Chapter 5
Programmed −1 Ribosomal Framesshifting in SARS Coronavirus

Jonathan D. Dinman

Abstract In coronaviruses such as the SARS coronavirus (SARS-CoV), programmed −1 ribosomal frameshifting (−1 PRF) is used to direct the synthesis of immediate early proteins, e.g., RNA-dependent RNA polymerase (RDRP) and proteases, that are thought to prepare the infected cell for takeover by the virus. Unlike other RNA viruses which make their structural proteins first, this class of proteins is synthesized after −1 PRF, from subgenomic mRNAs produced subsequent to production of RDRP. Also unique among the coronaviruses is the inclusion of mRNA structural elements that do not appear to be essential for frameshifting. Understanding the differences between −1 PRF signals from coronaviruses and other viruses will enhance our understanding of −1 PRF in general, and will be instructive in designing new classes of antiviral therapeutics. In this chapter we summarize current knowledge and add additional insight to the function of the programmed −1 ribosomal frameshift signal present in the SARS-associated coronavirus.

5.1 Introduction

The emergence of a new infectious disease, known as severe acute respiratory syndrome (SARS), became a global concern following an outbreak in southern China in late 2002. The World Health Organization released a global alert on the 12 March 2003 and at the close of the epidemic more than 8,000 people had been infected in 29 countries (reviewed in Lai 2003; Stadler et al. 2003). The mortality rate was approximately 10%, and morbidity was significantly greater (Stadler et al. 2003 and references within).
An international effort resulted in the rapid sequencing of many isolates, the first two published on May 1, 2003 (Marra et al. 2003; Rota et al. 2003). Using various portions of those sequences, phylogenetic analyses were performed with coronaviruses from the three previously described groups. The initial unrooted trees suggested that the SARS-CoV represented a new group (Holmes 2003; Lai 2003). Subsequent analyses using rooted trees indicated that the SARS-CoV is most likely an early split from the group 2 lineage (Stadler et al. 2003; Lio and Goldman 2004; Snijder et al. 2003). Although a body of research has been performed on some of the related viruses that impact economically important industries, and comparisons have been made based on the relationships, actual analysis of the SARS-CoV is more limited because of its recent emergence. Although there are no approved antiviral drugs that are highly effective against SARS-CoV, this virus has many unique steps in its replication that could be targeted (Holmes 2003). Various options for preventing additional outbreaks of the disease are addressed in other chapters in this book: these include methods of prevention, limiting spread, and targeting the virus after infection. Post-infection targets have been effective in limiting replication of other RNA viruses including HIV; for example, protease inhibitors are designed to target an early part of the viral lifecycle. In both HIV and SARS-CoV, polyproteins are synthesized as a consequence of programmed –1 ribosomal frameshift (–1 PRF) events, and the polyproteins are subsequently autocatalytically cleaved by their encoded proteases.

5.2 Programmed –1 Ribosomal Frameshifting

Programmed –1 ribosomal frameshifting is a mechanism in which cis-acting elements in the mRNA direct elongating ribosomes to shift reading frame by 1 base in the 5' direction. The use of a –1 PRF mechanism for expression of a viral gene was first published in 1985 for the Rous sarcoma virus (Jacks and Varmus 1985) and subsequently for other retroviruses (Jacks et al. 1988b). The first complete coronavirus sequence was published in 1987 (Boursnell et al. 1987) and later that year Brierley and co-workers used an in vitro translation system to demonstrate that a –1 PRF mechanism was used to translate ORF1ab (Brierley et al. 1987). The IBV frameshift signal was further analyzed by Brierley and co-workers and in the following years became one of the best characterized –1 PRF signals.

The –1 PRF signal can be broken down into three discrete parts: the “slippery site,” a linker region, and a downstream stimulatory region of mRNA secondary structure, typically an mRNA pseudoknot (reviewed in Baranov et al. 2002; Brierley 1995; Dinman and Berry 2006). Mutagenesis studies from many different laboratories have demonstrated that the primary sequence of the slippery site and its placement in relation to the incoming translational reading frame is critical: it must be X XXY YYZ (codons are shown in the incoming or 0-frame) where X is a stretch of three identical nucleotides, Y is either AAA or UUU, and Z is A, C, or U. Although less is known about the linker region, whose length and base composition
5.3 Programmed Frameshifting Rates and Virus Propagation

