1; RNase V_1 (Pharmacia) at 25°C for 20 minutes in 10 mM Tris (pH 8), 2 mM $MgCl_2$, 0.1 M KCl and 0-0.35 units of V_1; RNase U_2 (USB) on ice for 20 minutes in 20 mM sodium acetate

(pH 4.8), 2 mM $MgCl_2$, 100 mM KCl and 0-0.2 units of U_2.

Enzymatic reactions were stopped by the addition of 150 µl ethanol and the RNA recovered by centrifugation. RNAs were prepared for analysis on 10 or 15% polyacrylamide-7 M urea sequencing-type gels by dissolution in water, mixing with an equal volume of formamide gel loading buffer (95% (v/v) formamide, 10 mM EDTA, 0.1% (w/v) bromophenol blue, 0.1% (v/v) xylene cyanol) and heating to 80°C for three minutes. Chemical probing experiments were performed with lead acetate and imidizole in 10 µl reaction volumes. Lead probing was at 25°C for five minutes in 20 mM Hepes-NaOH (pH 7.5), 5 mM Mg acetate, 50 mM K acetate and 1-5 mM Pb acetate. Reactions were stopped by the addition of EDTA to 33 mM, the RNA recovered by precipitation in ethanol, redissolved in water and prepared for gel loading as above.

For imidizole probing, the end-labelled RNA was mixed with 10 µg carrier tRNA, dried in a desiccator and redissolved in 10 µl 2 M imidizole (pH 7) containing 40 mM NaCl and 10 mM $MgCl_2$. After incubation at 37°C for one to four hours, the reaction was stopped by the addition of 10 µl of a fresh solution of 2% (w/v) lithium perchlorate in acetone. The RNA was recovered by centrifugation, washed with acetone, dried, dissolved in water and prepared for gel loading as above. All the structure probing gels included an alkaline hydrolysis ladder of the relevant RNA as a size marker, prepared by dissolving the dried pellet from 3 µl of end-labelled RNA and 10 µg carrier tRNA in 3 µl of 22.5 mM $NaHCO_3$, 2.5 mM Na_2CO_3 and boiling for one to 2.5 minutes. After the addition of an equal volume of formamide gel loading buffer, the sample was heated to 80°C for three minutes and loaded immediately onto the gel.

Acknowledgements

This work was supported by the Medical Research Council, UK and the Biotechnology and Biological Sciences Research Council, UK. We thank Paul Digard for critical reading of the manuscript, Edwin ten Dam for helpful advice and Professor Alex Rich for drawing our attention to the work of Kielkopf and colleagues regarding G·C and C·G discrimination in the minor groove. J.L. and S.N. contributed equally to this work.

References

Benard, L., Mathy, N., Grunberg-Manago, M., Ehresmann, B., Ehresmann, C. & Portier, C. (1998). Identification in a pseudoknot of a U·G motif essential for the regulation of the expression of ribosomal protein S15. *Proc. Natl Acad. Sci. USA,* **95**, 2564-2567.

Brierley, I. (1995). Ribosomal frameshifting on viral RNAs. *J. Gen. Virol.* **76**, 1885-1892.

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.

Brierley, I., Jenner, A. J. & Inglis, S. C. (1992). Mutational analysis of the ''slippery sequence'' component of a coronavirus ribosomal frameshifting signal. *J. Mol. Biol.* **227**, 463-479.

Carter, E. S., II & Tung, C.-S. (1996). NAMOT2-a redesigned nucleic acid modeling tool: construction of non-canonical DNA structures. *CABIOS,* **12**, 25-30.

Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Kundrot, C. E., Cech, T. R. & Doudna, J. A. (1996a). Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science,* **273**, 1678-1685.

Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Szewczak, A. A., Kundrot, C. E., Cech, T. R. & Doudna, J. A. (1996b). RNA tertiary structure mediation by adenosine platforms. *Science,* **273**, 1696-1699.

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. Natl Acad. Sci. USA,* **89**, 713-717.

Chen, X., Chamorro, M., Lee, S. I., Shen, L. X., Hines, J. V., Tinoco, I., Jr & Varmus, H. E. (1995). Structural and functional studies of retroviral RNA pseudoknots involved in ribosomal frameshifting: nucleotides at the junction of the two stems are important for efficient ribosomal frameshifting. *EMBO J.* **14**, 842-852.

Chen, X. Y., Kang, H. S., Shen, L. X., Chamorro, M., Varmus, H. E. & Tinoco, I. (1996). A characteristic bent conformation of RNA pseudoknots promotes −1 frameshifting during translation of retroviral RNA. *J. Mol. Biol.* **260**, 479-483.

Dinman, J. D. (1995). Ribosomal frameshifting in yeast viruses. *Yeast,* **11**, 1115-1127.

