

Chapter 7

Pseudoknot-Dependent Programmed –1 Ribosomal Frameshifting: Structures, Mechanisms and Models

Ian Brierley, Robert J.C. Gilbert, and Simon Pennell

Abstract Programmed –1 ribosomal frameshifting is a translational recoding strategy that takes place during the elongation phase of protein biosynthesis. Frameshifting occurs in response to specific signals in the mRNA; a slippery sequence, where the ribosome changes frame, and a stimulatory RNA secondary structure, usually a pseudoknot, located immediately downstream. During the frameshift the ribosome slips backwards by a single nucleotide (in the 5′-wards/–1 direction) and continues translation in the new, overlapping reading frame, generating a fusion protein composed of the products of both the original and the –1 frame coding regions. In eukaryotes, frameshifting is largely a phenomenon of virus gene expression and associated predominantly with the expression of viral replicases. Research on frameshifting impacts upon diverse topics, including the ribosomal elongation cycle, RNA structure and function, tRNA modification, virus replication, antiviral intervention, evolution and bioinformatics. This chapter focuses on the structure and function of frameshift-stimulatory RNA pseudoknots and mechanistic aspects of ribosomal frameshifting. A variety of models of the frameshifting process are discussed in the light of recent advances in our understanding of ribosome structure and the elongation cycle.

Contents

7.1	Introduction	150
7.2	The Nature, Occurrence and Role of Ribosomal Frameshifting	152
7.3	The Structure of Ribosomal Frameshifting Signals	153
7.4	Stimulatory RNA Structures	153
7.4.1	The IBV Pseudoknot and Relatives	154
7.4.2	The RSV Pseudoknot and Interstem Elements	158

I. Brierley (✉)
Division of Virology, Department of Pathology, University of Cambridge, Tennis Court Road,
Cambridge CB2 1QP, UK
e-mail: ib103@mole.bio.cam.ac.uk

7.4.3	The “Kinked” Pseudoknots	159
7.4.4	The Luteoviral Pseudoknots and Loop–Helix Interactions	160
7.4.5	Long-Range Pseudoknots	160
7.5	Mechanistic Aspects of Ribosomal Frameshifting	160
7.5.1	tRNA Slippage	161
7.5.2	Ribosomal Pausing	161
7.5.3	The Stimulatory RNA Resists Unwinding by the Ribosome	162
7.5.4	A Role for trans-Acting Protein Factors?	164
7.5.5	Conceivable Points for Frameshifting During the Elongation Cycle	164
7.6	Models of Frameshifting	165
7.6.1	The Integrated and 9Å Models	165
7.6.2	The Simultaneous Slippage Model	167
7.6.3	The Dynamic Model	167
7.6.4	The Mechanical Model	168
7.6.5	The Three-tRNA Model	170
7.7	Perspective	171
	References	172

7.1 Introduction

The discovery of programmed -1 ribosomal frameshifting (hereafter frameshifting for brevity) is intimately entwined with studies of retrovirus gene expression. In the early 1980s, as the first complete retrovirus genomic sequences began to appear, it was clear that in many cases the open reading frames (ORFs) encoding the structural and enzymatic components of the virion, *gag* and *pol*, respectively, overlapped at the 3' end of *gag* with *pol* in the -1 reading frame (Fig. 7.1). As most retroviruses were known to express Pol solely as a C-terminal extension of Gag, it was thought that RNA processing would produce an in-frame spliced mRNA capable of expressing the Gag-Pol polyprotein. However, in papers describing the complete sequence of Rous sarcoma virus (RSV; Schwartz et al., 1983) and a bovine leukemia virus provirus (Rice et al., 1985) an alternative suggestion was put forward that Gag and Pol could be coupled at the translational level, through a frameshifting mechanism. Shortly afterwards, Tyler Jacks and Harold Varmus published their landmark paper confirming expression of the RSV Gag-Pol polyprotein by frameshifting (Jacks and Varmus, 1985). Since this time, frameshift signals have been described in viruses of all kingdoms and in a growing number of cellular genes. Research on frameshifting impacts upon diverse topics, including the ribosomal elongation cycle, RNA structure and function, tRNA modification, virus replication, antiviral intervention, evolution and bioinformatics. This chapter will focus on the mRNA structures involved in eukaryotic frameshifting and the mechanistic insights that have been gained over the last 20 years. Subsequent chapters in this book address specifically the human immunodeficiency virus (HIV) frameshifting signal, and also frameshifting in prokaryotes (including bacteriophages and insertion sequence elements), yeast and plants and so will not be discussed in detail here. Interested readers may also like to look at other reviews on this fascinating topic (Farabaugh, 1996;

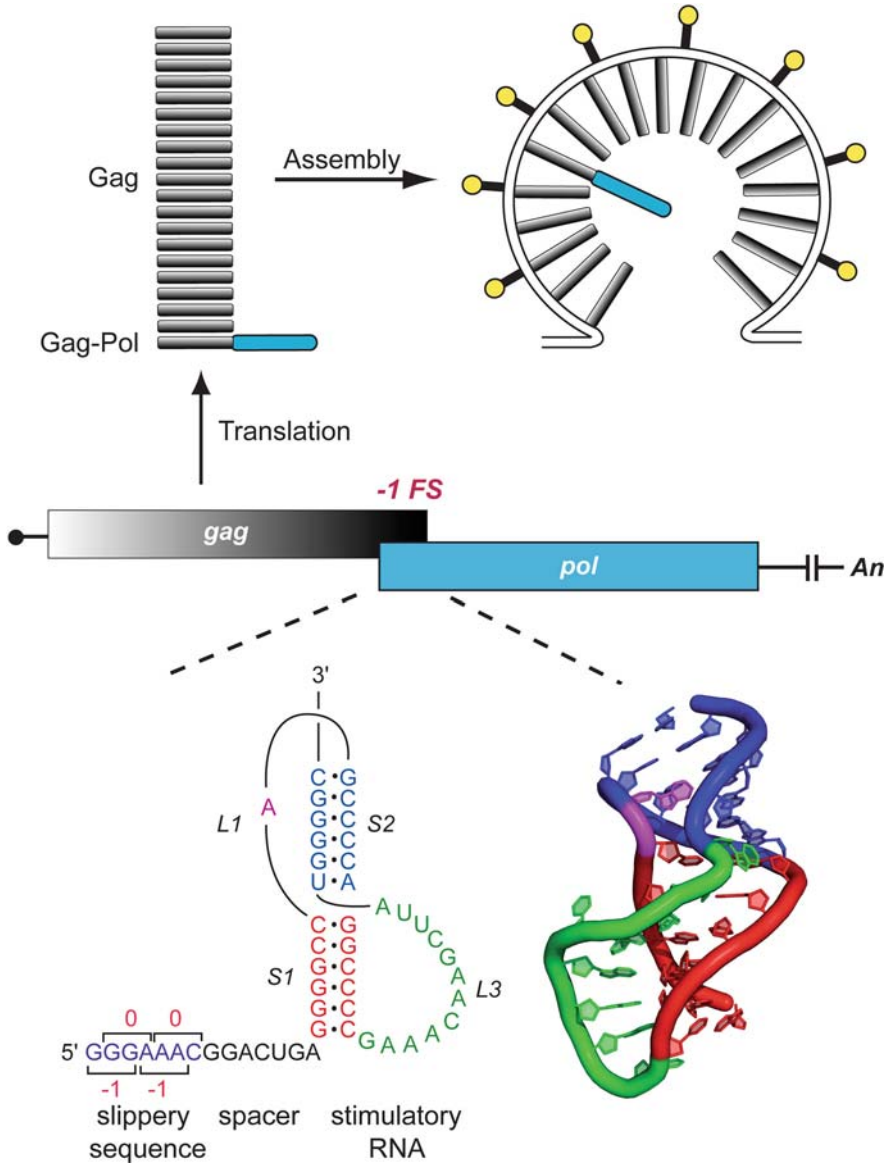


Fig. 7.1 Programmed -1 ribosomal frameshifting. In the centre, a schematic of a retrovirus genomic mRNA is shown with the 5'-proximal *gag* and overlapping *pol* coding sequences boxed. Translation of this mRNA (*upper panel*) yields mostly the Gag polyprotein, but occasional -1 frameshifting (-1FS) generates the Gag-Pol polyprotein. Assembly at the plasma membrane (*right*) includes the Pol domain in virions. The viral envelope glycoproteins are indicated as yellow and black spikes. Frameshifting (*lower panel*) is mediated by a signal composed of a slippery sequence (*purple*), a spacer region (*black*) and a downstream stimulatory RNA. The *gag* (0) and *pol* (-1) reading frames are indicated. This example shows the SRV-1 *gag/pro* stimulatory pseudoknot as a secondary structure (*left*) and as solved by NMR (*right*, Michiels et al., 2001; image prepared using PyMOL [www.pymol.org])

2000; Giedroc et al., 2000; Brierley and Pennell, 2001; Brierley and Dos Ramos, 2006; Dreher and Miller, 2006; Baranov et al., 2006; Giedroc and Cornish, 2008).

7.2 The Nature, Occurrence and Role of Ribosomal Frameshifting

Frameshifting is a translational recoding strategy that takes place during the elongation phase of protein biosynthesis. In response to signals in the mRNA, and at a certain frequency, ribosomes switch from the translation of one ORF (in the 0-frame) to an overlapping ORF (in the -1 reading frame) and continue translation, generating, in the majority of cases*, a fusion protein containing the products of both the upstream and the downstream ORFs (Fig. 7.1) (*frameshifting in the *dnaX* gene of *Escherichia coli* redirects ribosomes to a nearby termination codon in the -1 frame; in this case, and probably in some cellular examples, the frameshifting ribosomes produce a shorter product than that encoded by the 0-frame (reviewed in Chapter of Fayet and Prère, and see below). Frameshifting in eukaryotes is largely a phenomenon of virus gene expression and associated predominantly with the expression of viral replicases. Thus in retroviruses, frameshifting at the overlap of *gag* and *pol* permits expression of the viral reverse transcriptase (as a component of the Gag-Pol polyprotein) and in other RNA viruses, frameshifting mediates expression of RNA-dependent RNA polymerases (Brierley et al., 1995; Dreher and Miller et al., 2006). As the expression of distal ORFs on viral mRNAs can be achieved in numerous ways (Pe'ery and Mathews, 2000), it is pertinent to question why a frameshift strategy has evolved. One explanation likely to be of general relevance is that frameshifting allows the generation of a defined ratio of protein products. In HIV-1, for example, modulation of the cytoplasmic Gag:Gag-Pol ratio is detrimental to replication (Shehu-Xhilaga et al., 2001; reviewed in Brierley and Dos Ramos, 2006) and the same is true for the retrovirus-like dsRNA virus of yeast, L-A (Dinman and Wickner, 1992). Similarly, in other RNA viruses, changing the stoichiometry of non-frameshifted and frameshifted products is also likely to be disadvantageous. Thus frameshifting has emerged as a potential target for antiviral therapeutics (Dinman et al., 1998). The frameshifting process also generates a fusion protein which itself may be biologically relevant, for example, in the incorporation of retroviral reverse transcriptase into virions (Fig. 7.1).