In the best characterized examples of RNA viruses that utilize programmed ribosomal frameshifting (e.g., most retroviruses, totiviruses, and Ty elements), the open reading frame (ORF) encoding the major viral nucleocapsid proteins (e.g., Gag) is located at the 5′ end of the mRNA whereas the ORFs encoding proteins with enzymatic functions (typically Pro and Pol) are located 3′ of, and out of frame with, the Gag ORF. The mRNAs transcribed from these viral templates contain two overlapping ORFs. The enzymatic proteins are only translated as a result of a programmed ribosomal frameshift event that occurs with an efficiency of 1–40% depending on the specific virus and assay system employed (reviewed in Brierley 1995). Thus, the majority of translational events result in the production of the Gag protein, while a minority of frameshifts yield viral enzymatic proteins. The ratio of Gag to Gag–pol synthesized in viruses as a consequence of programmed
frameshifting varies between a narrow window of 20:1 to 60:1 (reviewed in Farabaugh 1997).

The importance of maintaining this precise ratio on viral propagation has been demonstrated using two endogenous viruses of the yeast *Saccharomyces cerevisiae* and with two retroviruses (reviewed in Dinman et al. 1998). Small alterations in programmed frameshifting efficiencies promote rapid loss of the yeast dsRNA L–A “killer” virus, and in inhibition of HIV-1 replication (reviewed in Dinman and Berry 2006). Similarly, increasing or decreasing the efficiency of the +1 ribosomal frameshift in the Ty1 retrotransposable element of yeast results in reduced retrotransposition frequencies (reviewed in Dinman 1995). In L–A, Gag–pol dimerization nucleates formation of the virus particle (reviewed in Wickner 1996). Increasing the amount of Gag–pol protein synthesized may cause too many particles to initiate nonproductively while producing too little may prevent efficient dimerization (Dinman and Wickner 1992). Proteolytic processing of the TyA–TyB (Gag–pol equivalent) polyprotein of Ty1 is more akin to the situation observed in retroviruses. In Ty1, increasing the amount of Gag–pol protein synthesized inhibited proteolytic processing of the polyprotein (Kawakami et al. 1993). As a consequence, formation of the mature forms of RNase H, integrase and reverse transcriptase is blocked (Kawakami et al. 1993). Similarly, changing the ratio of Gag to Gag–pol proteins in retroviruses like HIV or Moloney murine leukemia virus interferes with virus particle formation (reviewed in Baranov et al. 2002; Brierley 1995; Dinman and Berry 2006). In these cases, over-expression of the Gag–pol protein results in inefficient processing of the polyprotein and inhibition of virus production. To summarize, viral PRF efficiencies have been finely tuned to deliver the precise ratios of proteins required for efficient viral particle assembly: too much or too little frameshifting alters this ratio, with detrimental consequences. Based on these studies, it has been proposed that –1 may be a viable target for prevention of viral propagation (reviewed in Dinman et al. 1998).

5.4 Different Models, Different Assay Systems, Different Results

A number of models have been proposed to describe the mechanism by which –1 PRF occurs (Farabaugh 1996; Jacks et al. 1988a; Leger et al. 2007; Namy et al. 2006; Plant et al. 2003; Takyar et al. 2005; Weiss et al. 1989). All the models posit that the stimulatory element causes a pause in translation and that base-pairing is required at the non-wobble positions of at least two tRNA molecules to the mRNA after the frameshift. Differences arise in the timing of the frameshift: after peptide bond formation, before peptide bond formation, and before accommodation of the aminoacyl-tRNA. While there is strong genetic and biochemical evidence
supporting a co-accommodation/prepeptidyltransfer model of \(-1\) PRF, a co-translocational model is more intuitive, and there is both genetic and structural evidence supporting this concept as well (reviewed in Dinman and Berry 2006). When \(-1\) PRF is kinetically modeled in the context of the translation elongation cycle, it becomes clear that \(-1\) PRF is simply an endpoint that is potentially achievable through a number of different kinetic pathways. Indeed, the identification of two distinct frameshift products by protein sequencing supports the hypothesis that frameshifting can occur in at least two distinct phases of the elongation cycle (Jacks et al. 1988b). A complicating issue is the use of a large variety of assay systems that are used by different laboratories to monitor \(-1\) PRF. For example, prokaryotic, yeast, plant and mammalian ribosomes decipher coronavirus frameshift signals quite differently from one another (Brierley et al. 1997; Plant and Dinman 2006). Thus a suitable system must be used to draw purposeful conclusions from in vitro analyses of \(-1\) PRF.

