

Localization of an RNA-binding domain in the nucleocapsid protein of the coronavirus mouse hepatitis virus

P. S. Masters

Wadsworth Center for Laboratories and Research, New York State Department of Health, Albany, and Department of Biomedical Sciences, School of Public Health, State University of New York at Albany, Albany, New York, U.S.A.

Accepted December 11, 1991

Summary. The interaction between the nucleocapsid (N) protein of mouse hepatitis virus (MHV) and RNA was studied in an effort to define portions of the N molecule that participate in binding to RNA. N mRNAs transcribed from SP6 and T7 vectors were translated in a rabbit reticulocyte lysate. Analysis of synthesized N protein in a nondenaturing gel system showed that it bound *in vitro* to an endogenous RNA in the reticulocyte lysate but not to its own mRNA. A set of deletion mutants was constructed in order to localize the RNA-binding activity of the N protein. It was found that removal of as much as 135 amino-terminal or 57 carboxy-terminal amino acids from the molecule had little or no effect on RNA binding. Moreover, deletion mutants lacking both termini still retained RNA-binding ability. By contrast, internal deletions or truncations extending beyond these two limits effectively abolished RNA binding by N protein. Thus, the RNA-binding region of N has been mapped to the second (central) of the three structural domains of the molecule.

Introduction

Coronaviruses constitute a family of viruses whose single-stranded, positive-sense RNA genomes have the largest coding capacities among the RNA viruses [27, 31]. Paradoxically, coronaviruses have relatively few structural proteins. Their large, roughly spherical virions are enclosed by a membrane envelope containing two (or, in some cases, three) types of membrane-bound glycoproteins. Internal to this, multiple monomers of a nucleocapsid protein (N) encapsidate the viral genome, which, in the case of the prototype mouse hepatitis virus (MHV), is some 31,000 nucleotides in length [10, 19]. The nucleocapsids of coronaviruses are helically symmetric [12], a feature seen in only one other group of positive-stranded RNA viruses, those of the proposed family *Toro-*

viridae [25]. Consequently, these structures pose interesting problems with respect to genome translation as well as in understanding the processes of encapsidation and assembly.

Elucidation of coronavirus nucleocapsid structure and function will require description of the multiple interactions in which the N protein must participate: (i) the binding of N protein along the entire length of the RNA genome; (ii) contacts between adjacent N monomers on the RNA and between N monomers that neighbor each other per helical turn; and (iii) association between N protein and the membrane glycoprotein (M), which may stabilize virion assembly [32]. For at least some of these, it may be possible to map regions of the N molecule that are involved in a particular interaction. The work presented here focusses on the first type of interaction: the ability of N to bind to RNA.

Coronavirus N proteins are markedly basic, and they form complexes with RNA that tend to dissociate in high concentrations of salt and provide only limited protection against the action of ribonucleases [12, 22]. In these respects, they resemble the NP proteins of the orthomyxoviruses [1, 4]. By contrast, the N-RNA complexes of the rhabdoviruses and the paramyxoviruses, which also form helically symmetric nucleocapsids, are stable in high salt and are largely resistant to ribonuclease [1, 11]. Few details are known about the association between N protein and RNA in the coronaviral nucleocapsid. We have previously compared the amino acid sequences of the N proteins of five strains of MHV and, on the basis of this analysis, proposed a three domain structural model for the MHV N protein [20]. In this paper it is shown that the central of these three domains is responsible for the RNA-binding ability of the MHV N molecule.

Materials and methods

Construction of transcription vectors for the full-length MHV N gene and deletion mutants

Transcription vectors linking the N gene of MHV to the RNA polymerase promoter of either bacteriophage T7 or bacteriophage SP6 were constructed by standard recombinant DNA techniques [14]. A cDNA clone, pA15, containing a nearly full-length copy of the N mRNA of MHV-A59 [20] (EMBL/GenBank accession no. M35256) was restricted with *Sna*BI and *Sac*I, made blunt-ended, and inserted in either orientation into the vector pGEM3Zf (-) (Promega), which had been opened and blunt-ended at the *Kpn*I and *Sph*I sites of the polylinker. The resulting constructs, pA50 and pA48, contained the T7 or SP6 promoter, respectively, linked to 54 nt of the 78 nt 5' untranslated region (UTR) of RNA7 of MHV-A59 [8, 19], followed by the entire N protein coding region and 225 nt of the 301 nt 3' UTR (Fig. 1). *Hind*III-linearized pA50 encoded run-off transcripts that were bounded by polylinker sequences of 17 nt and 6 nt at their 5' and 3' ends; *Sac*I-linearized pA48 encoded run-off transcripts that were bounded by polylinker sequences of 15 nt and 2 nt at their 5' and 3' ends (Fig. 1). Full-length N mRNAs transcribed from these two plasmids had identical translation efficiencies *in vitro*. Truncated N mRNAs were transcribed from pA50 that had been linearized with *Bsm*I, *Sca*I or *Spe*I and from pA48 that had been linearized with *Acc*I or *Eco*RI (Fig. 1).

An internal deletion mutant of the N gene was constructed by restriction of pA50 with

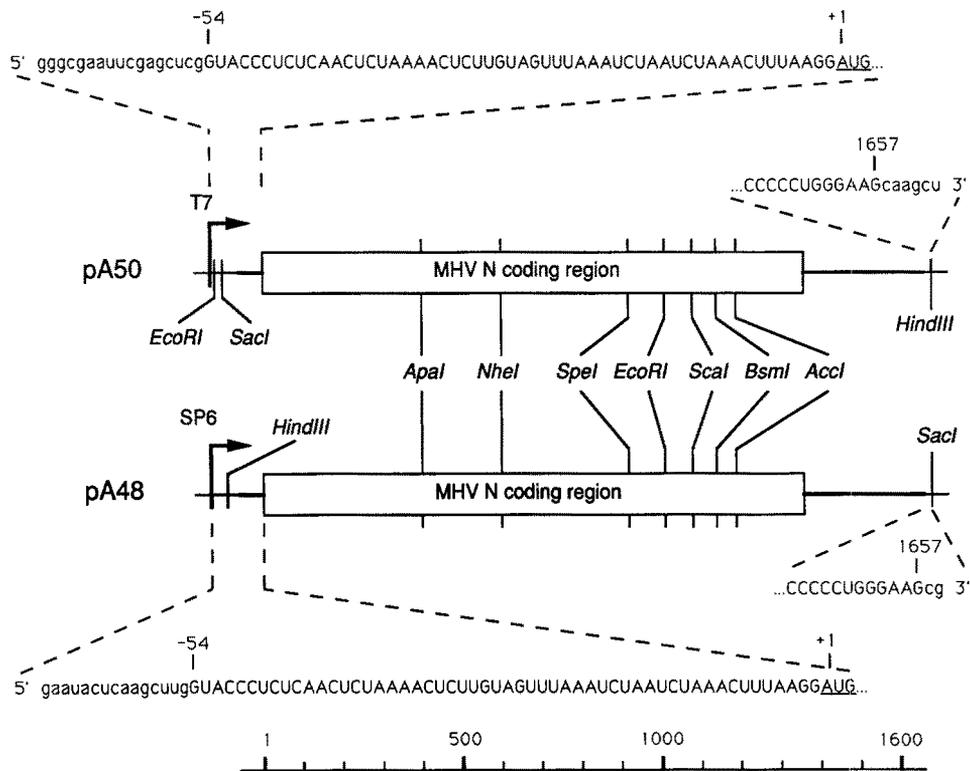


Fig. 1. MHV N gene transcription vectors. The plasmids pA50 and pA48 were constructed as described in Materials and methods. Relevant restriction enzyme sites are indicated in the MHV N coding region and in flanking regions unique to each vector. Shown in expanded view are sequences of the 5' and 3' extremes of the full-length run-off transcripts encoded by each. Capitalized nucleotides are identical to those of the authentic N mRNA; lower case nucleotides are those derived from vector polylinker sequences. Numbers above sequences and on the number line at the bottom indicate nucleotide positions from the start of the N coding region [20]. The N mRNA initiation codon is underlined

NheI and *SpeI* (Fig. 1), followed by religation of these compatible sites in the vector fragment. The resulting plasmid, pA60, lacked nt 603–920 of the N gene and encoded the mutant protein N/Nhe-Spe. A second internal deletion mutant was generated by cleavage of pA50 with *ApaI* and *NheI* (Fig. 1), followed by blunting and religation of the vector fragment. The resulting plasmid, pA66, lacked nt 401–598 of the N gene and encoded the mutant protein N/Apa-Nhe.

