Role of the Coronavirus Membrane Protein in Virus Assembly

by

Ariel L. Arndt

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved November 2010 by the
Graduate Supervisory Committee:

Brenda Hogue, Chair
Bertram Jacobs
Wilson Francisco
Tatiana Ugarova

ARIZONA STATE UNIVERSITY

December 2010

ABSTRACT

Coronaviruses are medically important viruses that cause respiratory and
enteric infections in humans and animals. The recent emergence through
interspecies transmission of severe acute respiratory syndrome coronavirus
(SARS-CoV) strongly supports the need for development of vaccines and
antiviral reagents. Understanding the molecular details of virus assembly is an
attractive target for development of such therapeutics. Coronavirus membrane
(M) proteins constitute the bulk of the viral envelope and play key roles in
assembly, through M-M, M-spike (S) and M-nucleocapsid (N) interactions. M
proteins have three transmembrane domains, flanked by a short amino-terminal
domain and a long carboxy-terminal tail located outside and inside the virions,
respectively. Two domains are apparent in the long tail - a conserved region (CD)
at the amino end and a hydrophilic, charged carboxy-terminus (HD). We
hypothesized that both domains play functionally important roles during
assembly. A series of changes were introduced in the domains and the functional
impacts were studied in the context of the virus and during virus-like particle
(VLP) assembly. Positive charges in the CD gave rise to viruses with neutral
residue replacements that exhibited a wild-type phenotype. Expression of the
mutant proteins showed that neutral, but not positive, charges formed VLPs and
coexpression with N increased output. Alanine substitutions resulted in viruses
with crippled phenotypes and proteins that failed to assemble VLPs or to be
rescued into the envelope. These viruses had partially compensating changes in

M. Changes in the HD identified a cluster of three key positive charges. Viruses could not be recovered with negatively charged amino acid substitutions at two of the positions. While viruses were recovered with a negative charge substitution at one of the positions, these exhibited a severely crippled phenotype. Crippled mutants displayed a reduction in infectivity. Results overall provide new insight into the importance of the M tail in virus assembly. The CD is involved in fundamental M-M interactions required for envelope formation. These interactions appear to be stabilized through interactions with the N protein. Positive charges in the HD also play an important role in assembly of infectious particles.

DEDICATION

This dissertation is dedicated to my husband and our daughter.

ACKNOWLEDGMENTS

together. I would like to thank Kelly for being a truly genuine person. You brought new light into my life. To my dear friend Pavi, I cannot imagine life without you. Your endearing words kept me sane. You are a wonderful, sincere person and I respect you for that. Thank you for being like a sister to me. Blake, who knew I would meet my brother from another mother in this lab. You brought so much happiness into the lab and my life. We went on a three year journey together in lab and it was a wonderful one. Thank you for getting me. To Dr. Brenda Hogue, thank you for allowing me to do research in your lab. I feel like you gave me the foundation to be a good scientist and move forward competitively in this field. I appreciate you always making me strive to do better and get "PQ" data. Also to all my committee members, Dr. Jacobs, Dr. Francisco and Dr. Ugarova, I appreciate all your questions and comments that allowed me to mature as a scientist.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Coronaviruses cause respiratory, enteric, and neurological disease in humans and a variety of animals. Many infections in animals can be severe and cause significant economic losses in the poultry, cattle and swine industries. Two human coronaviruses, 229E and OC43, are responsible for roughly 30% of common colds. However, human coronaviruses can cause severe disease as demonstrated by the severe acute respiratory syndrome (SARS-CoV) outbreak. Bat populations worldwide carry SARS-like CoVs and phylogenetic analysis indicates that interspecies transmission occurred, leading to infections in humans. Since the emergence of SARS-CoV, two new human coronaviruses have been identified, NL63 and HKU1. Thus, there is significant interest in understanding coronaviruses since they routinely circulate in animals as well as human populations, and cross species transfer can clearly happen. This provides a strong justification for antiviral therapeutics and vaccine development against the viruses.

Coronaviruses are enveloped, single-stranded, positive-sense RNA viruses with a genome of approximately 30kb. This is the largest genome known for RNA viruses. The viruses assemble and bud at intracellular membranes in the endoplasmic reticulum Golgi intermediate compartment (ERGIC). The virion envelope contains at least three structural proteins, the membrane (M), spike (S), and envelope (E) proteins. The genomic RNA is encapsidated by the nucleocapsid (N) phosphoprotein to form a helical nucleocapsid. Co-expression of the M and

the E proteins is sufficient for virus-like particle (VLP) assembly for most coronaviruses.

Coronavirus M proteins are type III proteins that all contain a short amino terminal domain located outside the virion, followed by three transmembrane domains, and a long carboxy tail located inside the virion. The tail is composed of a long amphipathic region and a hydrophilic, charged extreme carboxy tail. The M protein plays key roles in virus assembly, through M-M, M-S and M-N protein interactions.

A conserved domain (CD) is located at the amino end of the amphipathic region. To determine the functional significance of the CD, changes were introduced into the CD and the charged tail and the impact was studied in the context of the virus as well as during VLP assembly. Introduction of positive charges in the CD (SWWSFNPETNNL) in place of the negatively charged E residue gave rise to viruses with neutral residue substitutions that exhibited a wild-type phenotype. Expression of the mutant proteins showed that neutral, but not positive charge substitutions are competent for VLP assembly and coexpression with N increased output. Alanine substitutions for the first four (5'A) or last four residues resulted in viruses with crippled phenotypes and proteins that failed to assemble VLPs or to be rescued into the envelope. 5'A viruses had compensating changes in M that resulted in rescue into VLPs and increased VLP output when coexpressed with N. Overall the data show that the CD is important for virus assembly and suggests that it is involved in fundamental

2

M-M interactions required for envelope formation. Additionally, results indicate that N may help stabilize M-M interactions.

Previous studies have shown that the extreme carboxy end of the M tail ($VYVK_{205}SK_{207}VGNYR_{212}LPSNK_{217}PSGADTALLR_{227}T$) is important for virus assembly. With the exception of R_{227}, the significance of other charges in the domain have not been determined. To determine the possible role of these residues, amino acid substitutions were introduced by site directed mutagenesis for the positive charges. The results indicated that replacement of K_{205} or R_{212} and $K_{205}K_{207}$ or $R_{212}K_{217}$ with a negatively charged aspartic acid (D) resulted in non-viable viruses and no production of VLPs. Viruses harboring $K_{207}D$ or $R_{212}A$ changes had severely crippled phenotypes and could not form VLPs even in the presence of N. Additionally these mutants displayed a decrease in viral infectivity. The data suggests the charges in the tail of M are important for M-S interactions. Overall these studies show that positive charges in the tail of M are functionally important for virus assembly as well as infectivity.

Together the results of this work provide significant new insight into regions of the M protein that are vital for assembly of infectious particles. The study increases our understanding of a key player and its role in coronavirus assembly and provides insight that can form the basis for antiviral and/or vaccine development.

CHAPTER 1

LITERATURE REVIEW

History, taxonomy and medical importance of Coronaviruses

Coronaviruses were first described in the 1930's, the first of which was a
virus called avian infectious bronchitis virus (IBV) followed by murine hepatitis
virus (MHV) and transmissible gastroenteritis virus (TGEV) (10, 27, 27, 50, 162).
It was not realized these viruses were related until the 1960's when a human
coronavirus was described (191). The coronavirus genus was then defined. All
members were defined as having a distinct morphology of a crown-like
appearance due to surface projections and were given the name *corona* (Latin for
crown) (190).

Coronaviruses are in the *Coronaviridae* family along with the Torovirus
genus (Fig.1). Corona- and toro-viruses share similar replication strategies and
genome organization. However, they differ in their genome size and nucleocapsid
structure (22). The *Coronaviridae* family is in the *Nidovirales* order along with
the *Arteriviridae* and *Roniviridae* families (Fig.1). Viruses in this order are
morphologically different but all carry out RNA synthesis using discontinuous
transcription. Additionally, all Nidoviruses produce a 3'-coterminal nested set of
subgenomic mRNAs. All the viral structural proteins are translated from these
subgenomic mRNAs (24, 176, 177, 177).

Coronaviruses are separated into three groups based on various serological
assays including neutralization, immunofluorescence, and enzyme-

Nidovirales

Coronaviridae	**Roniviridae**	**Arteriviridae**

Coronavirus	**Torovirus**	**Ronivirus**	**Arterivirus**
Transmissible gastroenteritis virus	Equine torovirus	Gill associated virus	Equine arteritis virus
Feline coronavirus	Bovine torovirus		Procine reproductive & respiratory syndrome virus
Canine coronavirus	Porcine torovirus		Lactate dehydrogenase-elevating virus
Mouse hepatitis virus			Simian hemorrhagic fever virus
Bovine coronavirus			
Human coronavirus NL63			
Human coronavirus HKU1			
Human coronavirus 229E			
Human coronavirus OC43			
Porcine hemagglutinating encephalomyelitis virus			
Rat coronavirus			
Turkey coronavirus			
Infectious bronchitis virus			
Severe Acute Respiratory Syndrome coronavirus			

FIG.1. Taxonomy of the *Nidovirales* order. The order contains the *Coronaviridae*, *Roniviridae*, and *Arteriviridae* families. Below each family, the genus is listed. The *Coronaviridae* family consists of the *coronavirus* and *torovirus* genuses. Viruses associated with each genus are listed below the genus name.

linked immunosorbent assay (ELISA) (48, 176) (Fig.2). Each group is further characterized according to their host and diseases which they cause (Fig.2). It can be seen that coronaviruses infect a wide range of domesticated animals and humans and cause a variety of diseases. Evolutionary relationships between coronaviruses have also been studied using sequence analysis of different viral genes (169). The results from these analyses were similar to the serological studies.

Since the identification of the first coronavirus in the 1930's, several new coronaviruses have been identified in both animals and humans. The first human coronaviruses, 229E-CoV and OC43-CoV, were identified in the 1960's (74, 130, 191). Roughly 30% of common colds are caused by these human coronaviruses (57). In 2003, a novel human coronavirus emerged in China and was named severe acute respiratory syndrome coronavirus (SARS-CoV) (52). SARS-CoV caused more severe disease then other human coronaviruses and results in atypical pneumonia, fever, and shortness of breath (154). A total of about 8000 cases were reported to the World Health Organization (WHO) with a mortality rate of roughly 10% (154), (http://www.cdc.gov/ncidod/sars/reporting.htm). It is thought SARS-CoV originated from a wild animal reservoir, presumably from bats (104, 114). The emergence of SARS-CoV and the apparent transfer from animals to humans resulted in new interest in coronaviruses and the need to gain understanding of how these viruses cause disease. Since the identification of SARS-CoV, two other human coronaviruses have emerged, NL63 and HKU1 in

6

Group	Virus	Host	Respiratory	Enteric	Hepatitis	Neurological	Other
1	HCoV-229E	Human	✓			?	
	HCoV-NL63	Human	✓				
	TGEV	Pig	✓	✓			✓
	CCoV	Dog		✓			
	FECoV	Cat		✓			
	FIPV	Cat	✓	✓	✓	✓	✓
	RbCoV	Rabbit		✓			✓
2	HCoV-OC43	Human	✓	?		?	
	HCoV-HKU1	Human	✓				
	SARS-CoV	Human	✓	✓			✓
	MHV	Mouse	✓	✓	✓	✓	
	HEV	Pig	✓	✓		✓	
	BCoV	Cow	✓	✓			
	TCoV	Turkey	✓	✓			
3	IBV	Chicken	✓		✓		✓
	TCov	Turkey	✓	✓			

FIG.2. Antigenic groups, hosts and diseases of coronaviruses. Adapted from (98). The three groups are listed with representative viruses from each. The host of each virus is indicated. The disease(s) caused each virus are marked. Other diseases caused by coronaviruses are infectious peritonitis, immunological disorders, runting, nephritis, pancreatic, parotitis, myocarditis, and sialodacryoadenitis.

2004 and 2005, respectively (193, 207). Both NL63 and HKU1 cause mostly mild upper and lower respiratory tract infections similar to 229E-CoV and OC43-CoV (1, 105, 173). Although, both NL63 and HKU1 were reported to cause more severe symptoms as well, neither of these new human coronaviruses causes as severe disease as SARS-CoV (1, 105, 173).

Coronaviruses that infect domesticated animals have a significant economic impact globally. Many infections in animals can be severe and impose significant economic losses in the poultry, cattle and swine industries. For example, TGEV can cause vomiting, dehydration and severe and sometimes fatal diarrhea in piglets (148, 158). IBV in chickens can cause kidney nephritis as well as a decrease in egg production (34).

Virion morphology and structure

Coronaviruses are enveloped single stranded positive sense RNA viruses that are 100-120nm in size (101). The genomic RNA is 27-31 kb which is the largest RNA virus known. The envelopes of all coronaviruses contain three main structural proteins, the spike (S), membrane (M) and envelope (E) proteins (Fig.3). Inside the envelope the nucleocapsid (N) protein binds to the genomic RNA to form a flexible helical nucleocapsid. Some coronaviruses envelopes also contain the hemagglutinin esterase (HE) protein. The minimal requirements for envelope formation are the M and E proteins as they form virus-like particles (VLPs) (Fig.3).

8

Structural Proteins

Spike (S) Protein

The virion envelope is decorated with the spike (S) protein which gives coronaviruses their characteristic crown like appearance. The S protein is a type I large glycoprotein that is roughly 150-180kD in size and is thought to form trimers (49, 110). S contains a large N-terminal exodomain that extends outside the virion, followed by a transmembrane domain and a short C-terminal endodomain. The exodomain contains S1 and S2 subdomains. The globular S1 region is responsible for binding to host cell receptors (15, 15, 40). The sequences of S1 are variable between different coronaviruses. The S2 region forms the stalk region of S. S2 regions contain two heptad repeats that form a coiled-coil structure (40). After S1 binds to the receptor, the S2 region is responsible for mediating fusion of viral and host membranes. Additionally, S2 can mediate cell to cell fusion (15, 121). During maturation, S1 and S2 regions remain noncovalently associated after cleavage by cellular proteases (180). However for SARS, cleavage of S1 and S2 occurs during entry (170). Interestingly, the S proteins of group I coronaviruses are not cleaved (41).

The S protein has many functions during coronavirus infection. S is the major target of monoclonal neutralizing antibodies (66). Additionally, the S protein mediates fusion of viral and host membranes as well as cell to cell fusion, as mentioned above. Although the S2 region is thought to be responsible for fusion, changes in multiple regions of S1 and S2 affect the fusion process
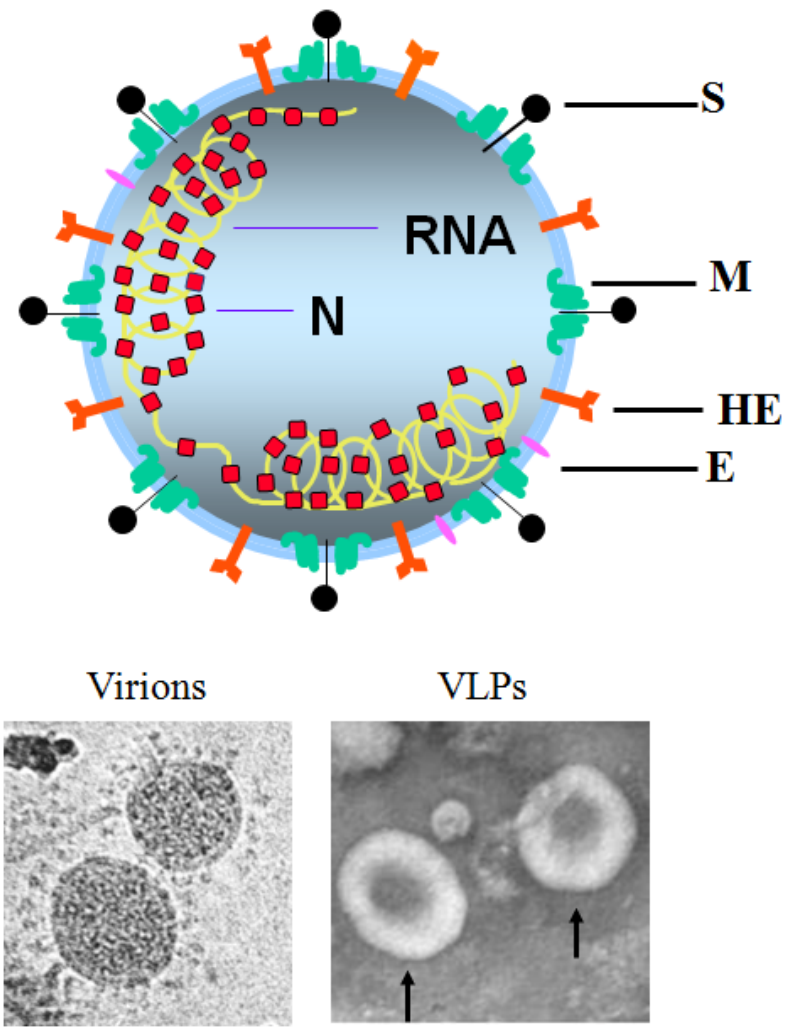
FIG.3. Schematic of the coronavirus virion. Top: The envelope contains the

membrane (M), spike (S) and envelope (E) proteins. Inside the virion, the

nucleocapsid (N) protein binds to the genomic RNA. Adapted from (210).

Bottom: Virions and virus-like particles (VLPs) visualized by electron

microscopy. The virions contain the characteristic "crown-like" morphology.

The VLPs are composed of only the E and M proteins.

(67, 70, 157). Furthermore, the S protein, when present at the cell surface, can bind to the Fc fragment of IgG (141). Since S makes cells susceptible to B-cell mediated cytotoxicity (78), binding to Fc fragments may protect S from being recognized by antiviral antibodies.

Hemagglutinin-esterase (HE) protein

The HE glycoprotein is found on the virion of some group II and group III coronaviruses (17, 97). It is thought the HE protein arose during a recombination event of a coronavirus ancestor and the influenza C virus (122). It is a type I membrane protein that forms short spikes on the virus envelope. HE associates with S and M proteins in infected cells (139). The HE protein is about 65kD in size; forms disulfide linked dimers and has hemagglutinating and esterase activity (17, 76). Due to the functions of HE, it is thought the protein may play a role in entry and/or release (97). Therefore it may be involved in host cell binding but S clearly is still required for attachment to the receptor and downstream entry events (62, 65, 97). Furthermore, it does not appear the HE is essential for replication as it becomes mutated or deleted after multiple passaging in cell culture (212). However, strains of MHV that express HE have an increase in neurovirulence *in vivo* (90).

Envelope (E) Protein

The E protein is present in low abundance in the virion envelope and is 9-12kD in size (68, 219). Coronavirus E proteins all share similar conserved characteristics. These include an amino terminal hydrophobic transmembrane

region, followed by a cysteine-rich region, conserved prolines, and a highly charged carboxy terminal domain (168). Two topologies have been reported for E proteins where it spans the membrane once or twice (35, 36, 126, 220). Additionally, E is palmitoylated on cysteine residues in the cysteine-rich region for several coronaviruses (36, 118, 220). When charged residues within the carboxy domain of E were mutated to neutrally charged alanines, virions displayed an aberrant morphology (61). The M and E proteins are required for budding of VLPs (194). However, when expressed alone the E protein is released from transfected and infected cells in the form of vesicles (125). Therefore, it is thought that the E protein triggers virus assembly. Although, the exact role of E has not been elucidated but it has been shown to play an important role in virus production (37, 61, 96, 116, 144, 219). MHV E can be deleted from the genome and viruses are viable, however these mutants grow to titers several logs lower compared to WT (96). Additionally, coronavirus E proteins have been shown to have ion channel activity (124, 205, 206) which presumably plays a role in viral entry and/or budding.

Nucleocapsid (N) protein

The N protein is a highly basic phosphoprotein that is 50-60kD in size (106). A three-domain structure for the N protein has been proposed based on early sequence comparisons of MHV strains (146). The first two domains, the amino terminal and central domains, of all coronavirus N proteins have an overall positive charge. Domain III, the carboxy terminus domain, is highly acidic. Each

12

domain is separated by sequences that are highly variable. The N protein is

phosphorylated and specific sites have been identified for TGEV, IBV, MHV and

SARS-CoV (20, 29, 202). The function of phosphorylation is not completely

known however it is speculated to play a role in RNA binding (discussed below)

or in localization (29, 137, 179, 182).

The N protein is a multifunctional protein. One primary role of N during

infection is encapsidation of the genomic RNA to form the helical nucleocapsid

structure (8, 39, 123). The RNA binding domain of different coronavirus N

proteins have been mapped. The RNA binding region for MHV N is Domain II.

Based on a recent model of N based on the structure of the protein ends, the IBV

and SARS-CoV N RNA binding domains have been mapped to the amino-

terminal region (81, 222). The N protein has been found to bind nonspecifically to

RNA as well as to the 5' leader, the 3' untranslated region (UTR) and the

packaging signal (33, 128, 132, 136, 137). The intrinsic property of N to bind

RNA allows for the formation of the nucleocapsid structure. Secondly, in addition

to N's ability to bind to RNA, it also plays a structural role during virus assembly.

The N protein interacts with the M protein (discussed below) which results in the

incorporation of the nucleocapsid into the viral particle (83, 95, 134, 196). Two

groups independently identified negative charges in Domain III of MHV N as

being functionally important in virus assembly (83, 195), suggesting N-M

interactions may be of electrostatic nature. Third, coronavirus N proteins play a

role in viral RNA synthesis and increased efficiency of replication (2, 26, 163,

214, 216). N colocalizes with some of the proteins involved in RNA synthesis (16). Finally, N proteins of MHV and SARS-CoV contain RNA chaperone activity (227). This activity may aid in transcription during template switching events.

Membrane (M) Protein

M protein, the focus of this dissertation, is the most abundant protein in the viral envelope and is about 25kD in size. M is a type III protein containing a short amino terminal domain, followed by three transmembrane (TM) domains, and a long carboxy terminal tail located inside the virion (Fig.4A) (77). For TGEV both amino and carboxy ends of the M protein are located on the outside of the virion (153). The carboxy tail is further divided into a roughly 100 residue amphipathic region that appears to be closely associated with the membrane (155) and a hydrophilic highly charged tail. M localizes in the Golgi region when expressed alone (91, 93). The amino terminus is O- or N-linked glycosylated depending on the coronavirus group (101, 142). Glycosylation does not appear to be important for correct localization of the protein or virus production (43). However, it has been implicated to play a role in interferon induction (42). The M protein is a key player in organizing the assembly process. M molecules interact with each other and also with the S protein and nucleocapsid during virus assembly (46, 47, 58, 95, 139, 143). M-M interactions constitute the overall scaffold for the viral envelope. The S protein and a small number of E molecules are interspersed in the M protein lattice in mature virions (Fig.4B). Previous

14

FIG.4. Schematic of the M protein and viral assembly complexes. A) The M protein is a Type III protein with a roughly 25 residue glycoylated amino terminal domain extending outside the virion, followed by three transmembrane domains, and a long carboxy tail located inside the virion. The carboxy tail contains a ~100 residue amphipathic domain that is closely associated with the membrane and a ~25 residue hydrophilic highly charged tail. Amino acid numbers are indicated at the N- and C-termini. B) The sequence of the WT M protein carboxy tail starting from after TM3. The amino acid numbers and charged residues are indicated. The CD and hydrophilic tail are underlined. Secondary structure predictions generated from (87) are shown underneath the protein sequence. E=strand C=coil C) Viral assembly complexes organized by the M protein. The M protein forms a lattice structure. The S and E proteins are interspersed in the M protein lattice. The M protein interacts with the nucleocapsid (N+RNA).