Examples of frameshifting in eukaryotic cellular genes have begun to emerge in the last few years. The signals present in the mouse embryonal carcinoma differentiation regulated (Edr, now known as PEG10) gene (Shigemoto et al., 2001; Manktelow et al., 2005) and the human paraneoplastic Ma3 gene (Wills et al., 2006) are clearly of retroviral origin, perhaps subsumed into the cellular genome from an endogenous retrovirus or retroviral relic. However, many candidates in the yeast genome that were identified by a bioinformatics approach appear to be genuinely unrelated to retroviruses, standing alone as the first true examples of novel cellular -1 ribosomal frameshifting signals (Jacobs et al., 2007). In contrast to viral frameshifting, the vast majority of these, and those predicted in other genomes

(Belew et al., 2008), direct the ribosome to a premature termination codon and may have a biological role in destabilising the mRNA (Plant et al., 2004).

7.3 The Structure of Ribosomal Frameshifting Signals

Eukaryotic ribosomal frameshift signals are composed of two essential elements: a “slippery” sequence, where the ribosome changes reading frame, and a stimulatory RNA secondary structure, often an RNA pseudoknot, located a few nucleotides (nt) downstream (Jacks et al., 1988; Brierley et al., 1989; ten Dam et al., 1990; see Fig. 7.1). A spacer region between the slippery sequence and the stimulatory RNA is also required and the precise length of this spacer must be maintained for maximal frameshifting efficiency (Brierley et al., 1989, 1992; Kollmus et al., 1994). The slippery sequence is a heptanucleotide stretch that contains two homopolymeric triplets and conforms in the vast majority of cases to the motif XXXYYYYZ (for example, UUUAAAC in the coronavirus infectious bronchitis virus (IBV) *Ia/Ib* signal). Frameshifting at this sequence has been proposed to occur by “simultaneous-slippage” of the ribosome-bound peptidyl- and aminoacyl-tRNAs, which detach from the zero frame codons (XXY YYZ) and re-pair in the -1 phase (XXX YYY) (Jacks et al., 1988). The homopolymeric nature of the slippery sequence seems to be required to allow the tRNAs to remain base-paired to the mRNA in at least two out of three anticodon positions following the slip. Frameshift assays, largely carried out *in vitro*, have revealed that XXX can be represented by any homopolymeric triplet (A, C, G or U), but the Y triplet must be AAA or UUU (Jacks et al., 1988, Dinman et al., 1991; Brierley et al., 1992). In addition to these restrictions, slippery sequences ending in G (XXXAAAG or XXXUUUG) do not function efficiently in *in vitro* translation systems (Brierley et al., 1992), yeast (Dinman et al., 1991) or mammalian cells (Marczinke et al., 2000). At naturally occurring frameshift sites, of the possible codons which are decoded in the ribosomal A-site prior to tRNA slippage (XXXYYYYZ), only five are represented in eukaryotes, AAC, AAU, UUA, UUC, UUU. Together with the *in vitro* data, it seems that the sequence restrictions observed are a manifestation of the need for the pre-slippage codon–anticodon complex in the A-site to be weak enough such that the tRNAs can detach from the codon during the process of frameshifting. G-C pairs are therefore avoided. Recent work suggests that the slippery sequence actually forms part of a somewhat larger motif, as certain bases immediately 5' of the heptanucleotide stretch can influence the efficiency of frameshifting (Bekaert and Rousset, 2005; Léger et al., 2007). While this has not been examined systematically, these studies suggest an involvement of the ribosomal E-site in frameshifting (see below and the Chapter by Pech and colleagues).

7.4 Stimulatory RNA Structures

Slippery sequences alone can direct frameshifting (typically to 0.1–0.2% in mammalian cells; e.g. Marczinke et al., 2000; Dulude et al., 2002) to levels significantly

higher than the estimated natural frameshift error rate (10^{-5} per codon for *E. coli* ribosomes; Kurland, 1992), indicating that these stretches are intrinsically frameshift-prone. However, to obtain the levels of frameshifting needed for their biological function (1–30% depending on the particular system), a downstream stimulatory RNA structure is also required (see Table 7.1). In eukaryotes, some stem-loop stimulatory RNAs have been described, such as those present at the *Ia/Ib* frameshift site of astroviruses (Marczinke et al. 1994) and the HIV-1 *gag/pol* overlap (discussed in more detail in the Chapter by Brakier-Gingras), but most commonly an H (hairpin)-type RNA pseudoknot is present (Brierley et al., 1989; ten Dam et al., 1990). These form when a single-stranded region in the loop of a hairpin base pairs with a stretch of complementary nucleotides elsewhere in the RNA chain (reviewed in Brierley et al., 2007; see Fig. 7.1). Such pseudoknots have two base-paired stem regions (S1 and S2) and, depending upon the number of loop bases that participate in the pseudoknotting interaction, two or three single-stranded loops (L1, L2 and L3). In almost all frameshift-promoting pseudoknots, L2 is absent (Table 7.1, Fig. 7.1) or very short (Fig. 7.2) and the base-paired stems stack to form a quasi-continuous helix. In these structures, L1 spans S2 and crosses the deep groove of the helix, whereas L3 spans S1 and crosses the shallow groove. While the precise folding of and, to some extent, the stability of pseudoknots has been shown to be critical to the frameshift process, few specific primary sequence requirements have been identified. This suggests that they are involved in few, if any, sequence-specific interactions, as evidenced by the failure to identify protein binding partners (see below). Where particular sequence requirements have been suspected, it may be that the bases concerned are involved in unanticipated tertiary interactions. The analysis of pseudoknot structure and function has involved a range of experimental approaches, including in vitro and in vivo structure–function assays, biophysical investigation of folding/unfolding and atomic resolution structure determination. Pseudoknots have been classified on the basis of secondary structural features (Brierley and Pennell, 2001) or by virus group (Giedroc and Cornish, 2008), but as our knowledge is incomplete, this is mostly for convenience of description. For the purposes of this review, examples of specific pseudoknot “classes” will be presented as they were described historically and linked, where possible, to related examples described more recently.

7.4.1 The IBV Pseudoknot and Relatives

A role for pseudoknots in frameshifting was first described for the coronavirus infectious bronchitis virus (IBV) *Ia/Ib* signal (Brierley et al., 1989; see Fig. 7.2). This pseudoknot is characterised by the possession of a long, stable S1, a coaxially stacked S2 and a capacity to direct high levels of frameshifting, up to ~60% in the rabbit reticulocyte lysate (RRL) in vitro translation system (Naphthine et al., 1999). Although no atomic resolution structural model is available, several features important for function have been identified. The most striking of these is the importance of maintaining S1 length. Naphthine and colleagues (1999) prepared a series of

Table 7.1 Details of established -1 ribosomal frameshift signals

Origin	System	Slippery site	DS	SP	S1	L1	S2	L2/ISE	L3	Comments
Stem-loop stimulatory RNA										
HAst-V1	orf1a/orf1b	Astrovirus	SL	6	7	10				A short stimulatory RNA
GLV	p100/p190	Totivirus	SL	5	9	11				Rare slippery sequence
HTLV-I	gag/pro	Deltaretrovirus	SL	7	10	3				HTLV-II gag/pro signal is almost identical
CfMV	orf2a/orf2b	Sobemovirus	SL	7	11	4				Flanking sequences may stimulate frameshifting (~twofold)
SIV	gag/pol	Lentivirus	SL	4 (9)*	16 (11)*	12				Single helix with ordered loop, NMR
HIV-1	gag/pol	Lentivirus	SL	0 (8)	19 (11)*	4				Two-stem helix with a purine-rich bulge, NMR
Pseudoknot stimulator with short S1										
PLRV	orf2a/orf2b	Polerovirus	PK	5	4	2	3	1	9	PK like that of BWYV, X-ray
BWYV	orf2/orf3	Polerovirus	PK	6	5	2	3	1	7	Compact, loop-helix, non-Watson-Crick, 1 st X-ray
ScYLV	P1/P2	Polerovirus	PK	6	5	2	3	1	9	PK like that of BWYV, NMR
PEMV	P1/P2	Polerovirus	PK	7	5	2	3	1	8	PK like that of BWYV, NMR
EIAV	gag/pol	Lentivirus	PK	9	5	4	5	0	9	Long SP, L3 may interact with S1
MMTV	gag/pro	Betaretrovirus	PK	7	5	1	7	1	8	Kinked PK, 1 st NMR
FIV	gag/pol	Lentivirus	PK	8	5	2	5	1	12	From MS3D, likely to be kinked like MMTV PK
SRV-1	gag/pro	Betaretrovirus	PK	7	6	1	6	0	12	Coaxially stacked, no kink, NMR
IAP	gag/pro	Endogenous RV	PK	8	6	1	8	0	9	PK not well characterised
HERV	K-10 gag/pro	Endogenous RV	PK	7	6	1	6	0	9	NMR, L3 may interact with S1
VMV	gag/pol	Lentivirus	PK	6	7	5	7	7	14	Contains unstructured ISE (cf. RSV, GAV)
Pseudoknot stimulator with long S1										
PEG10	(Edr)	Cellular gene	PK	5	10	3	9	0	9	1 st non-viral example of -1 frameshifting
RF1/RF2										Non-viral example identified by bioinformatics
Ma-3	orf1/orf2	Cellular gene	PK	6	11	3	5	1	10	

