Programmed −1 ribosomal frameshifting in the SARS coronavirus

F. Dos Ramos*[1], M. Carrasco*, T. Doyle† and I. Brierley†

*Department of Medicine, Addenbrooke's Hospital, University of Cambridge, Box 157, Cambridge CB2 2QQ, U.K., and †Division of Virology, Department of Pathology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QP, U.K.

Abstract

Programmed −1 ribosomal frameshifting is an alternate mechanism of translation used by coronavirus to synthesize replication proteins encoded by two overlapping open reading frames. For some coronaviruses, the mRNA *cis*-acting stimulatory structures involved in this process have been characterized, but their precise contribution to ribosomal frameshifting is not completely understood. Recently, a novel coronavirus was identified as the causative agent of the severe acute respiratory syndrome. This review describes the mRNA motifs involved in programmed −1 ribosomal frameshifting in this virus.

Introduction

Until recently, it was assumed that the standard mechanisms of protein synthesis were quite universal. Nevertheless, we now recognize that the genetic readout can be extended using alternate mechanisms of translation such as programmed −1 ribosomal frameshifting. This mechanism of translation has been described more frequently in viruses and prokaryotes, although a more widespread involvement in lower and higher eukaryotes seems also probable.

To date, animal viral pathogens described to use programmed −1 ribosomal frameshifting include the families *Retroviridae*, *Astroviridae*, *Flaviviridae* and *Coronaviridae*. Lately, a novel coronavirus was identified as the causative agent of the SARS (severe acute respiratory syndrome). From the study of its genome sequence, it is predicted that this virus uses −1 ribosomal frameshifting to translate enzymes involved in replication and RNA processing [1,2]. In the present review, we describe the *cis*-acting mRNA motifs involved in programmed −1 ribosomal frameshifting in SARS-CoV (SARS coronavirus) and discuss the implications of these structures as a potential antiviral target.

SARS-CoV genome

SARS-CoV virions are spherical enveloped particles, with a single molecule of linear, positive-sense and single-stranded RNA. On the basis of phylogeny clustering, SARS-CoV lineage is classified as a member of the group 2 coronaviruses. Its genome is approx. 29.7 kb in length and encodes 15 recognizable ORFs (open reading frames) flanked by 5′- and 3′-untranslated regions [1,2]. The replicase gene (265–21 485 nt, isolate Tor2) comprises two overlapping ORFs (*1a* and *1b*) that encode the polyprotein 1a and the fused polyprotein 1a/1b, translated by programmed −1 ribosomal frameshifting. Both polyproteins are synthesized directly from the genomic RNA and cleaved by two ORF *1a*-encoded viral proteinases (papain-like proteinase and 3C-like cysteine proteinase) to release 16 non-structural proteins [3,4]. Distinctively, segment 1b of the polyprotein 1a/1b is cleaved into five different products [5] predicted to be involved in replication and RNA processing {putative RNA-dependent RNA polymerase, putative 3′ to 5′ exonuclease, putative poly(U)-specific endoribonuclease, putative 2′-*O*-ribose methyltransferase [4] and superfamily 1 helicase [3]}. Thus programmed −1 ribosomal frameshifting is required for the translation of the segment 1b and, therefore, the production of new viral particles. Additionally, 13 ORFs downstream of *1a*/*1b* encode the spike (S), envelope (E), membrane (M), nucleocapsid (N) and nine putative proteins, predicted to be translated from eight subgenomic mRNAs synthesized through a discontinuous transcription process [3,4] (Figure 1).

Programmed −1 ribosomal frameshifting

Programmed ribosomal frameshifting involves ribosomes shifting during elongation to an overlapping reading frame in response to mRNA-stimulatory structures. Thus the downstream stop codon is avoided and a fused protein is translated at a defined ratio that can range from 1% to at least 50%. Ribosomal changes in reading frame can occur in either the 5′ (−1) or the 3′ (+1) direction. Both kinds of slippage have been described in lower and higher organisms. Specifically, −1 ribosomal frameshifting has been recognized more frequently in RNA viruses, although examples have been documented in an increasing number of systems, including bacterial insertion sequences, prokaryote genes and at least one gene from eukaryotes (reviewed in [6,7]). Currently, the mechanism of −1 ribosomal frameshifting stimulation is unknown. Extensive mutational analysis has shown that two *cis*-acting mRNA signals are essential for this process. First, a 'slippery' heptamer sequence (XXXYYYZ), which has been proposed as the segment where tRNAs

