Infectious Bronchitis Virus Nucleocapsid Protein Interactions with the 3' Untranslated Region of Genomic RNA Depend on Uridylate Bases

[1]ELLEN W. COLLISSON, [1]MINGLONG ZHOU, [2]PAUL GERSHON, AND [1]JYOTHI JAYARAM
[1]Department of Veterinary Pathobiology, Texas A&M University, College Station, TX; [2]Institute of Biosciences and Technology, Texas A&M University, Houston, TX

1. INTRODUCTION

Infectious bronchitis virus (IBV) is a highly contagious virus of worldwide economic concern in poultry. Infection causes respiratory disease, as well as kidney lesions, reproductive problems and gastrointestinal disfunction. The virion consists of four structural proteins: the membrane, spike, nucleocapsid (N), and the small envelope proteins. The IBV N protein, a basic, phosphorylated protein of 409 residues, is highly conserved especially within the central or middle region among IBV strains, (Williams *et al*, 1992).

In addition to being closely associated with the RNA genome, the coronavirus nucleocapsid protein has been suggested to have multiple functions (Robbins *et al.*, 1986; Compton *et al.*, 1987; Stohlman *et al.*, 1988; Baric *et al.*, 1988). The MHV nucleocapsid protein specifically binds to small leader-containing RNA. Within the cytosol of MHV infected cells, it can interact with membrane-bound small leader RNA in transcription complexes (Stohlman *et al.*, 1988). Anti-nucleocapsid monoclonal antibodies reportedly precipitate both full-length and subgenomic mRNA, as well as replicative intermediate RNA (Baric *et al.*, 1988). The amount of N protein found associated with the genome and the putative functions of this

The Nidoviruses (Coronaviruses and Arteriviruses).
Edited by Ehud Lavi *et al.*, Kluwer Academic/Plenum Publishers, 2001.

protein suggest that it should readily associate with additional regions of coronavirus RNA.

Overall, functions of the coronavirus N depend on its interaction with viral RNA. IBV N has been shown to bind to sequences representing the 3' untranslated region (UTR) of genomic or subgenomic RNA (Zhou *et al*, 1996). Specificity was suggested by preferential interactions with selected regions of the 3' UTR and the absence of interactions with ribosome and yeast tRNA. A region lying between 215 and 78 nt from the 3' end of genomic and subgenomic RNA did not shift when exposed to N, unlike the other similarly sized oligonucleotides spanning the 3' UTR. This region overlapped with CD, corresponding to the region lying 155 nt within the UTR at the ultimate 3' end. It was further shown that the amino and carboxyl regionw, but not the middle region, were able to bind to the 3' UTR (Zhou and Collisson, 2000). The binding site of the amino domain required more than the first 91 residues. In this report, sequences required for recognition by the IBV N and the UTR were investigated.

2. MATERIALS AND METHODS

2.1 Oligonucleotides

cDNA templates for RNA oligonucleotides corresponding to sequence within the 3' UTR of genomic IBV RNA were generated by PCR and incorporated the T7 RNA polymerase promoter. The oligonucleotides were synthesized by in vitro transcription with T7 RNA polymerase (Promega Corporation, Madison, WI). These RNA probes were labeled with ^{32}P by incorporating the ^{32}P-CTP during RNA transcription. A 40mer poly-U was synthesized and labeled with ^{32}P using T4 polynucleotide kinase (Promega Corporation, Madison, WI).

2.2 Polypeptides

Preparation of the N protein and polypeptides has been described in Zhou & Collisson (2000) and Zhou *et al* (1996). Briefly, the encoding nucleotide sequences were amplified by PCR from template plasmid and cloned into *Bam*HI and *Hin*dIII sites in the pQE8 vector (Qiagen Inc., Chatsworth, CA). Expression of protein from pQE8 resulted in fusion polypeptides with amino terminus six histidine tags. The polypeptides were purified first with a N^{+2}-NTA resin and then with Sephadex G200 or Bio-gel P60 (Ausubel *et al*, 1987).

