

Kissing interaction between 3 ' noncoding and coding sequences is essential for porcine arterivirus RNA replication

Verheije, M.H.; Olsthoorn, R.R.C.L.; Kroese, M.V.; Rottier, P.J.M.; Meulenberg, J.J.M.

Citation

Verheije, M. H., Olsthoorn, R. R. C. L., Kroese, M. V., Rottier, P. J. M., & Meulenberg, J. J. M. (2002). Kissing interaction between 3 ' noncoding and coding sequences is essential for porcine arterivirus RNA replication. *Journal Of Virology*, *76*(3), 1521-1526. doi:10.1128/JVI.76.3.1521-1526.2002

Note: To cite this publication please use the final published version (if applicable).

Kissing Interaction between 3' Noncoding and Coding Sequences Is Essential for Porcine Arterivirus RNA Replication

M. H. Verheije,^{1*} R. C. L. Olsthoorn,² M. V. Kroese,¹ P. J. M. Rottier,³ and J. J. M. Meulenberg^{1†}

*Department of Infectious Diseases and Food Chain Quality, Institute for Animal Science and Health, Lelystad,*¹ *Institute of Molecular Plant Sciences, Gorlaeus Laboratories, Leiden University, Leiden,*² *and Virology Division, Department of Infectious Diseases and Immunology, Veterinary Faculty, Utrecht University, Utrecht,*³ *The Netherlands*

Received 12 June 2001/Accepted 24 October 2001

We used an infectious cDNA clone of *Porcine reproductive and respiratory syndrome virus* **(PRRSV) to investigate the presence of essential replication elements in the region of the genome encoding the structural proteins. Deletion analysis showed that a stretch of 34 nucleotides (14653 to 14686) within ORF7, which encodes the nucleocapsid protein, is essential for RNA replication. Strand-specific reverse transcription-PCR analysis of viral RNA isolated from transfected BHK-21 cells revealed that this region is required for negative-strand genomic RNA synthesis. The 34-nucleotide stretch is highly conserved among PRRSV isolates and folds into a putative hairpin. A 7-base sequence within the loop of this structure was suggested to base-pair with a sequence present in the loop of a hairpin located in the 3 noncoding region, resulting in a kissing interaction. Mutational analyses confirmed that this kissing interaction is required for RNA replication.**

1521

Positive-strand RNA viruses replicate in infected cells by a process that is mediated by RNA-dependent RNA polymerase (RdRp). In this process, genomic RNA serves as a template for the production of negative-strand antigenomic RNA, which is used in turn as a template for the synthesis of new plus strands. The process of replication requires the recruitment of the RdRp to specific sequences or structures within the templates, also known as *cis*-acting replication elements. These elements are usually located in the noncoding regions at the termini of the viral RNA, where RdRp complexes initiate the synthesis of plus and minus strands (2). Only a few examples are known of *cis*-acting replication elements that reside within a coding region (7, 10, 11, 20).

Porcine reproductive and respiratory syndrome virus (PRRSV) is a positive-stranded RNA virus that belongs to the *Arteriviridae* family (reviewed in reference 24), together with *Equine arteritis virus* (EAV), *Lactate dehydrogenase-elevating virus* (LDV), and *Simian hemorrhagic fever virus* (SHFV) (17). On the basis of their similar genomic organization and replication strategy, the arteriviruses have been grouped into the order of *Nidovirales* together with the coronaviruses and the toroviruses (3).

The genome of PRRSV is 15.1 kb (17) , of which the $5'$ two-thirds is translated into two large polyproteins. These are subsequently cleaved by virus-encoded proteases to yield at least 12 nonstructural proteins, including the viral RdRp (reviewed in reference 24). In addition, a set of subgenomic mRNAs, collectively expressing the viral structural proteins, are produced through a process of discontinuous mRNA transcription (5, 8, 14, 16). This process has not been unraveled completely and may occur during plus- or minus-strand synthesis (reviewed in references 9, 23, and 26). Besides the coding regions, the PRRSV genome contains a 5' untranslated region (5'UTR) of 221 nucleotides (nt) (24), which carries a cap at its 5' end $(15, 21)$, and a 3'UTR of 114 nt to which the $poly(A)$ tail is attached (17).

