JOURNAL OF VIROLOGY, June 2007, p. 6771–6774 Vol. 81, No. 12
0022-538X/07/$08.00+0 doi:10.1128/JVI.02231-06

New Structure Model for the Packaging Signal in the Genome of Group IIa Coronaviruses▽

Shih-Cheng Chen,[1] Erwin van den Born,[2] Sjoerd H. E. van den Worm,[2] Cornelis W. A. Pleij,[1] Eric J. Snijder,[2] and René C. L. Olsthoorn[1]*

Leiden Institute of Chemistry, Leiden University, P.O. Box 9502, 2300 RA Leiden, The Netherlands,[1] and Molecular Virology Laboratory, Department of Medical Microbiology, Leiden University Medical Center, P.O. Box 9600, 2300 RC Leiden, The Netherlands[2]

Received 11 October 2006/Accepted 27 March 2007

A 190-nucleotide (nt) packaging signal (PS) located in the 3′ end of open reading frame 1b in the mouse hepatitis virus, a group IIa coronavirus, was previously postulated to direct genome RNA packaging. Based on phylogenetic data and structure probing, we have identified a 95-nt hairpin within the 190-nt PS domain which is conserved in all group IIa coronaviruses but not in the severe acute respiratory syndrome coronavirus (group IIb), group I coronaviruses, or group III coronaviruses. The hairpin is composed of six copies of a repeating structural subunit that consists of 2-nt bulges and 5-bp stems. We propose that repeating AA bulges are characteristic features of group IIa PSs.

For many plus-stranded RNA viruses, a specific structural element within the genome RNA that serves as a starting point in the assembly of the viral nucleocapsid has been identified. For the *Coronaviridae* family of RNA viruses (8), which includes important pathogens, such as the severe acute respiratory syndrome coronavirus (SARS CoV), the process of genome encapsidation, which precedes the assembly of the enveloped virus particle, is poorly understood. Previous studies of the mouse hepatitis virus (MHV), a group IIa CoV (4), have led to the identification of a so-called packaging signal (PS) in the 3′ end of replicase open reading frame 1b (ORF1b) (7, 12). A region of 190 nucleotides (nt) was sufficient to drive the encapsidation of the defective interfering RNAs of MHV. Subsequently, a 69-nt fragment of this PS was found to be able, albeit less efficiently, to direct the packaging of heterologous transcripts (3). A secondary structure for MHV PS was proposed (3), but its validity was never experimentally confirmed, nor was it supported by an analysis of the closely related bovine CoVs (2).

We have now used BLAST (1) to search for MHV PS homologs in currently available group IIa CoV sequences: human CoV (HCoV) OC43 (GenBank accession number NC_005147), porcine hemagglutinating encephalomyelitis virus (HEV; NC_007732), bovine CoV (NC_003045), HCoV HKU1 genotype A (NC_006577), and HCoV HKU1 genotype B (AY884001). The only common sequence appeared to be an 18-nt stretch that was previously predicted to form the top of the putative bovine CoV PS (2). Subsequently, we folded a 150-nt region surrounding this motif in all group IIa CoVs using Mfold (13) and adjusted the structure manually based on a phylogenetic analysis. The final model, shown in Fig. 1, is supported by at least nine naturally occurring covariations. It partly resembles the Cologna and Hogue structure (2) but contains several previously unnoticed features, such as the repeating AA and GA bulges on its 3′ side. This feature follows from conserved AGC/GUAAU motifs (Fig. 1), which are repeated four times at 10-bp intervals. Another, though less-well-conserved, element is the 2-nt bulge at the 5′ side of the stem-loop, which is also repeated every 10 bp. An internal loop seems to divide the PS into symmetric upper and lower parts in terms of the orientation and the number of the conserved motifs. Interestingly, the ORF1b sequences of HCoV OC43 and HEV have a GGU codon insertion, while the other group IIa viruses do not (Fig. 1). The extra three nucleotides are part of the internal loop and maintain the secondary structure of the PS.

