

AN INTRALEADER OPEN READING FRAME IS SELECTED FROM A HYPERVARIABLE 5' TERMINUS DURING PERSISTENT INFECTION BY THE BOVINE CORONAVIRUS

Martin A. Hofmann, Savithra D. Senanayake, and David A. Brian
Department of Microbiology
University of Tennessee
Knoxville, TN 37996-0845

INTRODUCTION

In an effort to unambiguously establish the 5' terminal nucleotides on bovine coronavirus (BCV) mRNAs, we developed a method that employs the head-to-tail ligation of single-stranded cDNA from extended primers before PCR amplification, cloning and sequencing^{1,2}. Only the mRNAs for the N, M, and S structural proteins were studied in this manner, and mRNAs isolated from cells at various times postinfection through 432 days were studied. We learned that the 5' terminal five nucleotides were hypervariable among four major terminus types, and that certain types predominated at various times postinfection. The type I terminus (GAUUGUG...) predominated within the first 4 days postinfection and was considered the wild type terminus. At 296 days postinfection and beyond, the type II terminus (GAUUAUG...) predominated. Types II, III and IV all possessed a G → A mutation that gave rise to an AUG codon and established an intraleader ORF for 11 amino acids. Although we have not yet been able to show the existence of the peptide in persistently infected cells, we have demonstrated that the intraleader ORF attenuates translation of downstream ORFs during translation *in vitro*.

We therefore hypothesize that the intraleader ORF serves to attenuate virus replication and that this may be a mechanism by which coronaviruses maintain a persistent infection.

MATERIALS AND METHODS

The infection of HRT-18 cells with the Mebus strain of BCV, establishment of persistent infection, and RNA extraction from cells were done as previously described³.

Primer extensions (using oligodeoxynucleotides that bind within the body of each mRNA studied), head-to-tail ligations (with RNA ligase), PCR amplification (using primers that amplify across the head-to-tail junction), cloning, and DNA sequencing, have all been described^{1,2}.

Preparation of RNA transcripts of the cloned N mRNA having the 5' type I terminus (GAUUGUG...) and *in vitro* translation procedures were carried out as previously described⁴. The type II terminus (GAUUAUG...) was prepared from the N-containing clone (pLN, having the type I terminus) by oligodeoxynucleotide-directed *in vitro* mutagenesis using the Amersham mutagenesis kit.

RESULTS

A total of 337 clones of the N, M, and S mRNA termini were sequenced and the results of 307 clones are summarized in Fig. 1. Thirty of the 337 clones had 5'-ward extensions of 1 to 6 nucleotides (i.e., 5'-ward of base #1 in a type I, II or III terminus) and were not included in the data on Fig. 1 except to determine percentage values. The origin of the additional 5'-terminal bases is unknown to us at this time. Several of the clones had the 5'-most base (i.e., base #1) missing and, for this study, we concluded that these were the result of premature termination of first-strand cDNA synthesis during the cloning procedure. With the first base missing, the identity of the terminus was still apparent and the truncated clones were scored as one of the four terminus types.

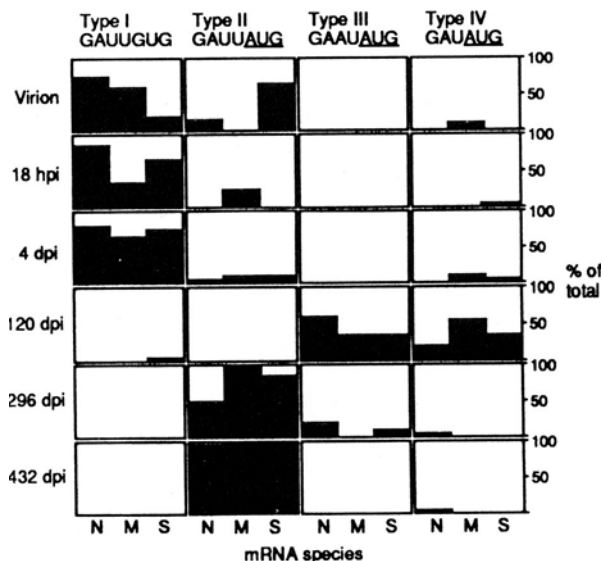


Fig. 1. Occurrence of the four types of 5' termini at different times postinfection. Results for N, M and S mRNA species are depicted separately. Variant types resulting from 5'-ward extension (described in the text) are considered in the percentage calculations.

The G → A mutation in termini types II, III and IV caused the appearance of a methionine codon which begins an 11 amino acid ORF within the leader (Fig. 2). The potential peptide is neutral and hydrophilic, and its synthesis would be terminated by a weak UGA (umbr) codon.

To (i) determine whether a peptide is translated from the ORF, and (ii) examine the effects of the intraleader ORF on the translation of a BCV mRNA, the cloned N

mRNA (pLN) that yields T7 polymerase-generated transcripts with a type I (GAUUGUG...) terminus was mutated to yield transcripts with a type II (GAUUAUG...) terminus. To date we have not been able to identify an 11 amino acid (1.2 kDa) peptide by *in vitro* translation and radiolabeling with ³⁵S-met or ³⁵S-cys. To examine the effect of the intraleader ORF on translation of downstream ORFs,

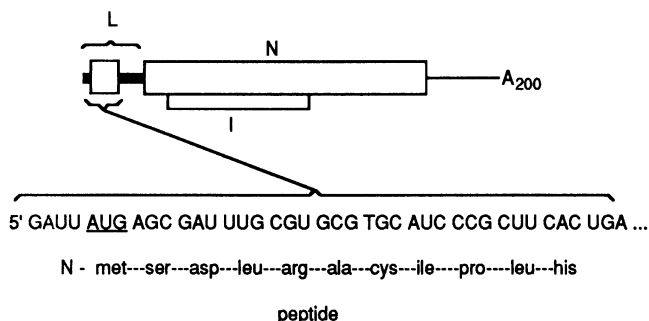


Fig. 2. Schematic diagram showing the nucleotide sequence and resulting amino acid sequence of the intraleader ORF on the BCV N mRNA. The I protein ORF (named for internal open reading frame) which exists on the bicistronic N mRNA⁴ is also shown.

transcripts of pLN-type I and pLN-type II were generated, quantitated by Northern analysis, and equal amounts were translated in parallel. Quantities of N and I proteins, both products of downstream ORFs on the N mRNA⁴, were determined by imaging on the Ambis Radioanalytic Imager (Fig. 3).

