Characterization of messenger RNA termini in Schmallenberg virus and related Simbuviruses

Damien Coupéau, François Claine, Laetitia Wiggers, Beer Martin, Nathalie Kirschvink and Benoît Muylkens

1Veterinary Integrated Research Unit, Faculty of Sciences, Namur Research Institute for Life Sciences, University of Namur, 5000 Namur, Belgium
2Friedrich-Loeffler-Institut, Greifswald-Insel-Riems, Germany

In summer 2011, a new arthropod-borne virus named Schmallenberg virus (SBV) emerged in northern Europe (Hoffmann et al., 2012). After its identification in cattle with innocuous clinical signs, a more sinister aspect was revealed during the lambing and calving seasons. SBV infection of pregnant ruminants caused miscarriage, or still-born or severely deformed offspring. Genetic analysis revealed that SBV belongs to the genus Orthobunyavirus within the Bunyaviridae family and is related to the Simbu serogroup viruses (Hoffmann et al., 2012).

To date, the characterization of this emerging pathogen has mainly focused on viral dissemination and prevalence within both mammalian and arthropod hosts (Beer et al., 2013; Claine et al., 2013; De Regge et al., 2012; Garigliany et al., 2012; Linden et al., 2012; Rasmussen et al., 2012; Saegerman et al., 2013). Data were also gathered regarding SBV pathogenesis and its interaction with the innate immune system (Elliott et al., 2013; Varela et al., 2013). In addition, two studies have analysed the genetic background of SBV suggesting that SBV is included in reassortment events together with Sathuperi virus (SATV) and Shamonda virus (SHAV) (Goller et al., 2012; Yanase et al., 2012).

The genome of Bunyaviridae consists of three single-stranded RNA molecules, small (S), medium (M) and large (L), that together encode six proteins. The L segment encodes the RNA-dependent RNA polymerase (RdRP) (Jin & Elliott, 1991; Obijeski et al., 1976) that produces three categories of RNA: genomic RNA (gRNA), messenger RNA (mRNA) and complementary RNA (cRNA). Numerous studies have focused on the NTRs, and on the mechanisms involved in the transcription initiation and termination in Bunyaviridae. NTRs are multifunctional. They serve as promoters for transcription (Barr & Wertz, 2004; Barr et al., 2005), contain transcriptional and translational termination signals (Blakgori et al., 2009, 2012; Ikegami et al., 2007; Lara et al., 2011), provide the encapsidation signal for the N protein, and are involved in the packaging of ribonucleoproteins into virions (Dong et al., 2013a, b). In the present study, we aimed to analyse SBV, SHAV and SATV mRNA termini. The sequence signatures identified at each mRNA terminus give an insight into the mechanisms used by Bunyaviridae for initiating and completing mRNA transcription.

To characterize mRNA 3′ ends, different 3′-RACE analyses were performed on total RNA extracted from SBV-12-Na1 (Coupéau et al., 2013)-, SHAV- and SATV-infected BHK-21 cells as described in the supplementary material available in JGV Online. First, in order to confirm that mRNAs of Simbuviruses do not contain poly(A) tails, 3′-RACE-PCR assays were performed on RNA without poly(A) polymerase pretreatment. In these assays, no viral RNA was found with a spontaneously acquired poly(A) tail. We next analysed the mRNA termini obtained after exogenous poly(A) polymerase treatment. In this analysis,
the S segments of SBV, SHAV and SATV showed similar profiles of transcription termination (Fig. 1a). The majority of mRNA aborted 51 nt before the cRNA termini (Fig. 1a). In silico analysis (http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi) revealed a conserved hairpin structure immediately upstream of the mRNA termini and a putative