A

NH₂ 1

228 COOH

B
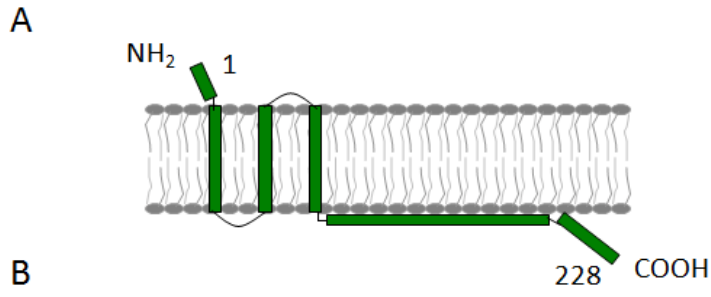
104
                +   +   −   − +   +   −−   +   +
NSIRLFIRTGSWWSFNPETNNLMCIDMKGTVYVRPIIEDYHTLTATIIRGHLYMQGVKLGTGFS
CEEEEEEECCCCEEECCCCCEEEEEEECCCEEEEECCCCCCEEEEEEEEEEEEEEEEEEECCCCC

        −   +   ++   −+ −   ++   +   +   −   +   228
LSDLPAYVTVAKVSHLCTYKRAFLDKVDGVSGFAVYVKSKVGNYRLPSNKPSGADTALLRT
CCCCCEEEEEECCCEEEEEEEEECCCCCEEEEEEEEEEECCCCCCCCCCCCCCCCCCCCCC

C

S

M

E

N

RNA
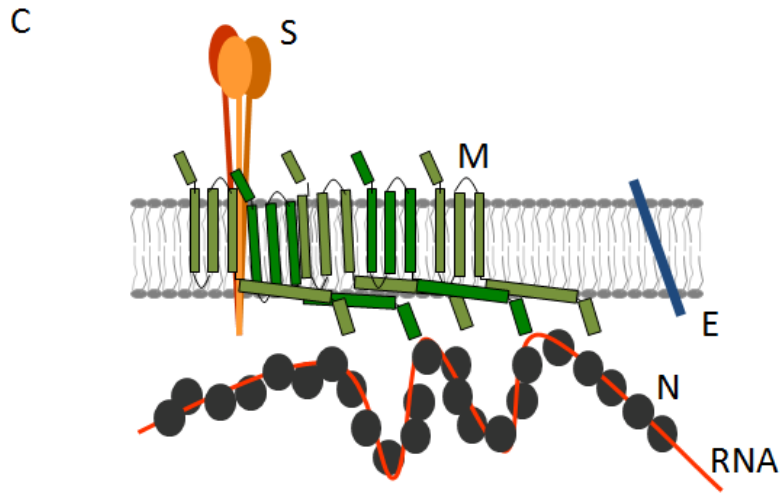
studies have implicated multiple M domains and residues to be important for coronavirus assembly. Regions of the amphipathic domain appear to play a role in M-M and M-S interactions (5, 46). Residues within the carboxy tail are important for interactions with the N protein (83, 95, 120, 134, 195) as well as the S protein (46, 129). This M-N interaction occurs in the absence of the S and E proteins (134). Interestingly, the M-N interaction forms the basis of an internal core structure for TGEV (152). The M protein also interacts directly with the RNA packaging signal (135) highlighting its role in efficient packaging of nucleocapsids. Coronavirus M proteins appear to also interact with host proteins. For example, IBV M has been shown to interact with β-actin and results suggest this interaction is important for assembly and budding of virions but not release (199). Additional host proteins M interacts with remain to be determined.

Coronavirus Genome and Nonstructural Proteins

Viral Genome

Coronavirus genomes are single stranded, positive sense RNA (Fig.5). The genomic RNA contains a 5' 7-methyl guanosine cap and a 3' poly(A) tail, is recognized as a large mRNA molecule by host cells and is infectious when purified (101). At the 5' end of the genome there is a so-called leader sequence 65-98 nucleotides in length. The leader is also located at the 5' ends of all the subgenomic mRNAs (discussed below) (Fig.5) (98). Both the 5' and 3' ends also contain untranslated regions (UTRs) that are 200-500 nucleotides long (Fig.5).

The UTRs play important roles in RNA replication and transcription (98).

Roughly the first two thirds of the genome contain open reading frame (ORF) 1,

which encodes ORF1a and ORF1b. The translation of ORF1b occurs after

ribosomal frameshifting at a region between the two ORFs (98). ORF1a and 1b

encode all the proteins necessary for transcription and replication. In the 3' end of

the genome, all the structural ORFs are present in the order 5'- S, E, M, N- 3'.

There are several nonstructural group specific genes that are located in between

these main structural ORFs (Fig.5). For MHV, these nonstructural genes are not

necessary for replication but deletion of some of these genes results in attenuation

in vivo (44). SARS-CoV contains the most of these nonstructural genes although

none of them have been found to be necessary for growth in tissue culture (217).

Nonstructural Proteins (nsps)

ORF1a and 1b are translated as a large polyprotein which is then

processed into all the proteins necessary for transcription and replication. ORF1

encodes for up to sixteen nonstructural proteins (223, 225). All coronaviruses

encode for one to three proteases, papain-like and chymotrypsin-like, in ORF1a

(98). The proteases then completely process the ORF1 polyprotein through

various processing intermediates. These processing events are just starting to be

studied (208, 224). Because of the proteases important role in viral replication,

many drugs being developed for SARS-CoV target this protein (98). Many gene

products produced from ORF1a have unknown functions. The nonstructural

proteins produced from ORF1a that are known are a protein containing ADP-

Group 1: TGEV

UTR Leader ORF1a ORF1b UTR
5' | S | 3a | 3b | E | M | N | 7 | 3'

Group 2: MHV

UTR Leader ORF1a ORF1b UTR
5' | 2a | HE | S | 4 | 5a | E | M | N | I | 3'

Group 2: SARS-CoV

UTR Leader ORF1a ORF1b UTR
5' | S | 3a | 3b | E | M | 6 | 7a | 7b | 8a | 8b | 9b | N | 3'

Group 3: IBV

UTR Leader ORF1a ORF1b UTR
5' | S | 3a | 3b | E | M | 5a | 5b | N | 3'
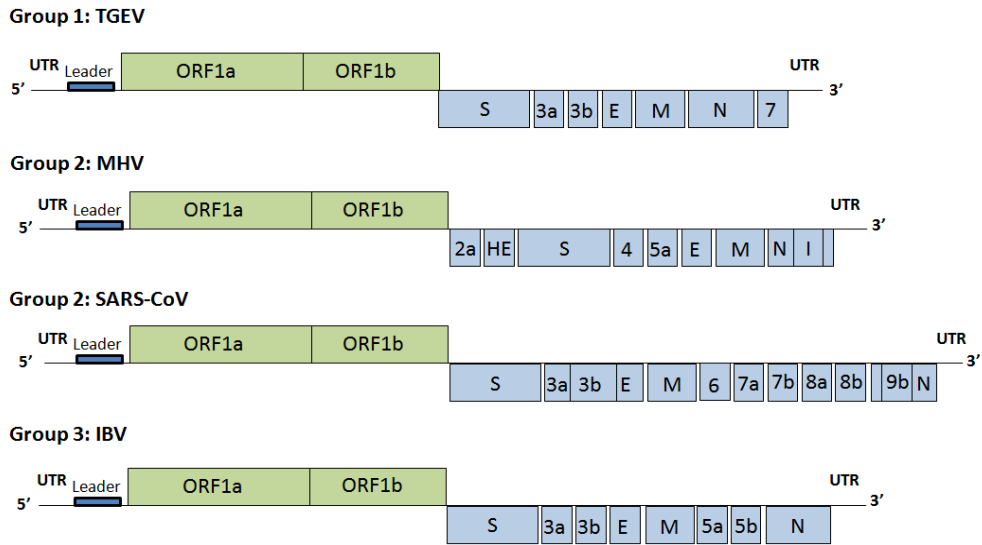
FIG.5. Organization of Coronavirus Genomes. Adapted from (98). The position of the leader, 5' cap, 3' poly(A) tail, and 5'/3' UTRs are indicated. Group 1, 2, and 3 representative virus genomes as well as SARS-CoV are depicted. The first two thirds of the genome contain ORF1a and ORF1b, which are translated into a large polyprotein. The structural and group specific nonstructural genes are located at the 3' end of the genome.

19

ribose 1'-phosphatase activity (nsp3), proteins with a cylinderlike structure which has been predicted to be important for replication (nsp7 and 8), and a single-strand RNA binding protein (nsp9) (55, 159). ORF1b encodes the viral RNA dependent RNA polymerase (RdRp) (nsp12) and a helicase (nsp13). The helicase protein also has other enzymatic properties including NTPase, dNTPase, and 5'-triphosphatase activities (166). Additional nonstructural proteins encoded by ORF1b include a 3'-to-5' exonuclease (ExoN), a uridylate-specific endoribonuclease (NendoU) and an S-adenosylmethionine-dependent 2'-O-ribose methyltransferase (2'-O-MT) (174). The ORF1 proteins localize to regions of intracellular membranes. Transcription and replication for MHV and SARS-CoV are thought to occur on double membrane vesicles (DMVs) that occur around perinuclear regions (71, 175).

Viral Life Cycle

The steps of the coronavirus life cycle are depicted in Fig.6.

Attachment and Entry

The virus life cycle begins with attachment of the virion to receptors on susceptible host cells. The S protein is responsible for binding to specific receptors on the cell surface. The receptors for several coronaviruses have been identified. The receptor for MHV is a biliary glycoprotein in the carcinoembryonic antigen family, Ig superfamily (CEACAM1) (54, 204). The MHV S protein binds to the extracellular Ig-like loops of the CEACAM1
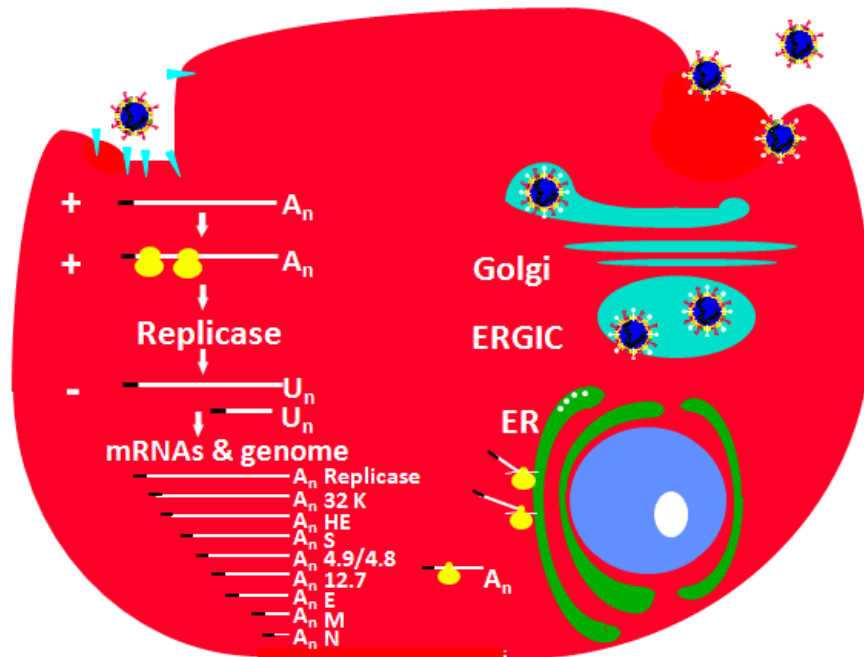
FIG.6. The Coronavirus Life Cycle. Virions attach to the cell via S protein interactions with specific cellular receptors. The viral envelope fuses with the plasma or endosomal membranes to release the genomic RNA into the cytoplasm. ORF1a and ORF1b immediately get translated to produce all the proteins necessary for transcription and replication including the RNA dependent RNA polymerase. A 3' nested set of subgenomic mRNAs are produced for subsequent translation of the structural proteins. Virus assembly and budding occurs in the ERGIC. Virions are transported in smooth-walled vesicles through the exocytic pathway for release from infected cells. From (210).

receptor. CEACAM1 molecules are expressed in the liver and gastrointestinal tract, on macrophages, dendritic cells, B cells, and activated T cells (101, 133, 189). Group I coronaviruses, such as HCoV-229E and TGEV, utilize a host specific cell membrane-bound metalloprotease called aminopeptidase N (APN) for their receptor (11, 101). APN is expressed in respiratory and intestinal epithelium and on myelocytic cells as well as at synaptic junctions (167). Group II coronaviruses, HCoV-OC43 and BCoV, use the N-acetyl-9-O-acetylated sialic acid as their receptor on host cells (94, 165). The host cell receptor for SARS-CoV and HCoV-NL63 is the angiotensin-converting enzyme 2 (ACE2) (75, 113). ACE2 is expressed in several tissues including the heart, lung, kidney and small intestine (73).

After the virus binds to host cells, the process of penetration and uncoating occurs. Once the S protein binds to cell surface receptors it undergoes a conformational change that induces fusion of viral and host membranes (221). Coronaviruses enter cells by fusion with the plasma membrane or endosomes. MHV, BCoV, and IBV induce fusion optimally at a neutral or slightly alkaline pH, which suggests these viruses fuse directly with the plasma membrane (111, 147, 181). In contrast, some MHV strains appear to enter cells by utilizing the pH-dependent endosomal pathway (67). For these viruses, infectivity is reduced in the presence of lysosomotropic drugs (67). After fusion, the nucleocapsid is released into the cytoplasm of host cells. The RNA is then uncoated before transcription and translation occur. The process of release and uncoating of the

22

genomic RNA is poorly understood. There is evidence of a phosphoprotein phosphatase that possesses activity against N and it is hypothesized that the dephosphorylation of N may lead to dissociation of N from the RNA (131). Other studies have suggested that cellular factors are involved in MHV entry (6).

Transcription and replication

After the genomic RNA is released into the cytoplasm, ORF1a and 1b are translated to produce all the proteins necessary for transcription and replication. First the so-called replicase-transcriptase complex generates negative-strand RNA molecules. These RNAs are then used for transcription of subgenomic mRNAs. A 3' coterminal nested set of subgenomic mRNAs are produced during viral infection. Therefore, they all contain the same 3' ends but have 5' ends of various lengths (Fig.7) (101). Although, each subgenomic mRNA contains more than one ORF, only the most 5' ORF gets translated. Each subgenomic mRNA contains the leader sequence at the 5' end, the same leader as is present on the 5' end of the genome.

Intergenic or transcription-regulatory sequences (TRSs) are present on the genomic RNA before each structural ORF. These TRSs play a role in regulation of subgenomic mRNA transcription. Importantly, the 3' end of the leader has similar sequences as the TRSs. Since each subgenomic mRNA contains the leader sequence as well as TRSs, a model of discontinuous transcription was proposed (9, 100, 178). It is still debated whether this discontinuous transcription mechanism occurs during positive- or negative-strand synthesis. The model for

23
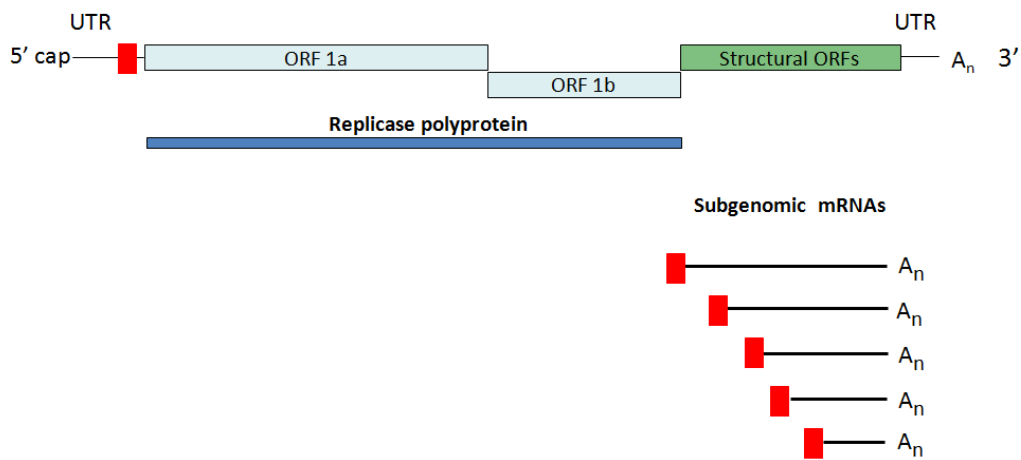
FIG.7. Coronavirus Genome and subgenomic mRNAs. A schematic of the full length RNA genome is shown on top. The 5' cap, poly(A) tail, and UTRs are depicted. The leader is represented as a red box. ORF1a and 1b are translated into the replicase polyprotein. A 3' coterminal nested set of subgenomic mRNAs are produced during infection. They all contain the leader at the 5' end. The subgenomic mRNAs are translated to produce the viral structural proteins.

transcription during positive-strand synthesis has been named leader-primed (99). During leader-primed transcription the leader RNA sequence is synthesized first from full-length negative-strand RNA and then the polymerase switches to the complementary TRS on the negative-stranded RNA (99). This results in a positive-stand mRNA molecule. The other model proposes that subgenomic mRNA transcription occurs during negative-strand RNA synthesis. The polymerase would stall at the TRS sequences and then jump to the 3' end of the leader, that contains homologous sequence to the TRS, to continue transcription (161). This results in negative-strand subgenomic mRNAs that would subsequently be used to generate positive-strand mRNAs. Currently the negative-strand discontinuous transcription mechanism is favored.

Full-length positive- and negative-strand genomic RNAs are also made during infection. A discontinuous synthesis mechanism may occur during full-length RNA synthesis as it does for subgenomic mRNA synthesis (127). The 5' and 3' UTRs as well as RNA secondary structures within these regions are necessary for RNA replication (80, 150).

Translation

ORF1 is translated immediately after infection as one large polyprotein 700-800kD in size. ORF1 contains ORF1a and 1b that are in different reading frames (18). Ribosomal frameshifting results in translation of the two separate ORFs. All the proteins in the transcriptase-replicase complex are encoded in ORF1. Each of the structural proteins is translated off of the subgenomic mRNAs.

Translation is initiated by a traditional cap-dependent mechanism. Even though the subgenomics contain more than one ORF, typically only the most 5' ORF is translated (98). However the MHV E protein, which is encoded on subgenomic mRNA 5, is translated from the second ORF (185). Translation of the E protein occurs through a cap-independent internal ribosomal entry site (IRES).

Assembly and Release

Virus assembly takes place at the ER/Golgi intermediate compartment (ERGIC) (69, 91, 187). The viral structural proteins are localized to the site of assembly. The S protein is cotranslationally inserted into the rough ER (RER). Although some S localizes to the plasma membrane, it primarily localizes at the site of assembly and interacts with the M protein (139, 143). The M protein is also cotranslationally inserted into the ER membrane and localizes primarily in the ERGIC in infected cells (91). When expressed alone, M localizes to the Golgi, beyond the site of virus assembly (91). The E protein localizes to the ERGIC region in cells (143). The HE protein is transported to the Golgi and a fraction is also localized to the plasma membrane (98). Finally, the N protein is located at intracellular membranes where it is associated with components of the RNA replication complex (171). However, N is found at the ERGIC region late in infection when assembly is taking place (19). Since E is the only viral protein that localizes at the site of assembly on its own, it is thought E determines the site of assembly.

The viral structural proteins interact with each other to form assembly complexes that help facilitate the assembly and budding process. The M protein is a key player in organizing the assembly process. M molecules interact with each other and also with the spike and nucleocapsid during virus assembly (Fig.4B) (46, 47, 58, 95, 139, 143). M-M interactions constitute the overall scaffold for the viral envelope. M associates with other M molecules in pre-Golgi regions, consistent with assembly occurring in the ERGIC (47). M-M interactions are thought to be mediated through multiple regions but the transmembrane domains appear to be especially important (47). In particular a conserved region at the amino end of the amphipathic region was found to be important for efficient M-M associations (5). Furthermore, cryo-tomography analysis of VLPs and virus particles revealed that the envelope is striated, presumably due to M-M associations (7). It has been suggested that the lattice structure actually precludes foreign proteins from being incorporated into the viral envelope (47).

During assembly, the M protein also interacts with the N protein (83, 95, 134, 195, 196). These interactions presumably facilitate the incorporation of the nucleocapsid into viral particles. Charged residues within the extreme carboxy tail of MHV (particularly the penultimate charge R_{227}), SARS-CoV and TGEV M, have been implicated to be important mediators for the M-N interaction (58, 83, 95, 120, 195, 196). Charged residues within the C-terminus of N have also been identified to be key residues for N's interaction with the M protein (83, 195). However, the data clearly suggests other regions of M are essential for the M-N

association. Due to the complexity of this interaction, further studies mapping regions in the M protein carboxy tail necessary for M-N association are of significant interest and could be integrated with those of previous studies to generate a more complete picture of M-N electrostatic charge interactions.

In addition to M-M and M-N interactions, the M protein interacts with the S protein as well during assembly. Interactions with S are important to ensure incorporation of the viral attachment protein. The amphipathic region of the M protein tail has been shown to be important in mediating M-S interactions (46). Regions within the hydrophilic tail of M have also been identified as being crucial for SARS-CoV and MHV M protein interactions with S (46, 129). Clearly there appears to be multiple contact points between the two proteins and there is a need for the specific residues on M that are involved in the interactions to be elucidated.

Although all these protein interactions must take place for efficient assembly of infectious particles, the M and E proteins are the minimal requirements for envelope formation for MHV, IBV, and BCoV, as they form VLPs (14, 35, 194). Therefore, it is thought these two viral proteins are crucial for assembly of virions. Most enveloped viruses require the nucleocapsid for assembly so coronaviruses are unique as they utilize nucleocapsid independent assembly. Mutations and deletions made in all regions of M, including the amino end, TM domains and carboxy tail, cause decreases in VLP assembly (43), highlighting the importance of the M protein in the assembly process. However,

since the M protein alone cannot efficiently drive the assembly process, it is thought the E protein plays an important role in assembly and release. Studies have suggested that E may induce membrane curvature or could be involved in the pinching off of budding virions (149, 194). Additionally, when E is expressed alone, vesicles are produced and released from cells (125). The ion channel activity of E may also play a role in budding events (124, 205, 206). After virions assemble and bud in the ERGIC region, they move through the Golgi to reach secretory vesicles. Virions accumulate in smoothed-walled vesicles which are trafficked to the plasma membrane via the cellular secretory pathway (98). Virions are then released extracellularly.

Reverse Genetics/Infectious Clones

Due to the large size of the coronavirus genome, it was initially difficult to generate full length cDNA infectious clones. However, full length cDNAs have now been developed for MHV (32, 216), IBV (21, 213), TGEV (3, 214), HCoV-229E (184) and SARS-CoV (215). The cDNAs are maintained using bacterial artificial chromosomes, vaccinia virus vectors, and multiple plasmids. One of the MHV infectious clones was used extensively in this dissertation (216). The viral genome is maintained in seven different plasmids. The seventh plasmid, called the G clone, contains all the structural genes. Mutagenesis can be done to any region of the genome. Each plasmid is isolated and ligated together to generate the full length cDNA clone. *In vitro* transcription is carried out and the infectious RNA is electroporated into susceptible cells.

CHAPTER 2

A CONSERVED DOMAIN IN THE CORONAVIRUS MEMBRANE PROTEIN

TAIL IS IMPORTANT FOR VIRUS ASSEMBLY

ABSTRACT

Coronavirus membrane (M) proteins play key roles in virus assembly, through M-M, M-spike (S) and M-nucleocapsid (N) protein interactions. The M carboxy-terminal endodomain contains a conserved domain (CD) following the third transmembrane (TM) domain. Importance of the CD (SWWSFNPETNNL) in mouse hepatitis virus was investigated with a panel of mutant proteins, using genetic analysis and transient expression assays. A charge reversal for negatively charged E_{121} was not tolerated. Lysine (K) and arginine (R) substitutions were replaced in recovered viruses by neutrally charged glutamine (Q) and leucine (L), respectively, after only one passage. $E_{121}Q$ and $E_{121}L$ M proteins were capable of forming virus-like particles (VLPs) when coexpressed with E, whereas $E_{121}R$ and $E_{121}K$ proteins were not. Alanine substitutions for the first four or the last four residues resulted in viruses with significantly crippled phenotypes and proteins that failed to assemble VLPs or to be rescued into the envelope. All recovered viruses with alanine substitutions in place of SWWS residues had second site, partially compensating, changes in the first TM of M. Alanine substitution for proline had little impact on the virus. N protein coexpression with some M mutants increased VLP production. The results overall suggest that the CD is important for formation of the viral envelope by helping mediate fundamental M-M interactions and that the presence of the N protein may help stabilize M complexes during virus assembly.