Little is known about the requirements for arterivirus RNA replication and transcription. In this study, we investigated whether genomic sequences encoding the structural proteins are essential for viral replication and/or transcription. Parts of this region were deleted from the infectious cDNA clone of the Lelystad virus (LV) isolate (15) of PRRSV by using available restriction sites (Fig. 1A). The RNA transcripts from these constructs were transfected into BHK-21 cells (15). Their ability to transiently express the remaining viral structural protein genes was tested by the immunoperoxidase monolayer assay (IPMA) (31) 24 h after transfection, as an indicator for replication and transcription.

RNA transcripts lacking open reading frame 2 (ORF2) through the 5' part of ORF6 (pABV594) expressed the N protein when transfected into BHK-21 cells (Fig. 1A), indicating that subgenomic mRNAs were still produced and that replication and transcription were not affected. In contrast, RNA transcripts lacking the entire ORF7 (pABV521) did not express any of the remaining structural proteins after transfection into BHK-21 cells (Fig. 1A). To determine whether sequences in the 3' part of ORF6 caused a difference, we tested a construct from which ORF2 through ORF6 and the transcription regulation sequence of ORF7 were deleted (pABV664). RNA transcripts lacking this region were still able to express the N protein, albeit to a lower level, as indicated by less intense immunostaining (data not shown).

To locate more precisely the region(s) in ORF7 responsible for the observed effects, we constructed a collection of additional mutants that contained smaller deletions (Fig. 1B). Us-

^{*} Corresponding author. Mailing address: Virology Division, Department of Infectious Diseases and Immunology, Veterinary Faculty, Utrecht University, Yalelaan 1, NL-3584 CL Utrecht, The Netherlands. Phone: 31 30 253 5094. Fax: 31 30 253 6723. E-mail: h.verheije @vet.uu.nl.

[†] Present address: Amsterdam Molecular Therapeutics, Amsterdam, The Netherlands.

FIG. 1. Design and analysis of deletion mutants of the recombinant cDNA clones of PRRSV. Parts of the genome were deleted by using restriction sites present in the cDNAs (A) or by replacement of regions by truncated fragments produced by PCR mutagenesis (B) into pABV437, a full-length cDNA clone containing a *Pac*I site directly downstream of ORF7 (15). PCR mutagenesis was performed as described previously (6). The outline of the constructs (boxes indicate the regions present; lines indicate the deleted regions), the deleted nucleotides, and the plasmid numbers are indicated. Full-length cDNA constructs were linearized and transcribed, and the resulting RNA transcripts were transfected into BHK-21 cells by the methods of Meulenberg et al. (15). Transfected cells were stained with IPMA (31) or with M- (MAb126.3) and N- (MAb122.17) specific monoclonal antibodies (28) at 24 h after transfection. +, positive staining; $-$, no staining. A superscript 1 indicates that identical results were obtained in IPMA with monoclonal antibodies against GP₃ and GP₄ (A) or monoclonal antibody 122.17 against the N protein (B). A superscript 2 indicates that staining with the N-specific monoclonal antibody was most probably due to the expression of a truncated N protein.

 $\overline{\mathbf{N}}$

4

1.of the used to introduceL L deletions \overline{a} ž. $\ddot{}$ pc to the introducedt. and forL ∌. $\frac{1}{2}$ bCE
B

Restriction sites are b Sp., specific.

underlined.