Table 7.1 (continued)

Origin	System	Slippery site	DS	SP	S1	L1	S2	L2/ISE L3	Comments	
IBV orf1a/1b	Coronavirus	UUUAAAC	PK	6	11	2	6	0	32	First identified frameshift-promoting PK
MHV orf1a/1b	Coronavirus	UUUAAAC	PK	5	11	2	11	0	30	Long S2 predicted
SARS-CoV orf1a/1b	Coronavirus	UUUAAAC	PK	5	11	3	7	0	29	Substructure in L3
BEV orf1a/1b	Arterivirus	UUUAAAC	PK	5	11	4	5	0	69	Most likely PK is shown
EAV orf1a/1b	Arterivirus	GUUAAAC	PK	6	11	2	6	1	69	Unusual slippery sequence
HCV orf1a/1b	Coronavirus	UUUAAAC	PK	5	12	3	5	0	164	S2 formed by kissing interaction
L-A gag/pol	Totivirus	GGGUUUA	PK	5	13	8	6	2	11	Yeast dsRNA virus signal used as model system
HTLV-I pro/pol	Deltaretrovirus	UUUAAAC	PK	6	13	6	10	0	20	Girmary & Brierley, unpublished
TGEV orf1a/1b	Coronavirus	UUUAAAC	PK	3	14	8	5	3	163	S2 formed as for HCV, precise SP/S1
RSV gag/pol	Alpharetrovirus	AAAUUUA	PK	1 (6)	14 (9)	11	8	53	17	length unknown
GAV orf1a/orf1b	Nidovirales	AAAUUUU	PK	8	22	3	7	64	41	Contains structured ISE
										A very large PK with substantial S1 and ISE
Pseudoknot stimulator with long-range interaction										
BYDV39K/60 K	Luteovirus	GGGUUUU	PK	5	8	57	6	0	3828	Unusual kissing PK, very long L3, see Fig. 7.2
RCNMV p27/p57	Dianthovirus	GGAUUUU	PK	5	12	46	6	0	2591	Similar to BYDV
Pseudoknot stimulator unclassified										
HIV gag/pol (group O)	Lentivirus	UUUUUUA	PK	8	8	2	8	0	20	PK features reminiscent of a viral PK that promotes stop codon suppression (Wills et al., 1991)

The table lists those established frameshift signals where details of the stimulatory RNA are available. Abbreviations used are DS, downstream stimulator; SP, spacer region (length in nt); S1, stem 1 (bp; unpaired residues within stems are not included in the total length); L1, loop 1 (nt); S2, stem 2 (bp); L2/ISE, loop 2/interstem element (nt); L3, loop 3 (nt); SL, stem-loop structure; PK, pseudoknot; NMR or X-ray, nuclear magnetic resonance or X-ray structure available; MS3D, three-dimensional mass spectrometric structural analysis; RV, retrovirus.

* Some or all of the SP may be incorporated into S1 (bracketed values show length of unpaired SP/S1).

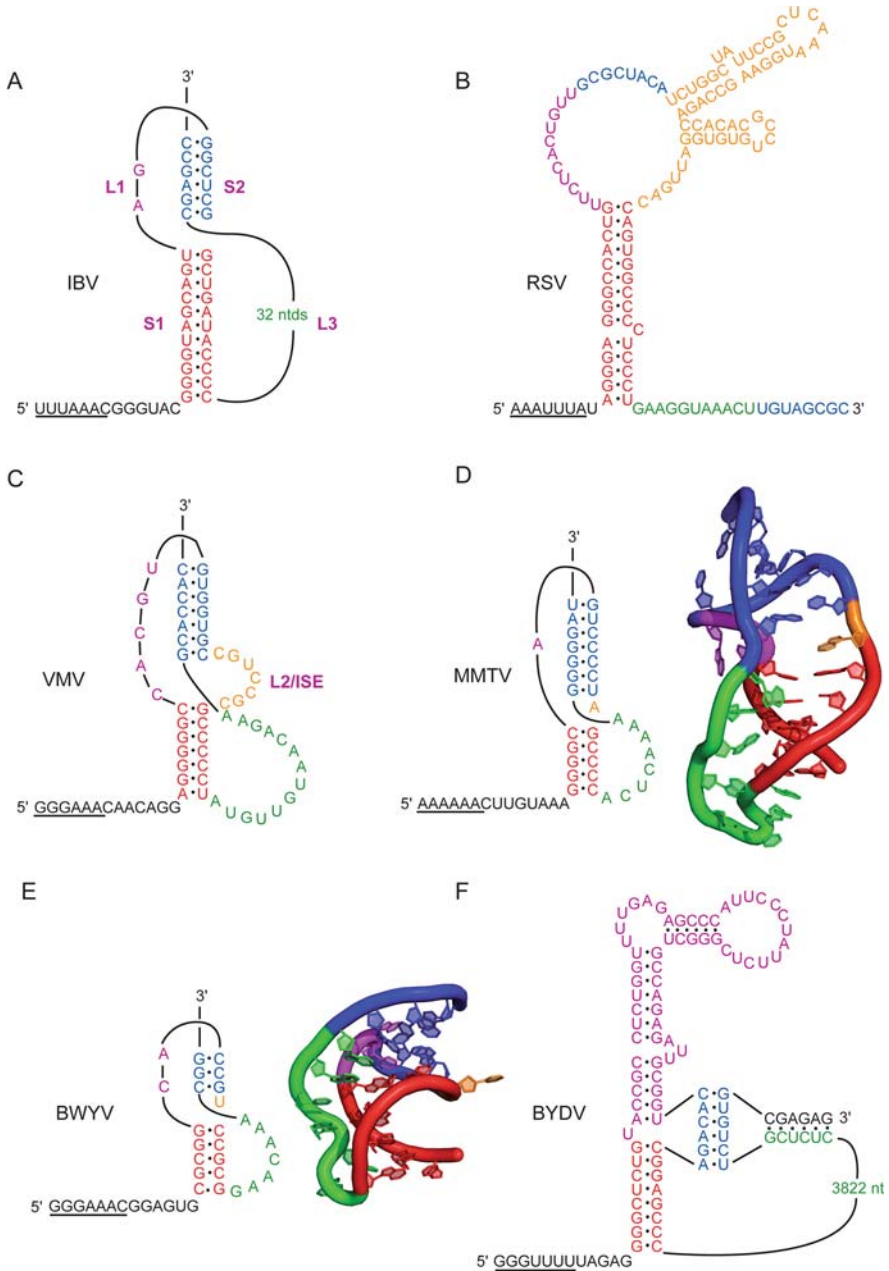


Fig. 7.2 Examples of RNA pseudoknot structures found at viral frameshift sites. Secondary structure models of a variety of pseudoknot stimulatory RNAs are shown, alongside three-dimensional models where available (S1, red; S2, blue; L1, purple; L2, orange; L3 green; details in Table 7.1). (A) Infectious bronchitis virus (IBV), (B) Rous sarcoma virus (RSV), (C) Visna-Maedi virus (VMV), (D) mouse mammary tumour virus *gag/pro* (MMTV), (E) beet western yellows virus (BWYV) and (F) barley yellow dwarf virus (BYDV). The slippery sequences are underlined. For simplicity of drawing, S2 of the RSV pseudoknot is not shown in base-paired representation

IBV-based pseudoknots in which this stem was varied from 4 to 13 base pairs (bp). Remarkably, frameshifting occurred very inefficiently until the wild-type length of 11 bp was reached, after which full function was retained. This phenomenon was unrelated to the thermodynamic stability of the stem, as the added bases were G-C pairs and the shorter stems of 7–10 bp were as stable as or more stable than the wild-type 11 bp stem. Whether there is any significance to the fact that 11 bp corresponds to a complete turn of an A-form RNA helix remains to be ascertained. A similar puzzle is the clear benefit of a G-rich stretch at the beginning of the 5' arm of S1. We assume that these features are important in terms of how the pseudoknot is recognised and unwound by the ribosome during the frameshift. The 10 and 11 bp variants have been studied in mechanical unfolding experiments but have produced somewhat contradictory data; in one study the 11 bp construct required twofold greater unfolding force (Hansen et al., 2007) but in another, there was no obvious correlation between frameshift efficiency and the mechanical force required to unwind the pseudoknots (Green et al., 2008). The latter study revealed, however, that pseudoknots have unusual mechanical properties (discussed below). L3 of the IBV pseudoknot can be reduced in length to eight bases (from 32), close to the minimum number required to span S1 (Pleij et al., 1985), with no loss of function (Brierley et al., 1991). A lack of obvious primary sequence requirements within L3 argues against the formation of a L3–S1 interaction as is seen in some other pseudoknots (see below). The frameshift signal of the severe acquired respiratory syndrome (SARS) coronavirus is very similar to that of IBV although L3 is mostly folded into a stem-loop, which has alternatively been named SL1 or S3 (Baranov et al., 2005; Plant et al., 2005). Unexpectedly, although most of this loop can be deleted without effect, maintenance of an appropriate conformation of the wild-type L3 seems to be required for optimal frameshift efficiency (Baranov et al., 2005; Plant et al., 2005). Pseudoknots containing an SL1 motif can be predicted for all group 2 coronaviruses (Plant et al., 2005). In group 3 coronaviruses (e.g. IBV), however, the region equivalent to SL1 is probably a single-stranded loop (Brierley et al., 1989, 1991; Plant et al., 2005). The pseudoknots present at the *1a/1b* overlap of group 1 coronaviruses (e.g. human coronavirus 229E) appear to form a more “elaborated” pseudoknot that can be viewed as “kissing” stem-loops separated by a long (~150 nt) L3 (Herold and Siddell, 1993). The stimulatory RNA is still a pseudoknot but in this case, the downstream sequence that pairs with the loop is itself constrained within a hairpin.