A construct fusing a portion of the N gene of vesicular stomatitis virus (VSV) to the MHV N gene was prepared from pA56, an intermediate derivative of pA50 in which a *KpnI* site had been inserted immediately preceding the start codon. Details of the construction of pA56 will be described elsewhere. The *KpnI*(blunted)-*HindIII* fragment of pA56 was exchanged for the *StuI*-*HindIII* fragment of the mutant VSV N SP6 transcription vector pN112 [16]. This yielded a plasmid, pA58, containing nt 1–69 and 370–420 of the coding region of the VSV N gene followed by the entire coding region of the MHV N gene; the encoded protein was designated N/fus1. A deletion mutant of pA58 was created by exchange of the *ApaI*(blunted)-*HindIII* fragment of pA50 for the *StuI*-*HindIII* fragment of pN112. The resulting plasmid, pA63, contained nt 1–69 and 370–420 of the VSV N gene followed by nt 406–1362 of the MHV N gene; the encoded protein was designated N/fus2.

During the construction of pA63, there serendipitously arose a similar plasmid, pA65, containing the same extent of MHV N sequence preceded only by nt 1–24 of the VSV N gene; the protein encoded by this construct was designated N/fus4.

All plasmid constructs were checked by restriction analysis, and all newly formed junctions were directly verified by dideoxy sequencing [24]. DNA plasmids to be used as transcription templates were purified by two cycles of equilibrium centrifugation in CsCl gradients in the presence of propidium iodide [14].

In vitro transcription reactions

Capped, run-off SP6 and T7 mRNAs [18] were typically synthesized in 100 μ l reactions containing 40 mM Tris hydrochloride (pH 7.9), 6 mM MgCl₂, 2 mM spermidine, 10 mM NaCl, 10 mM dithiothreitol, 100 U RNasin (Promega), 500 μ M each ATP and CTP, 100 μ M each GTP and UTP, 500 μ M m⁷GpppG (New England Biolabs), 10 μ Ci [5,6-³H]UTP (45 Ci/mmol; Amersham), 2 μ g linearized DNA template, and either 40 U SP6 RNA polymerase or 80 U T7 RNA polymerase (New England Biolabs). Reactions were carried out for 60 min at either 40 °C (SP6) or 37 °C (T7). Products were extracted twice with phenol-chloroform, twice with chloroform and were twice precipitated with ethanol. The incorporation of labeled UTP was determined by binding of product RNA to DEAE filter paper (Whatman DE 81); the size and homogeneity of synthesized RNA were verified by electrophoresis in 5% polyacrylamide gels containing 8 M urea, followed by fluorography. For the synthesis of high specific activity, ³²P-labeled N mRNA, all reaction conditions were identical, except that the labeling nucleotide was 100 μ Ci [α -³²P]UTP (3000 Ci/mmol; Amersham).

In vitro protein synthesis

Translation of synthetic mRNAs was carried out in a micrococcal nuclease-treated rabbit reticulocyte lysate (Amersham). In the standard reaction, 0.3 to 0.4 μ g of mRNA in 3 to 4 μ l of H₂O was used to program 12 μ l of reticulocyte lysate containing 14 U of RNasin and 8 to 10 μ Ci of [³⁵S]methionine (> 1100 Ci/mmol; New England Nuclear). Incubations were carried out for 90 min at 30 °C. In many experiments, where noted, 1 μ l of either H₂O or RNase A was then added, and samples were incubated an additional 30 min at 30 °C. Bovine pancreatic RNase A (ca. 100 U/mg; Sigma type XII-A) was rendered DNase-free by heat treatment [14]. For translation of authentic MHV mRNA, total RNA was purified [5] from MHV-A59-infected (as well as mock-infected) mouse 17 clone 1 cells at 20 h post-infection, and 6 μ g was translated as above.

Gel electrophoresis

In vitro-synthesized proteins were analyzed both by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) [7] and by nondenaturing PAGE employing the standard Laemmli discontinuous system in which SDS was omitted from all gels and buffers, and samples were not heated prior to electrophoresis [15]. Reticulocyte lysate samples (4 μ l) were run on SDS-PAGE gels (10% acrylamide-0.27% methylene bisacrylamide) and nondenaturing PAGE gels (7.5% acrylamide-0.2% methylene bisacrylamide) that were 19 cm long by 1.5 mm thick. Nondenaturing PAGE was performed at 80 V at ambient temperature (ca. 18 h); under these conditions there was no detectable heating of the gel for the duration of the run. Gels were fixed in 50% methanol-10% acetic acid and were impregnated with 1 M sodium salicylate prior to drying and fluorography at –70 °C. [³⁵S]Methionine-labeled bands were quantitated either by (i) excision and solubilization in 2 ml of 30% H₂O₂ for 48 h at 50 °C prior to counting in 8 ml of Aquasol (New England Nuclear) or (ii) direct counting with a Betascope 603 Blot Analyzer (Betagen).

Results

Binding of translated MHV N protein to RNA

In order to develop *in vitro* assays for MHV N protein function, the N gene was inserted into bacteriophage T7 and SP6 transcription vectors. Synthetic mRNA containing the entire coding region of the authentic N mRNA, as well as most of the 5' and 3' UTRs, was used to program a protein synthesizing system from rabbit reticulocyte lysate. The translation product from this message, labeled with [³⁵S]methionine, was analyzed by SDS-PAGE and had a mobility indistinguishable from that of N protein translated *in vitro* from total RNA purified from MHV-A59-infected cells (Fig. 2 A, lanes b and c). In addition, N protein translated from synthetic mRNA was immunoprecipitable by either anti-N or anti-MHV polyclonal antibodies (Peng and Masters, unpubl. data). *In vitro*-translated N protein had a slightly higher apparent molecular weight than N protein from [³⁵S]methionine-labeled virus of MHV (Fig. 2 A, lane a). This was possibly due to differential extents of N protein phosphorylation [30] *in vitro* and *in vivo*, or it may indicate that the mature N protein found in MHV virions had undergone amino-terminal modification, as has been observed for the N protein of the coronavirus avian infectious bronchitis virus [2]. This apparent difference does not bear on the present study.

To examine the properties of native N protein, the *in vitro* translation product was analyzed by nondenaturing PAGE. In this gel electrophoresis system, the major fraction (ca. 60%) of translated N protein migrated into the running gel, forming a discrete band with a mobility of 0.25 relative to bromophenol blue (Fig. 2 B, lane f). The MHV-A59 N protein contains a considerable excess of basic amino acid residues (63 lysines and arginines as opposed to 46 glutamates and aspartates) and has a calculated pI of 10.5 [20]. Initially it was surprising that such a molecule, expected to have a substantial net positive charge, was migrating toward the positive pole during electrophoresis. This result suggested that the N molecule was tightly bound to some polyanionic molecule and migrated as part of a complex having an overall negative charge. That this, in fact, had occurred was shown by incubation of translated N protein with RNase A prior to nondenaturing gel electrophoresis. RNase treatment abolished the ability of translated N protein to migrate into the nondenaturing gel (Fig. 2 B, lane g). Thus, synthesized N protein had bound to some RNA species present in the translation reaction, and as a consequence of this complex formation it was carried into the nondenaturing gel during electrophoresis. N protein translated from total RNA purified from MHV-A59-infected cells behaved in the same RNase-sensitive manner (Fig. 2 B, lanes h and i). Moreover, translated N proteins from MHV-1, MHV-3 and MHV-S also migrated to the same position on nondenaturing PAGE (Parker MM, Masters PS, unpublished data).