FIG. 2. RT-PCR strategy for (A) and results of (B) the detection of genomic positive-strand RNA (panels 1) and genomic negative-strand RNA (panels 2), subgenomic (sg) positive-strand mRNA7 (panels 3), and subgenomic negative-strand mRNA7 (panels 4). BHK-21 cells were electroporated with RNA transcripts from pABV437, pABV668, and pABV696, and cellular RNA was isolated 12 h after transfection (27). The viral RNA was reverse transcribed and amplified by PCR, as outlined in panel A. Products were analyzed in a 1% agarose gel. Numbers of the constructs from which the amplification products were derived are indicated below the lanes. The numbers on the left indicate the marker sizes (in kilobases). The nt positions of the primers are indicated in parentheses beneath the primers.

ing PCR mutagenesis, deletions increasing in length both at the 5' end and at the 3' end of ORF7 were introduced into the infectious cDNA clone of LV. The primers used are described in Table 1. After transfection of the RNA transcripts of these constructs into BHK-21 cells, we observed that large parts at the 3' end of ORF7 (up to position 14686) could be deleted without affecting M protein expression, whereas small deletions from the 5' end were already inhibitory (Fig. 1B). Detailed deletion analysis of this 5' end revealed that a stretch of 34 nt (14653 to 14686) was essential for the expression of the M protein (pABV696; Fig. 1B). As a control, the mutated ORF7 from pABV696 was replaced by ORF7 of the wild-type infectious cDNA clone of LV (pABV730; Fig. 1B). This restored the ability to express the M protein as well as the other viral structural proteins, demonstrating that the lack of structural protein expression by RNA transcripts from pABV696 had not been caused by unintended mutations elsewhere in the viral genome, possibly introduced during the cloning procedures.

Since deletions within the N gene outside the 34-nt stretch did not abolish RNA replication, subgenomic mRNA synthesis, or viral protein expression, the N protein itself does not play an essential role in any of these processes. These results are in accordance with a recent study of EAV, which showed that expression of the structural protein genes of EAV is not required for genomic RNA replication (19).

To further characterize the role of the 34-nt stretch, we analyzed the effects of its deletion on the synthesis of genomic and subgenomic mRNAs of both positive and negative polarity by strand-specific reverse transcription (RT)-PCR (Fig. 2A1 to A4). To obtain a control construct negative for replication, we deleted both ORF7 and the 3'UTR of the LV cDNA clone,

yielding pABV668 (Fig. 1A). BHK-21 cells were electroporated in parallel with RNA transcripts from pABV696, pABV437, and pABV668, and total RNA was isolated from these cells 12 h after transfection.

RT-PCR on viral RNA isolated from pABV668-transfected cells yielded an amplification product only in a test detecting positive-strand genomic RNA (Fig. 2B1). This product was probably derived from the input RNA, because pABV668 lacks the 3'UTR sequences and is therefore unlikely to yield RNA transcripts that are replication competent. RT-PCR on viral RNA isolated from pABV437-transfected cells yielded products of the expected sizes for both the genomic positive and negative strands (Fig. 2B1 and 2B2) and for the subgenomic mRNA7 positive and negative strands (Fig. 2B3 and 2B4). When we examined viral RNA from cells transfected with transcripts from pABV696, we obtained results similar to those for pABV668 (Fig. 2B1 to 2B4), suggesting that these transcripts do not replicate. The identity of the PCR products was verified by their size and by restriction enzyme analyses (data not shown).

To further confirm our RT-PCR data, an immunofluorescence assay was performed using an antiserum against the nonstructural precursor proteins nsp2 and nsp3 of PRRSV. These proteins are translated from genomic RNA, but their level produced from nonreplicating transcripts is too low to be detected by the antiserum (M. V. Kroese et al., unpublished data). Therefore, positive staining of nsp2 and -3 by the antiserum is dependent on RNA replication. In BHK-21 cells transfected with transcripts from pABV696, no expression of nsp2 or -3 was detected, as was the case with transcripts from our negative control pABV668. In contrast, when using transcripts from our positive control pABV437, we clearly detected

FIG. 3. (A) Schematic representation of the predicted secondary structures in the 34-nt stretch (nt 14653 to 14686) in ORF7 and in the 3'UTR of LV (GenBank M96262), predicted using Mfold (M. Zucker's Mfold server at www.ibc.wustl.edu/ \approx zuker/rna/), and the predicted kissing interaction between these hairpins. Nucleotides complementary between the loops of both hairpins are boxed. (B) Complementarity requirements in the loops of the predicted RNA hairpins within ORF7 and the 3'UTR. Mutated nucleotides are in lightface italic. Mutations were introduced into the infectious cDNA clone of LV by PCR mutagenesis as described previously (6). RNA transcripts of the full-length cDNA clones were transfected into BHK-21 cells by the methods of Meulenberg et al. (15). The expression of the M protein was analyzed 24 h later by IPMA (31) using the M-specific monoclonal antibody 126.3.

the expression of the nsp2 and -3 proteins (data not shown). In conclusion, the 34-nt stretch in ORF7 is essential for genomic minus-strand RNA synthesis.