31

INTRODUCTION

Coronaviruses are widespread medically important respiratory and enteric pathogens of humans and a wide range of animals. New human coronaviruses (HCoV), including severe acute respiratory syndrome (SARS-CoV), HCoV-NL63, and HCoV-HKU1, were recently identified (193, 207). The potential for emergence of other new viruses and the zoonotic nature of some coronaviruses strongly warrants understanding old and new viruses. Understanding vital interactions that take place during virus assembly and conserved domains that mediate these interactions can provide insight toward identification of targets for development of antiviral therapeutics and vaccines.

Coronaviruses are enveloped positive-stranded RNA viruses that belong to the *Coronaviridae* family in the *Nidovirales* order. The virion envelope contains at least three structural proteins, the membrane (M), spike (S), and envelope (E) proteins. The genomic RNA is encapsidated by the N phosphoprotein to form a helical nucleocapsid. The S glycoprotein is the viral attachment protein that facilitates infection through fusion of viral and cellular membranes and is the major target of neutralizing antibodies (66). The M glycoprotein is the most abundant component of the viral envelope that plays required, key roles in virus assembly (47, 91, 139, 143, 194). The E protein is a minor component of the viral envelope that plays an important, not clearly defined, role(s) during virus assembly and release (14, 35, 194).

Coronavirus M proteins are divergent in their amino acid content, but all share the same overall basic structural characteristics. The proteins have three TM domains, flanked by a short amino terminal glycosylated domain and a long carboxy terminal tail located outside and inside the virion, respectively (77) (Fig.8A). M localizes in the Golgi region when expressed alone (91, 93). M molecules interact with each other and also with the spike and nucleocapsid during virus assembly (46, 47, 58, 95, 139, 143). M-M interactions constitute the overall scaffold for the viral envelope. The S protein and a small number of E molecules are interspersed in the M protein lattice in mature virions. Previous studies from a number of labs implicated multiple M domains and residues to be important for coronavirus assembly (43, 46, 47, 83, 196). Coronaviruses assemble and bud at intracellular membranes in the region of the endoplasmic reticulum Golgi intermediate compartment (ERGIC) (93, 188). Co-expression of only the M and the E proteins is sufficient for VLP assembly for most coronaviruses (14, 194).

The long intravirion (cytoplasmic) tail of M consists of an amphipathic domain following the third TM and a short hydrophilic region at the carboxyl end of the tail. The amphipathic domain appears to be closely associated with the membrane (155). At the amino terminus of the amphipathic domain, there is a highly conserved 12 amino acid domain (SWWSFNPETNNL), consisting of residues 114-125 in the MHV A59 M protein (Fig.8B) (89). These residues are almost identically conserved across the entire *Coronaviridae* family. Because of

33

FIG.8. M protein conserved domain and mutants. (A) A linear schematic of the M protein is shown illustrating the relative positions of the three TM (black boxes) and position of the CD in the tail. (B) Alignment of CDs from representative coronaviruses is shown below. Full length amino acid sequences from transmissible gastroenteritis virus (TGEV), feline coronavirus (FeCoV), human coronavirus 229E, human coronavirus NL63, mouse hepatitis virus (MHV), bovine coronavirus (BCoV), human coronavirus OC43, porcine hemagglutinating encephalomyelitis virus (HEV), human coronavirus HKU1, SARS-CoV, infectious bronchitis virus (IBV), and turkey coronavirus (TCoV) were aligned by CLUSTAL W (103). (C) Mutations introduced into the MHV CD are shown at the bottom, with + and – symbols used to indicate VLP production and virus recovery for each mutant.

A. 114 **SWWSFNPETNNL** 125

NH₂—■ ■ ■——————COOH

B. Group I
TGEV   145 SWWSFNPETKAI 156
FeCoV   145 SWWSFNPETNAI 156
229E   107 TFWAWNPEVNAI 118
NL63   108 TFWAFNPETNAI 119
Group II
MHV   114 SWWSFNPETNNL 125
BCoV   113 SWWSFNPETNNL 124
OC43   113 SFWSFNPETNNL 124
HEV   113 SWWSFNPETNNL 124
HKU1   109 SWWSFNPETNNL 120
SARS   107 SMWSFNPETNIL 118
Group III
IBV   109 SWWSFNPESNAV 120
TCoV   107 SWWSFNPESSAV 118

C.

MHV	SWWSFNPETNNL		VLPs	Virus
E₁₂₁A		A	–	+
E₁₂₁K		K	–	–
E₁₂₁R		R	–	–
P₁₂₀A		A	+	+
5'A	AAAA		–	+
3'A		AAAA	–	+
5'A+3'A	AAAA	AAAA	–	–
ΔCD	———————		–	–

35

the crucial role that M plays in virus assembly and the high conservation of this domain, we hypothesized that the conserved domain (CD) is functionally important for virus assembly. To test this, a series of changes were introduced in the CD. The functional impact of the changes was studied in the context of the virus by genetic analysis and ability of the mutant M proteins to participate in VLP assembly. The results show that the CD is functionally important for M protein to participate in virus assembly. The domain may help mediate important lateral interactions between M molecules. The results suggest that the N protein helps stabilize M complexes during virus assembly.

MATERIALS AND METHODS

Cells and viruses. Mouse L2 and 17 clone1 (17Cl1) cells were maintained in Dulbecco's modified Eagle's medium (DMEM) containing 10% and 5% heat-inactivated fetal calf serum (FCS), respectively, plus penicillin, streptomycin, and glutamine. Baby hamster kidney cells expressing the MHV Bgp1a receptor (BHK-MHVR) (216) and BHK-21 cells (ATCC) were grown in Glasgow minimal essential medium (GMEM) containing 5% FCS, supplemented with antibiotics, glutamine and 10% tryptose phosphate broth. BHK-MHVR cells were maintained in the presence of Geneticin (G418) for continuous selection of cells expressing the receptor (216). Human 293T cells (ATCC) were grown on 0.25% gelatin coated flasks in DMEM containing 10% FCS, antibiotics and glutamine. Stocks of wild-type MHV A59 and cloned viruses were grown in mouse 17Cl1 or L2 cells. Virus titers were determined in L2 cells.

Construction of amino acid substitution and deletion mutants. pGEM-5Zf(+)M-N, a pGEM5Zf(+) vector (Promega) containing the M and N genes (EcoRV-SacI fragment), was used for mutagenesis (195). Site-directed amino acid substitutions were introduced by whole-plasmid PCR using appropriate primers and platinum Pfx DNA polymerase (Invitrogen). Mutations were confirmed by sequence analysis of the entire M-N insert before subcloning into the MHV G clone at the EcoV and BssH II sites (216).

Reverse genetics. All mutant viruses were made by reverse genetics using a MHV A59 infectious clone (216). Full-length cDNA clones were assembled,

transcribed, and electroporated into BHK-MHVR cells as described previously (195). Following electroporation BHK-MHVR cells were seeded alone or in some cases were seeded concurrently with L2 cells. At 24-46 h after electroporation media were harvested and an aliquot was used to infect 17Cl1 or L2 cells. The medium was removed from infected cells at 24 to 72 h p.i. Total cytoplasmic RNA from cells remaining on the flasks was extracted using RNAqueous-4PCR extraction buffers (Ambion) and treated with DNase before being reverse transcribed using the Superscript RT-PCR system from Invitrogen and an oligo (dT) primer. The RT product was subjected to 30 cycles of PCR amplification using SuperTaq Plus (Ambion) and appropriate primers to amplify the E, M, N and S genes. PCR products were cleaned with MiniElute PCR (Qiagen) before being sequenced directly.

Viruses were subsequently plaque purified from media taken directly off cells that had been electroporated. Multiple plaque isolates were passaged on 17Cl1 cells five times for all viable viruses, except for the 3'A mutant that was passaged on L2 cells. In some cases where viruses grew poorly and yielded small plaques from which it was difficult to recover isolates, the stock off electroporated cells was passaged five times to allow for selection of the most fit viruses. RNA was extracted from infected cells following passage of plaque purified viruses for RT-PCR and sequence confirmation of the E, M, N, endodomain of S genes and the packaging signal to determine the stability of introduced mutations and identify potential compensating changes.

Virus growth properties. Growth kinetic experiments were carried out in mouse 17Cl1 cells infected at a multiplicity of infection (MOI) of 0.01 or 0.001 PFU/cell with passage 5 virus stocks. Cell culture supernatants were collected at indicated times after infection and virus titers were determined by plaque assay on mouse L2 cells. Low melting agarose/medium overlays were removed at 72-96 h p.i. Cells were fixed and stained with crystal violet in ethanol to visualize plaques. Two independent growth kinetic experiments were performed for each set of mutant viruses in parallel with the wild-type virus. Growth kinetic curves represent exponential growth with saturation. For each experiment starting titer and saturation parameter were the same. Curves were distinguished by variation in doubling times. Relative error was assumed to be the same for all time points for each data set. Error estimate was computed by taking the standard deviation divided by mean for each of the time points and averaging over all points.

VLP analysis. Wild-type and mutant M genes were expressed in the pCAGGS vector under the control of the chicken β-actin promoter for transient expression (140). A Kozak sequence was included in the forward primer for all genes. All mutant M genes were shuttled into the pCAGGS vector from the MHV G fragment used to generate full-length MHV genomic cDNAs for virus construction. The 5'A M mutants with TM1 second site changes were subcloned by RT-PCR of RNA from mutant virus infected cells. Wild-type E and N genes were also expressed in the pCAGGS vector.

293T cells were transfected with the pCAGGS plasmids containing the wild-type or mutant M genes, E gene, and in some cases the N gene, using the TransIT-293 transfection reagent (Mirus). At ~24 h after transfection, the culture medium and intracellular cytoplasmic fraction were harvested. Cells were lysed on ice with 0.5% Triton X-100, 50 mM NaCl and 50 mM Tris-HCl (pH 7.5) in the presence of 1mM phenylmethylsulfonyl fluoride (PMSF). The medium was clarified at 14,000 x g for 10 min at 4°C. VLPs from clarified media were pelleted through a 30% sucrose cushion by ultracentrifugation for 3 h at 4°C in a Beckman SW55Ti rotor at 30,000 rpm. Pellets were resuspended in Laemmli sample loading buffer and analyzed by SDS-PAGE. Proteins were transferred to polyvinylidene difluoride (PVDF) membranes and analyzed with anti-MHV M A03 (kindly provided by Kathryn Holmes, University of Colorado Health Sciences) and anti-MHV E 9410 and anti-MHV N antibodies generated in our lab (33, 118). Following incubation with appropriate secondary antibodies, blots were visualized by chemiluminescence (Pierce). VLP release was quantified by densitometric scanning of fluorograms and analyzed using Image-Quant software (Molecular Dynamics).

Rescue of mutant M proteins. The A2A3 M was generated by site-directed mutagenesis of the second and third serines in the gene and rescue analysis was carried out basically as previously described (45). 293T cells were transfected with pCAGGS plasmids as described above with MirusTrans-293 reagent (Mirus). At 4 h after transfection cells were starved for 30 min at 37°C in

methionine- and cysteine-free DMEM, containing 5% FCS, prior to labeling with 150 μCi/ml of Expre^{35}S^{35}S labeling mixture (Perkin Elmer). At 24 h after transfection cells were washed with PBS and lysed on ice with radioimmunoprecipitation assay buffer (RIPA) containing 1% Triton X-100, 1% deoxycholate, 0.3% SDS, 150 mM NaCl, 50 mM Tris-HCl [pH 7.6], 20 mM EDTA and 1 mM PMSF. After clarification of the extracellular medium, VLPs were lysed by incubation with 2X RIPA buffer on ice and sonication for 1.5 min at 30 sec intervals. Lysates were precleared with protein A-sepharose CL-4B (GE Healthcare Life Sciences) by rocking for 1 h at 4° C. Each sample was divided in half and immunoprecipiated with anti-MHV M monoclonal J1.3 that recognizes the amino terminus of the protein (62) or an anti MHV-serum 488 (kindly provided by Kathryn Holmes, University of Colorado Health Sciences) antibody overnight at 4° C. Protein complexes were isolated by incubation with protein A-Sepharose for 2 h at 4° C while rocking. Pelleted complexes were washed 5 times with RIPA buffer followed by 1 wash with RIPA buffer minus detergents. Proteins were eluted in Laemmli SDS-PAGE sample loading buffer by heating at 95° C for 5 min and analyzed by SDS-PAGE. Gels were incubated in Amplify Fluorographic Reagent (GE Healthcare Life Sciences) for 30 min at room temperature prior to drying and subsequent fluorography.

Indirect immunofluorescence. Co-localization of mutant M genes with WT S was determined by expression of pCAGGS plasmids containing WT or mutant M and WT S genes in BHK-21 cells. Cells were plated on Lab-teck chamber slides

(Nunc) 1 day prior to transfection with MirusTrans-293T reagent (Mirus). At 12 h after transfection cells were washed two times with PBS and fixed in methanol for 15 min at -20°C. Following one additional wash with PBS, cells were blocked with 2% gelatin in PBS overnight at 4° C. Monoclonal antibodies J1.3 and 2.7 (62) and a polyclonal antibody, A04 (kindly provided by Kathryn Holmes, University of Colorado Health Sciences) were used to detect the M and S proteins, respectively. Slides were incubated for 2 h with primary antibodies at room temperature, washed with 2% gelatin in PBS before incubation with Alexa-labeled secondary antibodies. Cells were washed with 2% gelatin in PBS and then a final wash with PBS alone. Slides were mounted in ProLong Gold antifade reagent (Molecular Probes) containing 4,6-diamino-2phenylindole (DAPI) to stain nuclei. Images were collected by using an epifluorescence Nikon inverted microscope (Nikon, Inc., Melville, NY) with MetaMorph imaging soft-ware (Universal Imaging Corp., Downingtown, PA). Images were processed using Adobe Photoshop.

RESULTS

Construction of conserved domain mutants. To study the importance of the 12 amino acid CD in the tail of coronavirus M proteins, MHV-A59 virus was used as the model. Site directed mutagenesis was used to change or delete residues in the domain (Fig.8C). Two nucleotides were introduced for all codon changes to assist in identification of compensatory changes. A MHV-A59 infectious clone was used to study the impact of residue changes in the context of full virus assembly. Recovered viruses were plaque purified and passaged multiple times. Sequence analysis at various passage points was used to monitor genetic stability of the introduced mutation(s) and identify secondary compensatory changes. The M, N, and E genes, as well as the carboxy end of the S gene and packaging signal in gene 1b, were sequenced. All viruses were analyzed for their growth properties to determine the impact of genetic changes directly on assembly. Selected mutant M proteins were also analyzed for their ability to support VLP production by co-expression of the E and M genes, and in some cases the N gene was included, to provide insight into the possible role of the CD in envelope formation.

A positive charge within the conserved domain is not tolerated. To address the significance of the only charged residue, glutamic acid (E_{121}) was changed to neutrally charged alanine (A), as well as to positively charged arginine (R) or lysine (K) (Fig.8C). The negative charge was of interest for its potential to mediate protein-protein interactions through electrostatic means. Sequence analysis of the $E_{121}A$ virus from cells infected with medium taken directly from

43

electroporated cells showed that the introduced mutations were retained and that no other changes were present the structural genes. In contrast, the positive charges in both the $E_{121}K$ and $E_{121}R$ clones were replaced after the first passage with neutrally charged glutamine (Q) or leucine (L), respectively. Sequence analysis of the total virus population also revealed one additional change in the N gene, theonine (T) to A at position 428 in the $E_{121}Q$ mutant and glycine (G) to A at position 94, in the $E_{121}L$ mutant.

The recovered $E_{121}A$, $E_{121}Q$, and $E_{121}L$ charged residue viruses were plaque purified from media off cells that had been electroporated with full-length genomic RNAs and multiple isolates of each were passaged on 17Cl1 cells. After five passages the plaque purified viruses were reanalyzed by sequencing. Four of five P5 $E_{121}A$ viruses had no additional changes in any of the structural genes. The other P5 isolate contained an additional change, valine at position 410 to isolucine, in the N gene. Parallel growth analysis of the $E_{121}A$ virus with and without the change in N indicated that both grew comparably (data not shown). Therefore, the $E_{121}A$ mutant virus with no other changes was used for further analysis.

Four out of six $E_{121}Q$ plaque purified virus isolates maintained the $T_{428}A$ change that was seen in the population analysis following electroporation in the N gene, and two had only the $E_{121}Q$ change. To determine if the change in the N gene was providing a growth advantage, one of the plaque purified $E_{121}Q$ viruses without the $T_{428}A$ change and one with the change were analyzed for their growth

properties in parallel with WT virus. Both viruses exhibited comparable growth properties (data not shown). The $E_{121}Q$ virus with the $T_{428}A$ change in N was used for subsequent finer detailed growth analysis.

Sequence analysis of the $E_{121}L$ plaque purified viruses showed that all contained the $G_{94}A$ change seen immediately following electroporation in the N gene through P5. To determine the impact of the $G_{94}A$ second site change, an independent clone was constructed that contained only the $E_{121}L$ change. Recovered virus was plaque purified. Seven individual plaque isolates had only the $E_{121}L$ change. Interestingly however, an equal number had additional changes scattered across the N gene ($G_{23}E$, $E_{236}K + K_{395}E$, $N_{153}D + E_{173}K$, $E_{374}K$, $S_{417}N$, $R_{13}G$, $D_{346}N$) (Table 1). The virus isolate that contained the $N_{153}D + E_{173}K$ changes in the N gene also had a silent change in the E_{1321} codon in the S gene. One plaque isolate had an E to G change at amino acid position 71 in the E gene. One isolate had a change in the S gene ($E_{1321}K$). Since growth kinetic analysis indicated that the $G_{94}A$ change in the N gene was not providing additional compensation over the $E_{121}L$ change (data not shown), none of the other virus isolates with the additional changes were analyzed further.

Selected P5 viruses, as described above, were analyzed in parallel, along with the WT virus, for their growth properties (Fig.9). Viruses $E_{121}A$, $E_{121}Q$ with $T_{428}A$ and $E_{121}L$ with $G_{94}A$ change all produced plaques and exhibited growth properties similar to the WT virus. These results indicate that a negative charge at position 121 within the MHV M protein conserved domain is not absolutely

Table 1. Summary of recovered CD mutant viruses

Mutant Viruses	# Plaques Analyzed	# Passages	Mutation Retained	Change(s) in M	Change(s) in N
Membrane Protein					
$E_{121}A$	4	5	Yes	-	-
	1	5	Yes	-	$V_{410}I$
$E_{121}Q$ (K)	4	5	No	$K_{121}Q$	$T_{428}A$
	2	5	No	$K_{121}Q$	-
$E_{121}L$ (R) #1[a]	4	5	No	$R_{121}L$	$G_{94}A$
$E_{121}L$ #2[a,b]	7	2-5	Yes	-	$G_{23}E$ $E_{236}K$, $K_{395}E$ $N_{153}D$, $E_{173}K$[c] $E_{374}K$ $S_{417}N$ $R_{13}G$ $D_{346}N$
	7	2-5	Yes	-	-
$P_{120}A$	5	5	Yes	-	-
SWWS → AAAA #1[a]	2	5	Yes	$G_{31}R$	-
	Blind Passage	5	Yes	$Q_{42}R$	-
SWWS → AAAA #2[a]	Electroporation	1	Yes	$T_{38}N$	
	6	2-5	Yes	$Q_{42}R$ (3)[d] $Q_{42}R$ (2)[d] $L_{35}P$ (1)[d]	$N_{409}S$
TNNL → AAAA	Blind Passage	5	Yes	-	-

[a] #1 and #2 refer to independent constructions described in the text.

[b] Two additional plaque purified viruses were isolated. One had a change in E ($E_{71}G$) and another had a change in S ($E_{1321}K$).

[c] $N_{153}D$, $E_{173}K$ also contained a silent change in S in the $E_{1321}K$ codon.

[d] Numbers in parentheses refer to number of plaque isolates with indicated changes.

FIG.9. Growth properties of single substitution mutant viruses. (A) Mouse

17Cl1 cells were infected with mutant viruses at a MOI of 0.01 PFU/cell. Titers

were determined by plaque assay on L2 cells at the indicated times. Data points

are shown for all viruses, but the growth curve is included for only the WT virus.

Error bars and exponential growth curves are as described in the material and

methods. Estimated doubling time was ~0.62 h. (B) Plaque characteristics were

determined for the indicated viruses in L2 cells. Respective changes in the

recovered viruses with positive charge (R and K) substitutions and

corresponding noncompensating changes in the N gene are indicated.

A.

B.

WT   E$_{121}$A   P$_{120}$A

E$_{121}$L(R)   E$_{121}$Q(K)
N: G$_{94}$A   N:T$_{428}$A

48
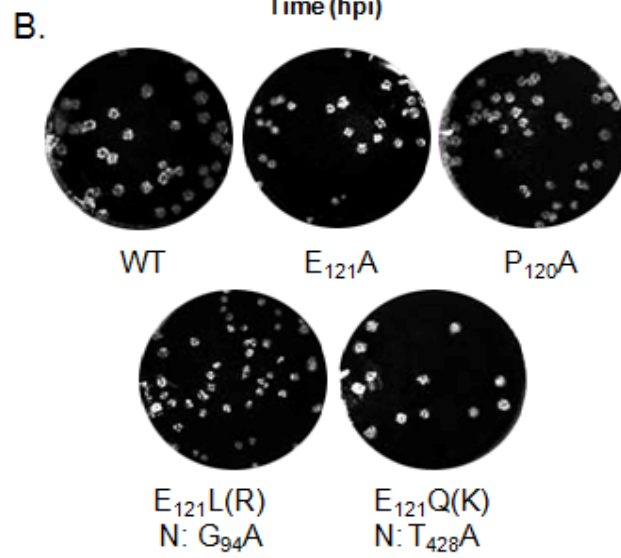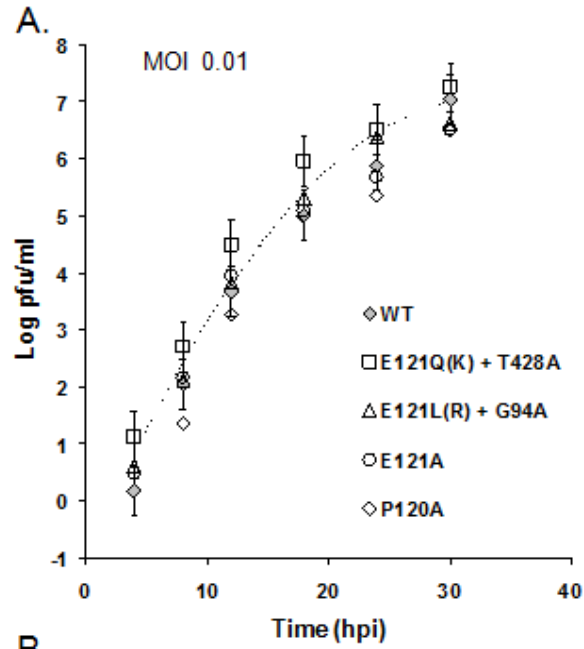
required. However, placement of a positive charge at this position is not tolerated. While additional changes within the N protein were identified, these appear to not provide any additional growth advantage.

