IMPORTANCE OF MHV-CoV A59 NUCLEOCAPSID PROTEIN COOH-TERMINAL NEGATIVE CHARGES

Valerie Bednar, Sandhya Verma, Andrew Blount, and Brenda G. Hogue*

1. INTRODUCTION

Coronaviruses are enveloped and have a single-stranded, positive-sense genome approximately 26–30 kb in length, the largest of all the RNA viruses. All coronoviruses contain at least four structural proteins: three envelope proteins, the membrane (M), spike (S), and envelope (E) proteins, and a 50–60 kDa phosphorylated nucleocapsid (N) protein.[1] The N proteins of all coronaviruses range between 375 and 455 amino acids and are phosphorylated. N protein is a multifunctional viral gene product. In virus-infected cells, N protein binds to the genomic RNA to form a helical ribonucleoprotein (RNP) complex. The N protein also plays a yet undetermined role(s) in transcription and/or replication, and possibly in translational control. N is a highly basic protein that contains a large number of potential phosphorylation sites. The protein has a high concentration of serine residues (7–11%). N consists of three conserved structural domains, two are basic and one, the carboxy terminal domain, is acidic.[2, 3] A number of conserved negatively charged amino acids are located in the carboxy-terminal domain III of the protein. These residues were previously hypothesized to play a role in N-M protein interactions during assembly.[4-6] The residues could alternatively serve as contributors to the general overall functional structure of the protein. Conceivably, the residues could be important for any of the functions that the protein provides during the virus life cycle.

As part of our goal to understand the functional importance of the charged residues in domain III, a series of N mutants were made and studied in the context of the viral genome using a mouse hepatitis coronavirus (MHV-CoV A59) infectious clone. We found that aspartic acids (D) 440 and 441 are functionally important. Viable viruses were recovered when either residue was changed singly to positively charged arginine (R) but not when both residues were changed to alanine (A). Analysis of a large number of plaque purified viruses from the panel of charged single and neutral double mutants revealed that, in addition to the introduced mutations at positions 440 and/or 441, nearly all had new amino acid changes within the N gene. All of these compensating changes were concentrated primarily in one region further toward the amino end of domain III. A

* Arizona State University, Tempe, Arizona.

few viruses were recovered that retained the arginine substitution at position 441 and no other changes. All of these $D_{441}R$ mutants exhibited a strongly crippled phenotype. Overall the results suggest that the negative charges at positions 440 and 441 are important for one or more of the functions of N.

2. MATERIALS AND METHODS

Plasmids containing seven cDNA fragments (A–G) constituting the entire MHV genome were kindly provided by Dr. Ralph Baric, University of North Carolina at Chapel Hill.[7] All original mutants were constructed using GeneEditor site-directed mutagenesis system (Promega) according to the manufacturer's instructions. The four reconstructed N double mutants were mutagenized by whole plasmid PCR using Pfu polymerase (Stratagene). Upon completion of the mutagenesis, all clones were confirmed by sequencing. All full-length infectious clones were assembled using a protocol basically as previously described.[7, 8] After electroporation, all mutant viruses were plaque purified and multiple plaques were picked for each mutation except $DD_{440-441}RR$, which failed to generate visible fusion foci on mouse L2 cells and was ultimately deemed not viable. Plaques were passaged on L2 cells and the presence of mutations was confirmed by extracting total RNA using Ambion's RNAqueous-4PCR, reverse transcribed using Invitrogen's Superscript RT kit, and amplified using Ambion's SuperTaq Plus polymerase according to manufacturer's directions. The entirety of the E, M, and N genes of all mutants was sequenced. All growth kinetics experiments were carried out in mouse 17Cl1 cells infected at a multiplicity of infection of 5. Cell culture supernatants were collected at 1, 4, 8, 12, and either 16 and 20 or 18 and 24 hpi. Titers were determined by plaque assay on L2 cells. Overlays were removed at 48 hpi, and plaques were fixed and stained with crystal violet in ethanol.