Dinman, J. D., Ruiz-Echevarria, M. J., Czaplinski, K. & Peltz, S. W. (1997). Peptidyl-transferase inhibitors have antiviral properties by altering programmed −1 ribosomal frameshifting efficiencies: development of model systems. *Proc. Natl Acad. Sci. USA,* **94**, 6606-6611.

Du, Z. & Hoffman, D. W. (1997). An NMR and mutational study of the pseudoknot within the gene 32 mRNA of bacteriophage T2: insights into a family of structurally related RNA pseudoknots. *Nucl. Acids Res.* **25**, 1130-1135.

Du, Z. H., Holland, J. A., Hansen, M. R., Giedroc, D. P. & Hoffman, D. W. (1997). Base-pairings within the RNA pseudoknot associated with the simian retrovirus-1 gag-pro frameshift site. *J. Mol. Biol.* **270**, 464-470.

Farabaugh, P. J. (1996). Programmed translational frameshifting. *Microbiol. Rev.* **60**, 103-134.

Gluick, T. C., Gerstner, R. B. & Draper, D. E. (1997). Effects of Mg^{2+}, K^+ and H^+ on an equilibrium between alternative conformations of an RNA pseudoknot. *J. Mol. Biol.* **270**, 451-463.

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.

Jackson, R. J. & Hunt, T. (1983). Preparation and use of nuclease-treated rabbit reticulocyte lysates for the translation of eukaryotic messenger RNA. *Methods Enzymol.* **96**, 50-74.

Kang, H. (1998). Direct structural evidence for formation of a stem-loop structure involved in ribosomal fra-

meshifting in human immunodeficiency virus type 1. *Biochim. Biophys. Acta,* **1397**, 73-8.

Kang, H. S. & Tinoco, I. (1997). A mutant RNA pseudoknot that promotes ribosomal frameshifting in mouse mammary tumor virus. *Nucl. Acids Res.* **25**, 1943-1949.

Kang, H. S., Hines, J. V. & Tinoco, I. (1996). Conformation of a non-frameshifting RNA pseudoknot from mouse mammary tumor virus. *J. Mol. Biol.* **259**, 135-147.

Kielkopf, C. L., Baird, E. E., Dervan, P. B. & Rees, D. C. (1998a). Structural basis for GC recognition in the DNA minor groove. *Nature Struct. Biol.* **5**, 104-109.

Kielkopf, C. L., White, S., Szewczyk, J. W., Turner, J. M., Baird, E. E., Dervan, P. B. & Rees, D. C. (1998b). A structural basis for recognition of AT andTA base pairs in the minor groove of B-DNA. *Science,* **282**, 111-5.

Kolchanov, N. A., Titov, I. I., Vlassova, I. E. & Vlassov, V. V. (1996). Chemical and computer probing of RNA structure. *Prog. Nucl. Acid Res. Mol. Biol.* **53**, 131-196.

Kolk, M. H., van der Graaf, M., Wijmenga, S. S., Pleij, C. W. A., Heus, H. A. & Hilbers, C. W. (1998). NMR structure of a classical pseudoknot: interplay of single- and double-stranded RNA. *Science,* **280**, 434-438.

Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946-950.

Krzyzosiak, W. J., Marciniec, T., Wiewiorowski, M., Romby, P., Ebel, J.-P. & Giege, R. (1988). Characterisation of the lead (II)-induced cleavages in tRNAs in solution and effect of the Y-base removal in yeast tRNA Phe. *Biochemistry,* **27**, 5771-5777.

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

Le, S. Y., Chen, J. H., Pattabiraman, N. & Maizel, J. V. (1998). Ion-RNA interactions in the RNA pseudoknot of a ribosomal frameshifting site: molecular modeling studies. *J. Biomol. Struct. Dynam.* **16**, 1-11.

Marczinke, B., Fisher, R., Vidakovic, M., Bloys, A. J. & Brierley, I. (1998). Secondary structure and mutational analysis of the ribosomal frameshift signal of Rous sarcoma virus. *J. Mol. Biol.* **284,**, 205-225.

Melton, D. A., Krieg, P. A., Robagliati, M. R., Maniatis, T., Zinn, K. & Green, M. R. (1984). Efficient *in vitro* synthesis of biologically active RNA and RNA hybridisation probes from plasmids containing a bacteriophage SP6 promoter. *Nucl. Acids Res.* **12**, 7035-7056.

Merritt, E. A. & Bacon, D. J. (1997). RASTER3D: photorealistic molecular graphics. *Methods Enzymol.* **277**, 505-524.

Napthine, S., Liphardt, J., Bloys, A., Routledge, S. & Brierley, I. (1999). The role of RNA pseudoknot stem 1 length in the promotion of efficient -1 ribosomal frameshifting. *J. Mol. Biol.* **288**, 305-320.

