

Predicted RNA Folding Suggests PRRSV Major and Heteroclit Subgenomic Transcripts Result from Polymerase Switching at Unpaired Nucleotides

KAY S. FAABERG, MICHAEL P. MURTAUGH, SHISHAN YUAN

*Department of Veterinary Pathobiology, University of Minnesota, Veterinary Science Building,
1971 Commonwealth Avenue, Saint Paul, MN 55108*

1. INTRODUCTION

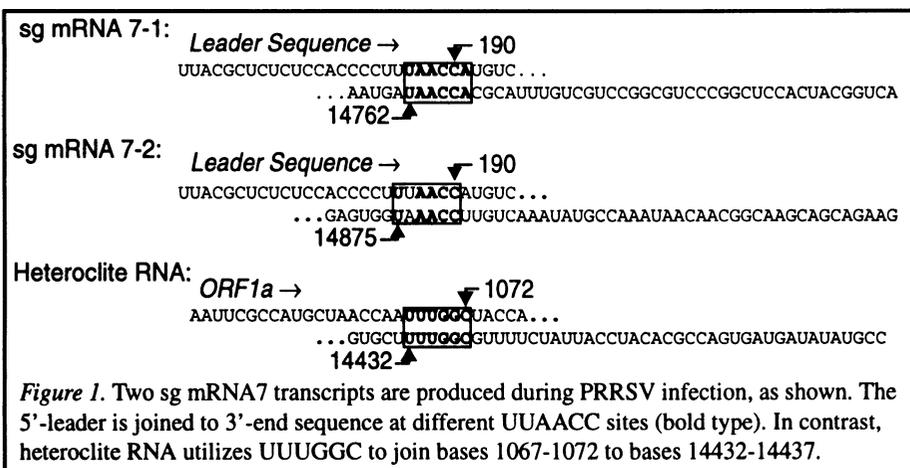
Porcine reproductive and respiratory syndrome virus (PRRSV) has emerged as a major disease of swine worldwide, causing stillbirths and abortions in pregnant sows and reproductive failure in young swine. Even though modified-live vaccines have been used for many years, PRRSV continues to thrive and spread. Previous work has shown that one of the mechanisms which PRRSV uses to evade host defenses is to undergo viral recombination at a high frequency (Yuan *et al* 1999). The high frequency nidovirus recombination mechanism occurs by discontinuous transcription (Baric *et al* 1987) that appears to be guided by basepairing between sense and antisense strands at sites of secondary structure. Support for this model has been recently been obtained for equine arteritis virus subgenomic (sg) mRNAs, which utilize a common leader body junction site sequence which undergoes sequential basepairing at two distinct sites on the genome of the opposite strand (Marle *et al* 1999). In this report, we document that besides standard PRRSV sg mRNAs, other subgenomic RNAs, which utilize alternative junction sites, are transcribed during PRRSV infection. We also provide evidence that these novel subgenomic RNAs are produced by a similar basepairing mechanism, yet the newly identified RNAs appear to be generated in a less-stringent manner. Thus, the PRRSV polymerase complex appears to be more promiscuous than other nidoviruses complexes in deciding where and when to switch templates during transcription. The identification of these novel subgenomic RNAs during normal PRRSV infection will allow further characterization of the specific mechanism for nidovirus discontinuous transcription.

2. NOVEL SUBGENOMIC RNA SPECIES

Defective RNA species have been identified for many coronaviruses and for one arterivirus (Makino *et al* 1984; Snijder *et al* 1991; Penzes *et al* 1994; Chang *et al* 1994; Mendez *et al* 1996; Marle *et al* 2000). In an effort to identify potential defective RNA species for PRRSV strain VR-2332, we passaged virus at both high and low multiplicity of infection (MOI) and surveyed each passage by Northern blot analysis with an ORF1a specific probe, assuming that probe would identify defective RNAs. To our surprise, regardless of the passage number or the MOI utilized, Northern analysis of infected cell RNA using the ORF1a oligonucleotide probe showed that many ORF1a-containing subgenomic RNA species were present during viral infection. They were found to consist of only two portions of the genome and contained the full 5'-leader and 3'-untranslated region (UTR). We have termed these species as heteroclite, meaning "deviating from common forms or rules".

3. TRANSCRIPTION OF SUBGENOMIC RNA

PRRSV strain VR-2332 is 15.4 kb in length, possesses a 190 base 5'-leader sequence ending in the sequence UUAACC and has a 154 base 3'-UTR. Strain VR-2332 produces two transcripts for mRNA7 during infection, termed mRNA7-1 and 7-2 (Nelsen *et al* 1999). The major species, mRNA7-1, utilizes a leader-body junction sequence (AUAACC) located 123 bases upstream from the start site for ORF7 translation, and the minor species, mRNA7-2, utilizes a junction sequence (UAAACC) located 9 bases upstream of the ORF7 initiation codon (Figure 1).



