New insights into RNA packaging in porcine reproductive and respiratory syndrome virus

Tayyba T. Baig and Alexander Zakhartchouk

Vaccine and Infectious Disease Organization-International Vaccine Centre (VIDO-InterVac), 120 Veterinary Road, University of Saskatchewan, Saskatoon, SK S7N 5E3, Canada

While identifying whether the smallest packaged heteroclite subgenomic RNA (S9) of porcine reproductive and respiratory syndrome virus (PRRSV) contains a packaging signal, we found that S9 was capable of binding to the basic amino acid-rich domain (synthetic peptide of aa 34–53) of the packaging protein (N). In addition, by using truncations at the 5’ and 3’ ends of S9, a minimal binding region of 35 nt was found to be essential for binding to both the synthetic peptide and to the full-length N protein. Furthermore, by using cell-culture experiments, we found that S9 was capable of packaging non-viral RNA sequence into PRRSV particles and that the 35 nt region was essential for this activity. Taken together, our data suggest that this 35 nt region might be an important element for packaging PRRSV genomic RNA into virus particles.

Porcine reproductive and respiratory syndrome virus (PRRSV) is a worldwide pathogen of swine and causes reproductive failure and respiratory problems in sows and piglets, respectively (Collins et al., 1992; Dea et al., 2000; Done & Paton, 1995; Music & Gagnon, 2010). PRRSV is a positive-sense ssRNA virus and belongs to the family Arteriviridae, which consists of other positive-sense RNA viruses including Lactate dehydrogenase-elevating virus, Equine arteritis virus, and Simian hemorrhagic fever virus (Conzelmann et al., 1993; Meulenberg et al., 1993; Plagemann & Moennig, 1992).

The full-length genomic RNA of PRRSV is ~15.1 kb in length, 5’ capped and 3’ polyadenylated. It contains ten ORFs flanked by 5’ and 3’ untranslated leader regions (UTRs) (Conzelmann et al., 1993; Johnson et al., 2011; Meulenberg et al., 1993, 1998; Nelsen et al., 1999; Snijder & Meulenberg, 1998; Wu et al., 2001). ORF1a and ORF1b are the overlapping ORFs and encode the large polyproteins, which are subsequently processed into functional non-structural proteins, including the RNA-dependent RNA polymerase (Bautista et al., 2002; Dea et al., 2000; Snijder & Meulenberg, 1998; Wootton et al., 2000). ORFs 2–7 encode the structural proteins including the glycoproteins (GP2–GP5), unglycosylated membrane protein (M) and the nucleocapsid (N) protein (Johnson et al., 2011; Mardassi et al., 1996; Meulenberg & Petersen-den Besten, 1996; Meulenberg et al., 1995; van Nieuwstadt et al., 1996; Wu et al., 2001).

The N protein is a packaging protein and contains a basic amino acid-rich domain (aa 37–57) (Wootton et al., 2002) that has been reported recently to bind to the full-length viral genomic RNA, thus providing information about the existence of the RNA-binding domain (aa 37–57) in the N protein (Yoo et al., 2003). However, its activity has not yet been unravelled for the RNA region(s) involved in binding to N protein to recruit it to mature PRRS viral particles.

Besides the full-length genomic RNA and a set of subgenomic mRNAs, the mature PRRSV particles produce additional RNA species, described as heteroclite subgenomic RNAs (sgRNAs). These heteroclite sgRNAs can be co-packaged inside the virus particles with genomic RNA and are composed of only two genomic regions. These include the sequences of the 3’ end of genome and the 5’ end of ORF1a, but exclude the sequences of the 5’ and 3’ UTRs. Furthermore, these heteroclite sgRNAs contain variable lengths of the 5’ end of ORF1a, but maintain a common sequence (nt 191–476), and were designated S1–S9. The packaging of these heteroclite sgRNAs into virus particles and their maintenance of the common sequence suggest that a putative packaging signal may lie within this sequence (Yuan et al., 2000, 2004). Therefore, the current study was conducted to analyse the roles of the smallest heteroclite sgRNA (S9, which consists of only nt 191–476 and is located downstream of the 5’ UTR) as an N protein-binding region and as a packaging signal in vitro and ex vivo, respectively.