Titration of the translated N protein with a range of concentrations of either RNase A or RNase T1 led to a stepwise decrease in the mobility of N protein

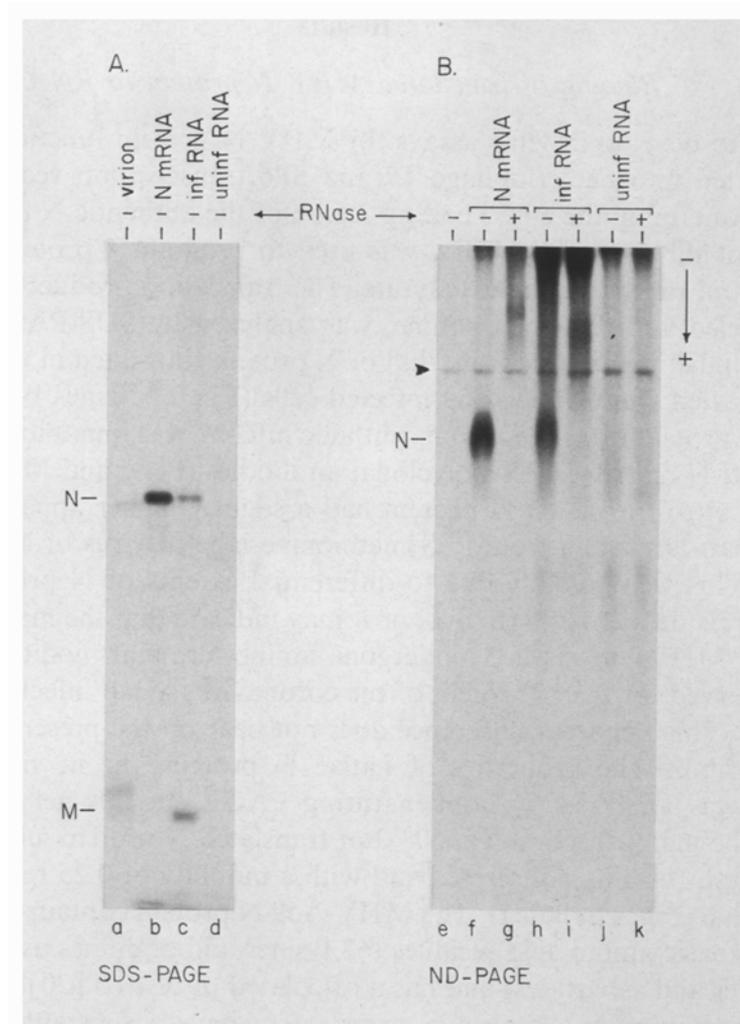


Fig. 2. Binding of RNA by translated N protein. **A** In vitro-synthesized full-length N mRNA was used to program a reticulocyte lysate, and translated [^{35}S]methionine-labeled N protein (*b*) was compared by SDS-PAGE to N protein from purified MHV-A59 virions that had been labeled in vivo with [^{35}S]methionine (*a*) and to N protein translated from total RNA from MHV-infected cells (*c*). *M* Unprocessed form (*a* and *c*) and processed form (*a*) of the MHV membrane protein. *d* Control of translated total RNA from uninfected cells. **B** Translated N protein was analyzed on nondenaturing (ND) PAGE, without or with prior incubation with 30 $\mu\text{g}/\text{ml}$ RNase A (*f*–*i*). Control translation reactions contained no exogenous mRNA (*e*) or total RNA from uninfected cells (*j* and *k*). The arrowhead (left) denotes an endogenous [^{35}S]methionine-labeled band in the reticulocyte lysate. The arrow on the right indicates the polarity and direction of nondenaturing electrophoresis

in the nondenaturing gel, followed by a complete inability of N to enter the gel at intermediate and higher RNase A concentrations (Fig. 3, lanes a–n, upper). This appeared to indicate that binding by N protein conferred a limited degree of protection to some parts of the bound RNA species, while other parts were

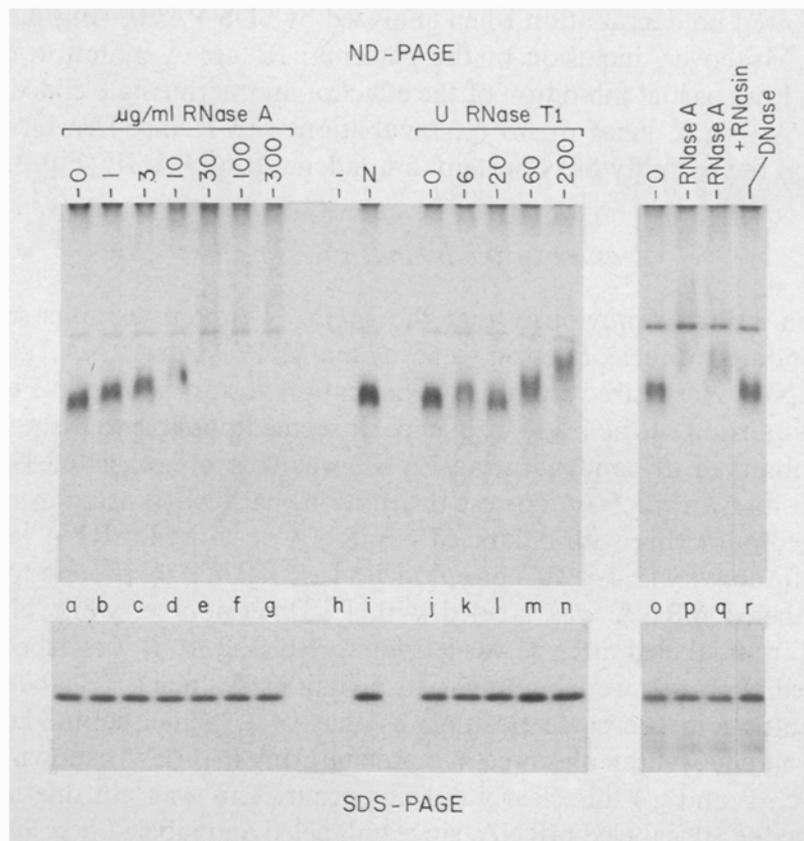


Fig. 3. RNase sensitivity of RNA binding by N protein. N mRNA was translated and reticulocyte lysates were subsequently incubated for 30 min at 30 °C with 0 to 300 µg/ml RNase A (*a-g*) or 0 to 200 U of RNase T1 (Boehringer Mannheim) (*j-n*). Control translation reactions contained no exogenous mRNA (*h*) or N mRNA with no post-translational incubation (*i*). In a separate experiment, N mRNA was translated and subsequently incubated with H₂O (*o*), 10 µg/ml RNase A (*p*), both 10 µg/ml RNase A and 80 U RNasin (*q*), or 1 U RNase-free DNase (Promega) (*r*). [³⁵S]methionine-labeled samples were analyzed by nondenaturing (ND)-PAGE (top) and by SDS-PAGE (bottom)

more immediately digested. The N-RNA complex was more sensitive to RNase A than to RNase T1, possibly due to the different nucleotide specificities of these enzymes (RNase A cuts 3' to pyrimidines; RNase T1 cuts 3' to G residues). Alternatively, the differential sensitivity may have reflected the relative accessibility of each enzyme to phosphodiester bonds within the N-RNA complex. The N proteins of coronaviruses provide their encapsidated genomes only a minimal degree of protection against the action of ribonucleases [12, 22]. Hence, in this respect, the N-RNA interaction in the nondenaturing gel assay resembled that found in the MHV nucleocapsid.

The RNase sensitivity of the ability of N protein to migrate into nondenaturing gels was not due to possible contaminating protease activity in the RNase preparations, since the same RNase-treated samples of translated N

protein showed no degradation when analyzed by SDS-PAGE (Fig. 3, lanes a–n, lower). Moreover, inclusion of the placental RNase A inhibitor, RNasin, allowed at least partial inhibition of the effect of an intermediate concentration of RNase A (Fig. 3, lanes p and q). Incubation with RNase-free DNase had no effect on the mobility of N protein in nondenaturing PAGE (Fig. 3, lane r).

Nature of the bound RNA species

It has been reported previously that the MHV N protein exhibits sequence-specific binding to nucleotides 56 to 65 of the MHV leader RNA [29]. Since the N mRNAs synthesized from the transcription vectors shown in Fig. 1 contained this portion of the leader sequence, it seemed possible that the N-RNA complex observed in nondenaturing PAGE was that of translated N protein binding to its own mRNA. To test this notion, parallel translation reactions were carried out either with unlabeled N mRNA or with N mRNA labeled to high specific activity (2.4×10^7 cpm/ μ g) with [α - 32 P]UTP. N protein translated from unlabeled mRNA was labeled with [35 S]methionine, while N protein translated from labeled mRNA was left unlabeled (Fig. 4). It was expected that if translated N protein were binding to its own mRNA, then a 32 P-labeled RNA band migrating to the same position as that of [35 S]methionine-labeled N protein should have been observed in nondenaturing PAGE. As shown in Fig. 4 (lanes b', c', f' and g') this clearly did not occur. This was not due to failure to translate the 32 P-labeled mRNA, since unlabeled and labeled N mRNAs had comparable translation efficiencies (Fig. 4, lanes b, e, b', and e').