The proline residue is not absolutely required. The proline (P) at position 120 was changed to alanine to determine if removal of the helix breaking residue, affects the function of M. Virus was recovered with the $P_{120}A$ change and no additional changes. Five plaque purified viruses were isolated and analyzed after five passages. All stably retained the introduced mutation and had no additional changes in any of the structural genes. The mutant viruses displayed a plaque phenotype and grew at a rate similar to the WT virus (Fig.9). This data interestingly strongly suggests that the proline does not play a crucial role in the structure or function of this region of M.

Replacement of multiple residues affects virus growth. The conserved four residues at the amino end of the CD consist of two that are large and nonpolar, flanked by two smaller polar amino acids (SWWS) (Fig.8C). The high conservation of the WW residues was of particular interest since these residues might participate in protein-protein interactions. To determine the significance of the SWWS residues, all four were replaced with alanines. After electroporation of full-length transcripts containing the alanine substitutions fusion was clearly visible, indicating that the full-length genomic transcripts were replication competent. However, recovery of viruses upon subsequent passage was difficult. Growth of the virus on L2 cells gave rise to very small plaques (Fig.10A). Only

FIG.10. Growth properties of 5'A and 3'A mutant viruses. (A) Plaque characteristics of WT, 5'A + $G_{31}R$ and 3'A viruses were analyzed in mouse L2 cells. (B) Summary of second site changes in the TM1 of recovered viruses from two independent virus constructions designated as 1 and 2 are shown under the sequence of WT TM1 (underlined). Assigned numbers for recovered plaque purified isolates and the passage numbers when analyzed are indicated. (C) Growth kinetics experiments were performed in mouse 17Cl1 cells infected at the indicated MOIs. Titers were determined by plaque assay on L2 cells at the indicated times. Data represent averages from two independent growth kinetic experiments as described for Fig. 2 and in the materials and methods. Estimated doubling times were ~0.62 h for WT virus, ~ 0.79 h for all 5'A viruses, except $L_{35}P$ which was ~1.2 h, and 1 h for the 3'A virus.

A.   WT   5'A+ G31R   3'A
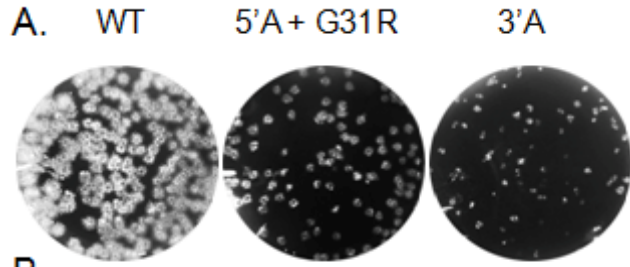
B.

```
                            30              40
WT M    ..KEWNFSLGIILLFITIILQFGYTSR..

Construct 1
Blind  (P5)                              R
#3     (P5)              R
#6     (P5)              R
Construct 2
Blind  (P1)                     N
#1     (P2)                              R
#2     (P5)          P
#3     (P5)                              R
#4     (P3)                              R
#5     (P3)                              R
#6     (P3)                              R
```

C.

51
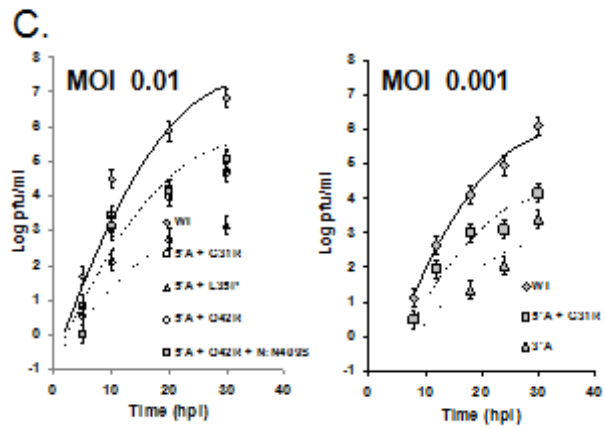
two plaques from a total of 8 could be recovered and subsequently passaged five times. Both of the plaque purified viruses contained the SWWS to AAAA change and each had an additional G to R change at position 31 in the M gene that was maintained through P5 (Fig.10B). Since viruses were recovered initially from only two isolated plaques, we blind passaged the medium from the cells that had been electroporated to determine if other second site compensatory changes might arise. Interestingly, after five passages, the virus population had stably maintained all of the alanine substitutions, but also contained a Q to R change in the M protein at position 42 (Fig.10B).

We then constructed a second independent clone with the SWWS to AAAA change. When the media from electroporated cells was used to infect new cells, the recovered virus population again contained, in addition to the introduced alanine substitutions, a second site change, $T_{38}N$, in the M gene. A mix of large and small plaques was subsequently observed when the virus population was analyzed. Sixteen representative plaques were isolated, but viruses could be recovered from only six of these, even after multiple passage attempts. All six of the recovered viruses retained the 5'A mutations and had one additional change located in TM1 as was seen with the first clone (Fig.10B). Five of the viruses had the $Q_{42}R$ change observed earlier. One of the viruses surprisingly had a $L_{35}P$ second site change. Two of the isolates (#5 and 6) with the additional $Q_{42}R$ changes also had a change in the N gene, $N_{409}S$.

All of the recovered 5'A viruses with the TM1 changes, except $T_{38}N$ that was only recovered during blind passage, were compared for their growth characteristics in parallel with the WT virus at a MOI of 0.01 pfu/cell (Fig.10C, left). All of the recovered viruses exhibited significant growth impairment with ~30% longer doubling time and production of ~100 fold less virus than the WT by 30 h p.i. The virus with the recovered change at $L_{35}P$ change actually grew even less well, producing at least 1,000 fold less virus at 20-30 h p.i. While we expected that the $N_{409}S$ second site change in the N gene might be a compensatory, this was not the case. The virus with the additional change in N grew comparably to the others.

The carboxy end of the CD exhibits overall conservation similar to that the amino end (Fig.8B). Alanine substitutions were introduced in place of the TNNL residues to also assess the importance of the region. Like the amino end mutations, the full length infectious clone RNA bearing the TNNL to AAAA changes was replication competent, as indicated by tiny centers of fusion following electroporation. However, recovery of the 3'A virus was even more difficult than for the 5'A virus. Multiple attempts to recover viruses from the tiny plaques were not successful. A virus stock with a very low titer was subsequently recovered after blind passage of the media from the electroporated cells, but no second site changes were present in the recovered virus population. The virus stock still produced very tiny plaques and growth kinetic analysis of at a MOI of 0.001 pfu/cell in parallel with WT and 5'A + $G_{31}R$, showed that the virus was

very crippled, reaching peak titers about 1000 times lower than the WT virus (Fig.10A and C).

Altogether, the results demonstrate that the conserved residues at the ends of the CD are important. Second site changes arose in the 5'A mutant virus, but none were recovered for the 3'A mutant virus. It is interesting that the majority of the second site changes clustered in the TM1 of M. Nonetheless, the results indicate that the changes are only partially compensating since all of the viruses were still significantly impaired in their growth.

Extensive changes to the conserved domain are lethal. Two additional more extensive mutations were also introduced during the study (Fig.8C). The four residues at each end of the CD were replaced with alanines in the 5'A + 3'A mutant. The entire CD was also deleted. Pinpoint size fusion foci were observed following electroporation, indicating that the infectious cloned RNA was replication competent. However, neither of the viruses with the more extensive changes could be recovered. As described above for the 5'A and 3'A mutants, blind passage directly from the media off electroporated cells was also attempted, but virus could not be recovered. These results strongly indicate that the 5'A+3'A and ΔCD mutations were lethal to the virus.

Since significant changes were introduced with the 5'A + 3'A and ΔCD mutations, the localization of the proteins were analyzed by immunofluorescence. Wild-type M colocalized with the Golgi marker giantin, as expected (91, 93). The mutant proteins colocalized instead with the endoplasmic reticulum marker,

calnexin (data not shown). Failure of the mutant M proteins to target to the Golgi provided an explanation for their lethal phenotype.

The majority of the CD mutants do not support VLP assembly. To gain further insight into how the CD mutations might be affecting virus assembly, the proteins were analyzed for their ability to form VLPs. Coexpression of the M and E proteins is sufficient for assembly of MHV VLPs (14, 35, 194). Thus, the mutant proteins were coexpressed with the wild-type E protein in 293T cells. In addition, the M genes with the $E_{121}Q$ and $E_{121}L$ changes were also analyzed. Intracellular and extracellular fractions were harvested at 24 h after transfection. Both fractions were analyzed by SDS-PAGE and Western blotting (Fig.11). The only M mutant to support VLP production was $P_{120}A$. Interestingly, even though the $E_{121}A$ change supported virion assembly in the context of the virus, the protein was not competent for VLP assembly. The E_{121} to R or K mutant M proteins did not support VLP assembly however, the two recovered changes, $E_{121}Q$ and $E_{121}L$, were competent for VLP production. As expected, the 5'A + 3'A and ΔCD mutant M proteins did not support VLP assembly.

During infection there are other viral and likely host proteins that participate in assembly of viral particles. The N protein encapsidates the genomic RNA and interacts with M (83, 95, 134, 196). Recent studies indicate that N contributes to efficient assembly of SARS VLPs (172, 200). Thus, we asked if the presence of N during VLP formation would impact VLP output of our mutant M proteins. We tested the three mutants that were capable of forming VLPs,
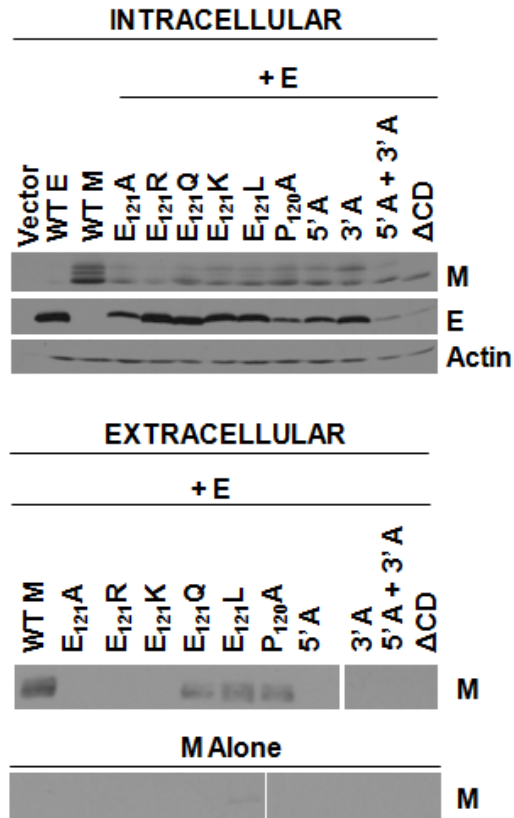
FIG.11. Effect of CD mutations on VLP production. 293T cells were transfected with pCAGGS vectors containing WT or mutant M genes singly or in combination with the pCAGGS containing the WT E gene. Control empty vector (vector) were analyzed in parallel. Intracellular cell lysates and pelleted extracellular VLPs were analyzed by SDS-PAGE and Western blotting using antibodies against M, E or actin as an internal loading control. The entire VLP pellet and 6% of the total intracellular fractions were analyzed.

$P_{120}A$, $E_{121}Q$ and $E_{121}L$. The 3'A mutant was used as a negative control. WT or mutant M proteins were coexpressed with WT E in the presence and absence of WT N clone (Fig.12). All of the mutant M proteins yielded less VLPs than the wild-type M, as expected from our initial experiments. Coexpression of the N protein with the mutant proteins and E resulted in ~25-45% increase in VLP output, compared with ~10% increase for WT M. This correlated well with the growth phenotypes of the $E_{121}Q$, $E_{121}L$ and $P_{120}A$ viruses. The 3'A M mutant was unable to form VLPs even in the presence of N. These results suggest that N may play a stabilizing role during envelope formation.

After determining that the N protein contributed to more efficient assembly of VLPs with the mutant M proteins described above, we also analyzed VLP output for the 5'A mutants that contained the alanine cluster replacement of SWWS at the amino end of the CD which gave rise to viruses with second site changes in the TM1 of M ($G_{31}R$, $Q_{42}R$, $T_{38}N$, and $L_{35}P$) (Fig.10B). At this point we assumed that these changes were providing some advantage, even if not fully compensatory, to the mutant virus with the 5'A mutations, since two independent virus constructions contained changes that were clustered in the domain. Wild-type M, 5'A and all of the mutant M proteins with the 5'A+TM1 changes were coexpressed with E, both in the absence and presence of the N protein (Fig.13). The 5'A and 5'A+TM1 proteins produced none or at most a very low amount of VLPs in the absence of N co-expression. Interestingly however, the 5'A+TM1

57

FIG.12. N protein enhancement of VLP production. 293T cells were transfected as indicated with pCAGGS vectors containing WT or mutant M genes in combination with pCAGGS-E and pCAGGS-N gene where indicated. Intracellular and extracellular VLP fractions were analyzed by SDS-PAGE and Western blotting as indicated in Fig.11. The entire extracellular pellet and 6% of the total intracellular fraction were analyzed. Protein bands were quantified by densitometric scanning and analyzed using ImageQuant software. VLP release was calculated as the percentage of the extracellular M of total M (intracellular plus extracellular) protein. The data represent deviations from the average of two experiments.

Intracellular

Extracellular
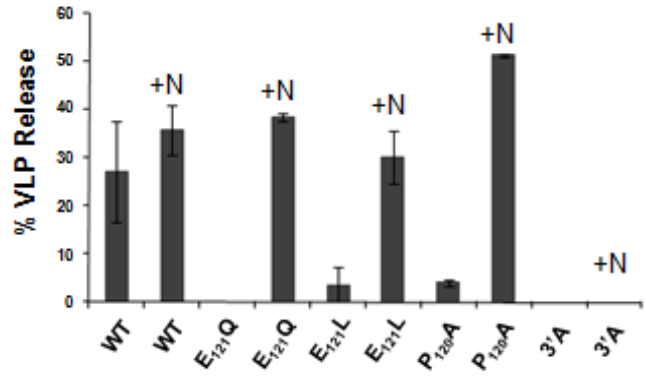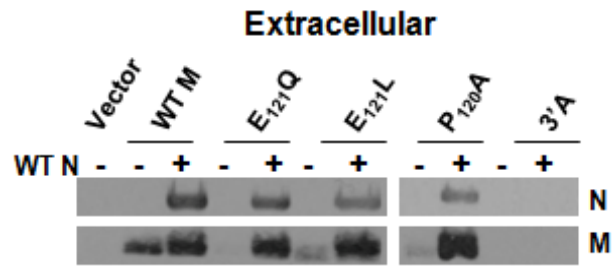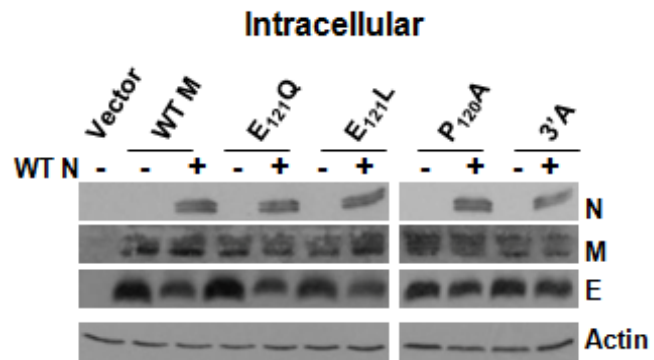
FIG.13. VLP analysis of 5'A M mutants with second-site changes in TM1 coexpressed with the N protein. 293T cells were transfected with WT and mutant genes in pCAGGS vectors. M genes were expressed with E, in the absence and presence of the N gene. Intracellular and extracellular fractions were analyzed by SDS-PAGE and Western blotting. Protein bands from the entire VLP pellet and 6% of total cytoplasmic lysates were quantified and release was calculated as described for Fig.12. Error bars represent the deviation of two independent experiments.

Intracellular
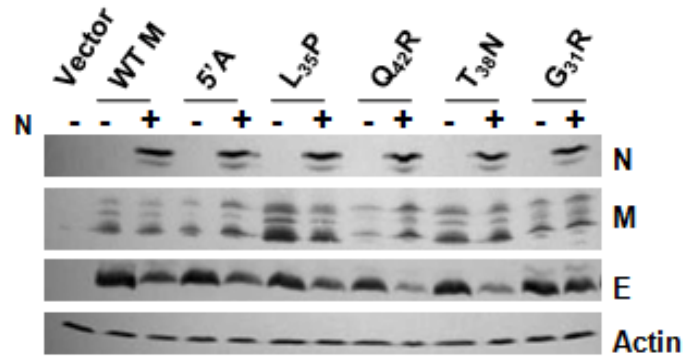
Extracellular
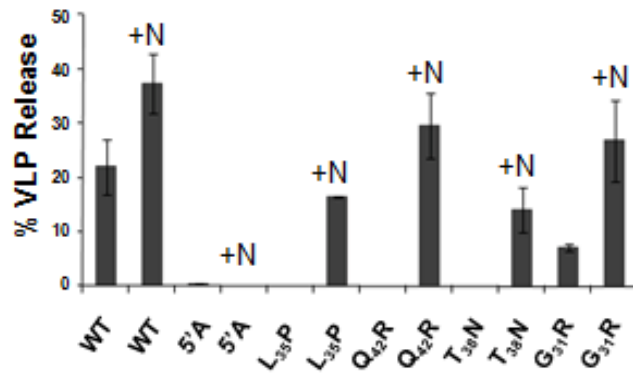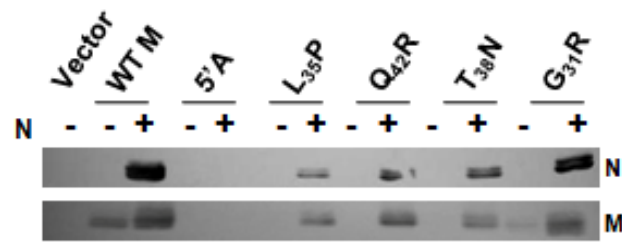
mutant M proteins produced ~20-30% more VLPs when coexpressed with N, even though the 5'A M was not capable of forming VLPs with or without N. Except for the 5'A + $G_{31}R$ M, the other increases represented changes from undetectable levels in the absence of N. This data suggests that the TM1 changes are allowing for more efficient envelope assembly when the 5'A mutations are present. The results support our initial observation that the N protein is likely helping to somehow stabilize the VLP particle during assembly with the mutant M proteins (Fig.12).

Analysis of M-M interactions. The M protein constitutes the bulk of the viral envelope. Since VLPs assemble for most coronaviruses when only M and E are coexpressed, the particles are presumably representative of the viral envelope at the most fundamental level. Thus, we reasoned that failure to form MHV VLPs might be due to a lack of or a decrease in M-M interactions. To test this idea, we asked if the mutant M proteins could be rescued into VLPs by WT M. We constructed a M protein called A2A3 that was previously described (43). The serine residues at positions 2 and 3 were replaced by alanines in A2A3 M. This destroys the epitope that is recognized by the monoclonal antibody J1.3, yet the protein localizes correctly and supports VLP assembly like wild-type M (43, 45). Our mutant M proteins were coexpressed with A2A3 M and wild-type E. A lower concentration of mutant plasmid DNA was transfected to assure efficient VLP formation as previously determined (43). Cells were metabolically labeled with Expre^{35}S^{35}S labeling mixture at 4-20 h after transfection. Intracellular and

extracellular fractions were divided into two equal aliquots and immunoprecipitated with monoclonal antibody J1.3 or an anti-MHV polyclonal antibody, F88. The ability of the A2A3 M protein to support VLP production and antibody recognition by F88, but not J1.3, was confirmed when the protein was coexpressed alone with WT E (Fig.14, lanes 5-6). When the 5'A or 3'A mutant M proteins were expressed in combination with A2A3 and E, both M proteins were detected in the intracellular fractions. However, neither the 5'A nor 3'A mutant M proteins could be rescued into VLPs along with the A2A3 protein (Fig.14, lanes 7-10). As a positive control and to ensure that the lack of detection of the 5'A and 3'A proteins was not the result of transfection of less plasmid, WT M was coexpressed with A2A3 M and E. The WT protein was clearly incorporated into VLPs (Fig.14, lanes 11-12). The M mutants $P_{120}A$, $E_{121}Q$, and $E_{121}L$ that supported VLP production when expressed alone with E were all also rescued into VLPs (data not shown). These data demonstrate that the 5'A and 3'A proteins are likely impaired in their ability to interact properly with other M molecules since they cannot be rescued into VLPs.

Additionally, the rescue experiment described above was carried out with the 5'A and one of the 5'A + TM1 mutants M proteins with the recovered $G_{31}R$ change. When expressed with A2A3 and E both mutant proteins were detected in the intracellular fractions with J1.3 and F88 antibodies (Fig.14, lanes 13-16). However, only the 5'A + $G_{31}R$ mutant could be rescued into VLP particles. This data further supports the conclusion that the second site TM1 changes are
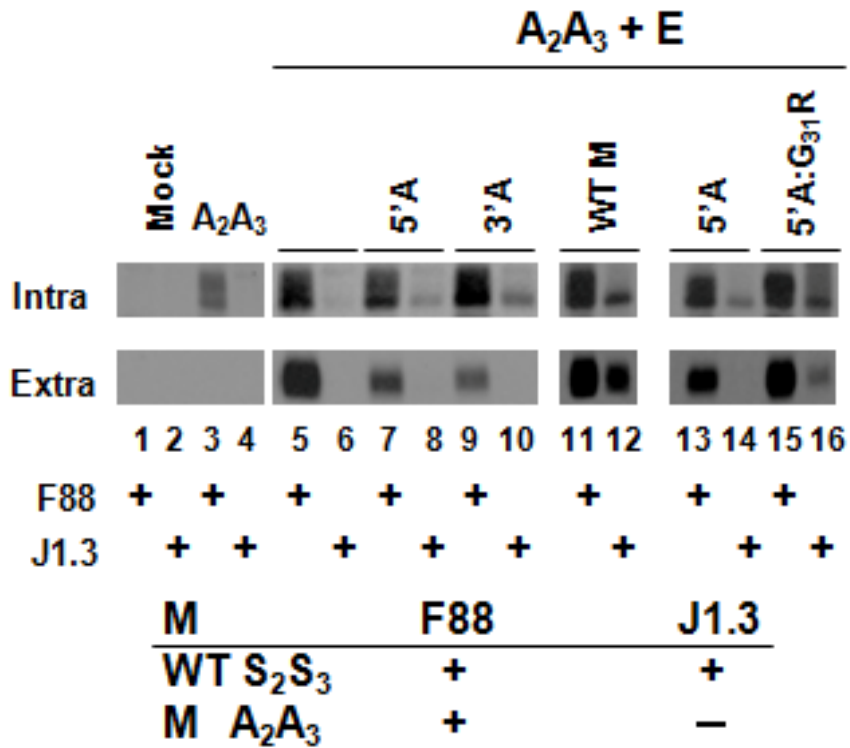
FIG.14. Rescue of CD mutants into VLPs. 293T cells were transfected with the pCAGGS vector containing WT or mutant M proteins as indicated. Cells were labeled for 20 h at 4 h after transfection with Expre^{35}S^{35}S labeling mixture. Intracellular lysates (Intra) and extracellular media (Extra) were divided in half and immunoprecipitated with monoclonal J1.3 or polyclonal F88 antibodies that recognize only WT and both WT and A_2A_3, respectively. Samples were analyzed by SDS-PAGE and autoradiography.

providing some advantage for the 5'A mutant proteins by allowing for more efficient M-M interactions.