Secondary-structure analysis using the program MFOLD predicted a hairpin within this 34-nt region of LV (Fig. 3A, lower part). Its existence is supported by the strong sequence conservation of this region in PRRSV isolates. The relatively large size of the loop of the hairpin prompted us to look for regions of sequence complementarity that might be involved in base-pairing interactions. Such complementarity was indeed observed with a 7-base sequence present in the 3'UTR. This sequence, again strongly conserved among PRRSV isolates, occurred in the loop of another predicted hairpin (Fig. 3A, upper part).

To probe for interaction between the loops (Fig. 3A), we first deleted a 32-nt region at the 5' end of the 3'UTR, removing almost the entire hairpin. Transcripts lacking this 32-nt region did not replicate (29). Next, we changed a 5-nt sequence from the loop of the putative ORF7 hairpin into its complement. This was predicted to abrogate the hypothesized kissing interaction. When tested, this mutant failed to replicate, as evidenced by the lack of M protein expression in BHK-21 cells transfected with its RNA transcripts (Fig. 3B, pABV769, compare upper left with upper right panel). Consistently, mutation of the corresponding 5-nt sequence in the loop of the 3'UTR hairpin into its complement was equally detrimental for replication (Fig. 3B, pABV768, lower left panel). Interestingly, when the 5-nt sequences in ORF7 and the 3'UTR hairpin were both mutated, thus restoring the opportunity for base-pairing between the loops, expression of the M protein was again observed (Fig. 3B, lower right panel), indicating that the ability to replicate had been reestablished. Apparently, it is the ability of the two loops to base-pair, not their primary sequences per se, that is functionally relevant. Similar kissing interactions in RNA viruses have only been reported for coxsackievirus (12, 13, 30) and human immunodeficiency virus (4), each consisting of 6 bp.

Though the 34-nt stretch is well conserved in PRRSV, no significant homology with the 34-nt stretch was found in the genomes of the arteriviruses EAV, LDV, and SHFV. However, a kissing interaction as demonstrated for PRRSV may still occur in these viruses, since putative sites of interaction between the 3'UTR and a sequence in the upstream ORF, containing at least seven complementary bases, were observed (data not shown). For EAV, the 353 3'-most nt are required for replication and/or transcription (19). Moreover, defective interference studies indicated that ORF7, which is partly contained within this sequence, contains elements involved in these processes (18). These observations indicate that a kissing interaction between the coding and noncoding sequences might occur in this 353-nt domain.

Arteriviruses encode at least 12 nonstructural proteins (reviewed in reference 24). In concert with host proteins, they assemble the viral replication complex, essential for replication of the viral RNA. It is conceivable that the kissing interaction stabilizes a three-dimensional conformation within the 3'-terminal region of the viral genome, onto which an RNA polymerase complex can be assembled for the initiation of negative-strand RNA synthesis (12). This same kissing interaction may, however, also be essential for the synthesis of negativestrand subgenomic mRNAs if produced from positive-strand genomic RNA during discontinuous transcription (1, 22, 23, 25). Binding of proteins of either host or viral origin might stabilize the kissing interaction or might prevent or melt the interaction, thereby shutting off minus-strand synthesis. Essentially nothing is yet known about the structure of such RNP complexes for any RNA virus. Elucidating these structures will eventually be required to shed light on their function and thereby to understand these fundamental processes of the viral life cycle.

We thank E. Snijder and J. Dobbe for the use of the anti-nsp2 and -3 antiserum and P. van Rijn for critical reading of the manuscript.