3. RESULTS AND DISCUSSION

As part of our studies to better understand the many functions of the N protein, we chose to focus on the high concentration of negative charges in domain III to gain insight into the role of the carboxy tail. We focused on this region because of the high conservation of the negative charges in this domain throughout the family, yet its inability to be exchanged between different group II coronaviruses.[9] Additionally, the protein participates in a number of different protein-protein interactions such as homo-oligermization and interactions with the M protein.[4, 5, 10-12] N may also directly affect host cell function, as SARS N protein has been shown to activate cellular transcription factors and affect signal transduction pathways.[13] In addition to its various protein-protein interactions, N also interacts with RNA.[14-16] Through N's interactions with the RNA, it is thought to have a role in both genome replication and/or transcription.[17, 18] The helical nucleocapsid's interactions with other viral proteins indicate that N also has an important role in packaging and assembly of the virion. Charged residues often mediate protein-protein interactions and can also affect the tertiary structure of a protein and therefore may affect any number of these functions that N serves.

To study the requirement of these negative charges, nine N mutant viruses were generated: $D_{440}R$, $D_{441}R$, $DD_{440-441}RR$, $DD_{440-441}EE$, $DD_{440-441}AA$, $D_{446}A$, $D_{451}A$, $D_{451}E$,

and $EDD_{449\text{-}451}AAA$. Mutant virus full-length cDNA clones were assembled. Viral RNA was transcribed and transfected into cells. All viable viruses were plaque purified, and the retention of the introduced mutations was confirmed by RT-PCR and sequence analysis.

Negatively charged residues 446 and 449–451 appear not to be absolutely required for N functionality. Following transfection, viable viruses were easily recovered, all mutants were plaque purified and multiple plaques of each mutant were followed for five passages. Sequence analysis of passage five viruses confirmed the stability of the introduced mutations and that no additional changes had arisen in the E, M, or N genes. Furthermore, $D_{446}A$, $D_{451}A$, $D_{451}E$, and $EDD_{449\text{-}451}AAA$ displayed growth characteristics and plaque size and morphology like the wild-type virus.[13]

However, negatively charged residues at positions 440–441 appear to be important for N protein function. Following transfection, centers of fusion were observed for all mutant viruses. Of the five mutant viruses, viable virus was easily recovered for four of them. Following electroporation of mutant virus $DD_{440\text{-}441}RR$, limited characteristic cytopathic effects (CPE), including fusion, was observed. The introduction of two positively charges residues at positions 440–441 appears to be lethal to the virus, as multiple attempts to recover the $DD_{440\text{-}441}RR$ virus were unsuccessful. Although the virus did not tolerate double mutations at residues 440–441 to positive charges, double neutral or other negative charge changes were tolerated. $DD_{440\text{-}441}EE$ grew to a titer comparable with wild-type and displayed plaque size, morphology and growth characteristics similar to wild-type virus, all without additional changes in the E, M, or N genes.[13] This strongly suggests that any negative charge in those positions is favorable, but that aspartic acid residues specifically are not required. The removal of both negative charges at positions 440–441, however, does not appear to be tolerated as well. After transfection and plaque purification of $DD_{440\text{-}441}AA$, multiple plaques were followed through five passages. All $DD_{440\text{-}441}AA$ plaques analyzed retained the introduced AA mutation and exhibited an additional change of $SR_{424\text{-}425}GG$ in the N gene (Table 1). Analysis of $DD_{440\text{-}441}AA$ plaques with $SR_{424\text{-}425}GG$ revealed that the mutant viruses exhibited plaque size, morphology, and growth characteristics indistinguishable from the wild-type virus.[13]

When either D_{440} or D_{441} was replaced by positively charged arginine, additional changes within N were observed. No $D_{440}R$ mutant viruses were recovered that did not contain additional changes within N. Some $D_{441}R$ mutant viruses were recovered that had no additional changes. The most prominent compensating change seen was the replacement of R_{425} with glycine. Also of note was the replacement of A_{436} with aspartic acid. Analysis of the growth characteristics and plaque morphology of the plaqued viruses strongly suggested that the new changes were compensating changes that were increasing the viability of the mutant viruses. Further analysis of the additional changes indicated that replacement of the R_{425} with glycine in the $D_{440}R$ and $D_{441}R$ single mutant backgrounds or replacement of A436 with aspartic acid are indeed important compensating changes.[13]