During the course of the 90 min translation reactions, the 32 P-labeled mRNA underwent limited nucleolytic degradation, presumably due to low levels of RNase activity in the reticulocyte lysate. Most of the 32 P-labeled material produced by this limited breakdown was retained by SDS-PAGE, but virtually all of the same species migrated near the dye front in nondenaturing PAGE. Some minor 32 P-labeled bands were seen in the nondenaturing gel at positions other than that expected for N protein-bound material. However, these could not have been due to complex formation with N protein, since the same bands were observed when N protein synthesis was abolished by inclusion of the protein synthesis inhibitor cycloheximide in translation reactions (Fig. 4, lane i').

Thus, the N-RNA complex observed in the nondenaturing gel assay must have been formed by the binding of N to some endogenous RNA species in the reticulocyte lysate. That the observed complex was not due to translated N protein binding to the leader portion of its own mRNA was also supported by the observation that full-length N protein translated from a construct in which the MHV leader sequence was entirely deleted was also able to migrate into nondenaturing gels (data not shown). The same was observed for VSV-N:MHV-N fusion constructs which had the 5' end of a heterologous mRNA substituted for the MHV leader (see below). The complex reported to occur between N protein and MHV leader RNA [29] may not have formed under

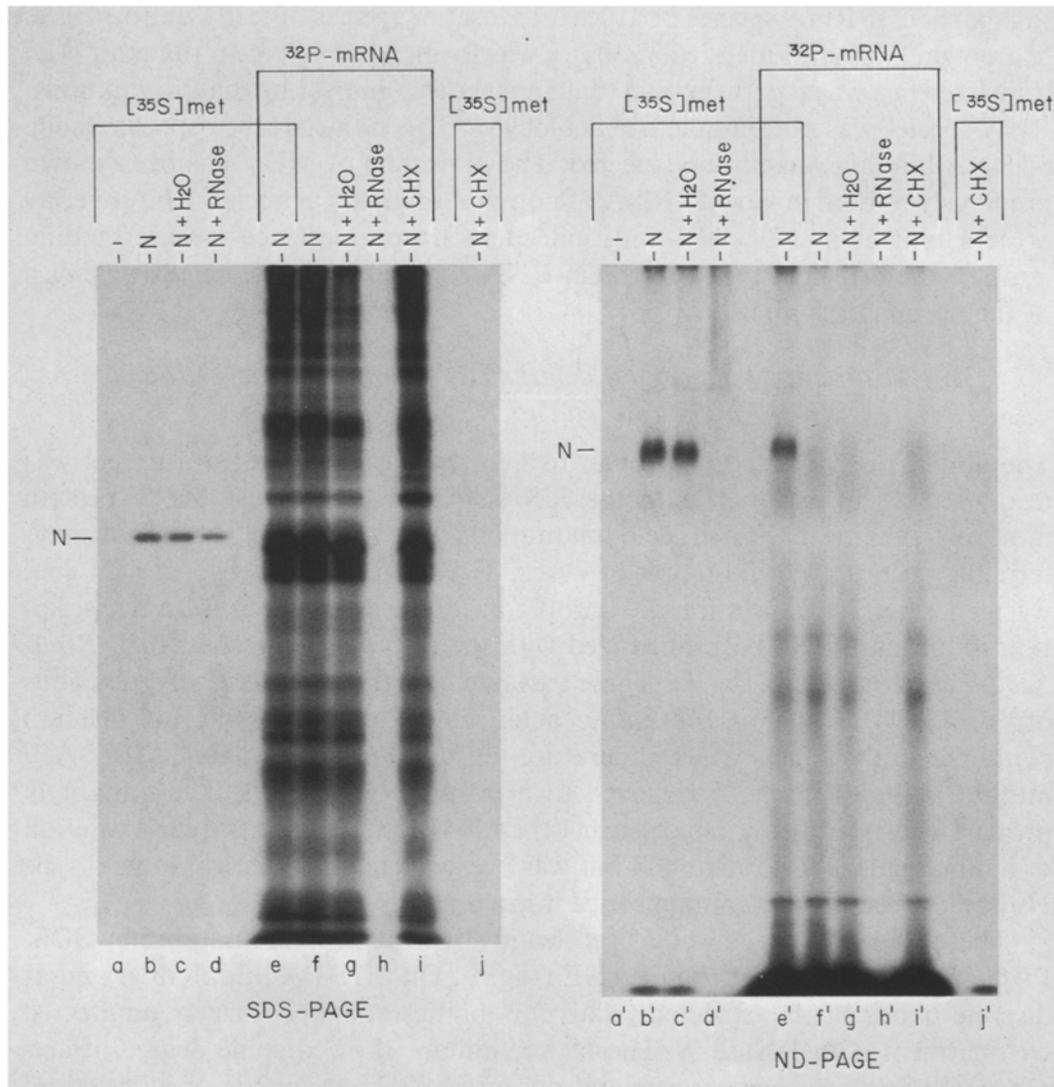


Fig. 4. Translated N protein does not bind to its own mRNA. Translation reactions were carried out with either unlabeled N mRNA (*a-d*, *j*, *a'-d'* and *j'*) or with ^{32}P -labeled N mRNA (2.4×10^7 cpm/ μg) (*e-i* and *e'-i'*), and protein product was either labeled with [^{35}S]methionine (*a-e*, *j*, *a'-e'* and *j'*) or was unlabeled (*f-i* and *f'-i'*). In control reactions, 100 $\mu\text{g}/\text{ml}$ cycloheximide (CHX) was included during the translation. Where indicated, reactions were subsequently incubated for 30 min at 30 $^{\circ}\text{C}$ with either H_2O or with 30 $\mu\text{g}/\text{ml}$ RNase A. A control reaction contained no exogenous mRNA (*a* and *a'*). Samples were analyzed by SDS-PAGE (*a-j*) and by nondenaturing (ND) PAGE (*a'-j'*)

the conditions of the translation assay or may not have been detectable in the electrophoresis system used here.

The identity of the RNA bound by N protein in the reticulocyte lysate has not yet been determined. N protein formed a discrete band rather than a heterogeneous smear in nondenaturing PAGE, suggesting that it was binding

to a particular RNA species or at least to a set of species of fairly uniform size. Moreover, a band of identical mobility was formed by N protein that had been translated in a wheat germ extract (data not shown), indicating that the unknown RNA species was not unique to reticulocytes. This non-sequence-specific mode of RNA binding was not unexpected. The N protein of MHV has been shown previously to bind *in vitro* to RNA of nonviral origin by a labeled RNA overlay protein blot assay [22, 29], and, indeed such non-sequence-specific binding must be the principal type of protein-RNA contact along the extensive length of the coronavirus nucleocapsid.

*Effect of carboxy-terminal deletions on the ability of N protein
to bind to RNA*

The ability of translated N protein to migrate into a nondenaturing gel was next used as an assay to map the RNA-binding function of the N protein molecule. A set of N protein deletion mutants with successively larger carboxy-terminal truncations (designated N/Acc, N/Bsm, N/Sca, N/Eco and N/Spe) was generated by translation of run-off transcripts synthesized from transcription vectors that had been linearized with the restriction enzymes *AccI*, *BsmI*, *ScaI*, *EcoRI*, or *SpeI* (Fig. 1). These variants contained, respectively, deletions of 57, 77, 94, 118, and 148 amino acids. In addition, the internal deletion contained in the N gene of a temperature-sensitive and thermolabile MHV-A59 mutant, Albany-4, was transferred to a transcription vector. The mutant N protein translated from this construct (N/Alb4) contained an in-phase deletion of amino acids 380 through 408 but was everywhere else identical to wild type N ([17]; Koetzner et al., manuscript in prep.).

These six mutants, as well as full-length N protein, were analyzed by SDS-PAGE and by nondenaturing PAGE (Fig. 5). On SDS-PAGE, all migrated as discrete bands of the expected relative mobilities and were largely unaffected by treatment with RNase A. The N/Acc mutant showed some degree of heterogeneity in both the presence and absence of RNase A (Fig. 5, lanes e and f). Since this protein product terminated in the basic amino acid-rich sequence KPQRKGR, it possibly was rendered sensitive to degradation by trypsin-like proteolytic activities within the reticulocyte lysate. When analyzed for the ability to bind to RNA, only two of the N mutants, N/Acc and N/Alb4, were able to enter a nondenaturing gel to a significant quantitative extent by comparison with full-length N protein (Fig. 5, lanes a' to p'). This migration was completely inhibited by prior incubation with RNase A. The remainder of the carboxy-terminal truncated N proteins were only minimally detectable or were undetectable by nondenaturing PAGE.