The S9 RNA, that we have named Ψ, was synthesized from pGEM-Ψ (Fig. 1) by using conventional cloning and a T7 transcription kit (Epicentre). Following this, Ψ (500 ng) was incubated with increasing concentrations of synthetic peptide (aa 34–53 of N protein; 0.7–1.3 μg) in binding buffer for 30 min at 4 °C, and RNA–peptide-binding was determined by electrophoretic mobility shift assay (EMSA). The buffer and electrophoresis conditions were described previously (Baig, 2008). Incubation of Ψ without peptide served as a negative control and as a molecular mass marker. As is shown in Fig. 1(b), in the absence of peptide RNA migrated faster as a free RNA band (lane 1). In

alex.zak@usask.ca

CORRESPONDENCE
Alexander Zakhartchouk
120 Veterinary Road, University of Saskatchewan, Saskatoon, SK S7N 5E3, Canada

Received 5 July 2011
Accepted 8 September 2011
contrast, the incubation of \( \Psi \) with increasing concentrations of peptide resulted in the appearance of a higher molecular mass RNA–peptide complex that migrated more slowly compared with free RNA (Fig. 1b, compare lanes 2–8 with lane 1).

In order to locate a minimal sequence responsible for peptide binding within \( \Psi \), a series of RNA fragments of different lengths (\( \Psi_1 \), \( \Psi_2 \), \( \Psi_3 \) and \( \Psi_4 \)) were synthesized from pGEM-\( \Psi_1 \), pGEM-\( \Psi_2 \), pGEM-\( \Psi_3 \) and pGEM-\( \Psi_4 \) (Fig. 1a) by progressively deleting 50 and 100 nt from either 5' or 3' ends of pGEM-\( \Psi \) by using conventional cloning and transcription. Following on from this, EMSA results demonstrated that \( \Psi_1 \), \( \Psi_2 \), \( \Psi_3 \) and \( \Psi_4 \) were still able to produce RNA–peptide complexes in a similar fashion to \( \Psi \) and indicated that peptide binding was not affected by deletions (compare Fig. 1b with Fig. 1c–f). However, we were interested in determining whether there was a difference in peptide binding caused by deletions and in locating a minimal sequence responsible for peptide binding. Therefore, two new RNA fragments (\( \Psi_5 \) and \( \Psi_6 \)) were synthesized from pGEM-\( \Psi_5 \) and pGEM-\( \Psi_6 \) (Fig. 1a) by deleting an additional 50 nt (150 nt total) from the 5' and 3' ends of pGEM-\( \Psi \), respectively, by using conventional cloning and transcription. Following this, EMSA results demonstrated that \( \Psi_5 \) was still able to produce an RNA–peptide complex in the presence of peptide as compared to free RNA (Fig. 1g, compare lane 1 with lanes 2–8). In contrast, \( \Psi_6 \) did not produce an RNA–peptide complex, and the fast migrating free-RNA band remained the same even in the presence of increased concentrations of peptide (Fig. 1h, compare lane 1 with lanes 2–8).

Next, we sought to determine whether a differential pattern of binding could also be seen between \( \Psi_5 \) and \( \Psi_6 \) with the use of full-length N protein. Therefore, a 5'His–N fusion protein expression vector (pET30a-5'His-N) was constructed by using conventional cloning and was transformed into BL21-Gold (DE3) pLysS cells (Stratagene). Protein was expressed, purified and analysed as described previously (Baig, 2008).

\[ \text{[\( \times^{32} \text{P} \)]UTP-radiolabelled} \] \( \Psi_5 \) and \( \Psi_6 \) were transcribed with a T7 transcription kit (Ambion) and urea-gel purified. This was followed by EMSA using increasing concentrations of
N (0×, 10×, 20× and 80×) with a constant concentration of the RNAs (1×). A distinctive mode of N-binding was observed between Ψ5 and Ψ6, and only Ψ5 showed a mobility shift (RNA–protein complex), whereas Ψ6 did not (compare Fig. 2a with Fig. 2b, lane 4), thus demonstrating the specificity of N towards Ψ5. Furthermore, the specificity of Ψ5–N binding was tested by competitive binding by using a 100-fold molar excess of unlabelled Ψ5; this showed that the shift observed (in Fig. 2a, lane 4 and Fig. 2c, lane 3) was caused by specific Ψ5–N binding and was abolished by competition with excess cognate unlabelled RNA (Fig. 2c, lane 4).