7.4.2 The RSV Pseudoknot and Interstem Elements

The RSV *gag/pol* frameshifting signal has the distinction of being the first example of both –1 frameshifting and pseudoknot-dependent frameshifting (Jacks et al., 1988), although the pseudoknot was confirmed some years after the original report (Marczinke et al., 1998). The RSV stimulatory RNA is one of the most complex found to date and belongs to a family of frameshift-promoting pseudoknots that possess a L2 and thus have an element at the junction of the two stems (an interstem

element, or ISE; Fig. 7.2). L2 of the RSV pseudoknot is 70 nt in length most of which is constrained within a substructure of two adjacent stem-loops (Fig. 7.2B). How the ISE folds within the pseudoknot is unknown. A similarly extensive ISE is predicted in the pseudoknot of gill-associated virus (GAV), a prawn-infecting member of the order *Nidovirales* (which includes coronaviruses; Cowley et al., 2000). The GAV pseudoknot is interesting both from the perspective of its sheer size (a total of ~170 nt, of which L1 and L3 contribute only 3 and 41 nt, respectively) and its resemblance to the RSV signal, with a huge S1 (22 bp), an extensive ISE and the use of an RSV-like slippery sequence (AAAUUUU; Table 7.1). The role of these large ISEs in frameshifting, if any, remains to be clarified, but they could add to pseudoknot stability or present additional challenges to a ribosome-associated heli-case. In principle, they could also offer a binding site for translation components or regulatory proteins of viral (or cellular) origin.

The coaxial stacking of stems to generate a quasi-continuous helix has long been considered important for pseudoknot function (ten Dam et al., 1992) and the paucity of L2-containing pseudoknots is consistent with the hypothesis. However, recent structural studies (see below) indicate that the key issue is more likely to be the maintenance of an appropriate architecture at the helical junction. For the RSV and GAV stimulatory RNAs, we predict therefore that the ISEs do not compromise this important region of the pseudoknot. In support of this, we recently characterised a shorter, unstructured ISE (7 nt) in the pseudoknot of the lentivirus Visna-Maedi virus (VMV) (Pennell et al., 2008). The ISE of the VMV pseudoknot proved to be essential for frameshifting and is therefore a key component of the folded, active pseudoknot (Fig. 7.2C).

7.4.3 The “Kinked” Pseudoknots

The first frameshift-promoting pseudoknot to be described in three dimensions was derived from the mouse mammary tumour virus (MMTV) *gag/pro* overlap (Shen and Tinoco 1995; Fig. 7.2D). This pseudoknot is characterised by the presence of an unpaired, intercalated adenosine residue between two short stems of similar length. The presence of this adenosine (A14) forces the stems to bend relative to each other and this “kink” is thought to be important for function (Chen et al., 1996, Kang et al., 1996). However, the general importance of the kink was challenged with the release of the NMR structure of the closely related pseudoknot from simian retrovirus 1 (SRV-1) *gag/pro* (Michiels et al., 2001). This pseudoknot was predicted from its primary sequence to have an organisation similar to that of MMTV and was presumed to have a kink. However, it was shown that the stems were in fact coaxially stacked, with a base predicted to be in L1 being present at the stem–stem junction and base pairing to the “intercalated” adenosine (Fig. 7.1). As detailed further below, this observation adds to the general belief that for most, if not all, pseudoknots, the specific interactions and resultant architectures of the helical junction that are required for frameshifting are strongly context dependent (Cornish et al., 2006).

7.4.4 The Luteoviral Pseudoknots and Loop–Helix Interactions

Luteovirus frameshift-promoting pseudoknots are small, typically encompassing only ~25 nt, with an especially short S2 of three base pairs. Their small size and unexpected stability have made them amenable to structural studies and numerous three-dimensional models are available, including the only frameshift pseudoknot structures solved by X-ray crystallography (Fig. 7.2E). The structure of the pseudoknot of beet western yellows virus (BWYV) reveals a tightly folded motif with extensive non-Watson–Crick base pairing, explaining how these structures have a greater stability than would be predicted from their secondary structure (Su et al., 1999). The main stabilising feature is the presence of an extended triplex formed between the minor groove of S1 and L3, mediated predominantly by A-minor interactions. In addition, the helical junctions of these pseudoknots are bounded by base triple and base quadruple interactions which stabilise the fold and are important for frameshifting (Kim et al., 1999). Several luteoviral pseudoknot structures have been solved, including pea enation mosaic virus (Nixon et al., 2002), potato leaf roll virus (Pallan et al., 2005) and sugar cane yellow leaf virus (ScYLV; Cornish et al., 2005) and all are similar structures. The structures of mutants of the ScYLV pseudoknot have been especially informative (Cornish et al., 2006) revealing that, at least in this class of pseudoknot, the global “ground-state” structure is not strongly correlated with frameshifting efficiency. Giedroc and Cornish (2008) propose that the helical junction is mechanically stable and functions as a kinetic barrier to force-induced unfolding. Frameshift efficiency would therefore be affected by different kinetics of unfolding, diminishing in those pseudoknots that unfold at lower applied forces. Luteoviral frameshift-promoting pseudoknots (and other plant stimulatory RNAs) are discussed in more detail elsewhere in this volume (Chapter of Miller and Giedroc).

7.4.5 Long-Range Pseudoknots

Perhaps the most remarkable example of a non-canonical frameshift pseudoknot comes from the recoding signal of barley yellow dwarf virus (BYDV) (Paul et al., 2001). Initially thought to be a stem-loop stimulator, the BYDV signal was later found to be a “kissing loop” pseudoknot, with the sequence required for S2 formation located some 4 kb downstream (Fig. 7.2F). In addition to being the longest L3 sequence identified to date, the position of the interaction near the 3' end of the genome directly links the control of translation to genomic replication (Barry and Miller 2002). This topic is discussed further in the Chapter by Miller and Giedroc.

7.5 Mechanistic Aspects of Ribosomal Frameshifting

The precise mechanism of programmed -1 ribosomal frameshifting remains unclear. On the surface it seems relatively straightforward; the ribosome encounters the stimulatory RNA structure while in the act of decoding the slippery sequence,

something perturbs normal frame maintenance and a proportion of the ribosomes enter the -1 frame. However, the details are entwined with the natural movements of tRNAs, ribosomal subunits, ribosomal proteins and elongation factors, the dynamics of which we have only recently begun to understand at the molecular level. The frameshift process is most often considered within the framework of the simultaneous tRNA slippage model, specifically how the movement of tRNAs into the -1 frame occurs and how this can be influenced by the stimulatory RNA. Within this context, a role of the stimulatory RNA in ribosomal pausing or in the recruitment of *trans*-acting components has also been considered, along with the specific RNA structural features that could promote these events. In the following sections, these general themes are debated prior to an examination of the different models of frameshifting that have been proposed.

7.5.1 *tRNA Slippage*

In the ribosomal elongation cycle, tRNAs occupy a series of specific positions in the intersubunit space, known as the aminoacyl- (A), peptidyl- (P) and exit- (E) sites. Translocation of tRNAs between these sites, while still attached to the mRNA due to codon-anticodon recognition, combined with the close fit of the tRNAs to the intersubunit space, is the basis of ribosomal processivity along the mRNA and the maintenance of a single reading frame (Moazed and Noller, 1989; Rodnina et al., 1997). During a -1 frameshift, the anticodons of the tRNAs (or tRNA) must detach from the mRNA and re-associate in the -1 frame, and this raises several issues for the design of models. For example, at which stage in the ribosomal elongation cycle does the frameshift occur? What are the parameters involved in promoting dissociation of the tRNAs? Do the tRNAs truly slip simultaneously or does it occur sequentially? In the re-binding phase, do the anticodons sample the -1 , zero or (even) the $+1$ frame and re-pair in the energetically most favourable phase, or are the tRNAs directed specifically to the -1 frame? Another relevant issue is how the mechanisms of frame maintenance inherent to the ribosome are overcome. In the discussion of putative models for -1 frameshifting detailed below, these issues will be addressed where possible. Many of these questions are also relevant to other recoding events and are debated extensively in other chapters of this book.

7.5.2 *Ribosomal Pausing*

Ribosomal pausing has been central to models of -1 frameshifting (and indeed, most recoding phenomena) where it may act to increase the time that ribosomes engage with the slippery sequence, promoting tRNA movements that would normally be unfavourable kinetically (Farabaugh, 2000). Polypeptide intermediates corresponding to ribosomes paused at RNA pseudoknots have been detected at the frameshift sites of RSV (unpublished data cited in Jacks et al., 1988), IBV (Somogyi et al., 1993) and L-A (Lopinski et al., 2000) and footprinting studies of elongating ribosomes have defined the site of pausing at the L-A (Tu et al., 1992; Lopinski

et al., 2000), IBV, SRV-1 and pKA-A signals (Kontos et al., 2001). Pausing at the frameshift signal in the *E. coli dnaX* gene has also been witnessed, again from examination of translational intermediates (Tsuchihashi, 1991). Consistent with a role for pausing in frameshifting, the footprinting studies reveal that the ribosome is stalled over the slippery sequence while in contact with, and perhaps having unwound partially, the pseudoknot (Tu et al., 1992; Kontos et al., 2001). One of the virtues of the pausing model is that it can accommodate the variety of stimulatory RNAs present at -1 frameshifting signals. Notwithstanding the range of secondary and tertiary features presented to the ribosome, as long as pausing occurs, frameshifting results. Unfortunately, the idea that a pause alone is sufficient to induce frameshifting is highly questionable. Simple provision of a roadblock to ribosomes in the form of stable RNA hairpins (Brierley et al., 1991; Somogyi et al., 1993), a tRNA (Chen et al., 1995) or even different kinds of RNA pseudoknot (Naphine et al., 1999; Liphardt et al., 1999) is not sufficient to bring about frameshifting. Furthermore, non-frameshifting pseudoknots and stem-loops exist that can still pause ribosomes (Tu et al., 1992; Somogyi et al., 1993; Lopinski et al., 2000; Kontos et al., 2001). These experiments, of course, do not rule out a contribution of pausing to the mechanism of frameshifting since there are no documented examples where frameshifting has occurred in the absence of a detectable pause. Indeed, one cannot ignore the possibility that a precise “kinetic pause” is required for frameshifting (Bidou et al., 1997), which only certain stimulatory RNAs can generate. For example, during a -1 frameshift, two pauses could occur, one productive (in terms of frameshifting) upon initial encounter of the stimulatory RNA structure, and a second, non-productive pause, corresponding to a delay in unwinding of the structure *after* the crucial event in frameshifting has taken place. The magnitude of the initial pause would perhaps influence the extent of the frameshift, whereas the second pause, occurring during the time that the ribosomal “unwinding activity” (see below) deals with the secondary structure, would be irrelevant. The pausing assays employed currently cannot distinguish between two such pausing events and a detailed analysis of the kinetics of pausing will require further experimentation, including the analysis of translational elongation at the level of individual ribosomes (Wen et al., 2008).