An estimate of the relative RNA-binding ability of each N protein construct was made by determining the ratio of cpm of translated protein that had banded in a nondenaturing gel compared to the cpm of an identical sample that had banded in SDS-PAGE (Table 1). By this analysis, the N/Acc construct retained the major part of the RNA-binding ability of full-length N protein, demon-

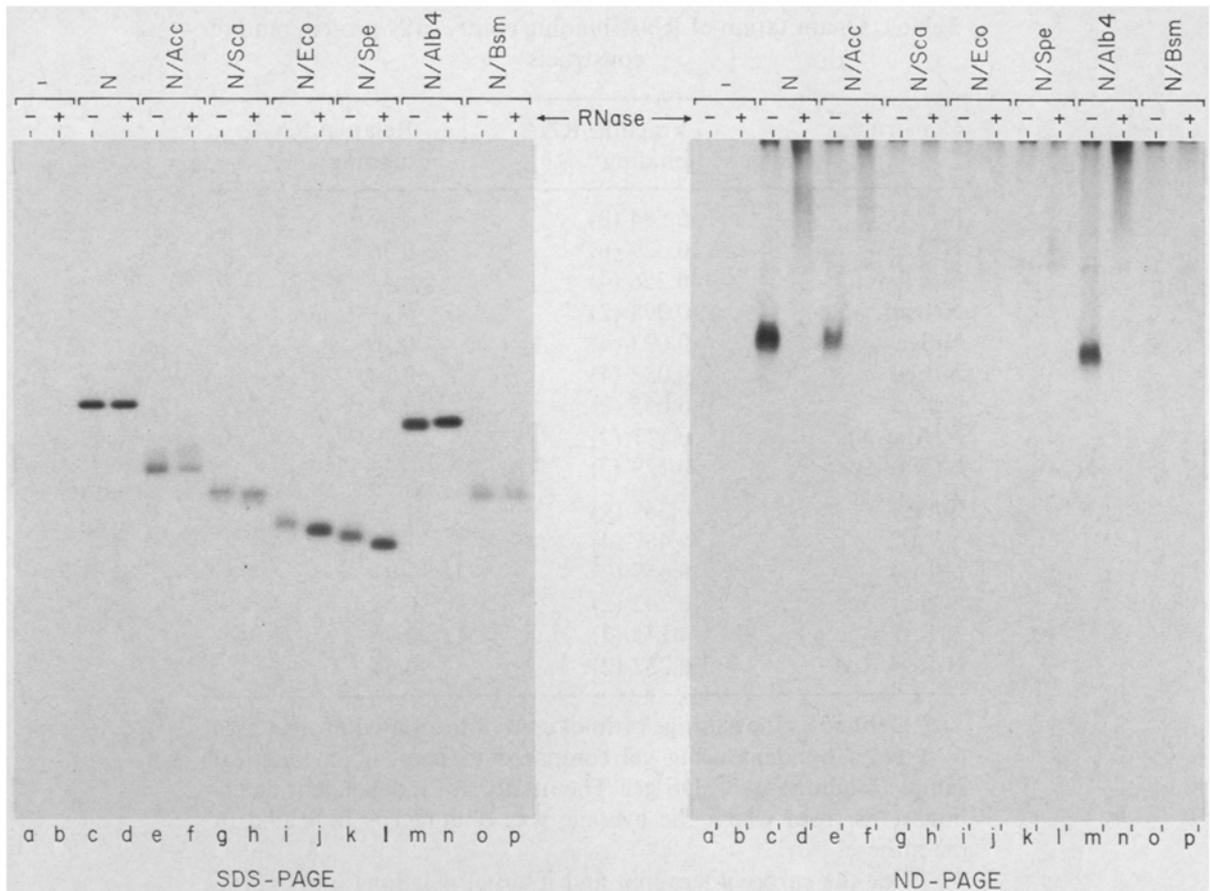


Fig. 5. Effect of carboxy-terminal deletions on the ability of N protein to bind to RNA. In vitro-synthesized mRNAs encoding full-length N protein or various carboxy-terminal deletion mutants (N/Acc, N/Sca, N/Eco, N/SpE, N/Alb4 and N/Bsm) were translated and subsequently incubated for 30 min at 30 °C with either H₂O or with 30 µg/ml RNase A. Control reactions contained no exogenous mRNA (a, b, a' and b'). [³⁵S]methionine-labeled samples were analyzed by SDS-PAGE (a–p) and by nondenaturing (ND)-PAGE (a'–p')

strating that at least the 57 carboxy-terminal amino acids of N were dispensable with respect to this function. However, carboxy-terminal truncation further into the N molecule, in the N/Bsm, N/Sca, N/Eco and N/SpE mutants, generally abolished the ability of N to bind to RNA. The relative RNA-binding ability of N/Alb4 was intermediate between these latter and that of N/Acc. Thus, the left end of the Albany-4 deletion may define the right-most boundary of the RNA-binding domain of the MHV N protein, since the N/Bsm mutant, which truncated two further residues, lost RNA-binding ability.

Effect of amino-terminal or centrally located deletions on the ability of N protein to bind to RNA

A set of deletion mutants collectively spanning the remainder of the N molecule was next constructed. Due to the paucity of convenient restriction sites near

Table 1. Quantitation of RNA-binding ability of N protein mutant constructs

Constructs	Fraction RNA binding ^a	Relative RNA binding ^b
N	0.584 (8)	1.00
N/Acc	0.329 (6)	0.56
N/Alb4	0.256 (4)	0.44
N/Bsm	0.098 (2)	0.17
N/Sca	0.093 (4)	0.16
N/Eco	0.082 (4)	0.14
N/Spe	0.055 (2)	0.09
N/Apa-Nhe	0.111 (2)	0.19
N/Nhe-Spe	0.129 (2)	0.22
N/fus1	0.363 (4)	1.00
N/fus2	0.487 (4)	1.34
N/fus4	0.657 (2)	1.81
N/fus1/Acc	0.242 (2)	0.67
N/fus2/Acc	0.275 (2)	0.76
N/fus4/Acc	0.232 (2)	0.64

^a Defined as the average ratio of cpm of translated protein able to enter a nondenaturing gel compared to cpm of an identical sample binding in an SDS gel. The number of independent determinations from which the average was obtained is indicated in parentheses

^b For the carboxy-terminal and internal deletions (upper set), this is calculated relative to full-length MHV N protein. For the VSV-N:MHV-N fusion constructs (lower set), this is calculated relative to the full-length construct N/fus1

the 5' end of the MHV N gene, a fusion construct incorporating a portion of the 5' end of a mutant VSV N gene was created in order to provide a start codon for a construct in which the 5' end of the MHV N gene was then removed. Consequently, both N fusion genes encoded proteins initiating with the VSV-derived sequence: MAPTVKRIINDSIIQPKL PANEDPDRSAEDDKWLPI-YILG, corresponding to amino acids 1–23 and 124–140 of that protein [16]. The full-length fusion protein, designated N/fus1, contained the entire MHV N protein following this sequence. The amino-terminal deletion mutant, N/fus2, lacking the first 135 amino acids of the MHV N protein, contained the VSV sequence fused to the remainder of MHV N. Two internal deletions mutants, encoding N/Apa-Nhe and N/Nhe-Spe, were generated by removal of the indicated restriction fragments from one of the original (nonfusion) N vectors (Fig. 1). The resulting protein products transcribed and translated from these lacked regions of 66 and 106 amino acids, respectively.