2.3 Gel shift assays

Protein-RNA interactions were analyzed by a modified gel-shift assay (Zhou *et al*, 1996). RNA and varying concentrations of nucleocapsid protein were co-incubated for 20 min at room temperature in 10 µl of gel shift buffer (25mM Hepes, 25 mM EDTA, 150 mM NaCl, 5 mM DTT, 10% glycerol and 20 units Rnase Inhibitor, Boehringer Mannheim Corp., Indianapolis, IN). After adding 1 µl 10x sample buffer, the reaction mixtures were loaded onto a 1% agarose gel and electrophoresed at 60V in 1x Tris-borate-EDTA (Sambrook *et al.*, 1989). Gels were then dried and autoradiographed.

3. RESULTS

cRNA was synthesized that corresponded to regions of the 3' UTR of IBV RNA. The sequences that were selected incorporated putative stem-loop structures (Fig 1). These oligonucleotides were used for gel shift assays to evaluate binding whole N and the amino, middle and carboxyl domains of N (Zhou and Collisson, 2000). In previous studies, 5 nM of IBV N protein and the amino and carboxyl regions had been found to shift RNA encoding the N gene to the 3' end of the genome and CD RNA, the ultimate 3' terminal 155 nucleotides (Zhou and Collisson, 2000). Interactions between the N polypeptides and RNA sequences within the 3' UTR RNA's were examined using 5 nM of IBV N protein (Fig 2) and N polypeptides (Table 1). The intact N protein, the carboxyl terminal polypeptide (C140) or the amino region (A171), shifted RNA1, RNA2, RNA3, RNA4, RNA5 and RNA7 probes very effectively. The shift of RNA7 with varying concentrations of the amino, middle and carboxyl regions of N are shown in Figure 3a. RNA 6 did not shift when exposed with N (Fig 3b) or the N polypeptides. This oligonucleotide corresponded to a region found within the larger 155 nt EF that also did not interact with N on gel shift assays.

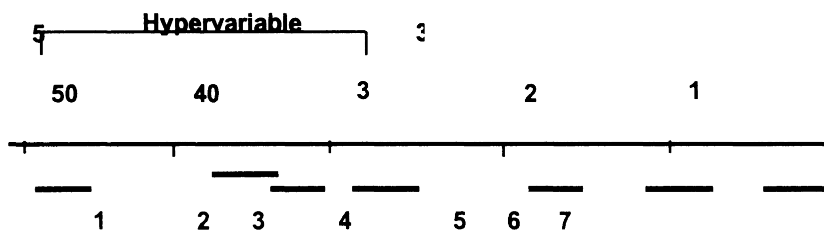

Figure 1. Schematic of synthesized 3' UTR sequences used for studies identifying interactions with the IBV N protein and polypeptides.

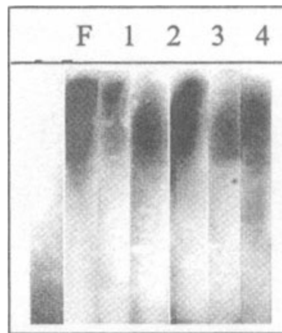

Figure 2. Gel shift of 1 ng of ^{32}P labeled oligonucleotide probe representing sequences from the 3' UTR of the IBV genome. The numbers above each lane indicate the oligonucleotide, as shown in Figure 1, used to interact with the 5nM of N protein. Free RNA1 in the absence of protein is shown in lane F.

Varying concentrations of polypeptide used in gel shift assays also demonstrated interactions with N and N polypeptides. The gel shifts of labeled RNA 7 when exposed to N, the amino region or the carboxyl region are shown in Fig 2. The polypeptide corresponding to the middle region of N did not shift any of these RNA probes (Table 1; Fig 2). In contrast, no interaction was observed between N or N polypeptides at any concentration of RNA 6 examined (Fig 3a). The single obvious difference between RNA 6 and the oligonucleotides that did bind to the N polypeptides was the low number of uridylates (U) in RNA 6. RNA 6 had only 7%, whereas from 24% to 58% of the other oligonucleotides consisted of U.

Table 1. Interactions of N polypeptides with oligonucleotides in Fig 1.