To support our model further, we PCR amplified a DNA fragment containing the MHV A59 PS with the primers T7MH (5′-TTAATACGACTCACTATAGACGATGTTATCT TCAGCCG-3′) and MHV1 (5′-CCGATCGAGGTGTGAAA GAAG-3′) and using an MHV A59 cDNA library as a template. Subsequent transcription by T7 RNA polymerase (RiboMAX, Promega) yielded a 169-nt MHV PS transcript, which was subjected to chemical and enzymatic probing as described in the legend to Fig. 2. Figure 2 shows RNase T_1-, RNase V_1-, RNase A-, and RNase S_1-susceptible positions and the results of dimethylsulfate (DMS) and diethylpyrocarbonate (DEPC) probing. The G residues recognized by RNase T_1 were G23 and G77 (Fig. 2B), in agreement with their bulging-out positions in the helix. Residues C46, A47, and A48 were susceptible to RNase S_1 (Fig. 2B) and DMS (Fig. 2C), in accordance with their exposed positions in the loop. The presence of the pentaloop, the internal loop, and several of the bulges was also supported by RNase A digestions (Fig. 2C) and by DMS, DEPC (Fig. 2C and D), and lead(II) probing (data not shown). Interestingly, A5, A13, A17, A29, and A41, which are supposedly involved in double-stranded regions, were susceptible to DMS probing at room temperature (Fig. 2C) but also at 0°C (Fig. 2D). Although this may point to "breathing" of these base pairs due to their location at the end of a helix, the absence of DEPC modification of these particular residues (Fig. 2D) strongly suggests the formation of non-Watson-Crick base pairs involving N7 rather than N1 of adenosine.

* Corresponding author. Mailing address: Leiden Institute of Chemistry, Leiden University, P.O. Box 9502, 2300 RA Leiden, The Netherlands. Phone: 31 71 527 4586. Fax: 31 71 527 4340. E-mail: olsthoor@chem.leidenuniv.nl.

▽ Published ahead of print on 11 April 2007.