        Analysis of M-S colocalization. In addition to M-M interactions, M also interacts with the S protein and helps retain it in the Golgi/ERGIC where viruses assemble (139, 143). WT M protein localizes to the Golgi when expressed alone. S localizes along the exocytic pathway and at the plasma membrane when singly expressed. To further analyze how the CD mutations might be affecting the functions of M, the WT and CD mutant proteins were coexpressed with S (Fig.15). When WT M and S proteins were expressed together, the latter colocalized with M. The S protein also colocalized with the 5' A and 3'A mutant M proteins (Fig.15). S localized to a large extent with the 5'A + 3'A and ΔCD protein that were retained in the ER (data not shown). These results suggest that the crippling effect of the 5'A and 3'A changes on the virus is not due to their failure to retain the S protein at the site of assembly.
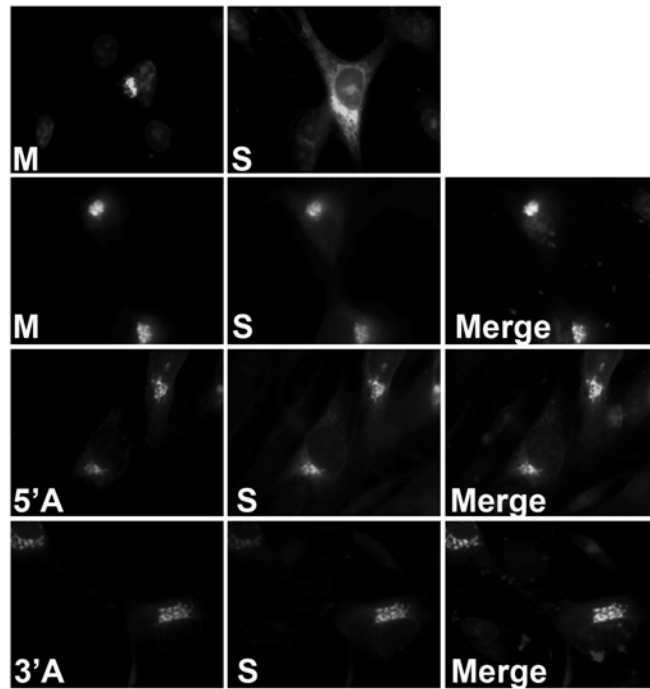
FIG.15. S protein colocalization with WT and mutant M proteins. 293T cells were transfected with pCAGGS vectors containing WT or mutant M proteins and the S gene. Cells were fixed at 12 h after transfection and analyzed by immunoflourescence using mouse and goat antibodies against the M and S proteins, respectively. Alexa Fluor 488-conjugated mouse and Alexa Fluor 594-conjugated goat secondary antibodies were used to visualize the localization of the M and S proteins, respectively. Nuclei were stained with DAPI. Singly expressed M and S proteins are shown in the top two panels. Colocalizations of the M and S proteins are shown in the merged images on the right.

DISCUSSION

The significance of the evolutionally conserved domain (CD) located at the amino end of the long amphipathic region in the coronavirus M carboxy tail was examined directly for the first time in this study. Using MHV A59 as the model virus we show that the CD (SWWSFNPETNNL) is important for M to function during both VLP and virus assembly. A positive charge substitution for the conserved negative charge was not tolerated. Residues in either the amino or carboxyl ends were very sensitive to changes, since recovered viruses exhibited crippled phenotypes, with 100 to 1000 fold reduction in virus yields. Removal of the conserved proline must not alter the domain since assembly of VLP and virus was not affected. The majority of the introduced mutations were tolerated, accompanied by second site changes in some cases, in the context of the virus, but most did not support VLP production. We conclude that the CD likely contributes to lateral M-M interactions during envelope formation. Our results strongly indicate that the N protein helps stabilize envelope complex formation, even in the absence of the viral genome. The study overall provides new insight into requirements of the key virion structural component and demonstrates the functional significance of the CD that appears to be involved in crucial protein-protein interactions that must take place for the fundamental process of envelope formation during virus assembly.

Viruses with the positive charge substitutions were selected against after only one virus passage, indicating strong intolerance for charge reversal at this

67

position. Since the $E_{121}R$ and $E_{121}K$ mutant proteins were not able to form VLPs, whereas the recovered $E_{121}Q$ and $E_{121}L$ changes in M were competent, this provides strong support for the significance of the CD in the fundamental process of envelope formation. Some recovered viruses with the Q and L changes at position 121 also had second site changes in the N protein (Table 1). We expected these to be important, but interestingly, viruses with only the recovered Q and R changes exhibited growth characteristics similar to the WT virus, suggesting that the changes in N are not providing additional advantage for the virus. The changes were located primarily toward the amino and carboxyl ends of the N protein. No structure has been determined yet for MHV N proteins, but recent nuclear magnetic resonance and crystal structure have been determined for parts of IBV and SARS-CoV N proteins (28, 59, 81, 86, 160, 183, 218). We did examine where our second site changes align and likely map based on the available structures. The majority would not fall within structured regions such as beta sheets or helices, which may reflect the fact that loss of the introduced positive charges appears to be the major contributor to the WT-like phenotype of the recovered viruses.

The 5'A viruses with changes in the amino end of the CD were particularly interesting. All of the recovered viruses had second site changes in the first transmembrane (TM1) of M. The TM1 changes were not fully compensating since the amount of virus output was significantly reduced compared with WT. Nonetheless, in both the context of envelope (VLP) and

virus assembly the TM1 changes are apparently providing some advantage for the mutant M proteins containing the 5'A mutation. The results suggest a possible interaction between the CD and first TM domain, which is consistent with the apparent intricate association of the amphipathic domain with the inner leaflet of the virion membrane. Tryptophan residues positioned at the cytoplasmic boundary have been suggested to possibly help stabilize TM helices or to provide vertical mobility relative to the lipid bilayer (102). Alternatively, it is possible that the tryptophan residues help mediate M-M or other important M interactions and that the second site changes in TM1 promote new interactions that compensate for loss of the SWWS residues. Tryptophan residues can promote protein-protein interactions. For example, hepatitis B virus small surface antigen (S-HBsAg) is able to assemble hepatitis delta virus (HDV) particles through interactions with HDV ribonucleoprotein (RNP) (198). Three W residues located in a small loop between two predicted S-HBsAg transmembrane domains mediate its interaction with the large HDV antigen, a component of the HDV RNP (92). Alanine substitutions prevent packaging of the RNP by the S-HBsAg.

Lateral interactions between M molecules are fundamental for organization of the coronavirus envelope. M-M interactions are thought to be mediated through multiple contact points, but the TMs appear to be especially important (47). Recent cryo-EM and cryo-EM tomography analysis of single virus particles provided new insight about the 3D structure of coronaviruses (7, 138). The tomography analysis revealed that the envelope is striated, presumably

69

due to M-M interactions (7). Based on the predicted size of the M carboxy tail and assumption that it is globular, it was suggested that at most, tetramers constitute the observed striations (7). Recent cryo-EM analysis of MHV particles suggests that the envelope lattice consists of M homodimers mediated by the globular amphipathic region of the M protein carboxy tail since the only contact points observed were between M densities in the endodomain (B. W. Neuman et al., submitted for publication). The structure of the amphipathic domain is not known at this time, but this part of the protein is largely protease resistant in virions, which suggests that it is very closely associated with the inner membrane, possibly entirely or partially embedded in the membrane (155, 156). Virions examined by cryo-electron tomography exhibited trilaminar-like or unusually thick membranes in the envelope, which was attributed to close association of the carboxy tail with the inner leaflet of the membrane bilayer (7). Since the CD is located at the amino end of the amphipathic domain, it may influence the interactions of the tail with the inner lipid bilayer, either locally or more distantly influencing positioning of the remainder of the tail relative to the inner leaflet of the membrane. This could affect M oligomerization and/or the assumed matrix-like function of the tail, which may account for the impact on virus assembly and the second site changes identified in our recovered viruses.

Our results strongly suggest that the N protein helps stabilize envelope assembly complexes during VLP assembly, most likely through interactions with M. We have some understanding about the role that N plays in the context of full

virus assembly, as part of the nucleocapsid, but what role might it play during

VLP assembly? Stabilization during envelope assembly, which in the case of

VLPs consists almost entirely of M, may be particularly important for our mutant

M proteins if indeed positioning of M molecules in the lattice framework and/or

the tail is altered. If so, then M-N interactions may help stabilize M in a

conformation that allows for more efficient assembly of the envelope. The N--M

interactions are also likely significant during complete virus assembly. This is

consistent with the thread-like densities previously observed in cryo-EM and

cryo-electron tomography images which appear to provide linkage points between

M and nucleocapsids in MHV particles (7, 138). Recent genetic analysis

demonstrated that important interactions occur between the carboxy ends of the N

and M proteins (83, 95, 195, 196). Tight interaction between the carboxy-terminal

regions of MHV N (amino acids 380-454) was also recently shown to play a role

in helping mediate N-N interactions (82). It was suggested that the domain of

each N molecule associates with either another N molecule or the carboxy

terminus of an M molecule in virions. This suggests that M-N/nucleocapsid

interactions may be important for optimal envelope formation. The stabilizing

effect of N may be of greater importance and more apparent with the mutant M

proteins we describe here, than during its normal role when expressed with WT

M.

The 5'A+3'A and ΔCD M proteins were not able to support virus

assembly. Both proteins localized in the ER, whereas replacement of only the

amino or carboxy residues localized like WT M in the Golgi. This suggests that the structure of the 5'A + 3'A and ΔCD proteins is altered or that a localization signal was disrupted. A MHV M deletion mutant lacking residues L_{108} through T_{112} (ΔLT) that are adjacent to the CD was also retained in the ER, but when residues E_{121} to D_{195} (ΔC) were removed the protein was retained in the Golgi (47). A specific localization signal has not been identified for MHV M. However, the cytoplasmic tail has been shown to be essential, but not sufficient for Golgi localization (117). Altogether, the results suggest that the region just beyond the third TM is important for Golgi retention. Interestingly, truncated SARS-CoV M consisting of the amino-terminal 134 residues, which includes the CD (S_{107}-L_{118}), is retained in the Golgi (197).

A panel of MHV M mutants was studied previously that provided important insight into which regions of M are involved in envelope assembly (43, 46, 47). Three deletion mutants from the earlier studies should be noted here. Amino acids 121-195 (ΔC) that constitutes a large portion of the amphipathic domain were deleted from MHV M. This deletion includes the last five residues (ETNNL) of the CD (Fig.8). Like our 3'A mutant M with alanine substitutions for TNNL residues, the ΔC M was not competent for VLP formation, but unlike our 3'A M, the protein was rescued into particles when coexpressed with wild-type M (47). Residues 108 to 112 just upstream from the CD were also deleted, but the protein could still be incorporated into VLPs when coexpressed with A2A3 M (47). Finally, M proteins with deletion of only I_{110} failed to participate

72

in VLP assembly (46). It is possible that the M proteins with these deletions might be able to support virus assembly, but they were not analyzed in this context. The amino acid substitutions in the present study are less likely to have as significant of an impact on the overall structure of the M tail as are residue deletions.

Through interaction with M, S is retained at assembly sites in the ERGIC region of the cell (139, 143). The 5'A and 3'A M proteins retained S in the Golgi, indicating that the alanine substitutions for SWWS or TNNL residues are not affecting M-S interactions. The amphipathic domain of MHV M was previously implicated to be involved in mediating interaction with S since the ΔC mutant described above did not retain S in the Golgi (46). A single $Y_{211}G$ change that is significantly distant from the CD has also been implicated to play a role in M-S interactions (46). It is interesting that the amino-terminal 134 amino acids of SARS-CoV are sufficient for retention of S in the Golgi (197). A single tyrosine (Y_{195}) in SARS-CoV M was recently shown to be necessary to retain S in the Golgi (129). Further mapping will be required to more precisely identify requirements for MHV M-S interactions. It is clear from the present study that changes in the M CD that do not affect localization of the proteins are still competent to mediate co-localization.

Overall, our results suggest that the CD functions in formation of the viral envelope by helping mediate lateral interactions between M molecules during virus assembly. The N protein clearly enhances or enables envelope formation,

which is likely reflected in the ability to recover virus when VLP assembly was compromised. It is likely that N helps stabilizes M complexes, possibly helping mediate a conformation in the tail that is important for efficient virus assembly. We know that interactions between the hydrophilic carboxy end of the M tail and nucleocapsids are important for encapsidation, a critical step in production of infectious virus (58, 95, 134, 143, 195, 196). Results from the present study indicate that the opposite end of the carboxy tail plays a complementary role during assembly, likely during organization of the envelope that subsequently encapsidates the nucleocapsid. Additional studies will be directed at understanding mechanistically how the second site changes in TM1 help partially compensate for the 5'A changes and how N provides the presumed stabilizing effect through interactions with M.

CHAPTER 3

ROLE OF CHARGED RESIDUES IN THE CORONAVIRUS MEMBRANE

PROTEIN CARBOXY TAIL IN VIRUS ASSEMBLY AND INFECTIVITY

ABSTRACT

The coronavirus (CoV) membrane (M) protein plays a key role in virus assembly. The terminal 25 amino acids of the long carboxy tail of M are hydrophilic and highly charged. This region plays an important role in mediating interaction with the nucleocapsid (N) protein. A panel of mutant viruses was generated with various single and double residue charge changes in the domain (VYVK$_{205}$SK$_{207}$VGNYR$_{212}$LPSNK$_{217}$PSGADTALLR$_{227}$T) of mouse hepatitis coronavirus A59 (MHV-A59). R$_{227}$ was shown in previous studies to be critically important for assembly, but other charges were not analyzed. All of the mutant M proteins were analyzed in the context of a MHV infectious clone and also in transient expression studies. Replacement of K$_{205}$ and R$_{212}$ individually or K$_{205}$ and K$_{207}$ or R$_{212}$ and K$_{217}$ in combination with a negatively charged aspartic acid (D) resulted in non-viable viruses and no production of virus-like particles (VLPs). Mutant viruses were recovered with K$_{205}$A, K$_{207}$A, K$_{217}$A, and K$_{217}$D charges that exhibited wild-type (WT) phenotypes. Viruses harboring K$_{207}$D or R$_{212}$A changes had severely crippled phenotypes and could not form VLPs in the presence of the envelope (E) and N proteins. Additionally, these mutants displayed a decreased infectivity. These results identify a cluster of three positive charges in the 25-amino acid hydrophilic domain of the M tail that are functionally important for virus assembly, as well as infectivity.

INTRODUCTION

Coronaviruses cause respiratory, enteric, and neurological disease in humans as well as several animals. The viruses are responsible for many infections in domesticated animals, including swine and chickens, and infection outbreaks can have a significant economic impact globally. In humans most coronaviruses cause mild upper respiratory tract infections. The emergence of the severe acute respiratory syndrome coronavirus (SARS-CoV) in 2003 demonstrated that coronaviruses can cause severe disease. Since 2003, two new human coronaviruses, NL63 and HKU1, have been identified (193, 207). Due to the high pathogenesis of SARS-CoV and the threat of re-emergence, research on coronaviruses and studies directed toward the generation of antivirals is strongly justified.

Coronaviruses are enveloped RNA viruses with a genome of approximately 30kb. This is the largest RNA genome known for RNA viruses. The viruses assemble and bud at intracellular membranes in the endoplasmic reticulum Golgi intermediate compartment (ERGIC) (93, 188). All coronavirus envelopes contain at least three structural proteins, the spike (S), envelope (E), and membrane (M) proteins. Some coronavirus envelopes also contain the hemagglutinin esterase (HA) protein. The genomic RNA is bound by the nucleocapsid (N) phosphoprotein and forms a helical structure. The S protein is the receptor binding protein and forms large spikes on the surface of the viral envelope (66). S induces neutralizing antibodies and induces cell mediated

77

immunity (66). The E protein, although present in low abundance in the viral envelope, plays an important role during assembly and release of coronaviruses (14, 35, 194). The M protein is the most abundant protein in the viral envelope. It plays absolutely key roles in virus assembly (47, 91, 139, 143, 194).

All coronavirus M proteins have a short amino terminus domain located outside the virion, followed by three transmembrane (TM) domains, and then a long carboxy tail extending into the virion (reviewed in (77) (Fig.16). The M protein organizes the assembly process by participating in M-M as well as M-S and M-Nucleocapsid interactions (46, 47, 58, 95, 139, 143). These fundamental interactions drive the efficient assembly of viral particles. For most coronaviruses, co-expression of the M and E proteins alone drives virus-like particle (VLP) assembly (14, 194).

The long carboxy tail of coronavirus M proteins contains an amphipathic domain of about 100 residues, followed by a hydrophilic highly charged tail of approximately 25 residues (Fig.16). M proteins of all coronaviruses contain several charges within their extreme carboxy tail (Fig.16). The charged region of the tail has been shown by genetic studies to be important for mediating interaction with the N protein (83, 95, 195, 196). Furthermore, analysis of M and N protein mutants has suggested the most C-terminal positive charge (R_{227}) on MHV M is particularly important for assembly but other residues appear to be involved as well (43, 95, 196). The studies points to the importance of charge-

charge interactions between these proteins that may contribute to the correct overall structure of M necessary during assembly.

The M protein also interacts with the S protein (139, 143). M and S protein heterotypic interactions allow specific S incorporation into virions (139, 143). S is absolutely necessary for infection since it is the viral receptor attachment protein and is responsible for cell-to-cell fusion (23). S incorporation into the viral envelope has also been shown to be mediated by charged residues in the endodomain of S (211).

Due to the key role M plays in the assembly process and the conservation of charged residues within the carboxy tail, we set out to determine the role of positively charged residues within the hydrophilic domain of the MHV M carboxy tail. A series of single and double changes were introduced into the carboxy tail of M. The functional impact of these changes was analyzed in the context of the virus and by several transient expression assays. The results show that a cluster of three positive charges (K_{205}, K_{207} and R_{212}) in the domain are important for the function of the M protein. Some mutants were not competent for VLP assembly and inclusion of the N protein did not increase VLP output. Importantly, loss of the positive charge at K_{207} and R_{212} results in a decrease in infectivity of viral particles. The results suggest charges within the hydrophilic tail of M are involved in assembly of infectious particles.

MATERIALS AND METHODS

Cells and viruses. Mouse L2 and 17 clone1 (17Cl1) cells were maintained in Dulbecco's modified Eagle's medium (DMEM) containing 10% and 5% heat-inactivated fetal calf serum (FCS), respectively, plus penicillin, streptomycin, and glutamine. Baby hamster kidney cells expressing the MHV Bgp1a receptor (BHK-MHVR) (216) and BHK-21 cells (ATCC) were grown in Glasgow minimal essential medium (GMEM) containing 5% FCS, supplemented with antibiotics, glutamine and 10% tryptose phosphate broth. BHK-MHVR cells were maintained in the presence of Geneticin (G418) for continuous selection of cells expressing the receptor (216). Human 293T cells (ATCC) were grown on 0.25% gelatin coated flasks in DMEM containing 10% FCS, antibiotics and glutamine. Stocks of wild-type MHV A59 and cloned viruses were grown in mouse 17Cl1 or L2 cells. Virus titers were determined in L2 cells.

Construction of amino acid substitution and deletion mutants. pGEM-5Zf(+)M-N, a pGEM5Zf(+) vector (Promega) containing the M and N genes (EcoRV-SacI fragment), was used for mutagenesis (195). Site-directed amino acid substitutions were introduced by whole-plasmid PCR using appropriate primers and platinum Pfx DNA polymerase (Invitrogen). Mutations were confirmed by sequence analysis of the entire M-N insert before subcloning into the MHV G clone at the BssHII and NheI sites (216).

Reverse genetics. All mutant viruses were made by reverse genetics using a MHV A59 infectious clone (216). Full-length cDNA clones were assembled,

transcribed, and electroporated into BHK-MHVR cells as described previously
(195). Following electroporation BHK-MHVR cells were seeded concurrently
with L2 cells. At 24-46 h after electroporation media were harvested and an
aliquot was used to infect 17Cl1 or L2 cells. The medium was removed from
infected cells at 24 to 72 h p.i. Total cytoplasmic RNA from cells remaining on
the flasks was extracted using RNAqueous-4PCR extraction buffers (Ambion)
and treated with DNase before being reverse transcribed using the Superscript
RT-PCR system from Invitrogen and an oligo (dT) primer. The RT product was
subjected to 30 cycles of PCR amplification using SuperTaq Plus (Ambion) and
appropriate primers to amplify the E, M, N and S genes. PCR products were
cleaned with MiniElute PCR (Qiagen) before being sequenced directly.
Viruses were subsequently plaque purified from media taken directly off cells that
had been electroporated. Multiple plaque isolates were passaged on 17Cl1 or L2
cells five times for all viable viruses. In some cases where viruses grew poorly
and yielded small plaques from which it was difficult to recover isolates, the stock
off electroporated cells was passaged five times to allow for selection of the most
fit viruses. RNA was extracted from infected cells following passage of plaque
purified viruses for RT-PCR and sequence confirmation of the E, M, N,
endodomain of S genes and the packaging signal to determine the stability of
introduced mutations and identify potential compensating changes.

Virus growth properties. Growth kinetic experiments were carried out in mouse
17Cl1 cells infected at a multiplicity of infection (MOI) of 0.01 or 0.001 PFU/cell

with passage 5 virus stocks. Cell culture supernatants were collected at indicated times after infection and virus titers were determined by plaque assay on mouse L2 cells. Low melting agarose/medium overlays were removed at 72-96 h p.i. Cells were fixed and stained with crystal violet in ethanol to visualize plaques. Two independent growth kinetic experiments were performed for each set of mutant viruses in parallel with the wild-type virus. Growth kinetic curves represent two-part exponential growth analysis. The initial rapid growth, followed by slower growth (or declined) after 8-12 hpi, was analyzed. Doubling times were estimated for the initial rapid growth slopes. Relative error was assumed to be the same for all time points for each data set. Error estimate was computed by taking the standard deviation divided by mean for each of the time points and averaging over all points.

Mfold analysis. A region of 622 nucleotides, encompassing the 190-nt PS region and the sites of the recovered mutations for the $K_{217}D$ and $R_{21}2A$ (20025-20648) was analyzed using Mfold (226). 14, 18 or 17 structures were predicted for WT, $K_{217}D$, and $R_{212}A$, respectively, will varying ΔG values. Only structures with the top five lowest ΔG values were chosen for further analysis. Any structures within these top five that did not contain a similar structure to the previously described 190-nt region were eliminated. For WT, $K_{217}D$ and $R_{212}A$, three, four or four structures, respectively, remained. Analysis was carried out using these structures that remained.

VLP analysis. Wild-type and mutant M genes were expressed in the pCAGGS

vector under the control of the chicken β-actin promoter for transient expression

(140). A Kozak sequence was included in the forward primer for all genes. All

mutant M genes were shuttled into the pCAGGS vector from the MHV G

fragment used to generate full-length MHV genomic cDNAs for virus

construction. Wild-type E and N genes were also expressed in the pCAGGS

vector.

293T cells were transfected with the pCAGGS plasmids containing the

wild-type or mutant M genes, E gene, and in some cases the N gene, using the

TransIT-293 transfection reagent (Mirus). At ~24 h after transfection, the culture

medium and intracellular cytoplasmic fraction were harvested. Cells were lysed

on ice with 0.5% Triton X-100, 50 mM NaCl and 50 mM Tris-HCl (pH 7.5) in the

presence of 1mM phenylmethylsulfonyl fluoride (PMSF). The medium was

clarified at 14,000 x g for 10 min at 4°C. VLPs from clarified media were

pelleted through a 30% sucrose cushion by ultracentrifugation for 3 h at 4°C in a

Beckman SW55Ti rotor at 30,000 rpm. Pellets were resuspended in Laemmli

sample loading buffer and analyzed by SDS-PAGE. Proteins were transferred to

polyvinylidene difluoride (PVDF) membranes and analyzed with anti-MHV M

9246 (generated in our lab), anti-MHV E 9410 and anti-MHV N antibodies

generated in our lab (33, 118). Following incubation with appropriate secondary

antibodies, blots were visualized by chemiluminescence (Pierce). VLP release

was quantified by densitometric scanning of fluorograms and analyzed using Image-Quant software (Molecular Dynamics).