These mutants were analyzed by the two gel electrophoresis systems and on

SDS-PAGE were observed to be homogeneous and of the anticipated relative mobilities (Fig. 6, lanes a to k). Nondenaturing PAGE revealed that the N/fus1 and N/fus2 proteins retained the ability to bind to RNA, whereas the deletions in N/Apa-Nhe and N/Nhe-Spe abolished RNA binding (Fig. 6, lanes a' to k'). Quantitation of the extent of RNA binding showed that the N/fus1 mutant was 62% as binding-competent as unaltered, full-length N protein (Table 1). Although the VSV N protein is, itself, an RNA-binding protein, the mutant VSV N protein used in the construction of N/fus1 and N/fus2 is unable to bind

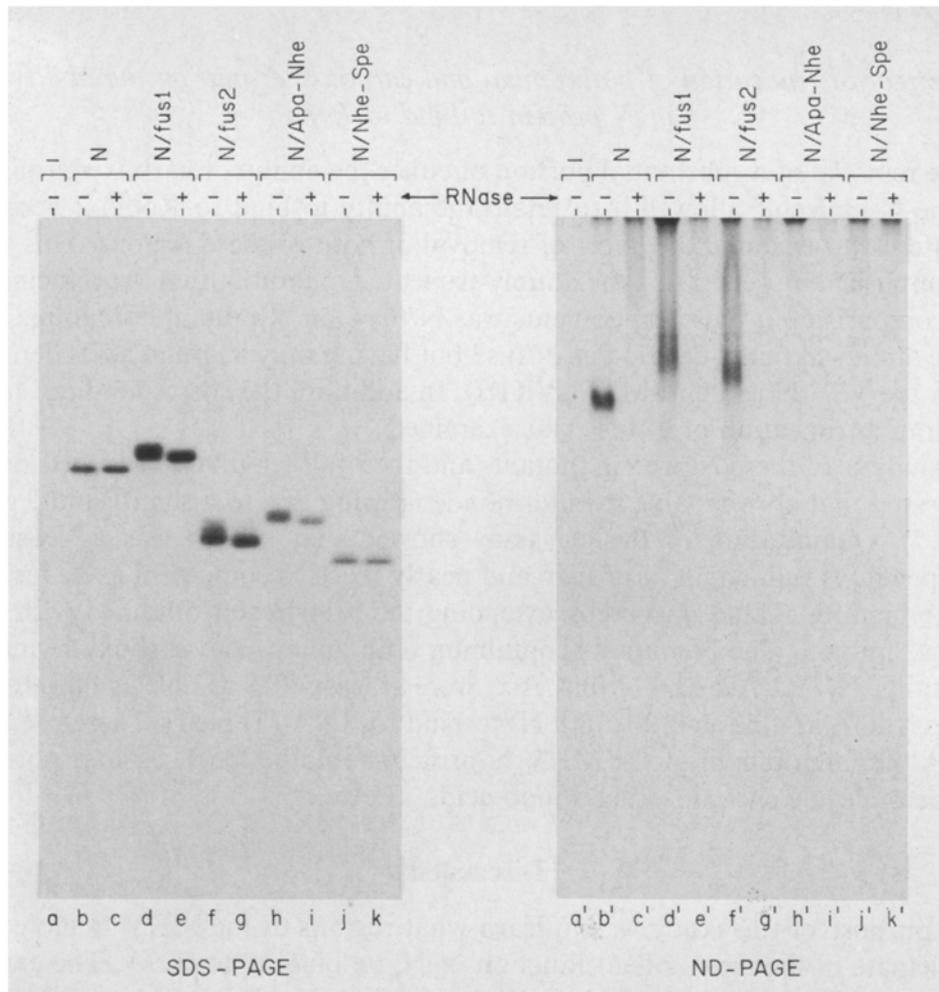


Fig. 6. Effect of amino-terminal or central deletions on the ability of N protein to bind to RNA. In vitro-synthesized mRNAs encoding full-length N protein, a VSV-N:MHV-N fusion protein (N/fus1), an amino-terminal deletion mutant (N/fus2) or central deletion mutants (N/Apa-Nhe and N/Nhe-Spe) were translated and subsequently incubated for 30 min at 30 °C with either H₂O or with 30 µg/ml RNase A. A control reaction contained no exogenous mRNA (a and a'). [³⁵S]methionine-labeled samples were analyzed by SDS-PAGE (a-k) and by nondenaturing (ND)-PAGE (a'-k')

to RNA [16]. Thus, the 40 heterologous amino acids of the two fusion constructs were not expected to contribute to the RNA-binding ability of the MHV N protein. Indeed, for N/fus1 the VSV N-derived portion of the molecule appeared to somewhat impede the entry of MHV N into the nondenaturing gel. Remarkably, the N/fus2 protein was a better RNA-binding molecule than its parent, N/fus1, in spite of the loss of a large portion of the MHV N amino terminus (Table 1). Clearly, this deleted portion of the N molecule can play little or no role in RNA binding. By contrast, the deletions in the N/Apa-Nhe and N/Nhe-Spe mutants reduced RNA binding to residual levels similar to those observed for the four shortest carboxy-terminal truncation mutants (Table 1).

Effect of truncation of both amino and carboxy termini on the ability of N protein to bind to RNA

Since removal of a substantial portion of either the amino or carboxy terminus of the N molecule allowed it to retain the ability to bind to RNA, it became of interest to examine the effect of removal of both of these regions. This was accomplished by generating the doubly-truncated mutant N/fus1/Acc. Included for comparison in these experiments was N/fus4, an N mutant containing the same amino-terminal deletion as N/fus2 but having only 8 amino acids derived from the VSV N protein (MAPTVKRI). In addition, the corresponding *AccI*-generated truncation of N/fus4 was examined.

Analysis of these N protein mutants and their full-length counterparts demonstrated that all were able to enter nondenaturing gels to a significant degree (Fig. 7). Quantitation of the gel assay showed that N/fus4 was at least as competent as full-length N protein and nearly twice as competent as N/fus1 in RNA binding (Table 1), thereby extending the prior result obtained with the N/fus2 mutant. The N mutants containing both amino- and carboxy-terminal deletions, N/fus2/Acc and N/fus4/Acc, were at least 60% as able as N/fus1 (or at least 40% as able as full-length N) to bind to RNA (Table 1). Therefore, the RNA-binding domain of the MHV N protein could be localized to a portion of the molecule encompassing amino acids 136 to 397.

Discussion

The purpose of this study was to learn what regions of the MHV N molecule participate in the most salient function of N, its binding to RNA. The extent of the deletions and the RNA-binding competence of the various MHV N protein mutants constructed for this study are summarized schematically in Fig. 8. The results indicate that the RNA-binding ability of the N molecule can be localized to an internal region comprising some 262 amino acids. Deletion of either 135 amino acids from the amino terminus or 57 amino acids from the carboxy terminus of N produced only a limited effect on the RNA-binding function of this protein. Moreover, mutants lacking both termini still retained