These two transcripts, along with one heteroclitite RNA which utilizes UUUGGC as a junction sequence to join bases 1067-1072 to base 14432-14437 during transcription (Figure 1), were utilized to examine potential mechanisms of discontinuous transcription. Toward this aim, we generated a model (Figure 2) to scrutinize predicted positive strand RNA folding structures and their function in putative negative strand RNA discontinuous transcription, as recent work has suggested subgenomic mRNAs are formed only during negative strand synthesis (Marle *et al* 1999).

We examined whether negative strand discontinuous transcription might be explained by the model presented in Figure 2. Sgro and Palmenberg (1998) have provided evidence that the genome of RNA viruses are folded three dimensionally such that the 5'-end is juxtaposed near the 3'-end. Thus, the junction sites for all subgenomic RNAs would be in close proximity to one another. In this model, transcribing negative strand subgenomic RNAs would stall at hairpins of stem loop structures, fall off of the 3'-end of the positive strand template and realign with similar sequences located in hairpins of stem loop structures located near the 5'-end of the positive strand template. The formation of both major subgenomic transcripts (mRNA 7-1 and 7-2) and heteroclitite subgenomic RNA would be explained by this model.

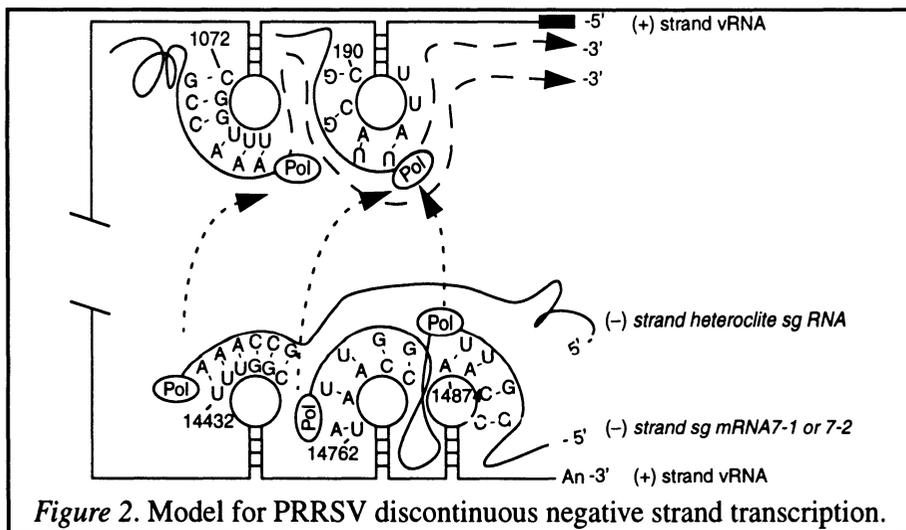
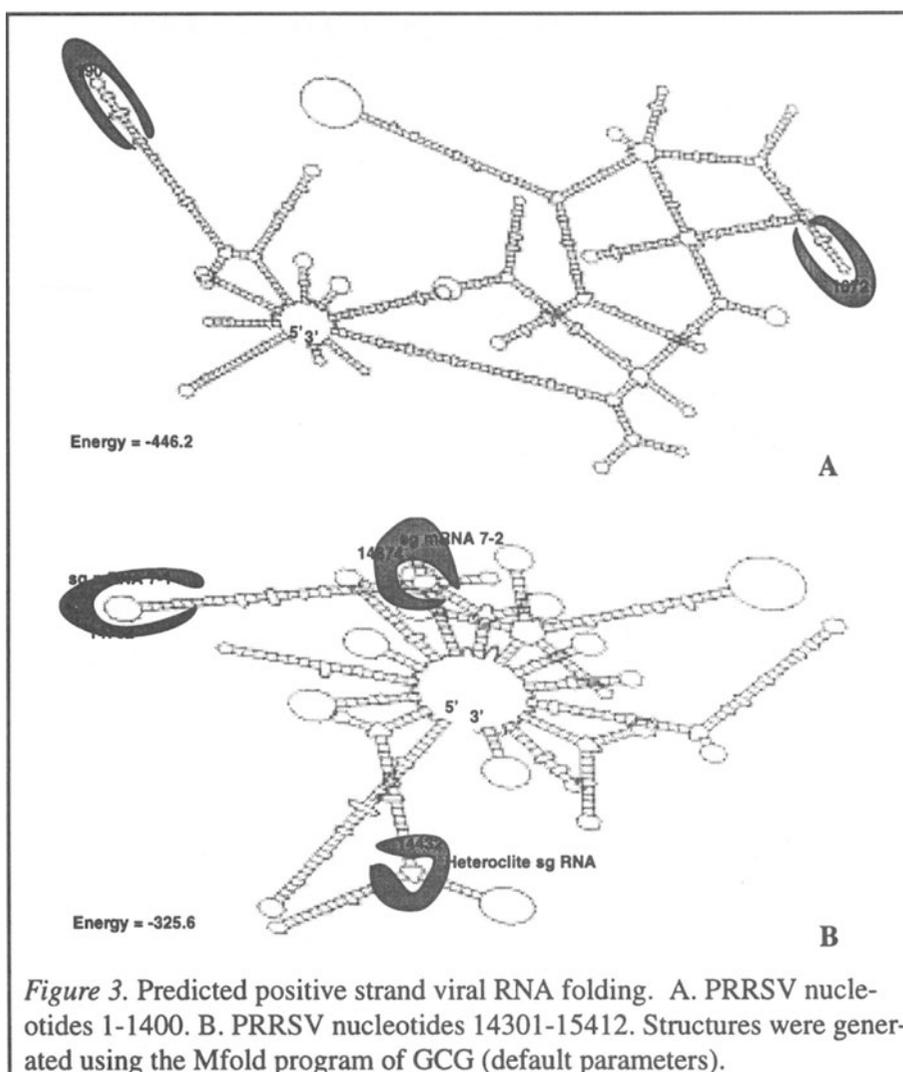