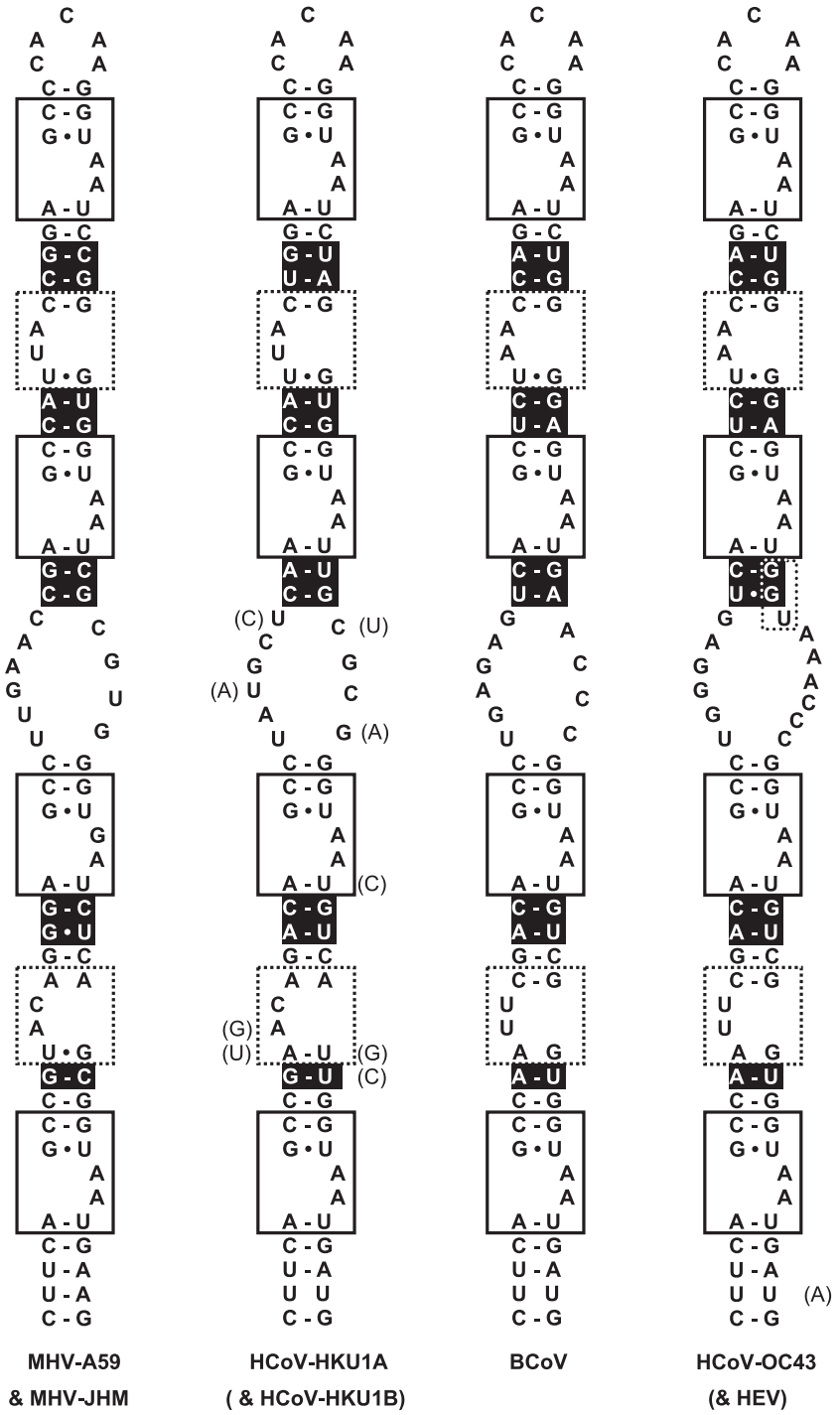


FIG. 1. Highly conserved secondary structure of the PS of group IIa CoVs. Covariations are highlighted in black boxes. The conserved AGC/GUAAU motifs are surrounded with a box of solid lines, while the other motif is surrounded with a box of dashed lines. The GGU insertion in the internal loop of HCoV OC43 and HEV is surrounded by the small stippled box. Residues in parentheses indicate the sequence of closely related species or isolates.

The absence of strong cuts in the internal loop suggests that it is more structured than presented here. In fact, one can draw three additional base pairs (Fig. 2A), U21-G73, A24-U72, and C25-G71, creating another bulge at the 5′ side of the hairpin that would complete the symmetry of alternating bulges. This is also supported by 1-cyclohexyl-3-(2-morpholinoethyl)-carbodiimide metho-*p*-toluene sulfonate probing, which modified U22 and G23 but not U21 (data not shown).

We were unable to detect RNase V_1 cuts in this transcript, although the enzyme was active on control RNAs (data not shown). The lack of V_1 cuts could be due to the relatively short stem regions of just 5 bp or to steric hindrance by the AA