Fig. 3 illustrates that in the presence of the intraleader ORF, synthesis of N is decreased by 40% and I is decreased by 10%.

DISCUSSION

We do not know the mechanism by which the 5' sequence heterogeneity was generated. We do not believe it arose as a function of the analytical procedure (i.e., by the action of enzymes used in the copying, amplifying and cloning of the RNA) since a cloned transcript of known sequence (i.e., type type I terminus, GUAAGUG...) when prepared, transfected into cells, recovered and carried through the amplification and cloning procedure, yielded no variants (in 18 of 18 clones).

We think it is very likely that a variety of 5' termini could have been present in the virus inoculum and passed on by infectious virus. Despite the fact that the virus inoculum used had been sequentially plaque purified three times, it is conceivable that more than one leader-containing genome, mRNA, or defective-interfering RNA molecule serving as potential leader donors could have been packaged in a single particle and carried through the plaque purification process. Such molecules would have become the source of the variant 5' termini, possibly by a leader switching mechanism⁵. We have demonstrated that subgenomic mRNA and defective-interfering molecules of BCV do become packaged³. To determine whether 5' terminus hypervariability arises *denovo* awaits production of a clone of BCV derived from an infectious genomic molecule.

To our knowledge, a coronavirus intraleader open reading frame has not been previously reported. But then, published coronavirus leader sequences have been determined only on RNA obtained from cells immediately after an acute infection⁶⁻¹¹.

Although the mechanism(s) by which variant termini become distributed among the mRNA species and establish predominance at certain times during persistent infection are not known, we propose that the selection of leaders with an internal open reading frame is causally related to the persistent infection. Namely, we suggest that the

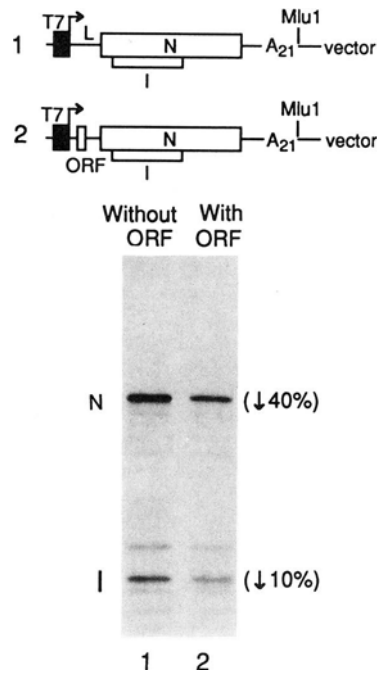


Fig. 3. Effect of the intraleader ORF on the translation of N and I proteins *in vitro*. Equal amounts (1 μ g each as quantitated by RNA blot hybridization) of transcripts made *in vitro* by T7 RNA polymerase were translated in wheat germ lysate in the presence of ³⁵S-methionine and the radioactive products were analyzed by SDS-polyacrylamide gel electrophoresis and quantitated by the Ambis Radioanalytic Imager.

intraleader ORF serves to attenuate virus gene expression (translation), which in turn attenuates virus replication. Since cells infected with a less virulent (i.e., attenuated) virus are less likely to die, cells infected with the attenuated virus are more likely to survive as persistently infected cells. Thus, selection of the intraleader open reading frame may be a mechanism by which coronavirus maintain a persistent infection.

REFERENCES

1. Hofmann, M. A., and D. A. Brian. 1991. *PCR Methods and Applications* 1:43-45.
2. Hofmann, M. A., and D. A. Brian. 1991. *BioTechniques* 11:30-31.
3. Hofmann, M. A., P. B. Sethna, and D. A. Brian. 1990. *J. Virol.* 64:4108-4114.
4. Senanayake, S. D., M. A. Hofmann, J. L. Maki, and D. A. Brian. 1992. *J. Virol.* 66:5277-5283.
5. Makino, S., and M. M. C. Lai. 1989. *J. Virol.* 63:5285-5292.
6. Brown, T. D. K., M. E. G. Boursnell, and M. M. Binns. 1984. *J. Gen. Virol.* 65:1437-1442.
7. Lai, M. M. M. C., R. S. Baric, P. R. Brayton, and S. A. Stohlman. 1984. *Proc. Natl. Acad. Sci. USA.* 81:3626-3630.
8. Kamahora, T., L. H. Soe, and M. M. C. Lai. 1989. *Virus Research* 12:1-9.
9. Schreiber, S. S., T. Kamahora, and M. M. C. Lai. 1989. *Virology* 169:142-151.
10. Sethna, P. B., M. A. Hofmann, and D. A. Brian. 1991. *J. Virol.* 65:320-325.
11. Shieh, C.-K., L. H. Soe, S. Makino, M.-F. Chang, S. A. Stohlman, and M. M. C. Lai. 1987. *Virology* 156:321-330.