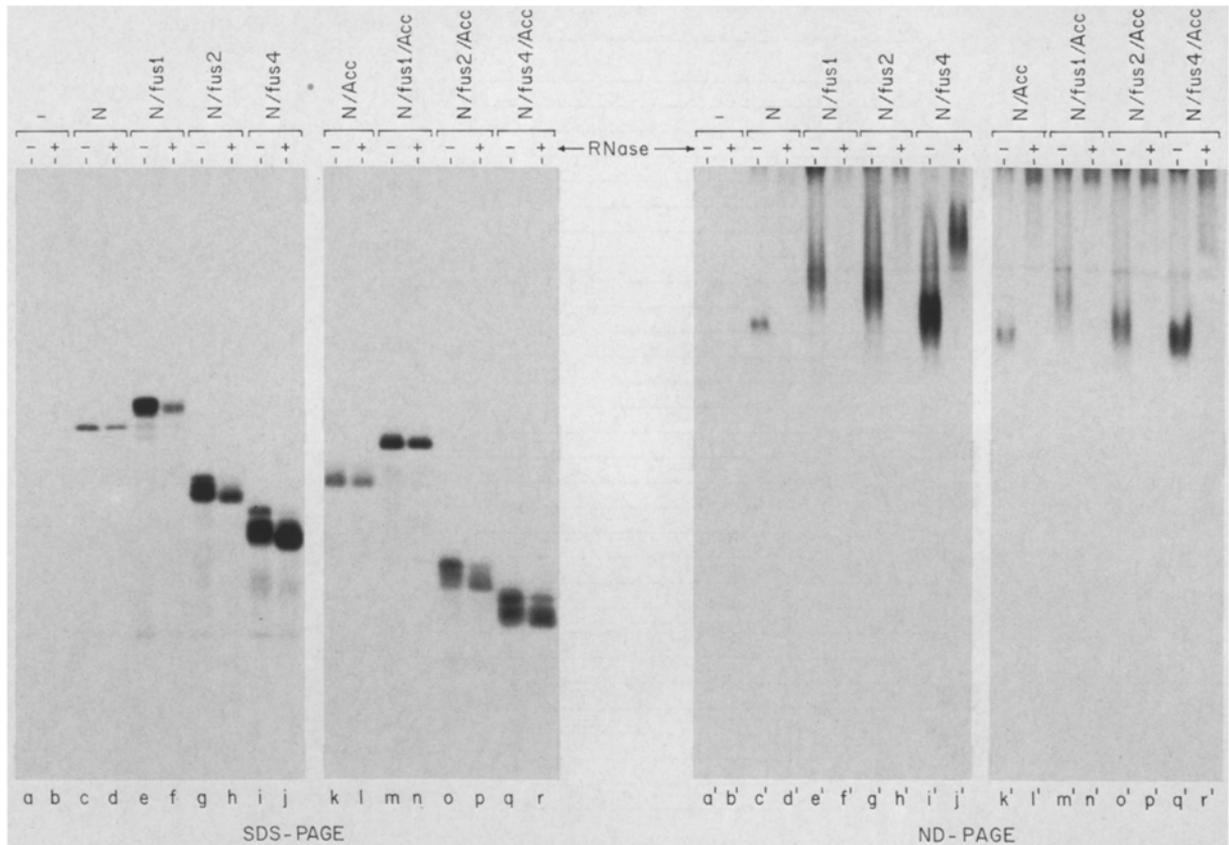


Fig. 7. Effect of combined amino-terminal and carboxy-terminal deletions on the ability of N protein to bind to RNA. In vitro-synthesized mRNAs encoding full-length N protein, a VSV-N:MHV-N fusion protein (N/fus1), amino-terminal deletion mutants (N/fus2 and N/fus4), carboxy-terminal deletion mutants (N/Acc and N/fus1/Acc) or combined amino-terminal and carboxy-terminal deletion mutants (N/fus2/Acc and N/fus4/Acc) were translated and subsequently incubated for 30 min at 30 °C with either H₂O or with 30 µg/ml RNase A. Control reactions contained no exogenous mRNA (a, b, a' and b'). [³⁵S]methionine-labeled samples were analyzed by SDS-PAGE (a-r) and by nondenaturing (ND)-PAGE (a'-r')

RNA-binding ability. By contrast, deletions occurring within the internal region or entering further into it via the carboxy terminus of the molecule effectively abolished RNA binding.

The regions of N defined by this analysis coincide strikingly with those of a three domain model for the molecule that we have proposed previously on the basis of a comparison of the sequences of the N proteins of five different strains of MHV [20]. The present work suggests that these three domains of N are not only structurally distinct but are functionally separable as well. The RNA-binding characteristic of the N molecule resides within domain II, whereas domains I and III are expendable with respect to this activity.

Since the deletions entering domain II are mutations to loss of function, it

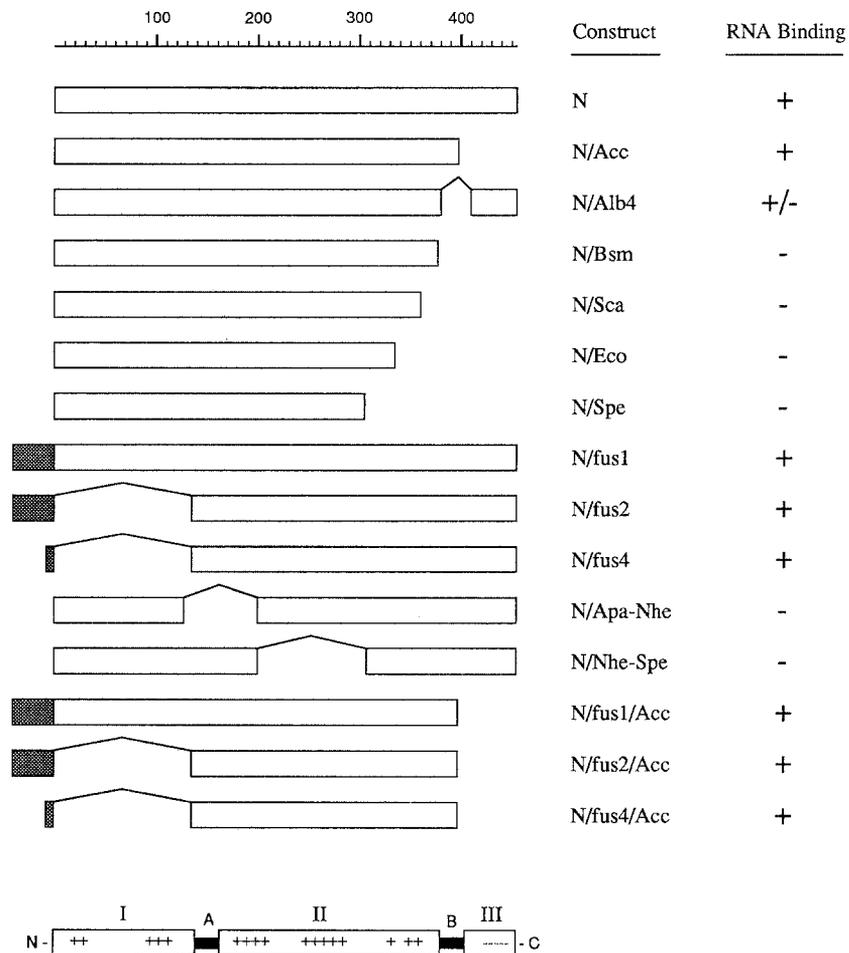


Fig. 8. Summary of the RNA-binding ability of all MHV N protein deletion mutant constructs in this study. The number line at the top denotes amino acid residues in the full-length N protein. Open rectangles show the position and extent of MHV N protein sequence in each of the translated proteins. Hatched rectangles indicate the 40 amino acids (N/fus1 and N/fus2) or 8 amino acids (N/fus4) derived from the VSV N protein in fusion constructs. At the bottom is shown the proposed three domain model for the MHV N protein [20] with clusters of positive and negative charge indicated

cannot be concluded that the entirety of this domain necessarily participates in RNA binding, and a finer mutational analysis of this region will be necessary to more precisely delineate the features of the RNA-binding site. Domain II of MHV N does not contain the RNP1 and RNP2 motifs found in numerous cellular RNA-binding proteins [3, 21], nor does it possess a zinc finger motif [6] or recognizable similarity to recently described arginine-rich [9] or methionine-rich [23] RNA-binding domains. Thus, the rules governing this particular mode of protein-RNA interaction remain to be elucidated.

In Fig. 8, the RNA-binding ability of N/Alb4 has been scored as equivocal, based on our recent finding that the ability of this mutant protein to enter a

nondenaturing gel is temperature-sensitive with respect to that of wild type N (Koetzner and Masters, unpubl. results). In addition, revertants of the Albany-4 mutant have been selected, and all of these have N proteins which no longer exhibit the RNA-binding temperature-sensitivity of N/Alb4. For the one revertant that we have analyzed in detail, the single point mutation that causes this restoration of RNA-binding ability resides in the distal third of domain II. This finding further underscores the role of domain II in RNA binding, and supports the notion that the nondenaturing gel assay measures a biologically relevant property of the N molecule.

Although the N protein may recognize a specific nucleation sequence or secondary structure on the MHV genome, it must also possess a more general non-sequence-specific RNA-binding capability, since it encapsidates the entire length of the 31,000 nucleotide genome, which does not appear to contain any iterated pattern of potential RNA recognition sequences. This general mode of RNA binding may be a common characteristic of the N (NP) proteins of RNA viruses having helical nucleocapsids. Bacterially expressed NP protein of influenza virus was shown to bind well to both viral and nonviral RNA substrates when assayed *in vitro* [4]. The N protein of VSV bound tightly to cellular RNAs when synthesized in the absence of other viral proteins in an *in vivo* expression system [28], and it bound to endogenous reticulocyte RNAs or added RNA species when expressed by *in vitro* translation [16]. The measles virus NP protein, expressed *in vivo* by a vaccinia virus recombinant, has recently been shown to assemble into structures very similar to authentic measles nucleocapsids, as determined by both electron microscopy and equilibrium sedimentation on CsCl gradients [27]. For MHV, it has been demonstrated previously that the N protein of this virus can bind to a variety of viral and nonviral RNA species following separation of N by SDS-PAGE and blotting onto nitrocellulose [22, 29]. Thus, it is not unexpected that the MHV N protein should bind *in vitro* to nonviral RNA species in the nondenaturing gel assay reported here.