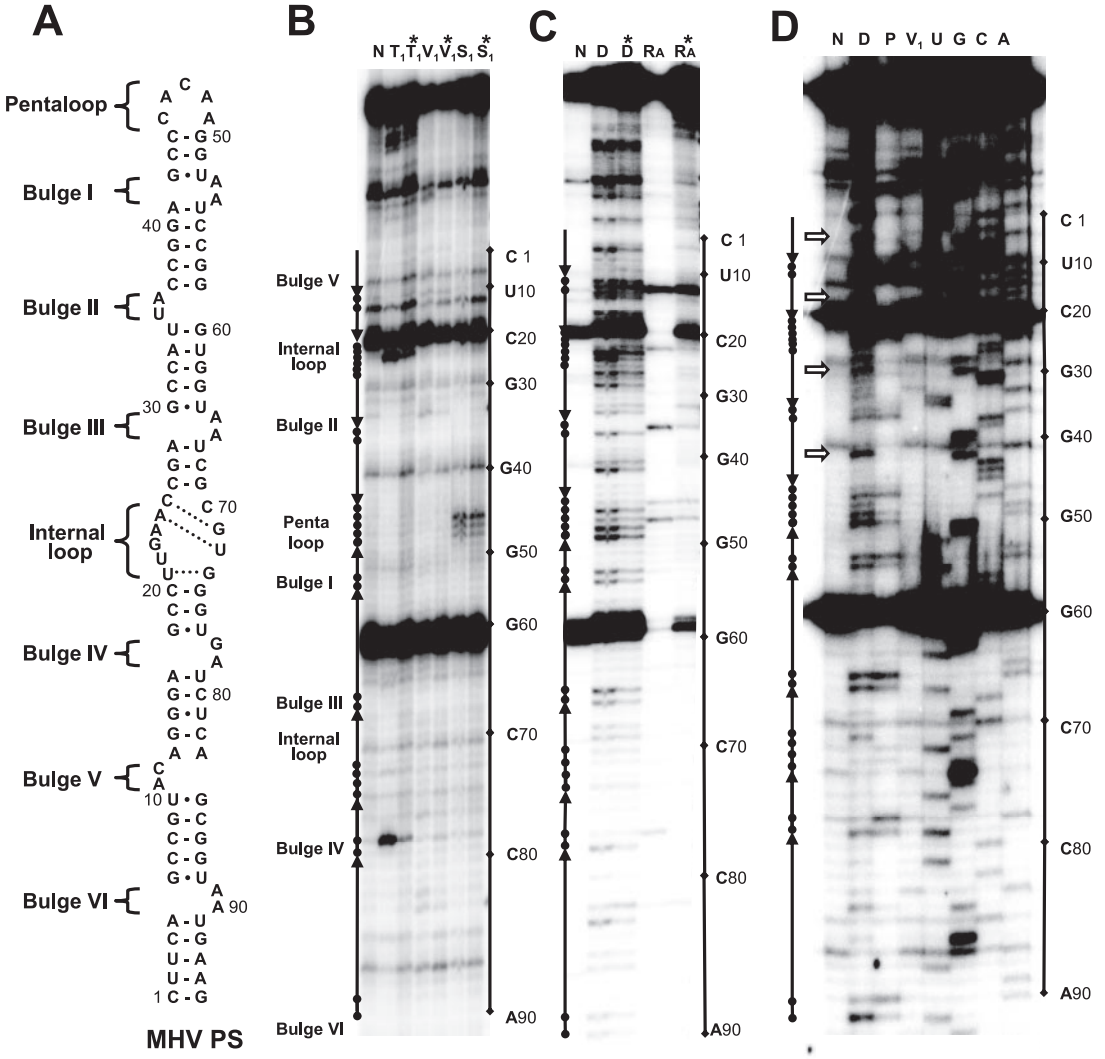
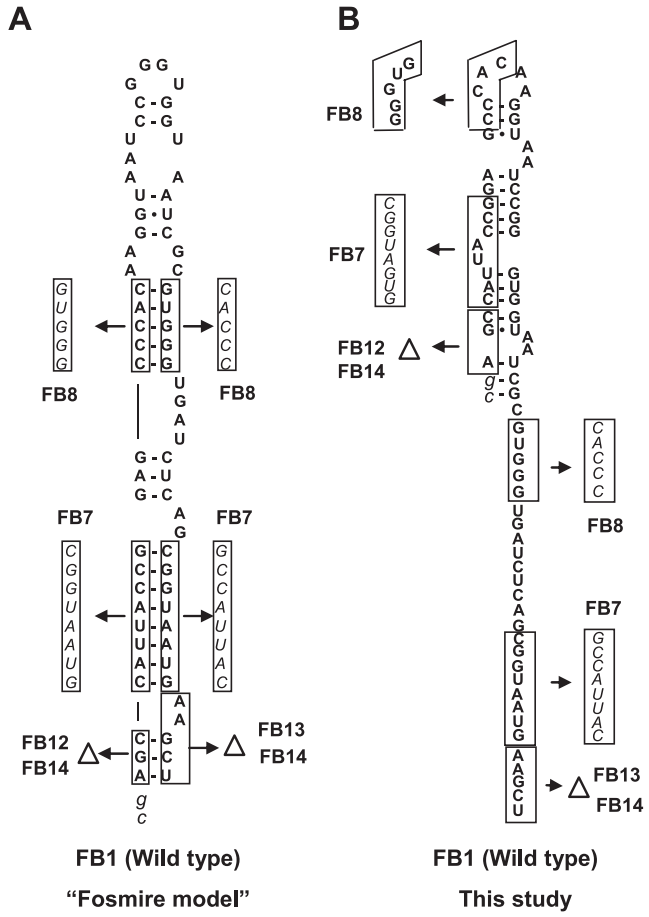