Philippe, C., Portier, C., Mougel, M., Grunberg-Manago, M., Ebel, J. P., Ehresmann, B. & Ehresmann, C. (1990). Target site of *Escherichia coli* ribosomal protein S15 on its messenger RNA. Conformation and interaction with the protein. *J. Mol. Biol.* **211**, 415-426.

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

Polson, A. G. & Bass, B. L. (1994). Preferential selection of adenosines for modification by double-stranded RNA adenosine deaminase. *EMBO J.* **13**, 5701-5711.

Rietveld, K., Van Poelgeest, R., Pleij, C. W., Van Boom, J. H. & Bosch, L. (1982). The tRNA-like structure at the 3′ terminus of turnip yellow mosaic virus RNA. Differences and similarities with canonical tRNA. *Nucl. Acids Res.* **10**, 1929-46.

Rijnbrand, R., van der Straaten, T., van Rijn, P. A., Spaan, W. J. & Bredenbeek, P. J. (1997). Internal entry of ribosomes is directed by the 5′ noncoding region of classical swine fever virus and is dependent on the presence of an RNA pseudoknot upstream of the initiation codon. *J. Virol.* **71**, 1451-457.

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

Shamoo, Y., Tam, A., Konigsberg, W. H. & Williams, K. R. (1993). Translational repression by the bacteriophage T4 gene 32 protein involves specific recognition of an RNA pseudoknot structure. *J. Mol. Biol.* **232**, 89-104.

Shen, L. X. & Tinoco, I. (1995). The structure of an RNA pseudoknot that causes efficient frameshifting in mouse mammary tumor virus. *J. Mol. Biol.* **247**, 963-978.

Somogyi, P., Jenner, A. J., Brierley, I. & Inglis, S. C. (1993). Ribosomal pausing during translation of an RNA pseudoknot. *Mol. Cell. Biol.* **13**, 6931-6940.

Su, L., Chen, L., Egli, M., Berger, J. M. & Rich, A. (1999). Minor groove RNA triplex in the crystal structure of a viral pseudoknot involved in ribosomal frameshifting. *Nature Struct. Biol.* **6**, 285-292.

Sung, D. & Kang, H. (1998). Mutational analysis of the RNA pseudoknot involved in efficient ribosomal frameshifting in simian retrovirus 1. *Nucl. Acids Res.* **26**, 1369-1372.

Tang, C. K. & Draper, D. E. (1989). Unusual mRNA pseudoknot is recognised by a protein translational repressor. *Cell,* **57**, 531-536.

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

ten Dam, E., Pleij, K. & Draper, D. (1992). Structural and functional aspects of RNA pseudoknots. *Biochemistry,* **31**, 11665-11676.

ten Dam, E., Brierley, I., Inglis, S. C. & Pleij, C. (1994). Identification and analysis of the pseudoknot-containing *gag-pro* ribosomal frameshift signal of simian retrovirus-1. *Nucl. Acids Res.* **22**, 2304-2310.

ten Dam, E., Verlaan, P. & Pleij, C. (1995). Analysis of the role of the pseudoknot component in the SRV-1 *gag-pro* ribosomal frameshift signal: loop lengths and stability of the stem regions. *RNA,* **1**, 146-154.

Tu, C., Tzeng, T-H. & Bruenn, J. A. (1992). Ribosomal movement impeded at a pseudoknot required for frameshifting. *Proc. Natl Acad. Sci. USA,* **89**, 8636-8640.

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

van Belkum, A., Verlaan, P., Bing, Kun J., Pleij, C. & Bosch, L. (1988). Temperature dependent chemical and enzymatic probing of the tRNA-like structure of TYMV RNA. *Nucl. Acids Res.* **16**, 1931-1950.

Vlassov, V. V., Zuber, G., Felden, B., Behr, J.-P. & Giege, R. (1995). Cleavage of tRNA with imidazole and spermine imidazole constructs: a new approach for

probing RNA structure. *Nucl. Acids Res.* **23**, 3161-3167.

Wang, C., Le, S. Y., Ali, N. & Siddiqui, A. (1995). An RNA pseudoknot is an essential structural element of the internal ribosome entry site located within the hepatitis C virus 5′ noncoding region. *RNA,* **1**, 526-537.

Wyatt, J. R., Puglisi, J. D. & Tinoco, I. (1990). RNA pseudoknots: stability and loop size requirements. *J. Mol. Biol.* **214**, 455-470.

Young, J. F., Desselberger, U., Graves, P., Palese, P., Shatzman, 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 D. E. Draper

(Received 22 January 1999; received in revised form 28 February 1999; accepted 11 March 1999)