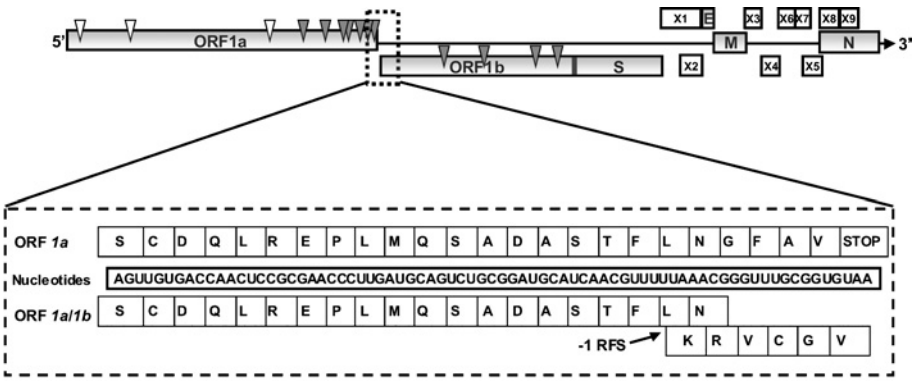
Key words: coronavirus, pseudoknot, recoding, ribosomal frameshifting, severe acute respiratory syndrome (SARS).

Abbreviations used: IBV, infectious bronchitis virus; ORF, open reading frame; SARS, severe acute respiratory syndrome; SARS-CoV, SARS coronavirus.

[1]To whom correspondence should be addressed (email fjd24@cam.ac.uk).

Figure 1 | Genome organization of the SARS-CoV

ORFs *1a* and *1b* are depicted. Arrows represent the cleavage sites for the papain-like proteinase (white) and the 3C-like cysteine proteinase (grey). ORFs encoding the spike (S), envelope (E), membrane (M) and nucleocapsid (N) structural proteins are indicated. ORFs X1–X9 encode nine putative proteins in the intergenic regions between *S* and *E* (two products) and *M* and *N* (seven products). The expanded view shows the overlapping region between ORFs *1a*/*1b* and the segment where −1 ribosomal frameshifting (−1 RFS) is predicted to occur.



dissociate from the mRNA and repair to another codon in an overlapping reading frame. Secondly, a 3′-stimulatory structure that generally starts 5–9 nt downstream of the heptanucleotide motif. In some systems, the stimulatory structure has been described as an mRNA stem loop, but in most cases it is an mRNA pseudoknot as first recognized in studies on the coronavirus IBV (infectious bronchitis virus) [8]. Several models have been proposed to explain −1 ribosomal frameshifting stimulation by pseudoknots. Although these structures induce ribosomal pausing, which could increase the slippage of ribosome-bound tRNAs, the presence of pseudoknots that can pause ribosomes and do not induce ribosomal frameshifting efficiently indicates that this is not sufficient. Other models suggest that pseudoknots might act as binding sites for cellular proteins, but no specific factor has been identified so far (reviewed in [9]). Important progress has been achieved in the field in recent years, including the publication of atomic structures of frameshifting stimulatory pseudoknots. Nevertheless, additional structural and functional studies, especially of mutant pseudoknots, are necessary to provide more insights into the mechanism of −1 ribosomal frameshifting.

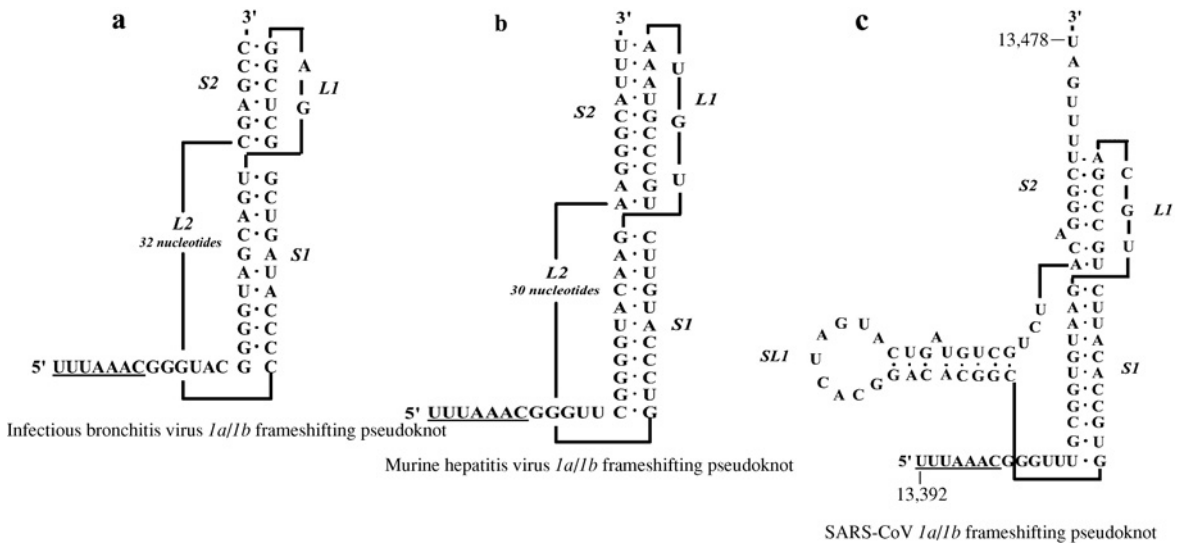
SARS-CoV −1 ribosomal frameshifting signal