FIG. 2. Structure probing of the MHV PS. (A) Secondary-structure model of the MHV PS. Nucleotides are numbered from the 5′ end of the first cytosine in the PS. Bulges, the internal loop, and the pentaloop are indicated. Additional base pairs in the internal loop are indicated by a dotted line. (B to D) Enzymatic and chemical probing. Susceptible nucleotides were identified by a reference sequencing ladder. For probing reactions, 0.1-μg aliquots of RNA transcripts were treated with nuclease-free water (N), 0.0001 U of RNase T_1 (T_1), 0.0001 U of RNase V_1 (V_1), 4 U of RNase S_1 (S_1), 10 pg of RNase A (RA), 0.01% (vol/vol) DEPC (P), and 0.05% (vol/vol) DMS (D) for 20 min at room temperature in a total volume of 50 μl. DMS and DEPC probing shown in Fig. 2D were done on ice. The incubation buffer contained 10 mM Tris, 100 mM KCl, 10 mM $MgCl_2$, and, optionally, 10 mM of zinc acetate (indicated by the asterisk above these lanes), pH 6.0. Primer extension reactions were carried out with 0.01 μg of treated transcripts, 0.5 μl of a 0.1 mM concentration of the MHV1 primer, 1 μl of 5 mM dGAT, 1 μl of 25 μM dCTP, 0.1 μl of α-^{32}P-labeled dCTP (10 mCi/ml), 1 μl of 5× reverse transcriptase buffer (Promega), and 20 U of Moloney murine leukemia virus reverse transcriptase (Promega).

bulges (6). Altogether, the probing data were in good agreement with the MHV PS model based on structure prediction and phylogenetic analysis (Fig. 1). Preliminary nuclear magnetic resonance data on a synthetic RNA corresponding to bases 37 to 59 is in agreement with the proposed structure (data not shown).

For comparison, Fig. 3A shows the secondary structure of the 69-nt core PS originally proposed by Fosmire et al. (3). Our proposed structure comprises these 69 nt and also falls within the original 190-nt PS identified by Makino et al. (7) and van der Most et al. (12). Fosmire et al. (3) used mutagenesis of the 69-nt PS to identify structural elements important for packaging. However, they obtained some unexpected results that were difficult to explain on the basis of their structure model. We have tried to reconcile these data with our revised PS model. For example, mutations (FB3 to FB6) that disrupted the secondary structure in the Fosmire model would also strongly interfere with the folding of the "AA bulge" hairpin identified in this study and, not surprisingly, interfere with defective interfering RNA packaging. However, stem inversion mutants, FB7 and FB8, which were designed to preserve the secondary structure of the Fosmire model (Fig. 3A) but would disrupt the "AA bulge" hairpin (Fig. 3B), showed a considerable reduction in packaging. Likewise, of three mutants lacking only 3 bp at the bottom of the stem according to the Fosmire model and exhibiting equal levels of thermodynamic stability, two (FB12



Relative encapsidation efficiency:
FB1 = FB13 >>FB12 = FB14
***FB1 > FB8 > FB7**

(*Separate experiment, Fosmire *et al.*, 1992)

FIG. 3. Predicted effects of mutations on the structure of the 69-nt MHV PS. Structure of the wild-type 69-nt MHV PS (A) adapted from the work of Fosmire et al. (3) and (B) redrawn to the new model. The boxed residues indicate the sequences of the FB series mutants FB7, FB8, FB12, and FB13, which were constructed by Fosmire et al. (3), where Δ stands for deletion. The two extra residues (g and c) at the 5′ end were vector derived and stabilized the structure shown in panel B.

and -14) showed significant packaging reduction, while one (FB13) showed nearly wild-type efficiency. However, according to our model, mutant FB13 adopts the same secondary structure as the wild type (FB1), while in mutants FB12 and FB14, the stem is truncated in a manner that likely allows the formation of only one of the AA bulges (Fig. 3B). Thus, reconciling these data with our model suggests that the 2-nt bulges at either side of the hairpin are important elements for the RNA encapsidation of MHV and, presumably, of other group IIa viruses in which these bulges are conserved. Finally, we note that the wild-type 69-nt PS used by Fosmire et al. (3) can form merely the upper half of the "AA bulge" hairpin and probably adopts the structure proposed by those authors most of the time. However, in the initially identified 190-nt PS domain, the entire "AA bulge" hairpin can be expected to form, as was also demonstrated by structure probing of our related 169-nt fragment. This discrepancy in conformations may well explain the inability of the 69-nt PS to interact with the MHV N protein (9) and its fivefold-lower packaging efficiency (3).