Infectivity Analysis. Each virus was pelleted through a 30% sucrose cushion and resuspended in HEPES Buffer (50mM HEPES, 100mM NaCl, 0.01% BSA pH 7.4). A fraction was used for a plaque assay to determine the total PFU obtained for each virus preparation. An equal number of PFUs based on the titer for each virus was run on a 5-20% SDS-PAGE gradient gel to visualize the particle protein profile. A WT standard of known amounts of PFUs was run in parallel (10^5 to 10^7 PFU). Proteins were transferred to polyvinylidene difluoride (PVDF) membranes and analyzed with anti-MHV M 9246 (generated in our lab), anti-MHV N (33) and anti-MHV S A04 (kindly provided by Kathryn Holmes, University of Colorado Health Sciences) antibodies. Following incubation with appropriate secondary antibodies, blots were visualized by chemiluminescence (Pierce). The N protein from images was quantified by densitometric scanning of the films and analyzed using ImageQuant software (Molecular Dynamics). The intensity of the mutant N proteins was compared to the WT standards to determine the relative amount of N protein for each.

RNA extraction and quantitation. To determine the amount of RNA present in the virus particles, slot blot hybridization was carried out. RNA was extracted from virus particles from a fraction of the same sample that was spun through the sucrose cushion (above) using Trizol LS reagent (Invitrogen) according to the manufacturer's instructions. RNA was also extracted from known amount of WT

particles (from 10^5 to 10^7 PFUs). 20μg Yeast RNA (Ambion) was added to each sample during the extraction to aid in RNA precipitation. The RNA pellets were resuspended in RNAase-free water and subsequently denatured with a formamide, formaldehyde and 3-(N-morpholino)-propanesulfonic acid pH 7.0 (MOPS) denaturing buffer at 65°C for 15 minutes. RNA was then blotted onto nitrocellulose charged membrane using a Vacu-slot-VS blotting apparatus (American Bionetics Inc.). The RNA was crosslinked to the membrane using a Stratalinker UV crosslinker (Strategene) and probed with a digoxigenin-labeled 357-nucleotide MHV-A59 N gene probe. The probe was transcribed using reagents for digoxigenin labeling (Roche Applied Science) according to the manufacturer's directions. To detect the DIG-labeled RNA, a DIG luminescent detection kit was used (Roche). Images were quantified by densitometric scanning of the films and analyzed using ImageQuant software (Molecular Dynamics). The intensity of the mutant samples was compared to the WT standards to determine the relative amount of RNA for each.

Indirect immunofluorescence. Co-localization of mutant M genes with WT S was determined by expression of pCAGGS plasmids containing WT or mutant M and WT S genes in BHK-21 cells. Cells were plated on Lab-teck chamber slides (Nunc) 1 day prior to transfection with MirusTrans-293T reagent (Mirus). At 12 h after transfection cells were washed two times with PBS and fixed in methanol for 15 min at -20°C. Following one additional wash with PBS, cells were blocked with 2% gelatin in PBS overnight at 4° C. Monoclonal antibodies J1.3 and 2.7

(62) and a polyclonal antibody, A04 (kindly provided by Kathryn Holmes, University of Colorado Health Sciences) were used to detect the M and S proteins, respectively. Slides were incubated for 2 h with primary antibodies at room temperature, washed with 2% gelatin in PBS before incubation with Alexa-labeled secondary antibodies. Cells were washed with 2% gelatin in PBS and then a final wash with PBS alone. Slides were mounted in ProLong Gold antifade reagent (Molecular Probes) containing 4,6-diamino-2phenylindole (DAPI) to stain nuclei. Images were collected by using an epifluorescence Nikon inverted microscope (Nikon, Inc., Melville, NY) with MetaMorph imaging soft-ware (Universal Imaging Corp., Downingtown, PA). Images were processed using Adobe Photoshop.

RESULTS

Construction of charged residue mutants. To study the functional significance

and role of conserved charged residues within the extreme end of the M protein

carboxy tail a panel of mutants were generated. Site directed mutagenesis was

used to change charged residues in the context of the full length MHV-A59

infectious clone. Four positive charges, K_{205}, K_{207}, R_{212}, and K_{217}, that are

separated from the functionally important, R_{227} by one negative charge, were

mutated in the MHV genome (Fig.16).

Recovered viruses were plaque purified and passaged multiple times.

Sequence analysis was used to monitor genetic stability of the introduced

mutation and identify any potential secondary compensatory changes. Because the

M protein interacts with the other viral proteins, the M, N, and E genes, as well as

the carboxy end of the S gene were sequenced (46, 47, 58, 95, 139, 143).

Additionally, the packaging signal in gene 1b was sequenced because the M

protein has been shown to associate with this region (135). Mutants were

analyzed for their growth properties to determine the impact of genetic changes

directly on assembly. Selected mutant M proteins were also analyzed for their

ability to support VLP production by co-expression with the E gene, to determine

the possible role of the charged residues in the fundamental process of envelope

formation. In some cases the N gene was included in VLP analyses, as the N

protein appears to enhance VLP assembly of SARS-CoV and MHV VLPs (5, 172,

200). Additionally, the infectivity of viral particles was analyzed for selected

87

FIG.16. Summary of charged tail mutant viruses. A linear schematic of the M protein is shown illustrating the relative positions of the three TM (black boxes) and the long carboxy tail. An alignment of the hydrophilic cytoplasmic tail is shown for below for MHV, bovine coronavirus (BCoV), human coronavirus OC43, SARS-CoV, transmissible gastroenteritis virus (TGEV), human coronavirus 229E, and infectious bronchitis virus (IBV) that was generated using CLUSTAL W (103). Below, the amino acid substitutions that were introduced into a full-length MHV-A59 infectious clone are indicated below the WT sequence and the corresponding names are indicated to the left. + and – symbols indicate virus recovery, growth and VLP production (+E and +E,N) for each mutant. For growth, two or one + signs indicate growth comparable to WT or crippled growth, respectively. For VLPs:+E,N, the designation of + and +/- indicates VLP release of above 20% and less than about 15%, respectively. n/d indicates no data.

NH₂ ▬▬ COOH

```
MHV    201 AVYVKSKVGNYRLPSNKPSGADTALLRT     228
BCoV   200 AVYVKSKVGNYRLPSTQKGSGMDTALLRNNI   230
OC43   200 AVYVKSKVGNYRLPSTQKGSGMDTALLRNNI   230
SARS   193 AAYNRYIGNYKLNTDHAGSNDNIALLVQ      221

TGEV   233 AYYVKSKAGDYSTEARTDNLSEQEKLLHMV    262
229E   195 VFYVRVKHGDFSAVSSPMSNMTENERLLHFF   225

IBV    199 ATFVYAKQSVDTGELESVAIGGSSLYT       225
```

MHV   VKSKVGNYRLPSNKPSGADTALLRT

MHV					Virus Recovered	Growth	VLPs:+E	VLPs:+E,N
K₂₀₅AK₂₀₇A	A	A			+	+	−	−
K₂₀₅DK₂₀₇D	D	D			−	−	−	n/a
R₂₁₂AK₂₁₇A			A	A	+	+	−	+/−
R₂₁₂DK₂₁₇D			D	D	−	−	−	n/a
K₂₀₅A	A				+	++	−	−
K₂₀₇A		A			+	++	+	+
R₂₁₂A			A		+	+	−	+/−
K₂₁₇A				A	+	++	+	+
K₂₀₅D	D				−	−	−	n/a
K₂₀₇D		D			+	+	−	+/−
R₂₁₂D			D		−	−	−	n/a
K₂₁₇D				D	+	++	+	+

89

mutants.

Double mutations are lethal or crippling. First, double mutations were introduced in the four targeted residues in pairs. Positively charged residues were replaced with a neutrally charge alanine (A) or negatively charged, aspartic acid (D) to determine if the charge might be playing a role in electrostatic interactions that mediate protein-protein or protein-nucleic acid interactions. The double mutants were designated $K_{205}AK_{207}A$, $K_{205}DK_{207}D$, $R_{212}AK_{217}A$, and $R_{212}DK_{217}D$ (Fig.16).

The double mutations were introduced into the MHV-A59 infectious clone in parallel with a WT virus control. The double neutral mutants, $K_{205}AK_{207}A$ and $R_{212}AK_{217}A$, were difficult to recover. Fusion, an indicator of viable virus, was only observed for the $R_{212}AK_{217}A$ mutant following electroporation, indicating genomic transcripts were replication competent. However, small plaques were observed for both double alanine mutants when grown on L2 cells (Fig.17C). Ten individual plaques for both mutants were passaged multiple times in an attempt to recover virus but only one $R_{212}AK_{217}A$ plaque was recovered. Therefore, medium from the electroporation was blind passaged directly onto L2 cells for both double alanine mutants. Sequence analysis of the blindly passaged mutants, as well as the one $R_{212}AK_{217}A$ isolate showed that the introduced double neutral mutation was retained with no additional changes in any of the other structural genes sequenced. The $K_{205}AK_{207}A$ and $R_{212}AK_{217}A$ mutants were analyzed for their growth properties along with the WT virus at a MOI of 0.001 (Fig.17A). Both
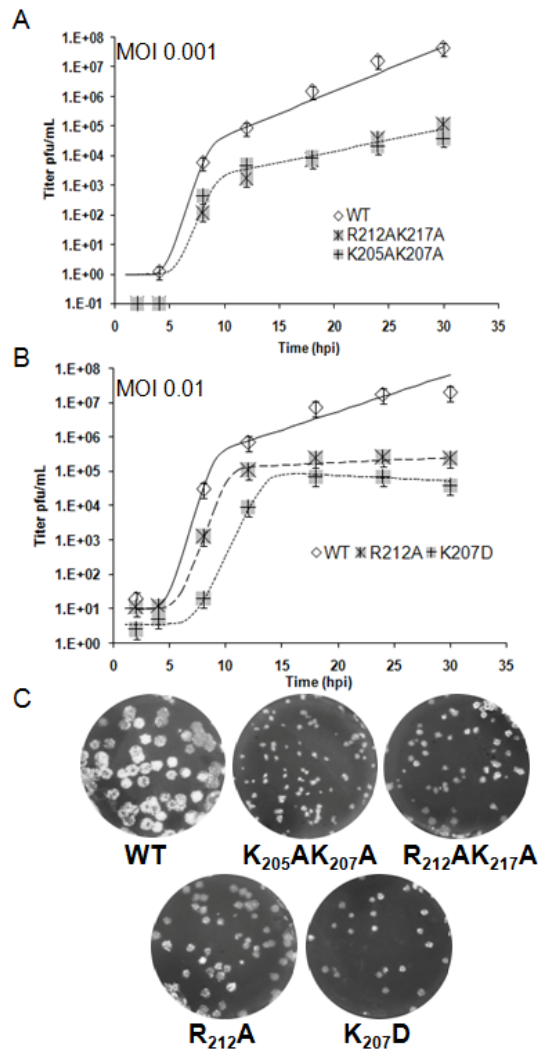
FIG.17. Growth properties of double and crippled single mutant viruses.

Mouse 17Cl1 cells were infected with mutant viruses at a MOI of 0.001

PFU/cell (A) or 0.01 PFU/cell (B). Titers were determined by plaque assay

on L2 cells at the indicated times. Error bars and exponential growth curves

are as described in the material and methods. Estimated doubling times were

~.27h for WT virus, ~.33h for $K_{205}AK_{207}A$, $R_{212}AK_{217}A$, and $R_{212}A$ viruses

and ~.44h for $K_{207}D$ virus. C) Plaque characteristics were determined for the

indicated viruses in L2 cells.

mutant viruses displayed a crippled phenotype, growing to titers 3 logs lower than WT virus. The negatively charged double mutants, $K_{205}DK_{207}D$ and $R_{212}DK_{217}D$, could not be recovered even after media taken directly off of electroporated cells was blindly passaged several times. This data strongly suggests that these mutations are lethal for the virus.

Neutral mutations that do not affect virus growth. Since viruses were sensitive to the introduction of paired charge changes, a panel of single charged residue mutants was generated to tease apart the significance of individual charges within the carboxy tail of M (Fig.16). When the $K_{205}A$ and $K_{207}A$ mutations were introduced, all four or six recovered viruses, respectively, contained the introduced mutation and no other changes. When these mutants were analyzed for their growth phenotype at an MOI of 0.01, they both grew similar to WT as they had peak titers of only 0.5-1 log lower than WT virus (Fig.18A and B). This data suggests that neutral charges at K_{205} and K_{207} are not detrimental for the virus and a positive charge is not required at these positions.

The K_{217} residue is most tolerant to changes. The K_{217} residue was changed singly to alanine (A) as well as aspartic acid (D). Virus was recovered readily with both mutations. Six isolated $K_{217}A$ plaque purified viruses maintained the introduced mutation with no additional changes and grew at a rate comparable to WT (Fig.18). Ten plaques were analyzed for the $K_{217}D$ mutation, all of which contained the introduced mutation. One of the plaques contained an additional change in the last residue of the M gene, a threonine (T) residue at position 228,

92
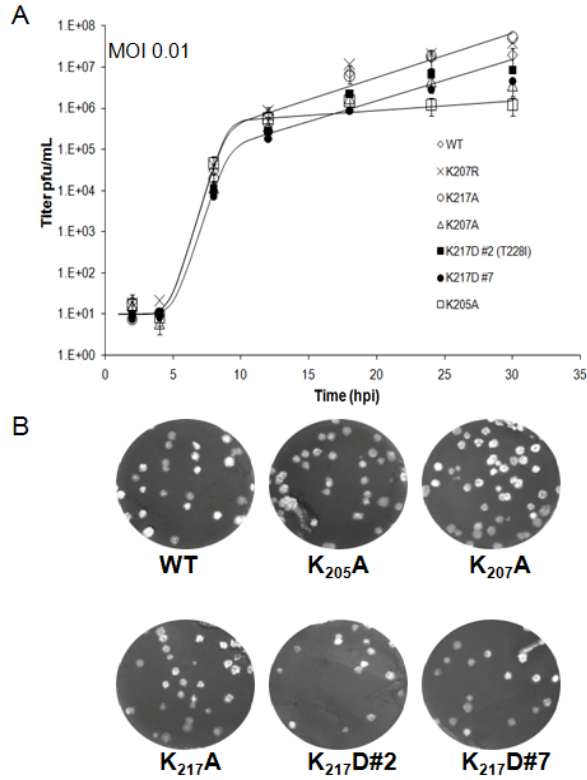
FIG.18. Growth properties of single substitution mutants. (A) Mouse 17Cl1 cells were infected with mutant viruses at a MOI of 0.01 PFU/cell (A). Titers were determined by plaque assay on L2 cells at the indicated times. Error bars and exponential growth curves are as described in the material and methods. Estimated doubling time was ~.28h for WT, $K_{207}R$, $K_{217}A$, and $K_{205}A$ viruses and ~.3h for $K_{207}A$, $K_{217}D\#2$ and #7 viruses. B) Plaque characteristics were determined for the indicated viruses in L2 cells.

to isoleucine (I). However, when growth analysis was carried out with one $K_{217}D$ mutant containing the $T_{228}I$ change and one without, both grew similarly to each other and to WT virus (Fig.18A), suggesting this change is not offering an advantage to the viruses. The data indicates that the positively charged K_{217} residue within the hydrophilic domain of M is not absolutely required.

Additionally, both isolated $K_{217}D$ mutants contained a change in the region near the packaging signal (PS) region. The MHV M protein has been implicated to bind to the PS in the absence of the N protein (135). For MHV originally a 190-nt RNA region, located in the nonstructural open reading frame (ORF)1ab, was identified to be important for packaging of the genomic RNA (64). Conserved 69-nt and 95-nt RNA stem loop structures were identified to be the crucial region, in the 190-nt area, involved in packaging the RNA (30, 64). The 190-nt, 95-nt and 69-nt regions are located at positions 20025-20158, 20064-20158 and 20099-201560 in ORF1ab, respectively. The change in the $K_{217}D$ mutants is at position 20546, an A to T change, which is beyond the conserved stem loop structures. Mfold analysis was done to compare the secondary RNA structure of WT compared to the $K_{217}D$ mutant (see materials and methods for criteria set for analysis) (226). The mutation in $K_{217}D$ viruses did not alter the PS loop structure in any prediction. There were some slight differences in structure in regions more distant from the PS in some predicted structures (Fig.19). However, whether these modest differences would actually influence the function of the PS will require further investigation.

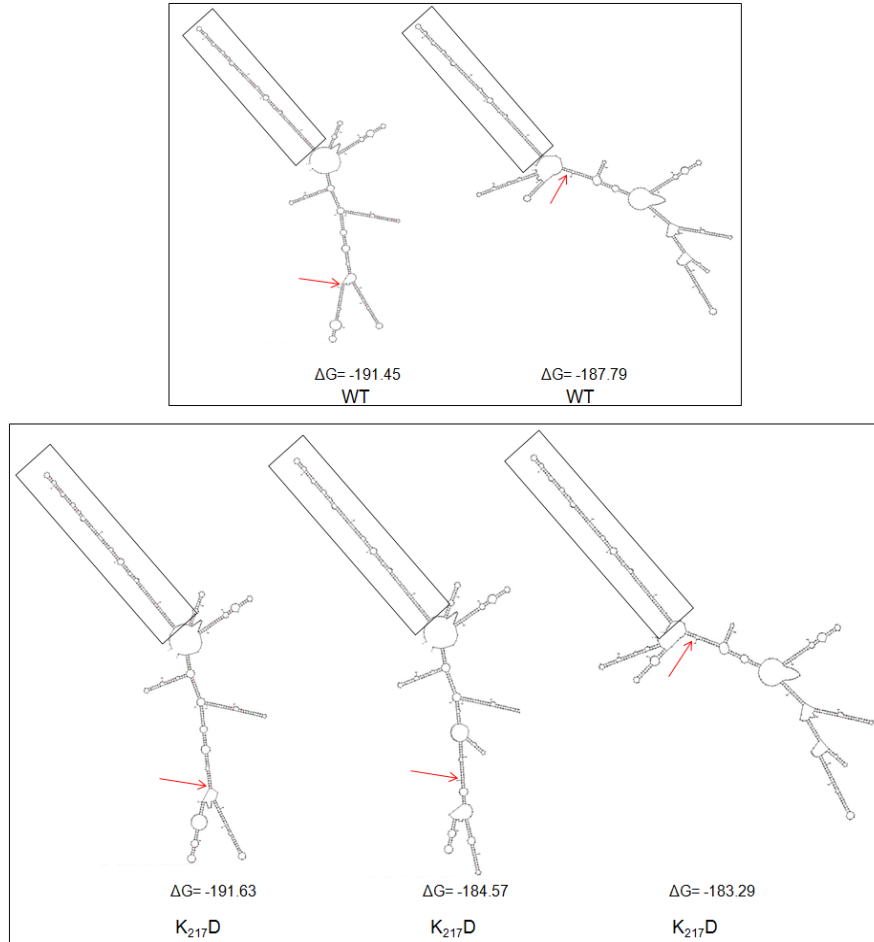FIG.19. Predicted packaging signal region structures for WT and $K_{217}D$. The Mfold version 4.4 (226) was used to predict RNA structures for a 622 base region encompassing the PS and a downstream region. For each prediction the ΔG value is listed and the known stem loop PS region is boxed. Arrows indicate the region where the mutation is located. Top: WT predicted structures Bottom: $K_{217}D$ predicted structures

The region of ORF1ab encompassing the packaging signal encodes the nsp15 protein. Nsp15 is a RNA processing enzyme with endoribonuclease activity (84, 151, 209). The change in the $K_{217}D$ mutants results in a lysine (K) to arginine (R) change at amino acid position 344 within nsp15. The nsp15 protein has been crystallized and key residues that are part of the catalytic domain have been identified and characterized (88, 209). The identified residues are H_{262}, H_{277}, L_{317} and G_{275}. Additional residues involved in stabilizing and substrate binding have also been described. The $K_{344}R$ change is clearly out of the catalytic region and was not specifically identified as a residue involved in other functions of substrate binding. The $K_{344}R$ change is a conservative change in terms of charge. An alignment of all group 2 coronavirus nsp15 proteins show that the K residue is conserved, suggesting a charge at this position may be important (data not shown). At this point there is no reason to think the change recovered in the nsp15 protein is compensating for the change in M, especially because these viruses have a WT-like phenotype. Thus, this analysis suggests that the change near the PS does not affect the packaging signal RNA structure or known key residues within the nsp15 protein.

Single residue mutations affect virus growth. Two of the introduced single changes did result in altered phenotypes. The first was the $R_{212}A$ mutation. Although fusion was clearly seen after electroporation, only one of ten passaged plaques could be recovered. The isolated virus did contain the introduced change

with no additional changes present in the other structural genes. The second

change that resulted in an altered phenotype was the $K_{207}D$ mutation. Five

individual $K_{207}D$ plaque isolates were passaged and analyzed by sequencing. All

the recovered mutants contained the introduced mutation with no other changes.

In growth kinetic studies, both the $R_{212}A$ and $K_{207}D$ mutants were severely

crippled, growing to titers 2 or 3 logs, respectively, below WT virus at an MOI of

0.01 PFU/mL (Fig.17). These results initially suggest the R_{212} residue is more

sensitive to neutral charge changes compared to the K_{205} and K_{207} residues.

Additionally, virus is crippled in the presence of an opposite charge at position

K_{207}. These residues may play a role in electrostatic interactions M participates in.

The $R_{212}A$ mutant did contain a change, an A to T at position 20372, just

beyond the PS region. This is beyond the conserved stem loop structures

described above. Mfold analysis was carried out for both the WT and $R_{212}A$

sequences (226). Three or four secondary RNA structures for WT and $R_{212}A$,

respectively, remained after the criteria were met (see Materials and Methods).

The change in the $R_{212}A$ virus did not cause alteration of the PS stem loop

structure in any prediction. There were some slight variations in the structure in

regions distant from the PS in some predictions (Fig.20). Since the PS region

itself in unaltered, it is presumed that these minor changes would not affect the PS

structure or function.

The change within $R_{212}A$ was then analyzed in the context of nsp15. The

mutation results in a Q to L change at position 286. This is just beyond the

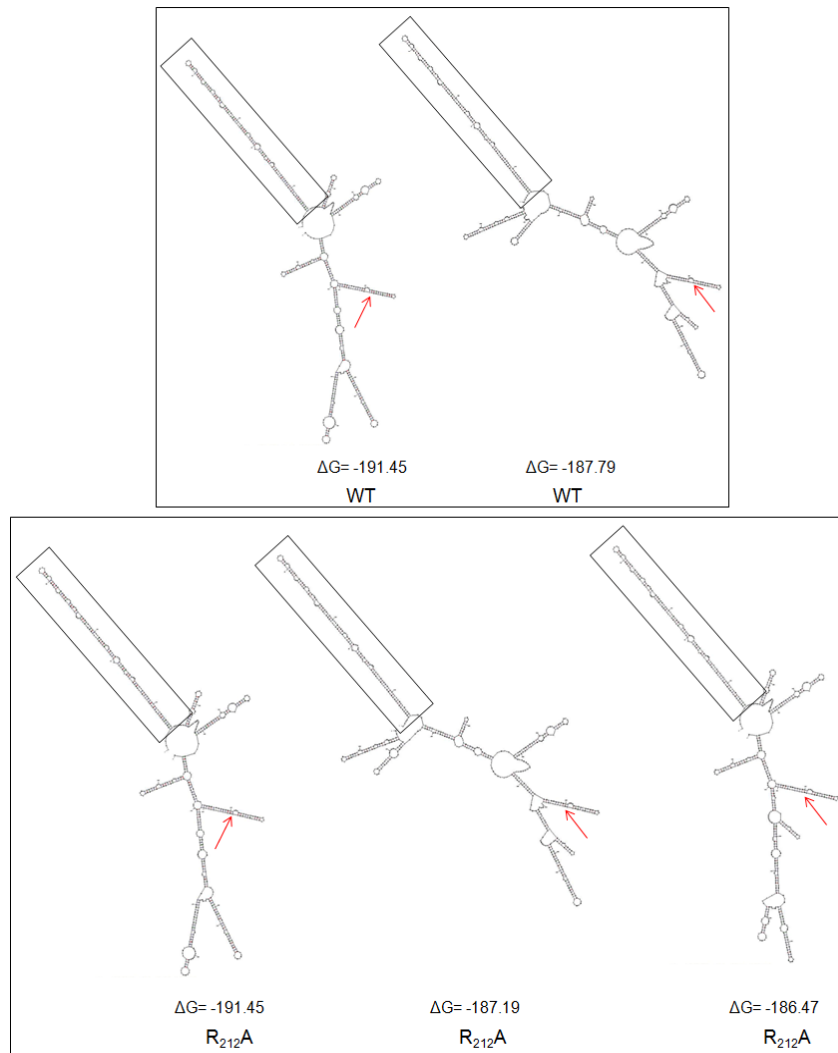FIG.20. Predicted packaging signal region structures for WT and R$_{212}$A. The
Mfold version 4.4 (226) was used to predict RNA structures for a 622 base
region encompassing the PS and a downstream region. For each prediction the
ΔG value is listed and the known stem loop PS region is boxed. Arrows indicate
the region where the mutation is located. Top: WT predicted structures Bottom:
R$_{212}$A predicted structures

identified key residues in the catalytic domain of nsp15, as described above (88, 209). The Q residue is conserved in some group 2 coronaviruses, however others have charged or polar residues in this position. The possible significance of the $Q_{286}L$ change in the nsp15 protein would require further analysis.