This work was supported by Boehringer, Ingelheim, Germany.

REFERENCES

- 1. **Baric, R. S., and B. Yount.** 2000. Subgenomic negative-strand RNA function during mouse hepatitis virus infection. J. Virol. **74:**4039–4046.
- Buck, K. W. 1996. Comparison of the replication of positive-stranded RNA viruses of plants and animals. Adv. Virus Res. **47:**159–251.
- 3. **Cavanagh, D.** 1997. Nidovirales: a new order comprising Coronaviridae and Arteriviridae. Arch. Virol. **142:**629–633.
- 4. **Chang, K. Y., and I. Tinoco, Jr.** 1994. Characterization of a "kissing" hairpin complex derived from the human immunodeficiency virus genome. Proc. Natl. Acad. Sci. USA **91:**8705–8709.
- 5. **de Vries, A. A., E. D. Chirnside, P. J. Bredenbeek, L. A. Gravestein, M. C. Horzinek, and W. J. Spaan.** 1990. All subgenomic mRNAs of equine arteritis virus contain a common leader sequence. Nucleic Acids Res. **18:**3241–3247.
- 6. **Groot Bramel-Verheije, M. H., P. J. M. Rottier, and J. J. M. Meulenberg.** 2000. Expression of a foreign epitope by porcine reproductive and respiratory syndrome virus. Virology **278:**380–389.
- 7. **Klovins, J., and J. van Duin.** 1999. A long-range pseudoknot in Qbeta RNA is essential for replication. J. Mol. Biol. **294:**875–884.
- 8. **Lai, M. M.** 1990. Coronavirus: organization, replication and expression of genome. Annu. Rev. Microbiol. **44:**303–333.
- 9. **Lai, M. M., C. L. Liao, Y. J. Lin, and X. Zhang.** 1994. Coronavirus: how a large RNA viral genome is replicated and transcribed. Infect. Agents Dis. **3:**98–105.
- 10. **Lobert, P. E., N. Escriou, J. Ruelle, and T. Michiels.** 1999. A coding RNA sequence acts as a replication signal in cardioviruses. Proc. Natl. Acad. Sci. USA **96:**11560–11565.
- 11. **McKnight, K. L., and S. M. Lemon.** 1998. The rhinovirus type 14 genome contains an internally located RNA structure that is required for viral replication. RNA **4:**1569–1584.
- 12. **Melchers, W. J., J. M. Bakkers, H. J. Bruins Slot, J. M. Galama, V. I. Agol, and E. V. Pilipenko.** 2000. Cross-talk between orientation-dependent recognition determinants of a complex control RNA element, the enterovirus oriR. RNA **6:**976–987.
- 13. **Melchers, W. J., J. G. Hoenderop, H. J. Bruins Slot, C. W. Pleij, E. V. Pilipenko, V. I. Agol, and J. M. Galama.** 1997. Kissing of the two predominant hairpin loops in the coxsackie B virus 3' untranslated region is the essential structural feature of the origin of replication required for negativestrand RNA synthesis. J. Virol. **71:**686–696.
- 14. **Meulenberg, J. J., A. P. Besten, E. P. De Kluyver, R. J. Moormann, W. M.**

Schaaper, and G. Wensvoort. 1995. Characterization of proteins encoded by ORFs 2 to 7 of Lelystad virus. Virology **206:**155–163.