7.5.3 The Stimulatory RNA Resists Unwinding by the Ribosome

As suggested several years ago (Brierley et al., 1991; Wills et al., 1991), the frameshift mechanism is likely to be linked to an unwinding activity of the ribosome, with the stimulatory RNA showing resistance to unwinding, perhaps by presenting an unusual topology (Yusupov et al., 2001; Yusupova et al., 2001; Plant et al., 2003, 2005; Takyar et al., 2005; Namy et al., 2006). Indeed, such resistance is likely to be responsible for the observed ribosomal pause. An examination of the path of the mRNA through the ribosome reveals a narrowing of the mRNA entry tunnel that would block entry of a folded mRNA (Yusupova et al., 2001). Recently, Takyar and colleagues (2005) demonstrated that the prokaryotic 70S ribosome can itself act as a helicase to unwind mRNA secondary structures before decoding, with the active

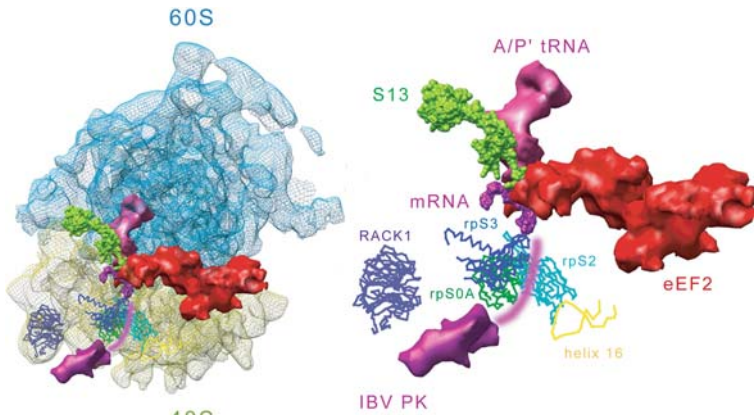


Fig. 7.3 Cryo-EM visualisation of 80S ribosomes stalled at the IBV pseudoknot. On the left is shown a representation of the 3D reconstruction of a pseudoknot-stalled rabbit 80S ribosome, with the large subunit coloured *blue* and the small *yellow*, bound with tRNA, mRNA and eEF2. On the right is shown an expanded view of these ligands, represented by either cryo-EM density or fitted atomic models. eEF2 is found between the subunits and coloured *red*, and the pseudoknot at the entry channel in the solvent face of the small subunit *purple*. The tRNA is distorted spring-like in an A/P' hybrid state and is shown in *pink*. The mRNA with a 45° kink found in the atomic structures for 70S prokaryotic ribosomes is shown in *purple*, linked by a sketched line in the same colour to the pseudoknot. The atomic structure of S13 (prokaryotic equivalent of rpS18) is shown in *light green*, its C-terminus interposed between the A- and P-sites. Also shown are the structures of RACK1, a ribosome-regulatory kinase, rpS0A and the helicase activity-associated proteins rpS3 (S3 in prokaryotes), and rpS2 (S5 in prokaryotes). The structure of the third helicase activity-associated protein rpS9 (S4 in prokaryotes) is not shown since it has not been identified as of yet in a eukaryotic reconstruction. However, it is known to be bound to 18S rRNA helix 16, which is shown. This image demonstrates how the pseudoknot is engaged with the proteins responsible for the helicase activity of the small subunit solvent face (Takyar et al., 2005)

site located between the head and the shoulder of the 30S subunit. Prokaryotic ribosomal proteins S3, S4 and S5 that line the mRNA entry tunnel are implicated in the helicase activity (Takyar et al., 2005), with the eukaryotic 80S counterparts rpS3, rpS9 and rpS2 thus likely to form important elements of an equivalent helicase. The cryo-EM structure of a stalled rabbit 80S ribosome–IBV pseudoknot complex (Namy et al., 2006; see below and Fig. 7.3) has features consistent with an interaction of this proposed helicase with the stimulatory pseudoknot. The specific structural features of the stimulatory RNAs that may affect helicase activity have not been determined. An examination of the confirmed and predicted stimulatory RNAs of Table 7.1 indicates that two basic pseudoknot groupings exist based on S1 length. In the first group, S1 is short (7 bp or less) and stabilised by interactions with a relatively short L3 (≤ 14 nt). The conformation of the helical junction is crucial in these pseudoknots and is stabilised by non-Watson–Crick interactions. In the second group, S1 is longer (~ 11 bp or more), the length and nucleotide composition of L3 seem to be less important to function and there is little evidence for loop–helix interactions (although the lack of high-resolution structures limits such

comparisons). Given these differences, it is unlikely that a single shared feature is responsible for function – more likely, each pseudoknot class has features that promote resistance to unwinding. For example, in those pseudoknots with an S1-L3 RNA triplex, the presence of the “third strand” may conceivably confound unwinding. The triplex may also provide unusual mechanical resistance to unwinding, as could a stable helical junction. In those pseudoknots with a longer S1, the length of the stem, in conjunction with the anchoring stem 2, may lead to torsional resistance as S1 begins to unwind (Plant and Dinman, 2005). Recently, mechanical unwinding studies have revealed that a functional IBV-based pseudoknot is a “brittle” structure, with a shallow dependence of the unfolding rate on applied force and a slower unfolding rate than component hairpin structures (Green et al., 2008). This greater mechanical stability and kinetic insensitivity to force is consistent with a role in resistance to unwinding. A caveat that must be mentioned here is that some viral stimulatory RNAs are stem-loops and have no obvious features that would suggest a specific capacity to resist helicase unwinding. In addition, it has been shown that simply annealing an RNA oligonucleotide at a suitable point downstream of a slippery sequence can in some circumstances promote efficient frameshifting, at least in vitro (Howard et al., 2004; Olsthoorn et al., 2004).

7.5.4 A Role for trans-Acting Protein Factors?

At present there is no evidence that frameshift-stimulatory RNAs are recognised in a specific functional manner by any cellular or viral proteins. This is surprising given the enormous number of known RNA-binding proteins (including pseudoknot-binding proteins) and the potential advantages that could accrue from a capacity to regulate frameshifting. Clearly, the extent to which researchers in the field have searched for frameshift-regulatory factors is hard to judge, given that a negative outcome of such research would likely remain unpublished. However, from our own perspective, we have been unable to identify specific binding partners of the IBV and MMTV pseudoknots in standard RNA-binding assays and in three-hybrid screens (Brierley, I. and colleagues, unpublished), and similar observations have been reported in at least one other laboratory (Dinman, J.D., unpublished). While this topic would benefit from a more extensive and rigorous analysis, the current information leads to the conclusion that the stimulatory RNA acts alone and is not targeted for regulatory purposes.

7.5.5 Conceivable Points for Frameshifting During the Elongation Cycle

The elongation cycle begins when a ternary complex of eEF1A-tRNA-GTP associates with a ribosome containing a P-site peptidyl-tRNA and an empty A-site. The newly delivered cognate tRNA is accommodated into the A-site and almost immediately, the peptidyl transfer reaction takes place, leaving a deacylated tRNA in the

P-site and a peptidyl-tRNA in the A-site (the pre-translocational, or PRE state). Subsequently, the acceptor ends of the A- and P-site tRNAs move, probably spontaneously, with respect to the large ribosomal subunit, but the anticodon ends remain in their original positions relative to the small ribosomal subunit, to yield hybrid state tRNA intermediates P/E and A/P. Complete translocation of the mRNA–tRNA complex into the E- and P-sites is catalyzed by a monomeric G protein, elongation factor 2 (eEF2), with associated GTP hydrolysis. Release of the eEF2–GDP complex and E-site tRNA results in the formation of the post-translocational complex (POST; peptidyl-tRNA in the P-site, empty A-site), competent for the next round of elongation. It is well established that the integrity of both homopolymeric triplets of the slippery sequence (X XXY YYZ) are required for efficient frameshifting, and this has led to the long-held view that the XXY and YYZ codons are present in the ribosomal P- and A-sites (respectively) prior to the frameshift (Jacks et al., 1988). Recent work also suggests the involvement of the 3 nt upstream of the slippery sequence, indicating that the key element is longer (denoted A BCX XXY YYZ) and occupies the E-site as well. There is currently a debate about whether the E-site tRNA anticodon forms an authentic complex with the mRNA but at present, the balance of evidence favours the view that such an interaction is likely. When considering tRNA slippage, it is also pertinent to ask how the tRNAs transit from A- to P- and P- to E-sites, as structural features have been noted that appear to define the boundary between the sites and may act as gates (Yusupov et al., 2001; Selmer et al., 2006; Giedroc and Cornish, 2008). For example, there is a $\sim 45^\circ$ kink in the mRNA between A- and P-sites (Yusupov et al., 2001), stabilised by a critically placed magnesium ion (Selmer et al., 2006) and additionally, the C-terminus of S13 (eukaryotic homologue is S18) can also form a barrier between the two sites (Selmer et al., 2006; Moran et al., 2008). During tRNA transit, it is also relevant to question whether this movement is most likely to occur when the ribosome is dynamic, for example in the so-called “unlocked” conformation (Valle et al., 2003). Within the elongation cycle, frameshifting could conceivably take place at one of five steps (at each step, a specific model that has been proposed is given in brackets): (1) during accommodation of the A-site tRNA (the 9 Å model), (2) subsequent to accommodation, but prior to peptidyl transfer (the simultaneous slippage model), (3) while tRNA hybrid-state intermediates are present (the dynamic model), (4) during the transition from hybrid state to POST state (the mechanical model) and (5) at the start of the next round of elongation (the three-tRNA model). Each has its merits and failings and aspects of each model are shared by others (Fig. 7.4).