5.5 The Biology of \(-1\) PRF in SARS-CoV is Different

In contrast to the examples discussed above, the genomic organization of coronaviruses is different. Instead of encoding viral structural proteins, the upstream ORF appears to encode immediate early proteins with enzymatic functions (e.g., proteases) thought to be involved in preparing the infected cell for takeover by the virus. These gene arrangements can be viewed in SARS-CoV genome maps elsewhere in this book. Similar to the viruses discussed above, a frameshift from this ORF directs ribosomes into a downstream ORF that encodes the viral RNA-dependent RNA polymerase (RDRP) along with other enzymes thought to play roles in replication. Once synthesized, this RDRP is used to produce a subgenomic mRNA (in addition to the antigenomic and subsequently new genomic RNAs), from which the viral structural proteins are synthesized. Thus, the biology of frameshifting is significantly different in coronaviruses. At present, we do not know if frameshifting rates play an important role in virus replication, and if so, we do not know what that role may be. The availability of a fully infectious reverse genetics system (Almazan et al. 2006; Yount et al. 2003) and of a noninfectious replicon (Almazan et al. 2006) are currently serving as the foundations for research into these questions in our laboratory. One hypothesis that we are currently testing is that, similar to a model proposed for barley yellow dwarf virus (Barry and Miller 2002), \(-1\) PRF in SARS-CoV may be involved in regulating the frequency of elongating ribosomes from the 3’ end of the mRNA or (+) strand, which in turn would affect the availability of this end to the viral replicase for (−) strand synthesis. Thus, we speculate that agents that alter \(-1\) PRF efficiency in SARS-CoV may also have antiviral activities by interfering with the switch from protein synthesis to viral genome replication.
5.6 A Unique Feature of the SARS-CoV Frameshift Signal: A Three-Stemmed mRNA Pseudoknot

As discussed above, −1 PRF signals are usually composed of a slippery site and a stimulatory structure separated by a short spacer region. The rules describing slippery sites appear to be conserved among all eukaryotic viruses, and it also appears that the nucleotides surrounding the heptameric slippery site may affect frameshifting efficiencies, albeit to a lesser extent. Experiments altering the spacer region between the slippery site and stimulatory element affect frameshifting frequencies, suggesting that there might be some optimal spacer sequence.

Where the coronaviruses in general, and SARS-CoV in particular, deviate from other viral −1 PRF signals is in the nature of their downstream stimulatory signals. While an mRNA stem-loop structure was first postulated to stimulate −1 PRF in HIV-1 (Jacks et al. 1988b), studies of IBV provided the first evidence for an H-type pseudoknot structure as a stimulatory elements (Brierley et al. 1989). Subsequently H-type pseudoknots were identified as the general rule in a number of frameshift signals of other plant and animal viruses. H-type pseudoknots are so called because they are composed of two coaxially stacked stem-loops where the second stem is formed by base-pairing between sequence in the loop of the first stem-loop, and additional downstream sequence (Fig. 5.2a). Given that SARS-CoV is a coronavirus, that coronaviruses use −1 PRF, and the general consensus that viral −1 PRF signals contain H-type pseudoknots, the first published analysis of the SARS-CoV −1 PRF signal depicted an H-type pseudoknot (Thiel et al. 2003). Simultaneously, our laboratory applied a computational analysis pipeline designed to identify −1 PRF signals from genomic databases to the SARS-CoV −1 PRF signal. This analysis identified the presence of a third stem-loop structure located 3′ of the end of the first stem, and 5′ of the 3′ half of the second stem (Fig. 5.2b). Molecular genetics, biochemical and biophysical studies confirmed the existence of this “three-stemmed pseudoknot,” and phylogenetic analyses revealed that this general feature is conserved among most coronaviruses (Plant et al. 2005). At the same time, two other groups of researchers used computational and molecular genetics

![Fig. 5.2 Comparison of typical H-type mRNA pseudoknot (a) and three-stemmed mRNA pseudoknot structure of the SARS-CoV (b)](image-url)
methods to identify the three-stemmed pseudoknot in the SARS-CoV –1 PRF signal (Baranov et al. 2005; Ramos et al. 2004).