RNA	Amino	Middle	Carboxyl	Whole N	%U*
1	+	-	+	+	41 (41)
2	+	-	+	+	52 (41)
3	+	-	+	+	47 (40)
4	+	-	+	+	24 (41)
5	+	-	+	+	58 (41)
6	-	-	-	-	7 (43)
7	+	-	+	+	36 (53)

* () indicate the number of bases within each oligonucleotide.

Gel shifts of additional oligonucleotides with N and the N domain fragments were examined. No detectable interactions were observed by gel shift assays between N polypeptides and an oligonucleotide synthesized with A, C, and G bases and no U's (Fig 3). However, another probe was synthesized with two U triplets placed within the ACG repeating 40-mer nucleotide sequence. Binding was observed with the ACG incorporated with the U triplets.

As the U bases seemed to be important in the interactions with the N protein, a 40mer poly-U sequence was used to examine potential interactions in the absence of A, G, C and the absence of any secondary structure. In the presence of N, the labeled poly-U sequence shifted within the agarose gel.

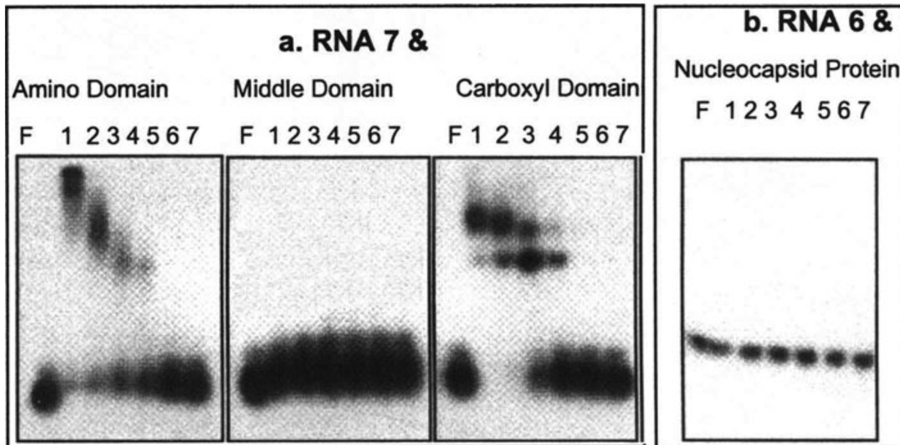

Figure 3. Gel shift of ²P-labeled RNA7 with 3µg of the three regions of the N protein and RNA6 with 3µg of whole N protein. RNA in lanes 1 to 7 were exposed to serial 2-fold dilutions of polypeptide with 1 ng of free probe in lane 1. F indicates labeled RNA that was not exposed to protein.

4. DISCUSSION

RNA structures recognized by various viral or cellular proteins are tremendously varied, ranging from single-strands, hairpins to seemingly complex irregular helices and cation dependent tertiary structures (Draper, 1995). A bulged stem-loop structure in the 3' non-translated region on the MHV genome was shown to be critical for negative RNA synthesis and cellular protein interaction, and a stem-loop structure in polymerase gene was shown to be critical for packaging (Hsue and Masters, 1997; Yu and Leibowitz, 1995; Fosmire *et al.,* 1992). In addition, stem-loops on the TAR of HIV-1 RNA, U1 snRNA, poliovirus 5' noncoding region and 16S ribosomal RNA have all been implicated as critical for recognition of proteins (Weeks and Crothers, 1991; Hall and Stump, 1992). It is anticipated that the interaction between N and the 3' UTR also involves secondary structure, such as a stem-loop. Although the buffer conditions in these studies were not the most favorable for maintaining secondary structure, these studies do provide evidence that the presence of uridylate bases contribute to the successful binding of N to sequences within the 3'

UTR. Uridylate-rich sequences have also been identified as protein binding sites for poly (A) polymerase from vaccinia virus (Gershon and Moss, 1993, Deng and Gershon, 1997). The eight-nucleotide sequence (UAUUUUCU) in the potato X 3' untranslated region (UTR) is required for both host protein binding and viral replication (Sriskanda *et al*, 1996). Proteins in cellular extracts bind to AU-rich repeats in the 3'-UTR of human TNF-alpha mRNA and the binding activity correlates the expression level of the TNF-alpha in human epithelial cancer cells (Wang *et al*, 1998).