7.6 Models of Frameshifting

7.6.1 *The Integrated and 9 Å Models*

That frameshifting could occur during accommodation of the A-site tRNA has its origins in the “integrated” model of ribosomal frameshifting (Harger et al., 2002). Here, the authors argued convincingly, on the basis of the biochemical,

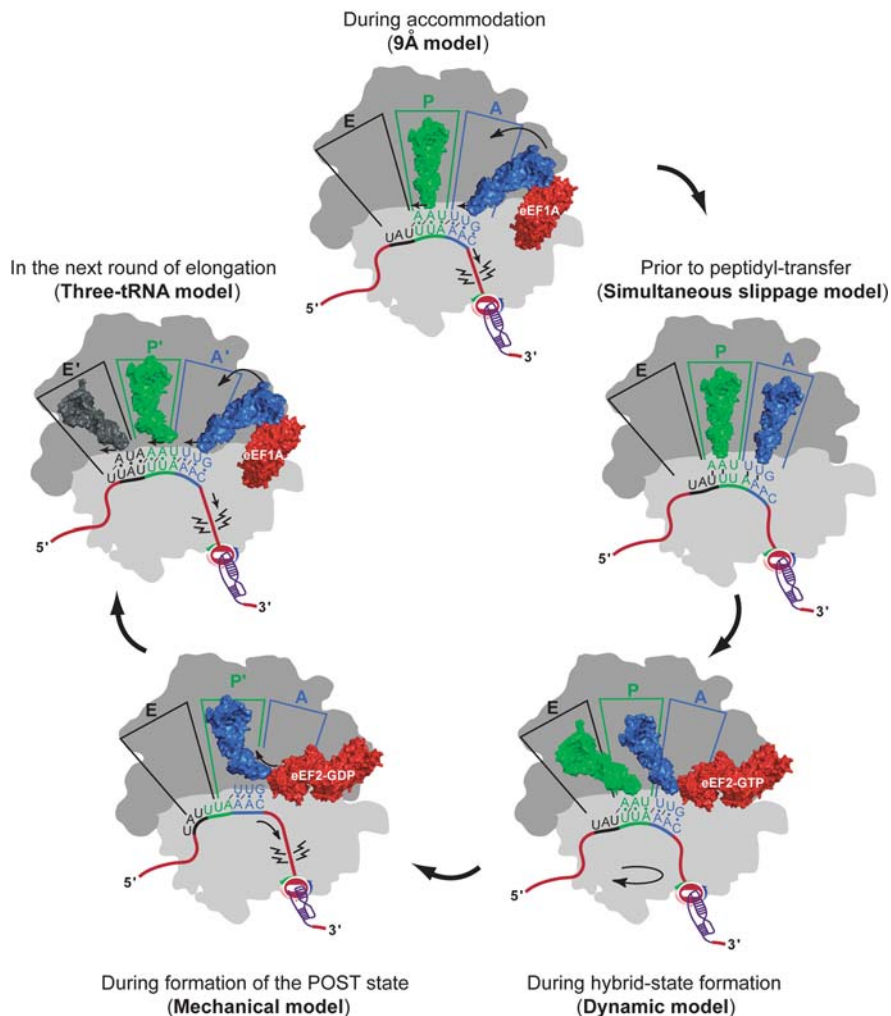


Fig. 7.4 Models of -1 ribosomal frameshifting. Cartoons of five models of frameshifting are shown, each occurring at a different stage of the elongation cycle. The 60S and 40S ribosomal subunits are in *dark grey* and *light grey*, the mRNA in *red*, the A-, P- and E-sites, codons and corresponding tRNAs in *blue*, *green* and *black*. The pseudoknot is in *purple*. The 40S helicase is shown as a *three-armed wheel* whose inhibition by the pseudoknot is indicated by a superimposed “ho-entry” symbol. Initial codon–anticodon base pairs are shown as dots, post-slippage pairs as lines. All models show the pre-slippage pairing except for the simultaneous slippage model. For those models where mRNA tension has been explicitly implicated, this is indicated as a taut mRNA between helicase and tRNA anticodon, highlighted by jagged lines. The curved arrow in the 40S subunit of the dynamic model signifies the ratchet-like movements (see text). Short arrows close to the anticodons or mRNA show directions of movement or force

pharmacological and genetic data available at the time, that the -1 frameshift must occur prior to peptidyl transfer. The evidence was manifold and included the observation that several eEF1A, but no eEF2, mutants had been described that could modulate -1 frameshifting (Dinman and Kinzy, 1997), that anti-ribosomal antibiotics like sparsomycin that inhibit peptidyl transfer should thus increase the residence time of ribosomes paused over the slippery sequence stimulated -1 frameshifting (Dinman et al., 1997) and that mutant ribosomal proteins that result in slower rates of peptidyl transfer also increased -1 frameshifting (Meskauskas et al., 2003). Within the context of the integrated model, the 9 Å model (Plant et al., 2003) explores a purely mechanistic theme and has its basis in modelling data (Noller et al., 2002) which suggests a 9 Å movement of the anticodon of the A-site tRNA upon accommodation. Plant and colleagues (2003) proposed that resistance to unwinding of the stimulatory RNA would hold the mRNA in place such that the movement of the anticodon (and thus the associated mRNA) during accommodation would introduce local “tension” into the mRNA that would be relieved by disruption of the codon–anticodon interaction and re-pairing in the -1 frame. While the 9 Å model satisfies many experimental observations, others indicate that it is not completely correct. The 9 Å movement of the anticodon predicted by Noller and colleagues (2002) remains unconfirmed and some eEF2 mutants that affect programmed frameshifting have now been described (Ortiz et al., 2006). Also, it is not clear how the “local” tension placed on the A-site tRNA is transmitted to the P-site tRNA, which is also thought to slip, nor how the documented effects of mutations within eEF1A on frameshifting could or would equate to different local “tensions”.

7.6.2 The Simultaneous Slippage Model

This original model of frameshifting (Jacks et al., 1988) postulated movement of P- and A-site tRNAs soon after accommodation and prior to peptidyl transfer. The model’s strength lies in the fact that it rationalises the role of both slippery sequence codons in the process, but it is not clear what would drive tRNA movement. Ribosomal pausing alone does not appear to be sufficient to promote tRNA realignment into alternative reading frames and there are no large-scale global movements of ribosomal subunits at this point in the elongation cycle. In addition, it is hard to see how the tRNA could passively shift over the mRNA kink and S13 (S18) between the A- and P-sites. Furthermore, there is little time for the shift to take place between accommodation and peptidyl transfer.

7.6.3 The Dynamic Model

Variants of the simultaneous slippage model have been proposed in which the tRNAs slip during the formation of hybrid-state intermediates on the ribosome, or during translocation itself (Weiss et al., 1989; Farabaugh, 1996). These models seek

ways to harness energy for tRNA dissociation from the molecular movements of the ribosome that occur during the elongation cycle. Recently, these ideas have been revisited and discussed in the light of the recent progress in understanding of ribosome dynamics (Giedroc and Cornish, 2008). In this review, Giedroc and Cornish highlight that the ribosome is most dynamic during the movement of the tRNA acceptor ends into the hybrid state conformation, a state in which -1 frameshifting has been shown to be promoted provided the translocating tRNA is deacylated (McGarry et al., 2005). Further, addition of EF-G.GDPNP to ribosomal complexes stabilises the P/E hybrid state tRNA, allowing for destabilisation of the P-site codon–anticodon hydrogen bonding interactions. Thus frameshifting could be promoted by the binding of eEF2-GTP to the hybrid state ribosome (tRNAs in P/E, A/P state), stabilising the P/E state and promoting frameshifting.

7.6.4 The Mechanical Model

This model was developed following cryo-EM studies of a RRL 80S ribosomal complex stalled at the IBV pseudoknot (80S_{PK}) and provided the first visualisation (at 16.2 Å) of a ribosome stalled at a frameshift-promoting pseudoknot (Namy et al., 2006; Moran et al., 2008). When compared with a reconstruction of an empty ribosome (80S_{Apo}; 14 Å), the 80S_{PK} complexes were found to contain additional densities corresponding to a tRNA (in a hybrid state, see below), the translocase eEF2 and the pseudoknot located at the entrance to the mRNA channel. Overall, these features reveal that encounter of the pseudoknot by the ribosome leads to stalling of the elongation cycle during the translocation step (Fig. 7.3). In support of this, eEF2 has not been released and thus the POST state has not been reached. In addition, the anticodon of the tRNA is not fully in the P-site and thus the tRNA is present in a hybrid state: we refer to this state as A/P' and consider it to be different from the “canonical” A/P state(s) that occur(s) spontaneously following the peptidyl transferase reaction (see above; Moazed and Noller, 1989). The A/P' state does, however, appear to be a necessary intermediate in passage of tRNA from the A to the P-site, given the obstacles on its path. A surprising feature of the A/P' tRNA is its location and structure, with the T-loop pushed towards the intersubunit face of the large subunit and the anticodon arm bent markedly towards the A-site, while retaining contact with domain IV of eEF2. On the basis of this bending, and the presence of the pseudoknot at the mRNA entry channel in close association with components of the putative 80S helicase, a mechanical model of frameshifting was proposed in which the pseudoknot resists unwinding during eEF2-mediated translocation such that tension builds up in the mRNA, subsequently placing strain on the tRNA and resulting in the adoption of a bent conformation (Namy et al., 2006). The opposing actions of translocation, catalysed by eEF2, and pulling from the mRNA strand account for the movement of the elbow of the tRNA into the roof of the P-site and the bending of the tRNA, spring-like, in a (+) sense (3' direction). These opposing forces place a strain on the codon–anticodon interaction that promotes breakage. Subsequent relaxation of the bent tRNA structure in a (–) sense direction

(5') provides an opportunity for the tRNA to re-pair with the mRNA in the -1 position. At the present resolution, it is uncertain whether the tRNA is still bound to the mRNA in these complexes, but the close association of domain IV of eEF2 with the anticodon loop of the tRNA supports this view. In these complexes, the slippery sequence was omitted in an attempt to minimise conformational heterogeneity in the ribosome population, but the consequence is that the tRNA does not have an option to repair in the -1 (or $+1$) phase. The sequence that replaced the slippery sequence was more GC rich and would stabilise mRNA-tRNA contacts and so prevent breakage. The complexes observed, however, are not "dead-end" products since in ribosomal pausing assays, almost all pseudoknot-stalled ribosomes continue translation after the delay (Somogyi et al., 1993; Lopinski et al., 2000; Kontos et al., 2001) and with this particular combination of pseudoknot and translation system, about 50% having entered the -1 reading frame (Brierley et al., 1992).