The EMSA results explain that Ψ1, Ψ2, Ψ3, Ψ4 and Ψ5 may contain a sequence that is absent in Ψ6, and that is why Ψ6 could not bind to either the synthetic peptide or N protein. Therefore, we performed a detailed sequence analysis and found that a region comprising 35 nt was present in all transcripts (Fig. 2d, dotted lines) except for Ψ6 (Fig. 2d, black box). Furthermore, analysis of the predicted secondary structure of this 35 nt region using the MFOLD program (Zuker, 2003; http://mfold.rna.albany.edu/?q=mfold/RNA-Folding-Form2.3), suggests the formation of a purine-rich bulge (nt 351–355; Fig. 2e). A similar purine-rich bulge has been found in a PRRSV-related coronavirus RNA that is conserved across several genera of coronaviruses (Robertson et al., 2005) and is proposed to be a likely site for interaction with N protein (Huang et al., 2004). Since PRRSV N protein contains sequence and structural similarities with coronavirus N proteins (Yu et al., 2006), this suggests, therefore, that it is likely to be involved in binding the purine-rich bulge of our identified sequence for packaging genomic RNA into virus particles. Since the nt 351–355 sequence of the RNA is also purine rich in European PRRSV (Lelystad virus) and different strains of North American PRRSV (data not shown), this, possibly, elicits a similar mechanism of RNA packaging in both types of PRRSV. Therefore, we chose to investigate the packaging capacity of Ψ in cell culture further. With this aim, we adapted an approach that has been described previously to study the packaging capacity of a defective-interfering (DI) coronavirus RNA by inserting a non-viral sequence upstream of this DI RNA (Woo et al., 1997). Therefore, an RNA transcript (CAT-Ψ), containing the sequence of non-PRRSV origin (CAT, chloramphenicol acetyltransferase) and of Ψ was synthesized from pGEM-CAT (Fig. 3a) by using conventional cloning and transcription. Additionally, to perform a negative-control experiment, another RNA transcript (CAT) that contained only CAT sequence and not Ψ was synthesized from pGEM-CAT (Fig. 3a) by using conventional cloning and transcription.

Next, MARC-145 (monkey kidney) cells that support the replication of PRRSV (Kim et al., 1993) were transfected with CAT and CAT-Ψ RNAs in independent and parallel experiments, followed by infection with PRRSV strain VR-2332; after 6 days virus was purified as described previously (Yuan et al., 2000). Following this, viral RNAs
(vRNAs) were extracted from 100 µl aliquots of suspended viral pellets (without any RNase treatment prior to vRNA extraction) by using TRIzol (Invitrogen) according to the manufacturer’s instructions. RT-PCR was performed by using a one-step RT-PCR kit (Qiagen) in the presence of CAT gene-specific primers, because CAT is a marker sequence in this study. As is shown in Fig. 3c, lanes 1 and 2, the RT-PCR product of the CAT gene (700 bp) could be detected in both experiments, thus indicating similar transfection efficiencies. Furthermore, we sought to ensure that an RT-PCR product should be detected for only intra-PRRSV-packaged RNAs and not from the supernatant non-packaged RNAs. Therefore, 100 µl aliquots of suspended viral pellets were treated with 1.7 µg RNase A ml⁻¹ (Invitrogen) prior to vRNA extraction. This concentration of RNase A was chosen because after 1.7 µg RNase A ml⁻¹ treatment, 100 µl aliquots of supernatants that contained only non-packaged RNAs did not show any RT-PCR product for both experiments (data not shown), thus indicating the complete digestion of residual RNAs. In contrast, vRNAs were resistant to RNase A digestion owing to being protected inside virus particles, and RT-PCR analysis of extracted vRNAs revealed that an amplified product of CAT could be detected reproducibly only in those virus particles that were produced in the cells transfected with CAT-Ψ RNA (Fig. 3d, lane 2) as compared with CAT RNA (Fig. 3d, lane 1). These results indicated that non-PRRS viral RNA could be detected in the virus particles only when attached to Ψ, and that Ψ may contain a packaging region that is responsible for packaging non-viral RNA into virus particles.