Our mutagenic studies described here have highlighted key negatively charged residues in the carboxy tail of domain III of MHV N protein. The aspartic acid residues at positions 440–441 appear to be critical residues as changes at these positions are tolerated less well than changes of other negatively charged residues in this region. This is further supported by the lethality of $DD_{440\text{-}441}RR$ mutant. Our inability to successfully passage $DD_{440\text{-}441}RR$ strongly suggests that at least one negative charge at position 440 or 441 is

Table 1. D$_{440}$R, D$_{441}$R, and DD$_{440\text{-}441}$AA recovered plaque purified mutant viruses with summary of additional amino acid changes observed in the N gene.[a]

WT	415					420					425					430					435					440					445					450				
	P	K	S	S	V	Q	R	N	V	S	R	E	L	T	P	E	D	R	S	L	L	A	Q	I	L	D	D	G	V	V	P	D	G	L	E	D	D	S	N	V
D$_{440}$R											G															R														
											G															R														
											G															R														
											G															R														
											G															R														
											G															R														
											G															R														
																					F					R														
														N												R														
														N												R														
														N												R														
											G															R														
D$_{441}$R																						D					R													
																						D					R													
											G																R													
											G																R													
											G																R													
											G																R													
																											R													
																											R													
											G																R													
											G																R													
		Q																									R													
								D																			R													
																											R				Q									
																											R													
																											D													
											G																R													
																											R													
											G																R													
																									V		R													
DD$_{440\text{-}441}$AA										G	G															A	A													
										G	G															A	A													
										G	G															A	A													
										G	G															A	A													
										G	G															A	A													
										G	G															A	A													
										G	G															A	A													

[a] Mutants with charge changes at positions 440 and/or 441 are indicated in the left column. Each line is representative of sequence analysis of a single plaque recovered from the parental mutant.

required for proper N protein function. Furthermore, analysis of the isolated plaques possibly indicates that the positioning of a negative charge at position 440 may be more important than at position 441. All of the plaques that were isolated from the D$_{440}$R mutant also had at least one new change that provided the virus with a growth advantage. Roughly one-quarter of the plaques analyzed for D$_{441}$R retained the original mutation with no additional changes. Although these plaques were genetically stable after five passages, they were nonetheless severely crippled and only grew to titers 3–4 orders of magnitude lower than the wild-type counterpart. Without any additional changes, D$_{441}$R produces plaques significantly smaller than the wild-type virus, and a rudimentary analysis of its kinetics again confirmed the severely crippling effect (data not shown). The larger number of additional changes observed for the D$_{441}$R mutant than for D$_{440}$R further suggests that replacement of a negative charge with a positive charge at position 441 may be more easily compensated for than when the charge is place at position 440.

Taken all together, our results indicate that the negatively charged carboxy tail is important for some aspect of the viral life cycle. Maintenance of the overall negative charge of the domain appears to be important because the vast majority of compensating changes seen reduced the net charge of the carboxy tail. The frequency with which R_{425} was replaced with glycine may indicate an effect on the tertiary structure. The removal of a positive charge at position 425 could affect the folding of the N protein. Due to its small size, glycines are implicated in affecting a protein's tertiary structure by allowing its surrounding environment more flexibility.[19]

4. ACKNOWLEDGMENTS

This work was supported by Public Health Service grant (AI54704), from the National Institute of Allergy and Infectious Diseases. We thank Ralph Baric for providing us with the MHV infectious clone.