- 15. **Meulenberg, J. J., J. N. BosDeRuijter, R. vandeGraaf, G. Wensvoort, and R. J. Moormann.** 1998. Infectious transcripts from cloned genome-length cDNA of porcine reproductive and respiratory syndrome virus. J. Virol. **72:**380–387.
- 16. **Meulenberg, J. J., E. J. de Meijer, and R. J. Moormann.** 1993. Subgenomic RNAs of Lelystad virus contain a conserved leader-body junction sequence. J. Gen. Virol. **74:**1697–1701.
- 17. **Meulenberg, J. J., M. M. Hulst, E. J. de Meijer, P. L. Moonen, A. den Besten, E. P. de Kluyver, G. Wensvoort, and R. J. Moormann.** 1993. Lelystad virus, the causative agent of porcine epidemic abortion and respiratory syndrome (PEARS), is related to LDV and EAV. Virology **192:**62–72.
- 18. **Molenkamp, R., B. C. Rozier, S. Greve, W. J. Spaan, and E. J. Snijder.** 2000. Isolation and characterization of an arterivirus defective interfering RNA genome. J. Virol. **74:**3156–3165.
- 19. **Molenkamp, R., H. van Tol, B. C. Rozier, Y. van der Meer, W. J. M. Spaan, and E. J. Snijder.** 2000. The arterivirus replicase is the only viral protein required for genome replication and subgenomic mRNA transcription. J. Gen. Virol. **81:**2491–2496.
- 20. **Paul, A. V., E. Rieder, D. W. Kim, J. H. van Boom, and E. Wimmer.** 2000. Identification of an RNA hairpin in poliovirus RNA that serves as the primary template in the in vitro uridylylation of VPg. J. Virol. **74:**10359– 10370.
- 21. **Sagripanti, J. L., R. O. Zandomeni, and R. Weinmann.** 1986. The cap structure of simian hemorrhagic fever virion RNA. Virology **151:**146–150.
- 22. **Sawicki, S. G., and D. L. Sawicki.** 1990. Coronavirus transcription: subgenomic mouse hepatitis virus replicative intermediates function in RNA synthesis. J. Virol. **64:**1050–1056.
- 23. **Sawicki, S. G., and D. L. Sawicki.** 1998. A new model for transcription, p. 215–219. *In* S. G. S. L. Enjuanes and W. J. M. Spaan (ed.), Coronaviruses and arteriviruses. Plenum Press, New York, N.Y.
- 24. **Snijder, E. J., and J. J. Meulenberg.** 1998. The molecular biology of arteriviruses. J. Gen. Virol. **79:**961–979.
- 25. **van Marle, G., J. C. Dobbe, A. P. Gultyaev, W. Luytjes, W. J. M. Spaan, and E. J. Snijder.** 1999. Arterivirus discontinuous mRNA transcription is guided by base pairing between sense and antisense transcription-regulating sequences. Proc. Natl. Acad. Sci. USA **96:**12056–12061.
- 26. **van Marle, G., W. Luytjes, R. G. van der Most, T. van der Straaten, and W. J. Spaan.** 1995. Regulation of coronavirus mRNA transcription. J. Virol. **69:** 7851–7856.
- 27. **van Marle, G., L. C. van Dinten, W. J. Spaan, W. Luytjes, and E. J. Snijder.** 1999. Characterization of an equine arteritis virus replicase mutant defective in subgenomic mRNA synthesis. J. Virol. **73:**5274–5281.
- 28. **van Nieuwstadt, A. P., J. J. Meulenberg, A. van Essen-Zanbergen, A. Petersen-den Besten, R. J. Bende, R. J. Moormann, and G. Wensvoort.** 1996. Proteins encoded by open reading frames 3 and 4 of the genome of Lelystad virus (*Arteriviridae*) are structural proteins of the virion. J. Virol. **70:**4767– 4772.
- 29. **Verheije, M. H., M. V. Kroese, P. J. M. Rottier, and J. J. M. Meulenberg.** 2001. Viable porcine arteriviruses with deletions proximal to the 3' end of the genome. J. Gen. Virol. **82:**2607–2614.
- 30. **Wang, J., J. M. Bakkers, J. M. Galama, H. J. Bruins Slot, E. V. Pilipenko, V. I. Agol, and W. J. Melchers.** 1999. Structural requirements of the higher order RNA kissing element in the enteroviral 3'UTR. Nucleic Acids Res. **27:**485–490.
- 31. **Wensvoort, G., C. Terpstra, J. Boonstra, M. Bloemraad, and D. Van Zaane.** 1986. Production of monoclonal antibodies against swine fever virus and their use in laboratory diagnosis. Vet. Microbiol. **12:**101–108.