To map more precisely the region(s) within Ψ responsible for packaging non-PRRSV RNA into virus particles, a series of RNA fragments of different lengths (CAT-Ψ₁, CAT-Ψ₂, CAT-Ψ₃, CAT-Ψ₄, CAT-Ψ₅ and CAT-Ψ₆) were synthesized from pGEM-CAT-Ψ₁, pGEM-CAT-Ψ₂, pGEM-CAT-Ψ₃, pGEM-CAT-Ψ₄, pGEM-CAT-Ψ₅ and pGEM-CAT-Ψ₆ (Fig. 3b) by progressively deleting 50 nt from either the 5’ or 3’ ends of pGEM-CAT-Ψ by using conventional cloning and transcription.

After transfection, infection and purification, the purified virions were either treated in the presence or absence of RNase A prior to vRNA extraction and RT-PCR analysis. As is shown in Fig. 3(e), all deletion mutants showed a RT-PCR product in the absence of RNase A treatment, thus indicating similar transfection efficiencies. In contrast, in the presence of RNase A treatment, RT-PCR product could be detected reproducibly (in three independent experiments) only in those virus particles that were produced in the cells transfected with CAT-Ψ₁, CAT-Ψ₂, CAT-Ψ₃, CAT-Ψ₄ and CAT-Ψ₅ RNAs (Fig. 3f, lanes 1–5), as compared with CAT-Ψ₆ RNA (Fig. 3f, lane 6).

These results indicated that packaging capacity was abolished only for CAT-Ψ₆ (the only RNA deficient in the 35 nt region), thus demonstrating the importance of this 35 nt region as an element of the PRRSV RNA-packaging signal.

Fig. 3. RT-PCR analysis indicating the packaging capacity of Ψ. (a, b) Schematic representation of the DNA constructs used for in vitro runoff transcription. (c) Without RNase A treatment the 700 bp product of the CAT gene was produced for both CAT and CAT-Ψ RNA transfections (lanes 1 and 2, respectively). (d) In the presence of RNase A the 700 bp product of the CAT gene was produced only for CAT-Ψ RNA transfection (lane 2) and not for CAT RNA transfection (lane 1). (e) Without RNase A treatment the 700 bp product of the CAT gene was produced for transfections of all the truncated RNAs (CAT-Ψ₁, CAT-Ψ₂, CAT-Ψ₃, CAT-Ψ₄, CAT-Ψ₅ and CAT-Ψ₆) (lanes 1–6, respectively). (f) In the presence of RNase, the 700 bp product of the CAT gene was produced for transfections of five of the truncated RNAs (CAT-Ψ₁, CAT-Ψ₂, CAT-Ψ₃, CAT-Ψ₄ and CAT-Ψ₅; lanes 1–5, respectively) and not for CAT-Ψ₆ (lane 6). L indicates the 700 bp product of 1 kbp DNA plus ladder (Fermentas).
However, a packaging signal(s) has been located at more than one region of the genomic RNAs of some RNA viruses (Pattanaik et al., 1995; Woo et al., 1997); we cannot rule out the possibility that multiple regions are involved in the packaging of PRRSV genomic RNA. It might be possible that our identified sequence is required to enhance the efficiency of genomic RNA packaging along with other regions. One of the candidates for PRRSV RNA packaging is the UTR, which has been shown to bind to the N protein (Daginakatte & Kapil, 2001); however, the biological relevance of this sequence has not been determined yet.

Investigations in the context of full-length genomic RNA might be an important step towards better understanding of the role of our identified sequence in virus replication and full-length genomic RNA packaging. Moreover, further studies are warranted to prove the existence of the predicted structure for the 35 nt region.

Acknowledgements

We gratefully acknowledge Dr Sam Attah-Poku, Dr Kay Faaberg and Wayne Connor for peptide synthesis, providing us the plasmid containing full-length PRRSV genome and technical assistance, respectively. This study was funded in part by Agriculture Development Fund, Saskatchewan Ministry of Agriculture and Alberta Livestock and Meat Agency (ALMA).

References