That frameshifting could take place during translocation has been suggested previously (Weiss et al., 1989; Yelverton et al., 1994; Horsfield et al., 1995; Farabaugh et al., 1996; Kontos et al., 2001). The mechanical model offers a mechanism for disruption of the slippery sequence codon-anticodon interactions which is similar, in some respects, to the 9 Å model (although the coupling with translocation would likely provide more force). However, it is difficult to tease apart the features of the stalled 80S_{PK} complex that are specific to frameshifting from those specific to translocation, as the molecular details of both processes are incompletely understood. Undoubtedly, the movement of the anticodon of the tRNA towards the A-site is suggestive of pseudoknot-induced mRNA tension, but the compression that places the anticodon above the plane of the mRNA and the movement of the elbow towards the top of the P-site is most likely a result of eEF2 action. It has been suggested recently that the 80S_{PK} complex could in fact be a previously undescribed natural intermediate of translocation that pseudoknot-induced stalling has allowed visualisation of (Moran et al., 2008). In this model, as the hybrid state tRNA is translocated from the (authentic) A/P state to the P/P-site proper, eEF2 action compresses the tRNA like a spring, with the anticodon sliding first up the face of eEF2 domain IV then over the "gate" between A and P-sites before the spring relaxes into the P-site, guided by the face of domain IV. This "molecular spring" model of natural translocation is not inconsistent with the mechanical model of frameshifting, although the precise point at which the pseudoknot acts to disrupt the codon-anticodon contacts is not clear. One possibility is that the disruption occurs as the tRNA is fully compressed at the "top" of the gate. At this point, eEF2 would guide the tRNA into the P-site, but as the anticodon is not engaged with the mRNA, the tRNA may be misplaced into the -1 frame. Alternatively, the tRNA may be placed accurately, but the mRNA may have slipped $+1$ (in the 3' direction, effectively a -1 frameshift). This would be due to the pulling action of the pseudoknot as tension is released following detachment of the anticodon, or as has also been suggested, if a partially unwound pseudoknot at the entry channel were to re-fold, dragging the mRNA in a 3' direction (Farabaugh, 1997). At present, however, there are too many uncertainties to favour a particular mechanism. For example, our complexes lack an E-site tRNA, a state that has been linked to frameshifting (Marquez et al., 2004; Sanders et al.,

2008), but we do not know whether this is a genuine absence. It may reflect the use of cycloheximide to stabilise complexes or structural changes in the L1 stalk, either of which could influence E-site occupancy (Pestova and Hellen, 2003). Further, the E-site tRNA may simply have been lost during purification.

While the features of the PK-stalled complexes are consistent with frameshifting during the translocation step, they could in principle represent ribosomes stalled during translocation that have already experienced the molecular events that lead to frameshifting, or as discussed in the final model below, have yet to frameshift. Further work, including the analysis of complexes containing a slippery sequence, and ribosomes stalled at alternative stimulatory RNAs would be of value in addressing some of these issues. A key question to answer is whether the presence of bound eEF2 is causally linked to frameshifting. A reconstruction of a complex comprising a ribosome stalled at a related stem-loop structure (80S_{SSL}; 15.7 Å) with reduced activity in frameshifting revealed a POST state ribosome lacking eEF2 and with a single tRNA in an authentic P/P state (Namy et al., 2006). No obvious density at the mRNA entry channel that could correspond to the stem-loop was observed, however, perhaps as a result of conformational flexibility. While this stem-loop has reduced activity in frameshifting, it is not negligible (~5- to 10-fold less efficient than the pseudoknot in RRL; Kontos et al., 2001). It may be that a sub-population of stalled ribosomes active in frameshifting are present that contain bound eEF2 but have evaded detection as such a species would represent only ~5% of the total ribosome population.

7.6.5 The Three-tRNA Model

Recent work by Léger and colleagues (2007) has provided evidence that the codon preceding the canonical slippery sequence can influence frameshifting. Mutations (underlined) within an “extended” HIV-1 frameshift signal (A BCX XXY YYZ; in HIV-1 U AAU UUU UUA) led to modest increases or decreases in frameshifting efficiency and this was also true for variant slippery sequences derived from other viruses. To account for these findings, the authors propose that the ribosomal P and A-sites are in fact occupied by BCX and XXY during encounter of the HIV-1 stimulatory RNA, with the YYZ codon (known to be crucial to frameshifting) becoming involved in the following cycle of elongation. In natural frameshift signals, nucleotides in the A BCX positions are not noticeably homopolymeric and could not form a stable interaction with the mRNA codon post-slippage (in most cases, only a single Watson–Crick pair can form). It may be that there are relaxed requirements for re-pairing in the E-site (if re-pairing occurs at all) and that the mutations act by changing the identity of tRNA (which in turn may also affect the capacity of the P-site tRNA to frameshift). The involvement of this extended frameshift signal led the authors to suggest a “three-tRNA” model of frameshifting which could also reconcile the conflicting arguments for frameshifting during translocation (see above) or accommodation (Farabaugh, 1997; Léger et al., 2004, Plant et al., 2005). In this model, the translocation of P- and A-site tRNAs decoding

the first two triplets of the slippery sequence (BCX XXY) is blocked with the tRNAs in a position between the PRE and POST states (based on the intermediates identified in cryo-EM [Namy et al., 2006; see above] and kinetic studies [Pan et al., 2007]). Upon release of eEF2, the elongation cycle is completed, but the tRNAs remain out-of-register in transition intermediate sites (referred to as E*/E*, P*/P* by Léger and colleagues). Subsequently, the next tRNA is brought to the ribosome as a ternary complex with eEF1A and during accommodation, all ribosome-associated tRNAs frameshift -1 such that they occupy the standard binding sites (E/E, P/P, A/A). The authors suggest that the tRNAs would slip sequentially, first E-site then P-site then A-site (as discussed by Baranov et al., 2004) and that the driving force for the frameshift would likely be derived from the more stable association of the tRNAs with their authentic sites on the ribosome.

A key question that emerges from the three-tRNA model is whether the elongation cycle can complete while translocation is in an intermediate state. Under normal circumstances, the answer would be no, but in the presence of a stimulatory RNA inducing a strained conformation, it remains possible and might, for example, involve the occupation of another intermediate/hybrid state by the A-to-P translocating tRNA, distinct from canonical A/P and our A/P – which one might call A/P'. More discussion of this intriguing model can be found in the chapter by Brakier-Gingras later in this book.

7.7 Perspective

In the last 20 years, significant progress has been made in the field: the number of confirmed frameshift signals has expanded, examples from cellular genes have been identified and we have a better understanding of frameshift site organisation and structure as they will influence ribosome function. These studies have made important general contributions to our understanding of protein synthesis, translational regulation, RNA structure and function and virus gene expression. However, in general, the biology of frameshifting has lagged behind structural studies such that the role of frameshifting in many viruses (and cellular genes) remains incompletely understood. Nevertheless, frameshifting offers a potential target for antiviral intervention and this will help biological understanding. Also, bioinformatics analyses will be invaluable in assessing just how extensively frameshifting is used in cellular gene expression, and whether there are likely to be side effects associated with drugs targeting the frameshift process. Meanwhile, mechanistic studies of frameshifting will continue to benefit from high-resolution structural and biophysical analysis of stimulatory RNAs and ribosomal complexes. Given the recent progress in ribosomal crystallographic analysis, it is not implausible that the crystal structure of a frameshifting ribosomal complex could be obtained. Further mechanistic insights will be gained from the comparison of RNA oligonucleotide-mediated frameshifting (e.g. Howard et al., 2004) and from studies involving other recoding signals. This is important since frameshifting mediated by RNA oligonucleotides raises the intriguing possibility of *trans*-signalling in vivo where the stimulatory RNA might

be synthesised elsewhere on the genome (e. g. as a miRNA) allowing frameshifting to be regulated temporally and spatially.