An overall influence of secondary or tertiary structure may also impact binding. Interactions may be demonstrated with small simple units of RNA, but in fact, be optimized with larger regions that, for example, confer more tertiary stability. Often, function depends on protein-induced alterations in RNA that create or expose sites necessary for binding to additional proteins (Draper, 1995).

Although RNA structures can be identified by computer analysis, predicting structures or sequences that are responsible for protein interactions is not always successful. The stability of the RNA structures may be directly related to the biological functions. The 3' UTR of mRNA has been shown to differentially regulate gene expression (Conne *et al*, 2000). The AUUUA pentanucleotide sequence is associated interactions between mRNA and RNA-binding proteins. Defective interactions lead to abnormal stabilization (Chen and Shyu, 1995; Peng *et al*, 1998). It is possible that N interacts with the U triplets so as to inhibit interactions with these proteins and stabilize viral RNA.

ACKNOWLEDGEMENTS

This research was supported by the U. S. Poultry & Egg Association, grant #142; USDA Animal Health (Section 1433), # TEXO-6824, Texas A&M University Interdisciplinary Program, 1998, and Texas Technology Research program #999902-212.

REFERENCES

Ausubel, F. M., Brent, R., Kingston, R., Moore, D. D., Seidman, J. G., Smith, J. A. and Struhl, K., 1994, *Current protocols in molecular biology*. 18.5.3-4. John Wiley & Son, inc.

Baric, R. S., Stohlman, S. A., Razavi, M. K., and Lai, M. M. C., 1985, Characterization of leader-related small RNAs in Coronavirus-infected cells. Further evidence for leader-primed mechanism of transcription. *Virus Res*. **3**: 19-33.

Masters, P. S., 1992, Localization of an RNA-binding domain in the nucleocapsid protein of the coronavirus mouse hepatitis virus. *Arch. Virol.* **125**: 141-160.

Chen, C.Y. & Shyu, A.B., 1995, AU-rich elements: characterization and importance in mRNA degradation. Trends Biochem. Sci. 20:465-470.

Compton, S. R., Rogers, D. B., Holmes, K. V., Fertsch, D., Remenick, J. and McGowan, J. J., 1987, In vitro replication of mouse hepatitis virus strain A59. *J. Virol.* **61**:1814-1820.

Conne, B., Stutz, A. & Vassalli, J.-D., 2000, The 3' untranslated region of messenger RNA : A molecular 'hotspot' for pathology? Nature Medicine 6:637-641.

Peng, S.S., Chen, C.Y., Xu, N. & Shyu, A.B., 1998, RNA stabilization by the AU-rich element binding protein, HuR, an EAV protein. EMBO J. 17:3461-3470.

Robbins, S. G., Frana, F. M., McGowan J. J., Boyle, J. F., Holmes, K. V., 1986, RNA-binding proteins of coronavirus MHV: detection of monomeric and multimeric N protein with an RNA overlay protein blot assay. *Virology.* **150**: 402-410.

Stohlman, S. A., Baric, R. S., Nelson, G. N., Soe, L. H., and Welter, L. M., 1988, Specific interaction between Coronavirus leader RNA and Nucleocapsid Protein. *J. Virology.* **62:** 4288-4295.

Williams, A. K., Wang, L., Sneed, L. W. and Collisson, E. W., 1992, Comparative analysis of the nucleocapsid genes of several strains of infectious bronchitis virus and other coronaviruses. *Virus Res.* **25**: 213-222

Williams, A. K., Wang, L., Sneed, L. W. and Collisson, E. W., 1993, Analysis of a hypervariable region in the 3' non-coding end of the infectious bronchitis virus genome. *Virus Res.* **28**: 19-27

Zhou, M. & Collisson, E.W., 2000, The amino and carboxyl domains of the infectious bronchitis virus nucleocapsid protein intereract with 3' genomic RNA. Virus Res, 67:31-39.

Zhou, M., Williams, A. K., Chung, S., Li, W. and Collisson, E. W., 1996, The infectious bronchitis virus nucleocapsid protein binds RNA sequences in the 3' terminus of the genome. *Virology.* **217:** 191-199.