It should be noted that an alternative interpretation of the nondenaturing gel assay is that the MHV N protein was not directly bound to RNA but, rather, to a protein constituent of an RNase-sensitive ribonucleoprotein in the reticulocyte lysate. This seems less likely in light of the previously demonstrated RNA-binding activity of N [22, 29], but it cannot be ruled out until an MHV N RNA-binding assay using purified components is established. If such an N protein-ribonucleoprotein complex is shown to occur, then this may indicate a previously unappreciated interaction between N protein and host cell components.

The work presented here does not address the problem of sequence-specific RNA recognition by N protein. Only full-length genomic RNA of MHV is packaged into virions. Recently, analysis of defective interfering particles of MHV has identified the RNA packaging signal as falling within a 400 nucleotide locus near the 3' end of the huge viral polymerase gene [13, 33], which is unique

to genomic RNA and is not found in subgenomic RNA species. Others have reported sequence-specific binding of N protein to nucleotides 56 to 65 of the MHV leader [29], which is common to all MHV genomic and subgenomic RNAs. This latter sequence was present on many of the synthesized mRNAs used in our study, but selective binding to this RNA region was not observed (Fig. 4). It is possible that the concentrations of N protein and non-MHV RNA in the reticulocyte lysate in our assay were too high to allow demonstration of a selective recognition of MHV RNA. We are currently seeking to establish the assay conditions and the correct RNA substrate to examine sequence-specific RNA binding by the MHV N protein.

Acknowledgements

I thank Lawrence Sturman for numerous valuable discussions and Cynthia Ricard, who first suggested the RNase experiment. I am grateful to Monica Parker for expert technical assistance and to Arlene Ramsingh for critically reading the manuscript.

This work was supported in part by Public Health Service grant GM31698 from the National Institutes of Health.

References

1. Bukrinskaya AG (1973) Nucleocapsids of large RNA viruses as functionally active units in transcription. *Adv Virus Res* 18: 195–255
2. Cavanagh D, Davis PJ, Pappin DJC, Binns MM, Bournnell MEG, Brown TDK (1986) Coronavirus IBV: partial amino terminal sequencing of spike polypeptide S2 identifies the sequence Arg-Arg-Phe-Arg-Arg at the cleavage site of the spike precursor polypeptide of IBV strains Beaudette and M41. *Virus Res* 4: 133–143
3. Dreyfuss G, Swanson MS, Pinol-Roma S (1988) Heterogeneous nuclear ribonucleoprotein particles and the pathway of mRNA formation. *Trends Biochem Sci* 13: 86–91
4. Kingsbury DW, Jones IM, Murti KG (1987) Assembly of influenza ribonucleoprotein in vitro using recombinant nucleoprotein. *Virology* 156: 396–403
5. Kingsman SM, Samuel CE (1980) Mechanism of interferon action. Interferon-mediated inhibition of simian virus-40 early RNA accumulation. *Virology* 101: 458–465
6. Klug A, Rhodes D (1987) “Zinc fingers”: a novel protein motif for nucleic acid recognition. *Trends Biochem Sci* 12: 464–469
7. Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227: 680–685
8. Lai MMC, Baris RS, Brayton PR, Stohlman SA (1984) Characterization of leader RNA sequences on the virion and mRNAs of mouse hepatitis virus, a cytoplasmic RNA virus. *Proc Natl Acad Sci USA* 81: 3626–3630
9. Lazinski D, Grzadzińska E, Das A (1989) Sequence-specific recognition of RNA hairpins by bacteriophage antiterminators requires a conserved arginine-rich motif. *Cell* 59: 207–218
10. Lee H-J, Shieh C-K, Gorbalenya AE, Koonin EV, La Monica N, Tuler J, Bagdzhadzhyan A, Lai MMC (1991) The complete sequence (22 kilobases) of murine coronavirus gene 1 encoding the putative proteases and RNA polymerase. *Virology* 180: 567–582
11. Leppart M, Rittenhouse L, Perrault J, Summers DF, Kolakofsky D (1979) Plus and minus strand leader RNAs in negative strand virus-infected cells. *Cell* 18: 735–747
12. Macnaughton MR, Davies HA, Nermut MV (1978) Ribonucleoprotein-like structures from coronavirus particles. *J Gen Virol* 39: 545–549

13. Makino S, Yokomori K, Lai MMC (1990) Analysis of efficiently packaged defective interfering RNAs of murine coronavirus: localization of a possible packaging signal. *J Virol* 64: 6045–6053
14. Maniatis T, Fritsch EF, Sambrook J (1982) *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
15. Masters PS, Banerjee AK (1988) Resolution of multiple complexes of phosphoprotein NS with nucleocapsid protein N of vesicular stomatitis virus. *J Virol* 62: 2651–2657
16. Master PS, Banerjee AK (1988) Complex formation with vesicular stomatitis virus phosphoprotein NS prevents binding of nucleocapsid protein N to nonspecific RNA. *J Virol* 62: 2658–2664
17. Masters PS, Parker MM, Ricard CS, Duchala C, Frana MF, Holmes KV, Sturman LS (1991) Structure and function studies of the nucleocapsid protein of mouse hepatitis virus. In: Cavanagh D, Brown D (eds) *Coronaviruses and their diseases*. Plenum, New York, pp 239–246 (*Advances in experimental medicine and biology*, vol 276)
18. Melton DA, Krieg PA, Rebagliati MR, Maniatis T, Zinn K, Green MR (1984) Efficient in vitro synthesis of biological active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. *Nucleic Acids Res* 12: 7035–7056
19. Pachuk CJ, Bredenbeek PJ, Zoltick PW, Spaan WJM, Weiss SR (1989) Molecular cloning of the gene encoding the putative polymerase of mouse hepatitis coronavirus, strain A59. *Virology* 171: 141–148
20. Parker MM, Master PS (1990) Sequence comparison of the N genes of five strains of the coronavirus mouse hepatitis virus suggests a three domain structure for the nucleocapsid protein. *Virology* 179: 463–468
21. Query CC, Bentley RC, Keene JD (1989) A common RNA recognition motif identified within a defined U1 RNA binding domain of the 70K U1 snRNP protein. *Cell* 57: 89–101
22. Robbins SG, Frana MF, McGowan JJ, Boyle JF, Holmes KV (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
23. Romisch K, Webb J, Lingelbach K, Gausepohl H, Dobberstein B (1990) The 54-kD protein of signal recognition particle contains a methionine-rich RNA binding domain. *J Cell Biol* 111: 1793–1802
24. Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain terminating inhibitors. *Proc Natl Acad Sci USA* 74: 5463–5467
25. Snijder EJ, Den Boon JA, Spaan WJM, Verjans GMGM, Horzinek MC (1989) Identification and primary structure of the gene encoding the Berne virus nucleocapsid protein. *J Gen Virol* 70: 3363–3370
26. Spaan W, Cavanagh D, Horzinek MC (1988) Coronaviruses: structure and genome expression. *J Gen Virol* 69: 2939–2952
27. Spehner D, Kirn A, Drillien R (1991) Assembly of nucleocapsidlike structures in animal cells infected with a vaccinia virus recombinant encoding the measles virus nucleoprotein. *J Virol* 65: 6296–6300
28. Sprague H, Condra JH, Arnheiter H, Lazzarini RA (1983) Expression of a recombinant DNA gene coding for the vesicular stomatitis virus nucleocapsid protein. *J Virol* 45: 773–781
29. Stohlman SA, Baric RS, Nelson GN, Soe LH, Welter LM, Deans RJ (1988) Specific interaction between coronavirus leader RNA and nucleocapsid protein. *J Virol* 62: 4288–4295
30. Stohlman SA, Fleming JO, Patton CD, Lai MMC (1983) Synthesis and subcellular localization of the murine coronavirus nucleocapsid protein. *Virology* 130: 527–532
31. Sturman LS, Holmes KV (1983) The molecular biology of coronaviruses. *Adv Virus Res* 28: 35–111

32. Sturman LS, Holmes KV, Behnke J (1980) Isolation of coronavirus envelope glycoproteins and interaction with the viral nucleocapsid. *Virology* 33: 449–462
33. Van der Most RG, Bredenbeek PJ, Spaan WJM (1991) A domain at the 3' end of the polymerase gene is essential for encapsidation of coronavirus defective interfering RNAs. *J Virol* 65: 3219–3226

Author's address: Dr. P. S. Masters, Wadsworth Center for Laboratories and Research, New York State Department of Health, Albany, NY 12201-0509, U.S.A.

Received September 2, 1991