Figure 2. Model for PRRSV discontinuous negative strand transcription.

4. PREDICTED FOLDING PATTERNS SUGGEST BOTH SUBGENOMIC RNA SPECIES ARE FORMED BY A SIMILAR MECHANISM

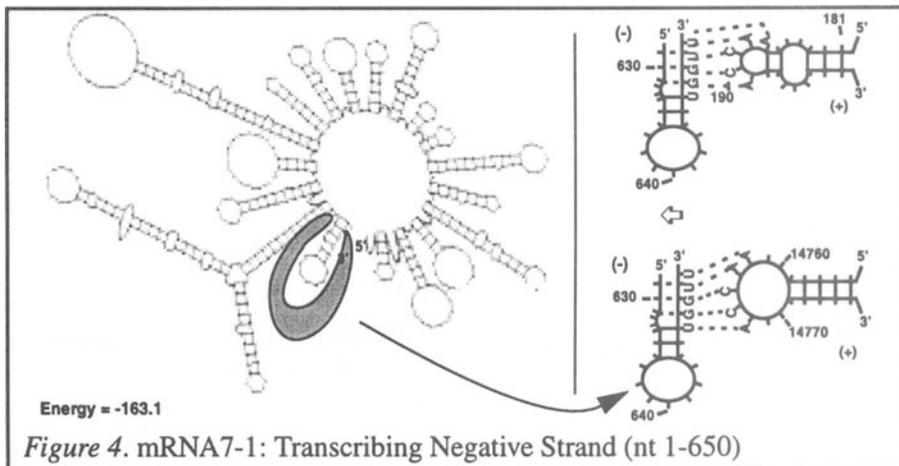
Portions of the positive-strand genome, representing 5'- and 3'-terminal nucleotides of PRRSV, were analyzed for predicted RNA folding structure (GCG Mfold; Zuker 1989). From the 5'- terminal predictions, it was apparent that the nucleotide sequence surrounding base 190 (consensus leader-body

junction site UUAACC) were predicted to form an extended stem-loop structure. Likewise, nucleotide sequence surrounding the heteroclitite junction site (UUUGGC) were predicted to form a similar structure. This suggested that both junction sequences were located at or near loop termini (hairpins) and might be readily available as a landing pad for a partially synthesized negative strand transcript bound to the PRRSV polymerase complex. Similarly, the relevant 3'-end leader-body junction sequences formed analogous hairpins, except for the heteroclitite RNA which was predicted to be part of a stem structure (Figure 3).