Interestingly, efficient frameshifting was observed when the third stem was deleted from the SARS-CoV pseudoknot, or when a similar region was deleted from the IBV stimulatory structure, suggesting that these regions are not required for –1 PRF per se (Brierley et al. 1991; Plant et al. 2005). However, it is clear from mutational analyses that when the third stem is present, it has an effect on –1 PRF (Baranov et al. 2005; Plant et al. 2005). It is also possible that the third stem may function as a binding site for a trans-acting factor that may have a role in regulating –1 PRF efficiency, or in some other aspect of the viral life cycle. Current research in our laboratory is focused on determining the biological role of the third stem with regard to the biology of SARS-CoV. Preliminary studies from our laboratory suggest that stabilization of the terminal loop of this stem, perhaps through dimerization, is critical for maintaining high levels of frameshifting. In addition, it has been shown that sequence upstream of the core frameshift signal has been shown to have an inhibitory effect on –1 PRF efficiency in SARS-CoV (Su et al. 2005). This sequence has the potential to form an extensive secondary structure. Although the effect of this region has been suggested to directly affect –1 PRF rates, it is also possible that its role may be indirect, e.g., by causing a fraction of translating ribosomes to dissociate from the mRNA before reaching the –1 PRF signal. Indeed, unpublished data from our laboratory indicate that this is indeed correct. To summarize, although core essential elements of the frameshift signal have been defined, we have just begun to scratch the surface with regard to our understanding of the influence of additional features influencing –1 PRF and the biology of SARS-CoV.

5.7 A Second PRF Signal in SARS-CoV?

A second potential PRF signal has also been identified in SARS-CoV (Wang et al. 2006). This is found in a series of variants of the ORF3a containing stretches of 7 –U residues located 14 nt downstream of initiation codon. Interestingly, runs of both 7 and 8 uridines were able to promote efficient –1 and +1 frameshifting respectively. Substitution of these polyU sequences for the native –1 PRF signal between ORF1a and ORF1b promoted efficient frameshifting even in the absence of a functional downstream pseudoknot structure, suggesting that hepta- and octo-uridine stretches can function as efficient frameshift elements by themselves. While interesting, these observations should be considered in light of two caveats. First, the notion that long stretches of identical nucleotides can promote ribosomal slippage is not new (reviewed in Atkins et al. 1991). Second, the hepta- and octa-U variants of ORF3a may simply be cloning artifacts consequent to reverse transcriptase slippage during the generation of RT-PCR products, and thus may not have any true biological significance. Thus, while provocative, the jury is still out with respect to the presence of a second PRF signal in SARS-CoV.
5.8 Summary and Perspectives

While progress has been made on elucidating the mechanism of –1 PRF and the RNA sequences involved in coronaviruses, the requirement for –1 PRF in the life cycle remains obscure. For other viruses –1 PRF efficiency directly determines the ratios between viral structural and enzymatic proteins. In contrast, –1 PRF in SARS-CoV would appear to affect the relative ratios between immediate early proteins, e.g., proteases and other uncharacterized proteins, with the viral replicase and enzymes that modify RNA. The downstream effect of –1 PRF on the abundance of coat proteins compared to viral RNA in coronaviruses has yet to be determined.

The reason for the presence of a third stem in the SARS-CoV –1 PRF signal also remains a mystery. Unpublished observations in our laboratory suggest that some alterations to the third stem in the SARS coronavirus pseudoknot inhibit virus infectivity without dramatically affecting frameshifting frequency. We have also identified a few proteins encoded by the SARS-CoV subgenomic RNAs that appear to interact with the SARS pseudoknot. This suggests that this region is vital for some part of the virus life cycle other than –1 PRF. Ongoing studies are focused on addressing this question.

An additional research challenge is the production of mutations that have a moderate affect on –1 PRF; these will provide the tools to probe more deeply into the biology of frameshifting in SARS-CoV. As these mutant viruses and replicons become available we will be able to correlate the efficiency of frameshifting with production of genomic and subgenomic RNAs and viral titers. It is expected that some of these mutations will result in defects that will give insight into the function of the extraneous sequence within the frameshift signal, and that this insight will provide an alternative starting point for dissecting the coronavirus replication system.

Acknowledgments The rapid emergence and severity of the SARS-associated coronavirus led to the sharing of unpublished information at conferences and meetings which, in turn, added vigor to the field. Because of the competitive nature of research, this sharing of resources is dependent on the ethical behavior of the researchers. We thank all those who are part of the community and have contributed important information that has the potential to prevent or control virus spread should a similar outbreak occur in the future. This work was supported by NIH RO1 grant AI064307.