The role of N protein in the encapsidation process is currently debated, as virus-like particles could be obtained in the absence of N protein (10), and a prominent role for M protein has been suggested instead (11). Also, Zúñiga and coworkers have recently shown that transmissible gastroenteritis virus N protein is a nonspecific RNA-binding protein that acts as an RNA chaperone (14). Indeed, we observed that purified MHV N protein is able to bind RNA containing a minimal PS (nt C27 to G69) but also tRNA (data not shown). Thus, how PS interacts with different structural proteins and initiates virion assembly is still poorly understood. The improved knowledge of the secondary structure of the MHV PS addressed in this study hopefully will lead to better insight into the encapsidation process for this group of CoVs.

Recently, Hsieh et al. (5) have proposed that a stem-loop structure similar to that proposed by Fosmire et al. (3) can be found in the 3′ part of SARS CoV ORF1b. They showed that this region could promote the encapsidation of a reporter RNA into virus-like particles. However, we were unable to identify a hairpin in this region that conforms to the model shown in Fig. 1. In fact, preliminary data suggest that the SARS CoV encapsidation signal is group I-like (R. C. L. Olsthoorn, unpublished data). In vivo studies to support this hypothesis are in progress.

This research was supported by The Netherlands Organization for Scientific Research (NWO), the M. W. Beijerinck Virology Fund (Royal Netherlands Academy of Arts and Sciences), and a VIDI grant awarded to R.C.L.O.

REFERENCES

1. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410.
2. **Cologna, R., and B. G. Hogue.** 2000. Identification of a bovine coronavirus packaging signal. J. Virol. **74:**580–583.
3. **Fosmire, J. A., K. Hwang, and S. Makino.** 1992. Identification and characterization of a coronavirus packaging signal. J. Virol. **66:**3522–3530.
4. **Gorbalenya, A. E., E. J. Snijder, and W. J. M. Spaan.** 2004. Severe acute respiratory syndrome coronavirus phylogeny: toward consensus. J. Virol. **78:**7863–7866.
5. **Hsieh, P.-K., S. C. Chang, C.-C. Huang, T.-T. Lee, C.-W. Hsiao, Y.-H. Kou, I.-Y. Chen, C.-K. Chang, T.-H. Huang, and M.-F. Chang.** 2005. Assembly of severe acute respiratory syndrome coronavirus RNA packaging signal into virus-like particles is nucleocapsid dependent. J. Virol. **79:**13848–13855.
6. **Lowman, H. B., and D. E. Draper.** 1986. On the recognition of helical RNA by cobra venom V_1 nuclease. J. Biol. Chem. **261:**5396–5403.
7. **Makino, S., K. Yokomori, and M. M. C. Lai.** 1990. Analysis of efficiently packaged defective interfering RNAs of murine coronavirus: localization of a possible RNA-packaging signal. J. Virol. **64:**6045–6053.
8. **Masters, M. S.** 2006. The molecular biology of coronaviruses. Adv. Virus Res. **66:**193–292.
9. **Molenkamp, R., and W. J. M. Spaan.** 1997. Identification of a specific interaction between the coroavirus mouse hepatitis virus A59 nucleocapsid protein and packaging signal. Virology **239:**78–86.
10. **Narayanan, K., C.-J. Chen, J. Maeda, and S. Makino.** 2003. Nucleocapsid-independent specific viral RNA packaging via viral envelope protein and viral RNA signal. J. Virol. **77:**2922–2927.
11. **Narayanan, K., and S. Makino.** 2001. Cooperation of an RNA packaging signal and a viral envelope protein in coronavirus RNA packaging. J. Virol. **75:**9059–9067.
12. **van der Most, R. G., P. J. Bredenbeek, and W. J. M. Spaan.** 1991. A domain at the 3′ end of the polymerase gene is essential for coronavirus defective interfering RNAs. J. Virol. **65:**3219–3226.
13. **Zuker, M.** 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. **31:**3406–3415.
14. **Zúñiga, S., I. Sola, J. L. Moreno, P. Sabella, J. Plana-Durán, and L. Enjuanes.** 2007. Coronavirus nucleocapsid protein is an RNA chaperone. Virology **357:** 215–227.