Programmed −1 ribosomal frameshifting in coronaviruses was first demonstrated in IBV, where a 'slippery' sequence of the type UUUAAAC and a H-type pseudoknot are sufficient for inducing translation of the segment 1b of the polyprotein 1a/1b [8]. Further studies in other coronaviruses have confirmed the presence of related structures [10,11]. Recently, the SARS-CoV genome was shown to encode an mRNA segment that induces −1 ribosomal frameshifting [3]. We have shown, using a plasmid reporter encoding a segment of the overlapping SARS-CoV ORFs *1a*/*1b* (nt 12 725–14 522, isolate Tor2), where the synthesis of the downstream protein (encoded by the segment *1b*) is dependent on −1 ribosomal frameshifting, that translation of a fused 1a/1b product occurs with an efficiency of approx. 27% *in vitro*. This proportion is similar to the 18–40% reported for other coronaviruses (i.e. IBV, murine hepatitis virus and human coronavirus 229E [8,10,11]). Furthermore, after comparing deletion mutants encoding truncated segments 1b, we have demonstrated that the SARS-CoV frameshifting signal is minimally composed of a UUUAAAC 'slippery' sequence and a stimulatory structure not longer than 80 nt. Modelling based on nearest-neighbour interactions and chemical and enzymic structure mapping suggests that the 'slippery' sequence is followed by a single-stranded spacer region of 5 nt and a stimulatory structure that could fold into a novel pseudoknot. This pseudoknot is composed of two stacked stems (S1 and S2) connected by a single-stranded loop (L1) and a large stem-loop of 28 nt (SL1) (F. Dos Ramos, M. Carrasco, T. Doyle and I. Brierley, unpublished work). This structure partially resembles those reported for group 2 (murine hepatitis virus) and group 3 (IBV) coronaviruses, whereas it differs from those reported for group 1 coronavirus (human coronavirus 229E), where a more 'elaborated' pseudoknot has been proposed [8,10,11] (Figure 2).

Currently, no atomic structure has been published for any of the coronavirus *1a*/*1b* ribosomal frameshifting pseudoknots. However, mutational analysis in IBV has identified essential elements for the −1 frameshifting process. For example, it is known that changes in the sequence or short deletions in the large loop L2 do not influence the induction of ribosomal frameshifting [12]. On the other hand, slight reductions in either total length or number of paired GC residues at the base of the stem S1 decrease the frameshifting stimulatory activity [13]. Similar to IBV, SARS-CoV pseudoknot conserves a stem S1 with GC paired residues at the base, which suggests that this region is

Figure 2 | Predicted secondary structures of *1a/1b* ribosomal frameshifting signals from group 2 and 3 coronaviruses

(**a**) IBV strain Beaudette (group 3) [8]. (**b**) Murine hepatitis virus strain MHV-A59 (group 2) [10]. (**c**) SARS-CoV (group 2). The SARS-CoV pseudoknot is composed of two double-stranded stems (S1 and S2), connected by a single-stranded loop (L1) of 3 nt and a long stem-loop (SL1) of 28 nt. Nucleotides of the long loop L2 in IBV and murine hepatitis virus cannot base-pair to form stable stems and were omitted from the diagram. The 'slippery' sequence is underlined. Numbers correspond to the nucleotide position in the SARS-CoV isolate Tor2.



probably important for ribosomal frameshifting in this virus. Notably, SARS-CoV pseudoknot is also distinguished by a stem-loop SL1 with a proximal double-stranded segment that is not present in other characterized coronavirus *1a/1b* pseudoknots (Figure 2). A systematic analysis on the impact of mutations on this SL1 is required to understand the function of this type of pseudoknot. Most importantly, the primary structure of this frameshifting signal is highly conserved, as indicated by comparison of 105 SARS-CoV sequences (including animal isolates), where mutations of 1–2 nt (localized within SL1 or S2 of the pseudoknot) are only present in two isolates (ZJ01 and GD69).