References

Almazan F, DeDiego ML, Galan C, Escors D, Alvarez E, Ortego J, Sola I, Zuniga S, Alonso S, Moreno JL, Nogales A, Capiscol C, Enjuanes L (2006) Construction of a severe acute respiratory syndrome coronavirus infectious cDNA clone and a replicon to study coronavirus RNA synthesis. J Virol 80:10900–10906

Atkins JF, Weills RB, Thompson S, Gesteland RFE (1991) Towards a genetic dissection of the basis of triplet decoding, and its natural subversion: programmed reading frame shifts and hops. Annu Rev Genet 25:201–228
Mayo M, McDonald H, Montgomery SB, Pandoh PK, Petrescu AS, Robertson AG, Schein JE, Siddiqui A, Smailus DE, Stott JM, Yang GS, Plummer F, Andonov A, Artsob H, Bastien N, Bernard K, Booth TF, Bowness D, Drebot M, Fernando L, Flick R, Garbutt M, Gray M, Grolla A, Jones S, Feldmann H, Meyers A, Kabani A, Li Y, Normand S, Stroher U, Tipples GA, Tyler S, Vogrig R, Ward D, Watson B, Brunham RC, Krajden M, Petric M, Skowronski DM, Upton C, Roper RL (2003) The genome sequence of the SARS-Associated Coronavirus. Science 300:1399–1404

Namy O, Moran SJ, Stuart DI, Gilbert RJ, Brierley I (2006) A mechanical explanation of RNA pseudoknot function in programmed ribosomal frameshifting. Nature 441:244–247

Plant EP, Dinman JD (2006) Comparative study of the effects of heptameric slippery site composition on -1 frameshifting among different translational assay systems. RNA 12:666–673

Plant EP, Jacobs KLM, Harger JW, Meskauskas A, Jacobs JL, Baxter JL, Petrov AN, Dinman JD (2003) The 9-angstrom solution: how mRNA pseudoknots promote efficient programmed -1 ribosomal frameshifting. RNA 9:168–174

Plant EP, Perez-Alvarado GC, Jacobs JL, Mukhopadhyay B, Hennig M, Dinman JD (2005) A three-stemmed mRNA pseudoknot in the SARS coronavirus frameshift signal. PLoS Biol 3:1012–1023

Ramos FD, Carrasco M, Doyle T, Brierley I (2004) Programmed -1 ribosomal frameshifting in the SARS coronavirus. Biochem Soc Trans 32:1081–1083

Rota PA, Oberste MS, Monroe SS, Nix WA, Campagnoli R, Icenogle JP, Penaranda S, Bankamp B, Maher K, Chen MH, Tong S, Tamin A, Lowe L, Frace M, DeRisi JL, Chen Q, Wang D, Erdman DD, Peret TC, Burns C, Ksiazek TG, Rollin PE, Sanchez A, Liffick S, Holloway B, Limor J, McCaustland K, Olsen-Rassmussen M, Fouchier R, Gunther S, Osterhaus AD, Drosten C, Pallansch MA, Anderson LJ, Bellini WJ (2003) Characterization of a novel Coronavirus associated with severe acute respiratory syndrome. Science 300:1394–1399

Snijder EJ, Breidenbeek PJ, Dobbe JC, Thiel V, Ziebuhr J, Poon LL, Guan Y, Rozanov M, Spaan WJ, Gorbalenya AE (2003) Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. J Mol Biol 331:991–1004

Stadler K, Masignani V, Eickmann M, Becker S, Abrignani S, Klénk HD, Rappuoli R (2003) SARS—beginning to understand a new virus. Nat Rev Microbiol 1:209–218

Su MC, Chang CT, Chu CH, Tsai CH, Chang KY (2005) An atypical RNA pseudoknot stimulator and an upstream attenuation signal for -1 ribosomal frameshifting of SARS coronavirus. Nucleic Acids Res 33:4265–4275

Takyar S, Hickerson RP, Noller HF (2005) mRNA helicase activity of the ribosome. Cell 120:49–58

Thiel V, Ivanov KA, Putics A, Hertzig T, Schelle B, Bayer S, Weissbrich B, Snijder EJ, Rabenau H, Doerr HW, Gorbalenya AE, Ziebuhr J (2003) Mechanisms and enzymes involved in SARS coronavirus genome expression. J Gen Virol 84:2305–2315

Wang XX, Liao Y, Wong SM, Liu DX (2006) Identification and characterization of a unique ribosomal frameshifting signal in SARS-CoV ORF3a. Adv Exp Med Biol 581:89–92

Weiss RB, Dunn DM, Shuh M, Atkins JF, Gesteland RF (1989) E. coli ribosomes re-phase on retroviral frameshift signals at rates ranging from 2 to 50 percent. New Biol 1:159–169

Wickner RB (1996) Prions and RNA viruses of Saccharomyces cerevisiae. Annu Rev Genet 30:109–139

Yount B, Curtis KM, Fritz EA, Hensley LE, Jahrling PB, Prentice E, Denison MR, Geisbert TW, Baric RS (2003) Reverse genetics with a full-length infectious cDNA of severe acute respiratory syndrome coronavirus. Proc Natl Acad Sci USA 100:12995–13000