5. REFERENCES

1. K. V. Holmes and M. M. C. Lai, in: *Coronaviridae: The Viruses and Their Replication*, edited by D. Knipe and P. Howley (Lippincott Williams & Wilkins, Philadelphia, 2001), pp. 1163-1185.
2. M. M. Parker and P. S. Masters, 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 (1990).
3. H. Laude and P. S. Masters, in: *The Coronaviridae*, edited by S. G. Siddell (Plenum Press, New York, 1995), pp. 141-163.
4. L. Kuo and P. S. Masters, Genetic evidence for a structural interaction between the carboxy termini of the membrane and nucleocapsid proteins of mouse hepatitis virus, *J. Virol*. **76**, 4987-4999 (2002).
5. K. Narayanan, A. Maeda, J. Maeda, and S. Makino, Characterization of the Coronavirus M protein and nucleocapsid interaction in infected cells, *J. Virol*. **74**, 8127-8134 (2000).
6. D. Escors, J. Ortego, H. Laude, and L. Enjuanes, The membrane m protein carboxy termins binds to transmissible gastroenteritis coronavirus core and contributes to core stability, *J. Virol*. **75**, 1312-1324 (2001).
7. B. Yount, M. R. Denison, S. R. Weiss, and R. S. Baric, Systematic assembly of a full-length infectious cDNA of mouse hepatitis virus strain A50, *J. Virol*. **76**, 11065-11078 (2002).
8. S. Verma, V. Bednar, A. Blount, and B. G. Hogue, Identification of functionally important negatively charged amino acids within domain III of mouse hepatitis coronavirus A59 nucleocapsid protein, *J. Virol.* **80**, 4344-4355 (2006).
9. D. Peng, C. A. Koetzner, T. MCMahon, Y. Zhu, and P. S. Masters, Construction of murine coronavirus mutants containing interspecies chimeric nucleocapsid protein, *J. Virol*. **69**, 5475-5484 (1995).
10. S. G. Robbins, M. F. Frana, J. J. McGowan, J. F. Boyle, and K. V. Holmes, RNA-binding proteins of coronavirus MHV: detection of monomeric and multimeric N protein with an RNA overlay-protein blot assay, *Virology* **150**, 402-410 (1986).
11. K. Narayanan, K. H. Kim, and S. Makino, Characterization of N protein self-association in coronavirus ribonucleoprotein complexes, *Virus Res*. **98**, 131-140 (2003).
12. I. M. Yu, C. L. T. Gustafson, J. Diao, J. W. Burgner, Z. Li, J. Zhang, and J. Chen, Recombinant SARS coronavirus nucleocapsid protein forms a dimer through its C-terminal domain, *J. Biol. Chem*. Accepted Manuscript (2005).
13. R. He, A. Leeson, A. Andonov, Y. Li, N. Bastien, J. Cao, C. Osiowy, F. Dobie, T. Cutts, M. Ballantine, and X. Li, Activation of AP-1 signal transduction pathway by SARS coronavirus nucleocapsid protein, *Biochem. Biophys. Res. Commun*. **311**, 870-876 (2003).
14. G. W. Nelson, S. A. Stohlman, and S. M. Tahara, High affinity interactioin between nucleocapsid protein and leader/intergenic sequence of mouse hepatitis virus RNA, *J. Gen. Virol*. **81**, 181-188 (2000).
15. R. Molenkamp and W. J. M. Spaan, Identification of a specific interaction between the coronavirus mouse hepatitis virus A59 nucleocapsid protein and packaging signal, *Virology* **239**, 78-86 (1997).

16. S. A. Stohlman, R. S. Baric, G. N. Nelson, L. H. Soe, L. M. Welter, and R. J. Deans, Specific interaction between coronavirus leader RNA and nucleocapsid protein, *J. Virol*. **62**, 4288-4295 (1988).
17. F. Alamazán, C. Galán, and L. Enjuanes, The Nucleoprotein is required for efficient coronavirus genome replication, *J. Virol*. **78**, 12683-12688 (2004).
18. B. Schelle, N. Karl, B. Ludewig, S.G. Siddell and V. Thiel, Selective replication of coronavirus genomes that express nucleocapsid protein, *J. Virol*. **79**, 6620-6630 (2005).
19. D. Oh, S. Y. Shin, S. Lee, J. H. Kang, S. D. Kim, P. D. Ryu, K. S. Hahm, and Y. Kim, Role of the hinge region and the tryptophane residue in the synthetic antimicrobial peptides, cecropin A(1-8)-magainin 1(1-12) and its analogues, on their antibiotic activities and structures, *Biochemistry* **39**, 11855-11864 (2000).