Fig. 1. Identification of transcription termini observed in the S segments of SBV, SHAV and SATV, and comparison between different viruses belonging to the Simbu serogroup. (a) Terminal nucleotides and read numbers identified are indicated above each corresponding nucleotide sequence. Putative hairpin structures are indicated below each sequence. Nucleotides in bold correspond to the putative termination signal G\textsubscript{CN1–3GC}. (b) Nucleotide alignment of structurally conserved regions between different viruses belonging to the Simbu serogroup. Underlined nucleotides represent complementary nucleotides involved in hairpin structures. The free energy for each secondary structure is indicated on the right of each sequence. Nucleotides in bold represent terminal nucleotides identified in this study. Highlighted nucleotides inside hairpin structures correspond to nucleotide modifications between SBV and related Simbuviruses. DOUV: Douglas virus (GenBank accession no. HE795092.1); AKAV, Akabane virus (GenBank accession no. NC_009896); SIMV: Simbu virus (GenBank accession no. NC_018477.1); SABOV: Sabo virus (GenBank accession no. AF362396.1); PEAV: Peaton virus (GenBank accession no. HE795095.1); AINOV: Aino virus (GenBank accession no. NC_018460.1); SANV: Sango virus (GenBank accession no. HE795101.1).
Characterization of Simbuviruses mRNAs termini

In comparison to the S segment, the mRNA termini of the M segment showed more variability in the three related ruminant Simbuviruses (Fig. 2a). mRNA transcription termination occurred in nine to 13 sites with no major termination site for the three Simbuviruses (Fig. 2a). In SBV, transcription terminated along a 53 nt region releasing mRNAs truncated at 32–84 nt before the cRNA end. Although the terminal nucleotides were different, a similar profile of distributed termination was observed for SHAV and SATV (Fig. 2a). Putative hairpin structures were predicted for SBV and SATV (Fig. 2a). However, no link between transcription termination and localization of RNA structure was observed. Messenger termini of the L segment differed from those observed in other segments. Indeed, no truncated mRNAs were found for SBV and SATV, and only two truncated mRNAs were found for SHAV, with 8 and 50 nt lacking (Fig. 2b). These data suggested differential termination between the three segments of the same virus. The S segment showed a conserved termination pattern, M segment termination occurred in a distributed fashion and mRNA of the L segment was not truncated.

To characterize the 5′ termini of SBV mRNAs, 5′-RACE-PCR assays were performed as described in the supplementary material. Total RNAs were extracted in vitro from SBV-12-Na1-infected BHK-21 cells and in vivo from the central nervous system of a SBV-infected lamb (Coupeau et al., 2013). In these assays, two sets of 49 (in vitro) and 24 (in vivo) viral mRNA 5′ termini were sequenced and analysed. With regard to the non-template nucleotides found at the 5′ end, these termini were all characterized by the presence of short capped host RNA leaders used for transcription initiation of mRNA during in vitro and in vivo infection (Fig. 3). Analysis of RNA leaders with exclusion of the viral gRNA templates revealed the presence of 11–16 or 11–18 nt RNA leaders during in vitro or in vivo infection, respectively.

To better characterize these RNA leaders, logos of two sequences were created from the whole sets of aligned leader sequences obtained in vitro and in vivo (Fig. 3a). A higher frequency of AGU terminal trinucleotides was found in the leaders. The trinucleotide repeats corresponded to the first complementary nucleotides of the gRNA that were most likely used as a template to produce several repeats (two to four) of the trinucleotide motif following a putative prime-and-realign phenomenon. This model was supported by identifying the putative origin of several host RNA leaders that revealed the absence of these supplementary nucleotides on the cellular transcripts (Table S2).

Two mechanisms of transcription initiation with or without the prime-and-realign phenomenon were suggested. (i) In a first series of transcripts, transcription elongation occurred directly after priming of the host RNA leader on the gRNA template by using one or two nucleotide complementarities, as observed in clones
BH-S-2 or BH-L-3 (Fig. 3b). This mechanism occurred with relatively long (11–17 nt) host RNA leaders and was found in 30% of the analysed transcripts. (ii) The second mechanism used the prime-and-realign phenomenon before elongation of viral transcripts, as observed in clones BH-S-5 or BH-M-6 (Fig. 3b). In this scenario, one or two nucleotide complementarities were used for initial priming. After adding the initial nucleotides (1–5 nt) complementary to the gRNA terminus, the RNA polymerase stopped and realigned at the gRNA terminus. The completion of polymerization produced an expansion of the initial repeat (AGUAGU). This option occurred with relatively short host RNA leaders (9–14 nt) and was found in the remaining 70% of the analysed transcripts.