Considering that −1 ribosomal frameshifting is exploited by SARS-CoV for synthesis of the replication–transcription complex and that the *1a/1b* ribosomal frameshifting signal is highly conserved, we have proposed that this mRNA structure is a potential target for drug design. In support of this, some domains of this motif, such as the 'slippery' sequence and the stem S1 of the pseudoknot, are segments expected to resist mutation due to selective pressure. Indeed, studies of IBV ribosomal frameshifting have shown that variations in either length or nucleotide content of these regions produce important modifications in the efficiency of ribosomal frameshifting [13,14], which would probably prove to be deleterious for the virus. Additionally, the development of escape mutants could be reduced by targeting different motifs of this structure. Along with the 1a/1b polyprotein proteolytic processing, −1 ribosomal frameshifting is also a mechanism that could be targeted to inhibit simultaneously the production of more than one viral protein. Experimental evidence in other viruses (i.e. HIV) has shown that the ratio of structural and enzymic proteins translated by programmed −1 ribosomal frameshifting is critical for RNA dimerization and viral infectivity [15]. Thus a drug aimed at modulating this process could have application as an antiviral agent.

References

1 Marra, M.A., Jones, S.J., Astell, C.R., Holt, R.A., Brooks-Wilson, A., Butterfield, Y.S., Khattra, J., Asano, J.K., Barber, S.A., Chan, S.Y. et al. (2003) Science **300**, 1399–1404
2 Rota, P.A., Oberste, M.S., Monroe, S.S., Nix, W.A., Campagnoli, R., Icenogle, J.P., Penaranda, S., Bankamp, B., Maher, K., Chen, M.H. et al. (2003) Science **300**, 1394–1399
3 Thiel, V., Ivanov, K.A., Putics, A., Hertzig, T., Schelle, B., Bayer, S., Weissbrich, B., Snijder, E.J., Rabenau, H., Doerr, H.W. et al. (2003) J. Gen. Virol. **84**, 2305–2315
4 Snijder, E.J., Bredenbeek, P.J., Dobbe, J.C., Thiel, V., Ziebuhr, J., Poon, L.L., Guan, Y., Rozanov, M., Spaan, W.J. and Gorbalenya, A.E. (2003) J. Mol. Biol. **331**, 991–1004
5 Fan, K., Wei, P., Feng, Q., Chen, S., Huang, C., Ma, L., Lai, B., Pei, J., Liu, Y., Chen, J. et al. (2004) J. Biol. Chem. **279**, 1637–1642
6 Baranov, P.V., Gesteland, R.F. and Atkins, J.F. (2002) Gene **286**, 187–201
7 Namy, O., Rousset, J.P., Napthine, S. and Brierley, I. (2004) Mol. Cell **13**, 157–168
8 Brierley, I., Digard, P. and Inglis, S.C. (1989) Cell (Cambridge, Mass.) **57**, 537–547
9 Brierley, I. and Pennell, S. (2001) Cold Spring Harb. Symp. Quant. Biol. **66**, 233–248
10 Bredenbeek, P.J., Pachuk, C.J., Noten, A.F., Charite, J., Luytjes, W., Weiss, S.R. and Spaan, W.J. (1990) Nucleic Acids Res. **18**, 1825–1832
11 Herold, J. and Siddell, S.G. (1993) Nucleic Acids Res. **25**, 5838–5842
12 Brierley, I., Rolley, N.J., Jenner, A.J. and Inglis, S.C. (1991) J. Mol. Biol. **220**, 889–902
13 Napthine, S., Liphardt, J., Bloys, A., Routledge, S. and Brierley, I. (1999) J. Mol. Biol. **288**, 305–320
14 Brierley, I., Jenner, A.J. and Inglis, S.C. (1992) J. Mol. Biol. **227**, 463–479
15 Shehu-Xhilaga, M., Crowe, S.M. and Mak, J. (2001) J. Virol. **75**, 1834–1841

Received 6 July 2004