References

- Baranov PV, Gesteland RF, Atkins JF (2004) *RNA* 10:221–230
- Baranov PV, Henderson CM, Anderson CB, Gesteland RF, Atkins JF, Howard MT (2005) *Virology* 332:498–510
- Baranov, PV, Fayet O, Hendrix RW, Atkins JF (2006) *Trends Genet* 22:174–181
- Barry JK and Miller WA (2002) *Proc Natl Acad Sci USA* 99:11133–11138
- Bekaert M, Rousset JP (2005) *Mol Cell* 17:61–68
- Belew AT, Hepler NL, Jacobs JL, Dinman JD (2008) *BMC Genomics* 9:339
- Bidou L, Stahl G, Grima B, Liu H, Cassan M, Rousset J-P (1997) *RNA* 3:1153–1158
- Brierley I, Digard P, Inglis SC (1989) *Cell* 57: 537–547
- Brierley I, Rolley NJ, Jenner AJ, Inglis SC (1991) *J Mol Biol* 220:889–902
- Brierley I, Jenner AJ, Inglis SC (1992) *J Mol Biol* 20:463–479
- Brierley I (1995) *J. Gen. Virol.* 76:1885–1892
- Brierley I, Pennell S (2001) *Cold Spring Harbor Symposium of Quantitative Biology LXV*: 233–248
- Brierley I, Dos Ramos FJ (2006) *Virus Res* 119:29–42
- Brierley I, Pennell S, Gilbert RJC (2007) *Nat Rev Microbiol* 5:598–610
- Chen X, Chamorro M, Lee SI, Shen LX, Hines JV, Tinoco I Jr, Varmus HE (1995) *EMBO J.* 14:842–852
- Chen X, Kang H, Shen LX, Chamorro M, Varmus HE, Tinoco I Jr (1996) *J Mol Biol* 260:479–483
- Cornish PV, Hennig M, Giedroc DP (2005) *Proc Natl Acad Sci USA* 102:12694–12699
- Cornish PV, Stammner SN, Giedroc DP (2006) *RNA* 12:1959–1969
- Cowley JA, Dimmock CM, Spann KM, Walker PJ (2000) *J Gen Virol* 81:1473–1484
- Dinman JD, Icho T. and Wickner RB (1991) *Proc Natl Acad Sci USA* 88:174–178
- Dinman JD, Wickner RB (1992) *J Virol* 66:3669–3676
- Dinman JD, Kinzy TG (1997) *RNA* 3:870–881
- Dinman JD, Ruiz-Echevarria MJ, Czaplinski K, Peltz SW (1997) *Proc Natl Acad Sci USA* 94:6606–6611
- Dinman JD, Ruiz-Echevarria MJ, Peltz SW (1998) *Trends Biotech* 16:190–196
- Dreher TW, Miller WA (2006) *Virology* 344:185–197
- Dulude D, Baril M, Brakier-Gingras L (2002) *Nucleic Acids Res* 30:5094–5102
- Farabaugh PJ (1996) *Microb Rev* 60:103–134
- Farabaugh PJ (1997) *Programmed Alternative Reading of the Genetic Code*, pp. 69–102. Landes Bioscience, Austin, Texas and Springer, Heidelberg, Germany
- Farabaugh PJ (2000) *Prog Nucleic Acid Res Mol Biol* 64:131–170
- Giedroc DP, Theimer CA, Nixon PL (2000) *J Mol Biol* 298:67–185
- Giedroc DP, Cornish PV (2008) *Virus Res E pub*
- Green L, Kim CH, Bustamante C, Tinoco I Jr (2008) *J Mol Biol* 375:511–528
- Hansen TM, Reihani SN, Oddershede LB, Sørensen MA (2007) *Proc Natl Acad Sci USA* 104:5830–5835
- Harger JW, Meskauskas A, Dinman JD (2002) *Trends Biochem Sci* 27:448–454
- Herold J, Siddell SG (1993) *Nucleic Acids Res* 21:5838–5842
- Horsfield JA, Wilson DN, Mannering SA, Adamski FM, Tate WP (1995) *Nucleic Acids Res* 23:1487–1494
- Howard MT, Gesteland RF, Atkins JF (2004) *RNA* 10:1653–1661
- Jacks T, Varmus HE (1985) *Science* 230:1237–1242
- Jacks T, Madhani HD, Masiarz FR, Varmus HE (1988) *Cell* 55:447–458
- Jacobs JL, Belew AT, Rakauskaitė R, Dinman JD (2007) *Nucleic Acids Res* 35:165–174
- Kang HS, Hines JV, Tinoco I (1996) *J Mol Biol* 259:135–147

- Kim YG, Su L, Maas S, O'Neill A, Rich A (1999) *Proc Natl Acad Sci USA* 96:14234–14239
- Kollmus H, Honigman A, Panet A, Hauser H (1994) *J Virol* 68:6087–6091
- Kontos H, Naphthine S, Brierley I (2001) *Mol Cell Biol* 21:8657–8670
- Kurland CG (1992) *Ann Rev Genet* 26:29–50
- Léger M, Sidani S, Brakier-Gingras L (2004) *RNA* 10:1225–1235
- Léger M, Dulude D, Steinberg SV, Brakier-Gingras L (2007) *Nucleic Acids Res.* 35:5581–5592
- Liphardt J, Naphthine S, Kontos H, Brierley I (1999) *J Mol Biol* 288: 321–335
- Lopinski JD, Dinman JD, Bruenn JA (2000) *Mol Cell Biol* 20:1095–1103
- Marczinke B, Bloys AJ, Brown TD K, Willcocks MM, Carter MJ, Brierley I (1994) *J Virol* 68:5588–5595
- Marczinke B, Fisher R, Vidakovic M, Bloys AJ, Brierley I (1998) *J Mol Biol* 284:205–225
- Marczinke B, Hagervall T, Brierley I (2000) *J Mol Biol* 295:179–191
- Manktelow E, Shigemoto K, Brierley I (2005) *Nucleic Acids Res* 33:1553–1563
- Marquez V, Wilson DN, Tate WP, Triana-Alonso F, Nierhaus KH (2004) *Cell* 118:45–55
- McGarry KG, Walker SE, Wang H, Fredrick K (2005) *Mol Cell* 20:613–622
- Meskauskas A, Harger JW, Jacobs KL, Dinman JD (2003) *RNA* 9:982–992
- Michiels PJ, Versleijen AA, Verlaan PW, Pleij CW, Hilbers CW, Heus HA (2001) *J Mol Biol* 310:1109–1123
- Moazed D, Noller HF (1989) *Nature* 342:142–148
- Moran SJ, Flanagan JF IV, Namy O, Stuart DI Brierley I, Gilbert RJC (2008) *Structure* 16:664–672
- Namy O, Moran SJ, Stuart DI, Gilbert RJC, Brierley I (2006) *Nature* 441:244–247
- Naphthine S, Liphardt J, Bloys A, Routledge S, Brierley I (1999) *J Mol Biol* 288:305–320
- Nixon PL, Rangan A, Kim YG, Rich A, Hoffman DW, Hennig M, Giedroc DP (2002) *J Mol Biol* 322:621–633
- Noller HF, Yusupov MM, Yusupova GZ, Baucom A, Cate JHD (2002) *FEBS Lett* 514:11–16
- Olsthoorn RC, Laurs M, Sohet F, Hilbers CW, Heus HA, Pleij CW (2004) *RNA* 10:1702–1703
- Ortiz PA, Ulloque R, Kihara GK, Zheng H, Kinzy TG (2006) *J Biol Chem* 281:32639–32648
- Pallan PS, Marshall WS, Harp J, Jewett FC 3rd, Wawrzak Z, Brown BA 2nd, Rich A, Egli M (2005) *Biochemistry* 44:11315–11322
- Pan D, Kirillov SV, Cooperman BS (2007) *Mol Cell* 25:519–529
- Paul CP, Barry JK, Dinesh-Kumar SP, Braut V, Miller WA (2001) *J Mol Biol* 310:987–999
- Pe'ery T, Mathews MB (2000) In Sonenberg N, Hershey J Mathews M (eds), *Translational control of gene expression*, 371pp. Cold Spring Harbor Laboratory Press, New York
- Pennell S, Manktelow E, Flatt A, Kelly G, Smerdon SJ, Brierley I (2008) *RNA* 14:1366–1377
- Pestova TV, Hellen CU (2003) *Genes Dev* 17:181–186
- Plant EP, Jacobs KL, Harger J.W, Meskauskas A, Jacobs JL, Baxter JL, Petrov AN, Dinman JD (2003) *RNA* 9:168–174
- Plant EP, Wang P, Jacobs JL, Dinman JD (2004) *Nucleic Acids Res* 32:784–790
- Plant EP, Dinman JD (2005) *Nucleic Acids Res* 33:1825–1833
- Plant EP, Perez-Alvarado GC, Jacobs JL, Mukhopadhyay B, Hennig M, Dinman JD (2005) *PLoS Biology.* 3:e172
- Pleij CW A, Rietveld K, Bosch L (1985) *Nucleic Acids Res* 13:1717–1731
- Rice NR, Stephens RM, Burny A, Gilden RV (1985) *Virology* 142:357–377
- Rodnina MV, Savelsbergh A, Katunin VI, Wintermeyer W (1997) *Nature* 385:37–41
- Sanders CL, Lohr KJ, Gambill HL, Curran RB, Curran JF (2008) *RNA* 14:1874–1881
- Selmer M, Dunham CM, Murphy FV IV, Weixlbaumer A, Petry S, Kelley AC, Weir JR, Ramakrishnan V (2006) *Science* 313:1935–1942
- Schwartz DE, Tizard R, Gilbert W (1983) *Cell* 32:853–869
- Shehu-Xhilaga M, Crowe SM, Mak J (2001) *J Virol* 75:1834–1841
- Shen LX, Tinoco I (1995) *J Mol Biol* 247:963–978
- Shigemoto K, Brennan J, Walls E, Watson CJ, Stott D, Rigby PW, Reith AD (2001) *Nucleic Acids Res* 29:4079–4088
- Somogyi P, Jenner AJ, Brierley I, Inglis SC (1993) *Mol Cell Biol* 13:6931–6940

- Su L, Chen L, Egli M, Berger JM, Rich A (1999) *Nat Struct Biol* 6:285–292
- Takyar S, Hickerson RP, Noller HF (2005) *Cell* 120:49–58
- ten Dam E.B, Pleij CWA, Bosch L (1990) *Virus Genes* 4:121–136
- ten Dam E, Pleij K, Draper D (1992) *Biochemistry* 31:11665–11676
- Tsuchihashi Z (1991) *Nucleic Acids Res* 19:2457–2462
- Tu C, Tzeng T-H, Bruenn JA (1992) *Proc Natl Acad Sci USA* 89:8636–8640
- Valle M, Zavialov A, Li W, Stagg SM, Sengupta J, Nielsen RC, Nissen P, Harvey SC, Ehrenberg M, Frank J (2003) *Nat Struct Biol* 10:899–906
- Weiss RB, Dunn DM, Shuh M, Atkins JF, Gesteland RF (1989) *New Biol* 1:159–169
- Wen JD, Lancaster L, Hodges C, Zeri AC, Yoshimura SH, Noller HF, Bustamante C, Tinoco I (2008) *Nature* 452:598–603
- Wills NM, Gesteland RF, Atkins JF (1991) *Proc Natl Acad Sci USA* 88:6991–6995
- Wills NM, Moore B, Hammer A, Gesteland RF, Atkins JF (2006) *J Biol Chem* 281:7082–7088
- Yelverton E, Lindsley D, Yamauchi P, Gallant JA (1994) *Mol Microbiol* 11:303–313
- Yusupov MM, Yusupova GZ, Baucom A., Lieberman K, Earnest TN, Cate JH, Noller HF (2001) *Science* 292:883–896
- Yusupova GZ, Yusupov MM, Cate JH, Noller HF (2001) *Cell* 106:233–241