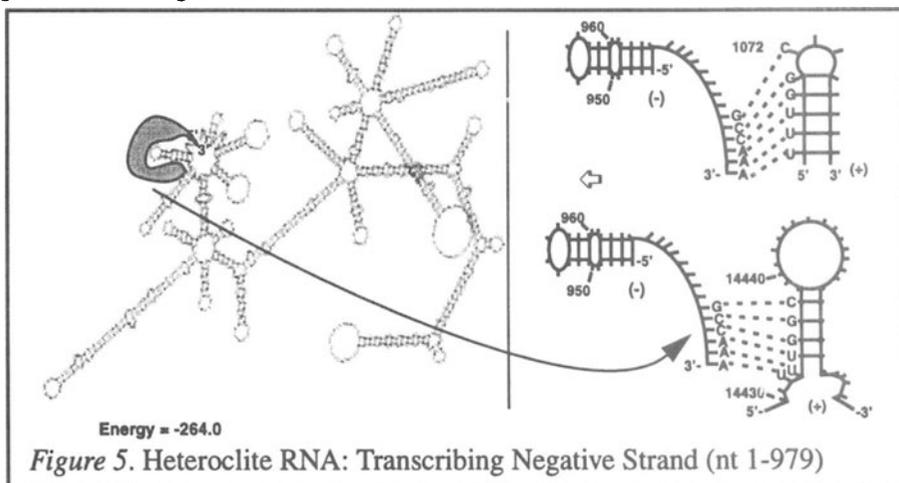


According to the model, negative strand sgmRNA7-1 would stall in transcription at nucleotide 650 (UGGUU), where the positive strand nucleotides ACCAA were predicted to be located at loop termini by Mfold

analysis. At this juncture, the transcribing negative strand RNA was predicted to form a loose stem structure, held together by just three nucleotide interactions. Thus, two short sequences of unpaired nucleotides would be predicted to interact at the point where the nascent transcript falls off of the template. The transcript would then realign itself with unpaired nucleotides located in another hairpin stem loop on the positive strand template (186-190, ACCAA) (Figure 4).



However, during the transcription of negative strand heteroclite RNA, the polymerase complex may stall at nucleotide 979, where the nucleotides (GCCAAA) are predicted to be unpaired with other nascent negative strand bases. The predicted positive strand RNA structure suggested that nucleotides 14432-14437 (UUUGGC) were not located in hairpin loops, but rather were involved in stem structure interactions with other nucleotides (Figure 5). Thus, heteroclite RNA formation could not be explained by the model presented in Figure 2.



5. CONCLUSION

Predicted RNA folding patterns suggested that the major subgenomic mRNAs are formed by interaction between unpaired nucleotides on the transcribing negative strand RNA with unpaired nucleotides located at hairpin termini of extended stem-loop structures of positive strand RNA. The interaction between these short sequences might be unstable due to secondary structure. Thus, transcription terminates and then reinitiates at a similar structure of positive strand RNA located at position 186-190. The model of transcription could not explain heteroclit RNA formation, however. This suggests that other factors may be involved in junction site formation, such as RNA-protein interactions and alternative RNA secondary and tertiary structures which are not yet elucidated.

REFERENCES

- Baric, R.S., Shieh, C.K., Stohlman, S.A., and Lai, M.M., 1987, Analysis of intracellular small RNAs of mouse hepatitis virus: evidence for discontinuous transcription. *Virology* **156**, 342-54.
- Chang, R. Y., Hofmann, M. A., Sethna, P. B., and Brian, D. A., 1994, A cis-acting function for the coronavirus leader in defective-interfering RNA replication. *J. Virol.* **68**, 8223-8231.
- Makino, S., Taguchi, F., and Fujiwara, K., 1984, Defective interfering particles of mouse hepatitis virus. *Virology* **133**, 9-17.
- Mendez, A., Smerdou, C., Izeta, A., Gebauer, F. and Enjuanes, L., 1996, Molecular characterization of transmissible gastroenteritis coronavirus defective interfering genomes: packaging and heterogeneity. *Virology* **217**, 495-507.
- Nelsen, C.J., Murtaugh, M.P., and Faaberg, K.S., 1999, Porcine reproductive and respiratory syndrome virus comparison: divergent evolution on two continents. *J. Virol.* **73**, 270-280.
- Penzes, Z., Tibbles, K., Shaw, K., Britton, P., Brown, T. D. K. and Cavanagh, D., 1994, Characterization of a replicating and packaged defective RNA of avian coronavirus infectious bronchi-tis virus. *Virology* **203**, 286-293.
- Sgro, J.-Y. and Palmenberg, A.C., 1998, 2D folding of large RNA genomes: confidence level, information content and evolutionary selection. 16th Annual Meeting of the American Society for Virology, W26-7.
- Snijder, E. J., den Boon, J. A., Horzinek, M. C. and Spaan, W. J. M., 1991, Characterization of defective interfering RNAs of Berne virus. *J. Gen. Virol.* **72**, 1635-1643.
- van Marle, G., Dobbe, J.C., Gultyaev, A.P., Luytjes, W., Spaan, W.J., and Snijder, E.J., 1999, Arterivirus discontinuous mRNA transcription is guided by base pairing between sense and antisense transcription-regulating sequences. *Proc. Natl. Acad. Sci. U S A* **96**, 12056-12061.
- Yuan, S., Nelsen, C.J., Murtaugh, M.P., Schmitt, B. J., and Faaberg, K. S., 1999, Recombination between North American strains of porcine reproductive and respiratory syndrome virus. *Virus Res.* **61**, 87-98.
- Zuker M., 1989, On finding all suboptimal foldings of an RNA molecule. *Science* **244**, pp. 48-52.