Charge changes result in lethal phenotypes. Residues at position K_{205} and R_{212} were replaced with the oppositely charged D residue. For both mutants, the infectious cloned RNA was replication competent, as subgenomics could be detected following electroporation, however, no virus could be recovered after multiple passage attempts of media directly off of electroporation. These results strongly indicate that a negative charge at these two positions is not tolerated by the virus.

The majority of the HD mutants do not support VLP assembly. To gain further insight into how the CD mutations might be affecting virus assembly, the proteins were analyzed for their ability to form VLPs. Coexpression of the M and E proteins is sufficient for assembly of MHV VLPs (14, 35, 194). All double and single mutant proteins were coexpressed with the wild-type E protein in 293T cells. Intracellular and extracellular fractions were harvested at 24 h after transfection. Both fractions were analyzed by SDS-PAGE and Western blotting and extracellular samples are shown in Fig.21. The only mutants able to support VLP production were $K_{207}A$, $K_{217}A$ and $K_{217}D$. Interestingly even though several other mutations were supported in the context of the virus, they were not able to

99

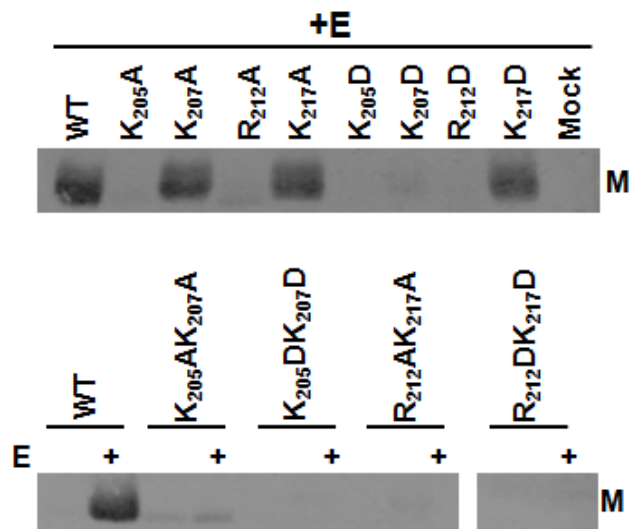FIG.21. Effect of charged residue mutations on VLP production. 293T cells were transfected with pCAGGS vectors containing WT or mutant M genes singly or in combination with the pCAGGS containing the WT E gene. Control empty vector (mock) were analyzed in parallel. Pelleted extracellular VLPs were analyzed by SDS-PAGE and Western blotting using antibodies against M. The entire VLP pellet was analyzed.

generate VLP particles suggesting other viral and possibly host proteins stabilize the assembly of particles when charged residue mutations are present.

During infection there are other viral and likely host proteins that participate in assembly of viral particles. The N protein encapsidates the genomic RNA and interacts with M (83, 95, 134, 196). Recent studies indicate that N contributes to efficient assembly of SARS-CoV and MHV VLPs (5, 172, 200). We wanted to determine if the presence of N during VLP formation would impact VLP output of our mutant M proteins. We tested the all the viable mutants, which included $K_{205}A$, $K_{207}A$, $R_{212}A$, $K_{217}A$, $K_{207}D$, $K_{217}D$ and the double mutants, $K_{205}AK_{207}A$ and $R_{212}AK_{217}A$. WT or mutant M proteins were coexpressed with WT E in the presence and absence of a WT N clone (Fig.22). The results show that the addition of N in WT M coexpressions results in a modest increase in VLP production (Fig.22). The mutants that were capable of forming VLPs in the presence of just E ($K_{207}A$, $K_{217}A$, and $K_{217}D$) displayed a roughly similar increase in VLP output with N compared to WT. Mutant M proteins that result in crippled growth ($R_{212}A$, $K_{207}D$, $K_{205}AK_{207}A$ and $R_{212}AK_{217}A$) yielded no VLPs when coexpressed with just E protein. The addition of the N protein resulted in little or no increase in VLP production for these mutants (Fig.22). Interestingly for the $K_{205}A$ mutant, no VLPs were produced in the presence or absence of although virus was produced that grew similar to WT. These results suggest that the N protein cannot dramatically enhance VLP production in the presence of crippling mutations in M proteins.

FIG.22. VLP analysis of viable charged residue mutants in the presence of the N protein. 293T cells were transfected as indicated with pCAGGS vectors containing WT or mutant M genes in combination with pCAGGS-E and pCAGGS-N gene where indicated. Intracellular and extracellular VLP fractions were analyzed by SDS-PAGE and Western blotting (top). The entire extracellular pellet and 6% of the total intracellular fraction were analyzed. Protein bands were quantified by densitometric scanning and analyzed using ImageQuant software. VLP release was calculated as the percentage of the extracellular M of total M (intracellular plus extracellular) protein (bottom). The data represent deviations from the average of two experiments.
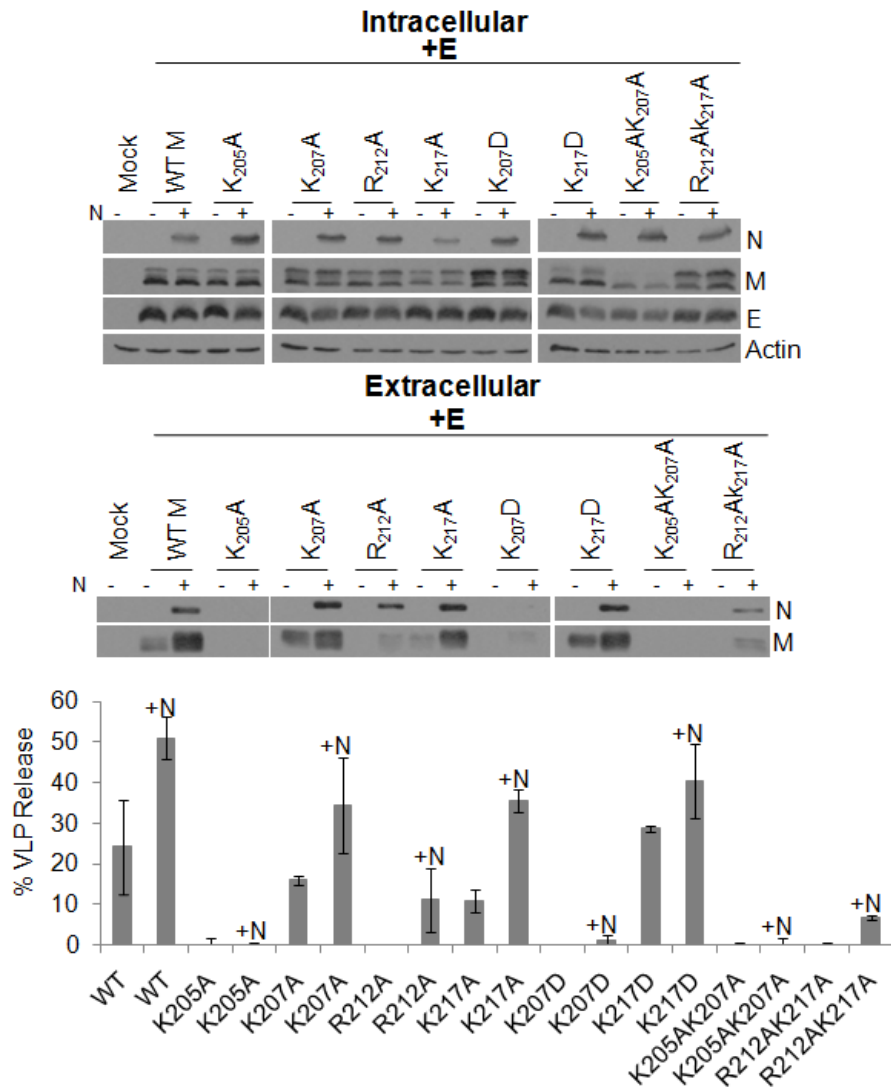
Infectivity of Crippled Mutants. It is well established that the M and N proteins interact and this interaction presumably facilitates the incorporation of the nucleocapsid (N+RNA) into the viral particle. If our charged residue mutations are effecting the M-nucleocapsid interaction, non-infectious (empty) particles would be produced during virus infection. To begin to determine if mutant infections were resulting in non-infectious particles production, we first analyzed the protein profiles of virions for the two crippled single mutants, $K_{207}D$ and $R_{212}A$. We hypothesized that if we visualized the protein profiles of a known amount of PFU and empty particles are produced then we would see more relative amount of protein than expected for the PFUs present. The PFU was determined by standard plaque assay on L2 cells. An equal number of PFUs (10^5) for each mutant was run on a SDS-PAGE gel and Western blot analysis done (Fig.23A). In parallel, a WT dilution series from 10^5-10^7 PFUs was run. The approximate number of particles for each mutant was determined by comparing the intensity of the N protein to the WT standards. The $K_{217}A$ mutant, which has a growth phenotype similar to WT, was analyzed in parallel. The results show that both the $K_{207}D$ and $R_{212}A$ mutants have the equivalent of about 10^6 particles for 10^5 PFUs, which is equivalent to a relative protein-to-PFU ratio of about 10 (Fig.23A). However, 10^5 PFUs of the $K_{217}A$ mutant correspond to roughly 10^5 particles. This data suggests that the crippled phenotype of the $K_{207}D$ and $R_{212}A$ mutants may be due to the production of non-infectious particles.

FIG.23. Infectivity of $K_{207}D$ and $R_{212}A$ mutants. A) Protein profiles of $K_{207}D$, $R_{212}A$ and $K_{217}A$ mutants in parallel with a WT standard dilution of known PFUs. Virions were purified through a sucrose cushion, tittered to obtain the PFU and then analyzed by SDS-PAGE and Western blotting using antibodies against S, M and N. A lighter exposure of the N protein is shown to the right. The relative amount of N protein was obtained by comparing the intensity of N to the WT standards. The data represent deviations from the average of two experiments. B) Visualization of RNA from $K_{207}D$, $R_{212}A$ and $K_{217}A$ mutant particles in parallel with a WT standard dilution of known PFUs. RNA was subjected to slot blot hybridization after extraction from purified particles and was detected using a digoxigenin-labeled N gene probe. The relative amount of RNA was obtained by comparing the intensity of the mutants to the WT standards. The data represent deviations from the average of two experiments.
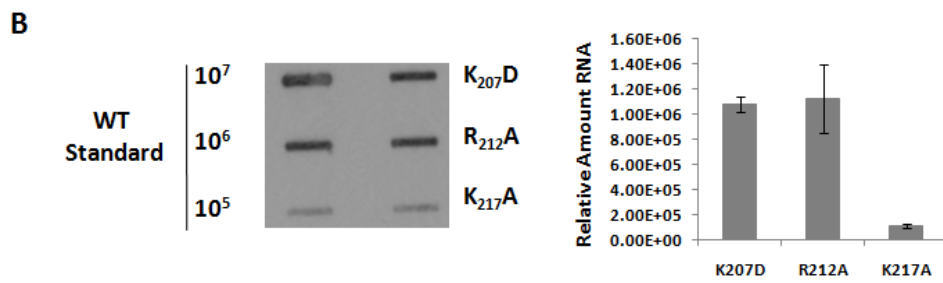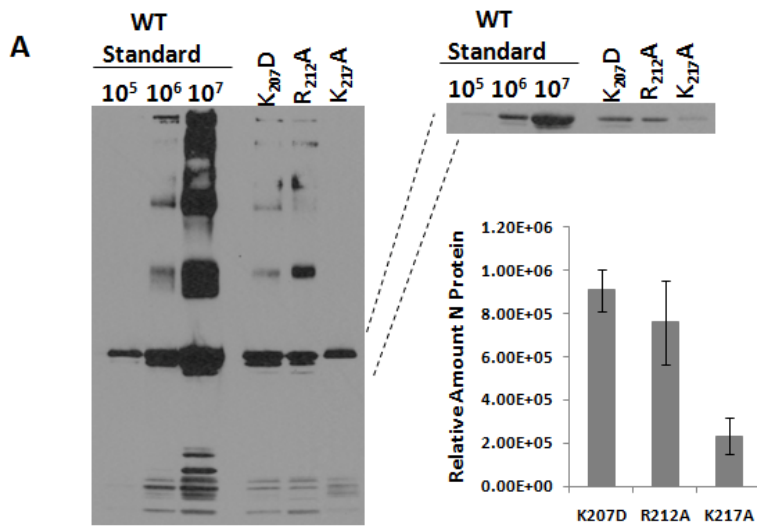
To determine if the production of non-infectious particles is due to empty particles being produced, we analyzed the levels of RNA in virus particles. Again, 10^5 PFU each for the $K_{207}D$, $R_{212}A$, and $K_{217}A$ mutants and a WT dilution series were analyzed. A fraction of the same samples analyzed by Western blotting above was taken for each. Slot blot hybridization was carried out with each sample and RNA was detected using a digoxigenin-labeled N gene probe. The amount of RNA for each mutant was determined by comparing to the WT standards. The results show that the $K_{207}D$ and $R_{212}A$ have a level of RNA equivalent to 10^6 PFU (Fig.23B). In contrast, the $K_{217}A$ mutant has an RNA level of 10^5 PFU which is roughly equivalent to the number of PFUs analyzed. This data suggests that the non-infectious particles produced during $K_{207}D$ and $R_{212}A$ mutant infection is not due to formation of particles that lack the RNA.

Since, the decrease in infectivity of the $K_{207}D$ and $R_{212}A$ mutants is not due to the production of particles lacking the RNA we began to analyze the M-S interaction. The S protein is the viral attachment protein and is necessary for infectivity (66). It is well established M-S interactions aid in retaining S in the Golgi/ERGIC, which is the site of coronavirus assembly. Charged residues within the endodomain of S have been implicated to be important for assembly into viral particles (211). Thus we reasoned that the decrease in infectivity may be due to a reduction in M's ability to retain S intracellularly. The M protein localizes in the Golgi when expressed alone. S localizes along the exocytic pathway and at the plasma membrane when singly expressed. When coexpressed, the M protein will

colocalize with S. WT, $K_{207}D$, and $R_{212}A$ M proteins were coexpressed with S.

Both mutants were able to retain S intracellularly (Fig.24). However, the level of

S incorporation into the envelope is not directly addressed in these experiments.
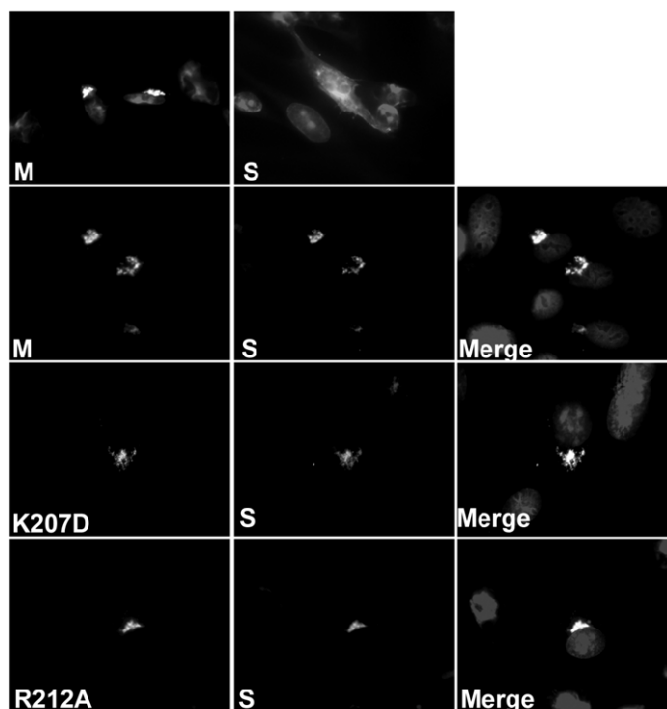
FIG.24. S protein colocalization with WT and mutant M proteins. 293T cells were transfected with pCAGGS vectors containing WT or mutant M genes and the S gene. Cells were fixed at 12 h after transfection and analyzed by immunoflourescence using mouse and goat antibodies against the M and S proteins, respectively. Alexa Fluor 488-conjugated mouse and Alexa Fluor 594-conjugated goat secondary antibodies were used to visualize the localization of the M and S proteins, respectively. Nuclei were stained with DAPI. Singly expressed M and S proteins are shown in the top two panels. Colocalizations of the M and S proteins are shown in the merged images on the right.

DISCUSSION

In this study, we examined the importance of four positively charged residues in the MHV M hydrophilic tail. A cluster of three residues (K_{205}, K_{207} and R_{212}) were identified to play a role in the function of the M protein, as introduction of oppositely or neutrally charged residues were not tolerated or resulted in viruses displaying poor growth. In contrast, the K_{217} residue does not appear to play a crucial role in the function of M because replacement with neutral or negative residues results in viruses that have a growth phenotype similar to WT. Most mutations did not support VLP assembly with WT E protein. The addition of the N protein did not greatly enhance VLP production for crippled M mutants. The data suggests that some of the charge changes result in a loss of infectivity of viral particles. The study overall provides new insight into residues within the M protein that are important for virus envelope assembly, as well as infectious particle production.

Most of the mutations in M resulted in a lack of VLP assembly with WT E protein. This is consistent with previous work on the M protein where several deletions were made from the extreme carboxy tail of MHV M, ranging from 1 up to 18 residues ($\Delta 1$, $\Delta 2$, $\Delta 3$, $\Delta 5$, $\Delta 11$ and $\Delta 18$) (43). The mutant with 18 residues deleted ($\Delta 18$) lacks the R_{212} and K_{217} residues but still contains the K_{205} and K_{207} residues. All of these mutants were not competent for VLP formation. This data suggests that the residues located in the extreme tail of M play a role in formation of the envelope. However, it is possible that the deletions affected the structure of

110

the end of the tail. Additionally, previous work has shown that the N protein

offers a stabilizing role for assembly complexes when mutations are present in a

conserved domain in the amino end of the amphipathic region of MHV M (5).

Similarly, an increase in VLP release has been observed for several RNA viruses

when the nucleocapsid proteins are coexpressed with the matrix proteins (72, 112,

115, 164, 192). However, this is not the case with our charged residue mutants.

The affect the C-terminus mutations are having on the formation of the envelope

cannot be aided by the presence of the N protein. It is possible N protein binds to

the M mutants less efficiently. The data does suggest that M-N interactions are

not completely abolished, as all (except $K_{205}A$) mutants incorporated N into the

VLP envelope to some extent. However, a direct assessment of M-N interactions

was not done.

The M protein interacts extensively with itself to form the lattice of the

envelope (77). Could our mutations be affecting M-M interactions during VLP

envelope formation? Based on past studies we do not expect that our mutations

are affecting M-M interactions. A mutant lacking the last 22 residues can still

form M-M oligomers, although it localizes to the plasma membrane (117). This

mutant lacks K_{207}, R_{212} and K_{217}. M-M interactions were studied for the $\Delta 18$

mutant mentioned above using co-immunoprecipitation (coIP) and VLP rescue

assays (47). The $\Delta 18$ mutant was able to form M-M complexes and be rescued

into VLP particles by an assembly competent M protein. However, this mutant

could not form VLPs with just WT E, similar to our mutants. This data suggests

111

that the extreme carboxy tail of the M protein is not involved in mediating M-M interactions. Although M-M interactions are required for envelope assembly clearly other requirements must be met. For example, for Semliki Forest virus, although lateral interactions between envelope proteins are a requirement for budding of particles, without nucleocapsid binding efficient assembly is reduced (56). The effect of the introduced charged residue mutations on VLP assembly seem to be due to other fundamental problems with M forming the envelope and generating particles. These residues could be involved in intra- or inter-molecular protein interactions M participates in or possibly M-host protein interactions. Although, direct M-E interactions have not been shown for MHV, it is possible the mutations are affecting the complex interplay between these two proteins necessary for efficient envelope assembly. Alternatively, the mutations could be affecting the conformation of M (discussed below), which could result in less efficient envelope formation.

The $K_{205}D$, $R_{212}D$, $K_{205}DK_{207}D$ and $R_{212}DK_{217}D$ mutations were not able to support virus assembly. All of these mutants did localize correctly to the Golgi region (data not shown) suggesting the lethality of the introduced mutations was not due to protein not being in the correct location for virus assembly to occur. These opposite charge changes presumably abolished one of the protein-protein interactions M participates in or resulted in an aberrant M protein structure that did not allow for one of the multi-functions of M during virus assembly. Since the

single mutations $K_{205}D$ and $R_{212}D$ resulted in a lethal phenotype it is not surprising that the double mutants harboring these changes were also lethal.

The $K_{207}D$ and $R_{212}A$ charge changes resulted in viruses with crippled phenotypes, growing to peak virus titers 100 fold less than WT virus. We speculated that these mutants might produce empty particles. However, these viruses were found to produce non-infectious particles that contained the RNA. The $K_{217}A$ mutant did not appear to produce a significant amount of non-infectious particles. Although, initial analysis of M-S interactions suggest the $K_{207}D$ and $R_{212}A$ mutants are able to retain S at the site of assembly, the extent of S incorporation in the envelope was not determined.

Our initial hypothesis based on previous genetic studies was that the targeted charged residues of the M tail are interacting with the N protein electrostatically (83, 95, 195, 196). Residues 201-224 of Mouse hepatitis virus (MHV) M (particularly R_{227}), residues 237-252 of transmissible gastroenteritis virus (TGEV) and residues 197-221 of SARS-CoV M are regions identified to be involved in the interaction with N, all of which are highly charged regions (Fig.16) (43, 58, 83, 95, 195, 196). Genetic studies on the N protein identified charged residues as well (D_{440} and D_{441}) as being functionally important key residues involved in virus assembly (83, 195). Additionally, recent cryo-EM studies have shown that the extreme carboxy tail of M is the contact point of the viral envelope with the internal helical nucleocapsid structure (7, 138). It was reasonable to speculate the charges on M may contribute to direct charge-charge

interactions between the two proteins. However, the data shows the non-infectious particles produced during mutant virus infection contain the nucleocapsid suggesting there is no major defect M-N interactions.

How might the introduced mutations result in a decrease in infectivity? The charged residues in the tail of the M protein appear to be involved in more than just interactions with the nucleocapsid. The particles could be non-infectious due to a reduction of S, the viral attachment protein, on the surface. It is known M interacts with the S protein and helps retain it in the Golgi/ERGIC where viruses assemble (139, 143). Cyro-EM work has suggested that the placement of S in the envelope is restricted by interactions with the organized M lattice (138). Genetic studies have analyzed the regions of MHV S and M necessary for assembly into virions. The S endodomain consists of an 18 aa membrane proximal cysteine-rich region and a 27 aa charge-rich region. Even though there is little primary sequence conservation within the S endodomain, the cysteine-rich and charge-rich regions are conserved among all coronaviruses S proteins (14, 25). Three negatively charged residues within the charge-rich region were found to be important for the inclusion of S into virions (211). It is possible the targeted positive charges in this study interact electrostatically with these negative charges in S.