As with all segmented negative RNA viruses, SBV RdRP uses cap-snatching to initiate mRNA transcription. Here, we have shown that SBV used 1 or 2 nt of the host RNA leaders to prime with the viral gRNA template. As seen in the genus *Tospovirus*, the transcription initiation required at least a single-base complementarity between the 3′ end of the RNA leader and one of the 3 nt found at the 5′ end of gRNA template (Duijsings *et al.*, 2001). Thus, the terminal 3 nt (A, G or U) of the host RNA leaders could prime with the ultimate U, the penultimate C or antepenultimate A of gRNA.

A preference for RNA leader cleavage has been proposed in different viruses known to use the cap-snatching mechanism. Within the *Bunyaviridae* family, a strong preference for cleavage after a U, C or A has been observed in Bunyamwera, Dugbe and Tomato spotted wilt virus, respectively (Duijsings *et al.*, 2001; Jin & Elliott, 1993a, b). Our results do not suggest any cleavage preference for RNA leaders. Indeed, the 3 nt (A, G and U) were used to prime with gRNA (Fig. S1). In the SBV cap-snatching model (Fig. S1), the RdRP does not select any nucleotide at the cleavage site and all the 4 nt can be found on the terminal position of the host RNA leader. However, only the RNA leaders that possess the terminal 3 nt (A, G or U) are able to initiate transcription and RNA leaders possessing a C as the 3′ terminal nucleotide have not been identified in our data.

A prime-and-realign mechanism has been proposed for the Hantaan virus in order to explain the presence of extra

---

**Fig. 2.** Identification of transcription termini observed in the M and L segments of SBV, SHAV and SATV. Terminal nucleotides and read numbers identified are indicated above each corresponding nucleotide sequence. Putative hairpin structures are indicated below each sequence. Nucleotides in bold correspond to the putative termination signal GCN₁₋₃GC.
Characterization of Simbuviruses mRNAs termini

Fig. 3. Alignment of host RNA leaders identified during in vitro and in vivo SBV infection. (a) Sequence logos of host RNA leaders with exclusion of the viral gRNA template. The relative frequency of each nucleotide is indicated at each position by proportionally sized text. (b) Alignment of host RNA leaders identified in vitro during BHK-21 cell infection by SBV. (c) Alignment of the host-derived primer identified in vivo from the SBV-infected central nervous system (CNS) of sheep. Underlined nucleotides represent nucleotides used to prime the gRNA template. *Host RNA leaders whose predictions for the putative cellular origin of mRNA were identified (referenced in Table S1). #Similar origin of host-derived primer (similar host RNA leaders are indicated in parentheses).

nucleotides at the 5′ end of the viral mRNA (Garcin et al., 1995). This phenomenon does not seem to be exclusive to the Hantavirus genus, and extra nucleotides have also been observed following a similar process for Germiston, La Crosse (genus Orthobunyavirus) or Rice Stripe virus (genus Tenuivirus) (Bouloy et al., 1990; Dobie et al., 1997; Yao et al., 2012).
This phenomenon, observed in 70% of SBV mRNA, is suggested to increase transcription efficiency by the addition of complementary nucleotides at the 3' end of the RNA leader. Indeed, RNA leaders that display extended base complementarities to the viral template have been shown to be a predominant criterion and are preferentially used to initiate transcription (Geerts-Dimitriadou et al., 2011; van Knippenberg et al., 2005b). As observed in Tenuiviruses (Yao et al., 2012), the prime-and-realign phenomenon occurs preferentially with short host RNA leaders. Indeed, prime-and-realign may convert short leaders (9–14 nt) into longer leaders (11–18 nt). Use of extended RNA leaders could be more suitable for transcriptional elongation and extended RNA might be used to stabilize the transcriptional complex.

In this study, RNA termini of SBV were characterized, and the mechanisms occurring in mRNA initiation and termination were investigated. Exploration of these mechanisms improves our understanding of Bunyaviridae family biology and shows that RNA termini play an important role for each Bunyavirus. Common features are shared by the whole viral family; however, different regulation strategies were also found, especially at the mRNA 3' termini of the three segments of the same virus.

Acknowledgements

We thank Dr R. B. Tesh (University of Texas Medical Branch) and Dr K. Wernike (Friedrich-Loeffler-Institut, Insel Riems) for providing us with the two Simbuvirus isolates SHAV and SATV.

References


Characterization of Simbuviruses mRNAs termini