Previous research on M protein mutants led to initial mapping of the domains within M that are responsible for the M-S interactions (46). In these studies, it was found when the last 18 residues were deleted there was a

114

significant reduction in M-S interactions. Importantly, when the last 15 residues were deleted ($\Delta15$), M-S interactions were not grossly reduced. The difference between these two mutants are the Y_{211}, the targeted R_{212}, and L_{213} residues. This is also consistent with our changes to the K_{217} residue that did not affect virus production, as the $\Delta15$ mutant does not contain K_{217}. A single $Y_{211}G$ change also resulted in a reduction of M-S interactions (46). Furthermore, a single tyrosine (Y_{195}) in SARS-CoV M was recently shown to be necessary to retain S in the Golgi (129). Therefore it is reasonable to think the charged residues in M's tail, especially R_{212}, may influence proper M-S interactions. Because S was detected in the protein profiles of purified $K_{207}D$ and $R_{212}A$ virions, the charge changes introduced may decrease, but not abolish, crucial M-S interactions leading to a subsequent decrease in infectivity.

Cryo-tomography and EM analyses of have shown that M forms local ordered networks of M molecules in the envelope and this protein lattice formation is crucial for organization of the envelope (7, 138). Furthermore, another recent cryo-tomography and EM study has shown M may exist in two forms in the viral envelope: M_{short} and M_{long} (Fig.25A) (B.W. Neuman, submitted). Regions of M_{long} appear to help mediate S protein incorporation in the envelope, as well as contact with the nucleocapsid, and appear to be associated with membrane curvature (Fig.25B). The binding of S and/or the nucleocapsid may help to stabilize M_{long} as well. It is speculated that the change in the
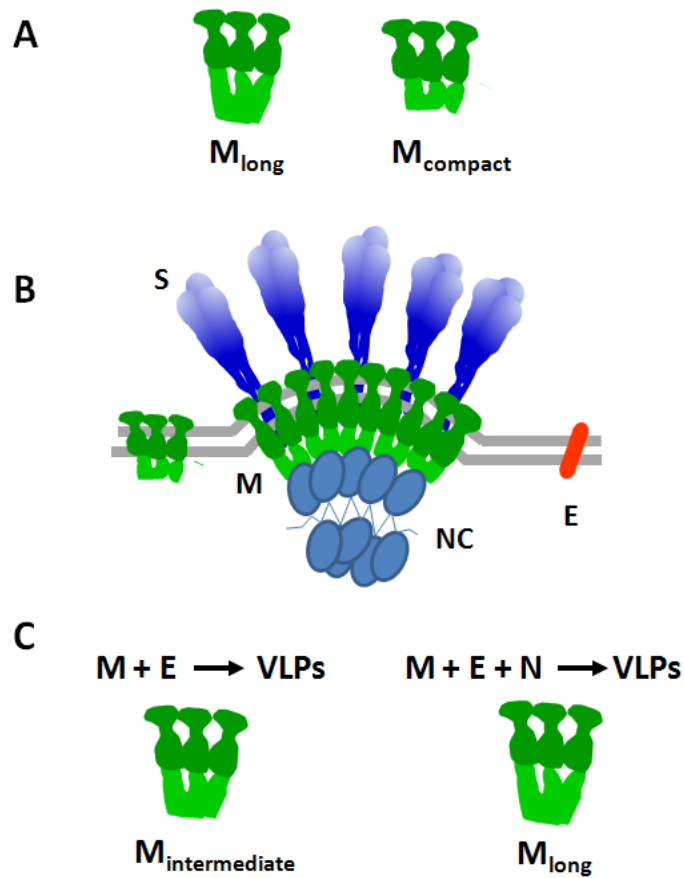
FIG.25. M protein conformation in the envelope. A) M may exist in two forms in the envelope, M_{long} and $M_{compact}$, which differ in a lengthening or shortening of the endodomain of M. B) In virus assembly, regions of M_{long} are where S clusters and the membrane appears to bend around the nucleocapsid (NC). C) In VLPs consisting of only M and E proteins, M exists in an intermediate form between M_{long} and $M_{compact}$. However, in VLPs consisting of M, E and N proteins, M exists primarily in the M_{long} conformation, suggesting N may mediate conversion to M_{long}. Images modified from B.W. Neuman.

conformation of M is due to the lengthening or shortening of the endodomain. It is possible the charge mutations in the extreme carboxy tail of M somehow affect the conversion to M_{long}. This could be due to loss of or new electrostatic interactions as there are numerous charged residues located in the endodomain of M. This change in the ratio of M_{long} to M_{short} could result in less incorporation of S, which would clearly affect the infectivity of mutant particles. This thinking would also be consistent with our mutated M proteins resulting in little or no VLP output as the M protein in E+M VLPs is in an intermediate form between M_{long} and M_{short} (Fig.25C) (B.W. Neuman, submitted). In the context of infection when various other interactions stabilize M_{long}, viral particles are produced more efficiently.

The S protein must undergo refolding events during entry to allow for fusion of the viral and host cell membranes. Cysteine residues within the cysteine-rich region mentioned above, are palmitoylated (13). One study analyzed the role of these S protein acylations and suggested they force M to be an optimal distance away from itself to allow for refolding events to occur to allow for cell-to-virus fusion (186). It is possible the introduced mutations are modifying the interactions between M and S in such a way that the fusion process is being affected. This could be due to M_{long}:M_{short} ratio changes or someother structural alteration in M. The Sindbis virus envelope glycoprotein, E2, C-terminus tail interacts with the capsid (C) protein during assembly for incorporation of the nucleocapsid into the particle (31, 107, 109, 119, 145, 203). Mutations were made to residues within the

endodomain of E2 (201). It was found that the targeted residues were not necessary for assembly but were important viral infectivity and more specifically for fusion (201). How the mutations resulted in a loss of fusion activity is not described.

Alternatively, the affect of our mutations on the infectivity of viral particles could be even more complex than M-S interactions or entry and fusion events. The decrease in infectivity could be due to any downstream process from entry. How the charge changes introduced into the M protein are reducing viral infectivity by influencing protein-protein interactions, entry or other life cycle processes remains to be determined.

Our results show that charged residues in the tail of the M protein are important for virus assembly and infectivity. The effects on assembly may be distinct or intertwined with the effect on viral infectivity. It is likely that assembly complexes must be organized to allow for incorporation, as well as the correct spacing and placement, of all the necessary components to ensure infectious virus production. Additional studies will aim to determine the mechanism by which the charge changes are affecting viral infectivity.

CHAPTER 4

SUMMARIZING DISCUSSION

This dissertation focuses on two regions of the coronavirus M protein carboxy tail. Chapter one describes a study of a highly conserved domain (CD) at the amino end of the long amphipathic region of the tail. The CD was examined directly for the first time in this study. The results suggest that this region is involved in fundamental M-M interactions. These interactions appear to be stabilized through interactions with the N protein. Chapter two focuses on the hydrophilic, highly charged extreme carboxy tail of the M protein. The data identifies a cluster of three key positive charges that play a role in assembly of infectious particles. Both studies provide new insight into requirements of the key virion envelope structural component and demonstrate the functional significance of the CD and charged residues. The new information is important for our basic understanding of the molecular biology of coronaviruses. It may also provide the basis for exploring the domains as potential targets for the development of antiviral therapeutics or vaccines that prevent virus assembly.

The Membrane Protein and its Role in Virus Assembly

The M glycoprotein is the most abundant component of the viral envelope that plays required, key roles in virus assembly (47, 91, 139, 143, 194). Coronaviruses assemble and bud at intracellular membranes in the region of the endoplasmic reticulum Golgi intermediate compartment (ERGIC) (93, 188).

Coronavirus M proteins are divergent in their amino acid content, but all share the same overall basic structural characteristics. The proteins have three transmembrane ™ domains, flanked by a short amino terminal glycosylated domain and a long carboxy terminal tail located outside and inside the virion, respectively (77) (Fig.4). The long intravirion (cytoplasmic) tail of M consists of an amphipathic domain following the third TM and a short hydrophilic region at the carboxyl end of the tail (Fig.4). The amphipathic domain appears to be closely associated with the membrane (155). M molecules interact with each other and also with the spike and nucleocapsid during virus assembly (Fig.4) (46, 47, 58, 95, 139, 143). These fundamental interactions drive the efficient assembly of viral particles. M-M interactions constitute the lattice structure that is apparent in the viral envelope. The S protein and a small number of E molecules are incorporated in the M protein lattice. Co-expression of only the M and the E proteins is sufficient for VLP assembly for most coronaviruses (14, 194).

Previous studies from a number of labs implicated multiple M domains and residues to be important for coronavirus assembly (43, 46, 47, 83, 196). Mutations and deletions made in all regions of M, including the amino end, transmembrane domains (TMs) and carboxy tail, effect VLP envelope assembly (43), again highlighting the importance of the M protein in the assembly process. M-M interactions are thought to be mediated through multiple contact points, but the TMs appear to be especially important (47). Cryo-tomography analysis revealed that the envelope is striated, presumably due to M-M interactions (7).

120

The amphipathic region of the M protein tail has been shown to be important in mediating M-S interactions (46). Additionally, particular residues within the hydrophilic domain have been identified as being crucial for SARS-CoV and MHV M protein interactions with S (46, 129). Finally, genetic studies have shown that charged residues in the extreme carboxy tail of M are mediating interactions with the N protein (83, 95, 195, 196). These interactions presumably facilitate the incorporation of the nucleocapsid into viral particles.

The M Protein Conserved Domain

At the amino terminus of the amphipathic domain, there is a highly conserved 12 amino acid domain (SWWSFNPETNNL), consisting of residues 114-125 in the MHV A59 M protein (Fig.8) (89). The results from work described in this dissertation suggest a possible interaction between the CD and first TM domain (TM1), since all of the recovered viruses with changes in the amino end of the CD had second site changes in the TM1 (Fig.10). No viruses could be recovered without second site changes. In both the context of VLP and virus assembly the TM1 changes are apparently providing some advantage for the mutant M proteins containing the changes in the amino end of the domain. Additionally, the second site changes in TM1 were required for rescue into VLP particles (Fig.14). Although the structure of the amphipathic domain of M is not known, it appears to be associated with the inner leaflet of the virion membrane. An interaction between the CD and the TM1 suggests the amphipathic region may

be more intimately associated with the membrane than previously recognized. Indeed, virions examined by cryo-electron tomography exhibit trilaminar-like membranes in the envelope, which was attributed to close association of M's tail with the inner leaflet of the membrane bilayer (7). Most of the TM1 changes were polar to charged residue changes. It is possible the changes, particularly the charges, result in a shortening or shifting of the TM region to allow for more efficient M-M interactions when alanine substitutions were present in the amino end of the CD. This may be particularly relevant since tryptophan (W) residues positioned at the membrane:cytoplasm boundaries have been suggested to possibly help stabilize TM helices or to provide vertical mobility relative to the lipid bilayer (102). The replacement of the W residues may destabilize the TM3. The changes in TM1 may then compensate for alternations in TM3 to allow for efficient M-M interactions (Fig.28A).

To further analyze where the TM1 changes occur in the context of the transmembrane helices itself, helical wheel analysis was used to determine the position of the recovered compensating changes. Residues consisting of the TM1, W_{26} to F_{43}, were used for the analysis. For the WT sequence there is a region where hydrophilic residues (N_{27}, T_{38}, and Q_{42}) lie next to each other on the same face of the α-helix (Fig.26). All the remaining residues, but one serine (S_{29}), are hydrophobic (Fig.26). Three of the four changes that arose in TM1 were in this region of hydrophilic residues ($G_{31}R$, $T_{38}N$, and $Q_{42}R$). The L_{35} residue, which

122
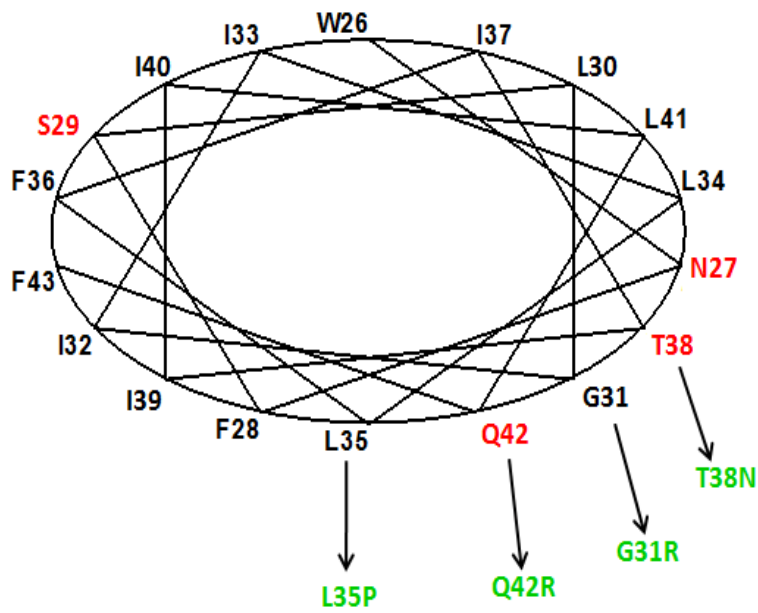
FIG.26. Helical Wheel Analysis of the M protein TM1. Top view from the amino end of WT TM1. Hydrophilic polar residues, N_{27}, S_{29}, T_{38} and Q_{42} are in red. Most of the hydrophilic residues lie along one face of the α-helix. The compensating changes that arose in 5'A mutant viruses ($G_{31}R$, $L_{35}P$, $T_{38}N$ and $Q_{42}R$) are indicated in green. Helical wheel plot was prepared using Gene Runner version 3.05.

changed to a P, is just adjacent to the hydrophilic region (Fig.26). This suggests

that the changes that arose in the TM1 of M are not randomly positioned

throughout the transmembrane α-helix. The hydrophilic region of the TM may be

positioned at the interface between the membrane and cytoplasm. This would be

consistent with the TM1 interacting with the CD of an adjacent M protein or with

the TM1 compensating for potential positioning changes in the TM3. These

conclusions provide a more detailed picture of the requirements of M-M

interactions and help us better understand the assembly process as a whole. It will

be of interest in future studies to determine mechanistically how the second site

changes in TM1 are compensating for the 5'A changes. Possibly cryo-

tomography studies to visualize VLP particles containing 5'A or 5'A+TM1 M

proteins would allow for determination of differences in the M lattice formation.

Our results strongly suggest that the N protein helps stabilize assembly

complexes during VLP assembly, most likely through interactions with M. We

know that the N protein plays important roles in encapsidation of the genomic

RNA (through N-RNA and N-M interactions) as well as in RNA synthesis in the

context of virus infection (2, 26, 39, 123). However, N's role in VLP, as well as

virus assembly, has not been completely elucidated. Our results suggest that N

stabilizes the VLP envelope, which consists almost entirely of M. This previously

undefined role of N may be particularly important for our mutant M proteins if

positioning of M molecules in the lattice framework and/or the tail is altered.

Stabilization by N might allow for a confirmation of M that promotes more

efficient assembly. Indeed, recent cryo-tomography and EM work has shown that M may exist in two forms in the viral envelope: M_{short} and M_{long} (Fig.25A and discussed in Chapter 3 Discussion) (B.W. Neuman et al, submitted). Regions of M_{long} are shown to be where contact is made with the nucleocapsid, where membrane curvature is seen, and where the S protein is incorporated into the envelope (Fig.25B). Therefore, M_{long} is thought to be important for efficient virus assembly. The change in M's conformation is thought to be due to the lengthening or shortening of the endodomain. Furthermore, M protein in E+M VLPs is in an intermediate form between M_{long} and M_{short} whereas E+M+N VLPs contain mostly M_{long} (Fig.25C). This suggests that the N protein, when present during VLP assembly, may convert M to a form that is more efficient in assembly and budding. The N-M interactions are also likely significant during complete virus assembly. This is consistent with the thread-like densities observed in cryo-EM and cryo-electron tomography images which provide the contact points between M and nucleocapsids in MHV particles (7, 138) and also genetic analysis pointing to the importance of the M-N interaction for encapsidation of the genomic RNA (83, 95, 195, 196). The stabilizing effect of N may be of greater importance and more apparent with the mutant M proteins, than during its normal role when expressed with WT M.

The M Protein Hydrophilic Tail

The terminal 25 amino acids of the long carboxy tail of M are hydrophilic and highly charged. This region has been shown to be important for interactions with the N and S proteins (83, 95, 129, 134, 196, 211). The results presented here suggest that viruses containing the $K_{207}D$ and $R_{212}A$ changes produce non-infectious particles. These particles do contain the RNA, suggesting M-N/nucleocapsid interactions are taking place. Several previous studies are consistent with the thinking that regions of the M tail are interacting with the S protein. It is known that M interacts with S and helps retain it in the Golgi/ERGIC where viruses assemble (139, 143). Cyro-EM work has suggested that the placement of S in the envelope is restricted by interactions with the M lattice structure (138). Additionally, genetic studies have analyzed the regions of MHV S and M necessary for assembly into virions. Three negatively charged residues within the S protein charge-rich region (discussed in Chapter 3 Discussion) were found to be important for the inclusion of S into virions (211). It is possible the key positive charges identified in our studies interact electrostatically with these negative charges in S.

Previous research on M protein mutants led to initial mapping of the domains within M that are responsible for the M-S interactions (46). It was found that residues within the hydrophilic carboxy tail of M appear to help mediate M-S associations. These include Y_{211}, the targeted R_{212}, and L_{213} (46). For SARS-CoV M, a single tyrosine, Y_{195} was recently shown to be necessary to retain S in the

Golgi (129). Therefore it is reasonable to think the charged residues in M's tail

that were studied may influence proper M-S interactions.

We initially began to look at M-S associations by determining the ability

of WT and mutant M proteins to retain S at the site of assembly (Fig.24). Both

$K_{207}D$ and $R_{212}A$ mutants were able to retain S intracellularly. However, the

extent of S incorporation into the envelope was not determined. Therefore, we

analyzed purified WT and $K_{207}D$ virions by negative staining to visualize S in the

viral envelope. WT particles contain the characteristic S proteins decorating the

envelope (Fig.27A and B). However, the $K_{207}D$ particles appear to be primarily

spikeless (Fig.27A and B). Additionally, there appeared to be a lot of broken

particles and membranous debris present for the $K_{207}D$ mutant (Fig.27B and C).

First of all, this data suggests that the introduced mutation within the C-terminus

of M may result in less S protein being incorporated into the envelope. These

results are consistent with previous work indicating that M-S interactions may be

mediated by electrostatic means. Further studies must be directed at quantifying

the amount of S in the mutant virus envelopes to determine the ratio of M to S

proteins. Cryo-EM studies have estimated that there is 1 S trimer for every 16 M

molecules (138). It would be interesting if our mutant virions had a different ratio

of S to M. Furthermore, the EM results also point to $K_{207}D$ particles possibly

being less stable. This could be due to the lattice of M being altered or maybe the

ratio of $M_{long}:M_{short}$ is effected in such a way to influence the stability of the viral

envelope. Possibly the particles are more sensitive to preparation methods for EM

FIG.27. Negative Staining of WT and $K_{207}D$ virus particles. 17Cl1 mouse cells were infected with WT and $K_{207}D$. Supernatant was clarified and virus was purified through two- 20- 60 % sucrose gradients and then concentrated through a 30% sucrose cushion. Pellets were resuspended in TMEN pH 6.0. Particles were airfuged onto Formvar coated 50 mesh Ni grids and stained with 1% phosphotungstic acid pH 6.5. Images were collected at 53,000X on a Philips STEM electron microscope.

A) Digitally magnified images of WT and $K_{207}D$ particles. WT particles retained S peplomers in the viral envelope. Most $K_{207}D$ particles appeared spikeless. B) WT samples contained intact particles with S. Identically prepared $K_{207}D$ samples contained damaged particles (thin arrow) and other membranous structures (thick arrow) that co-purified with the virus. C) WT sample preparations were relatively clean with respect to cellular debris. $K_{207}D$ samples contained large membranous structures as well as debris. Scale bar for panel A is 50 nm and panel B is 100 nm and and panel C is 200 nm. Electron microscopy and image analysis by Pavithra Venkatagopalan

A

WT

K$_{207}$D

B

C

WT   K$_{207}$D
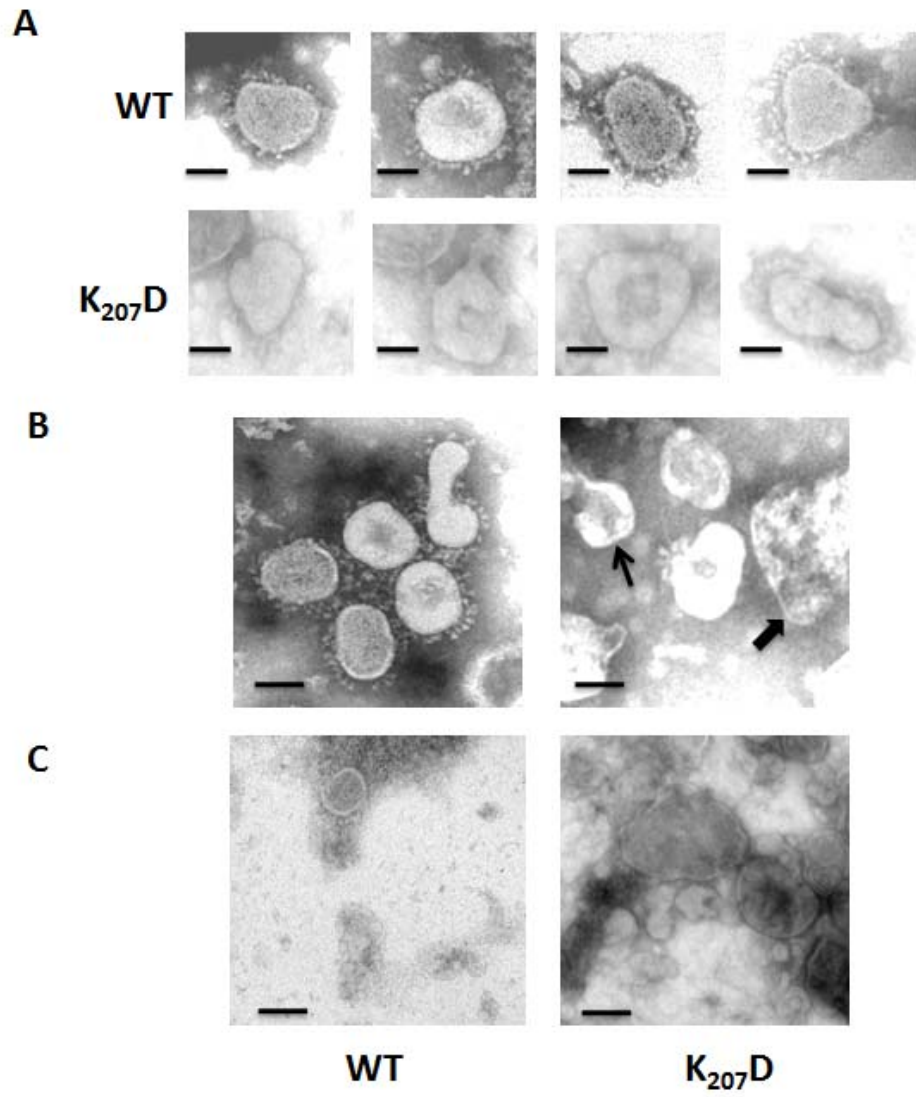
analysis. Further experiments could determine the level of sensitivity of the $K_{207}D$ particles to such conditions as pH, temperature and buffer composition.
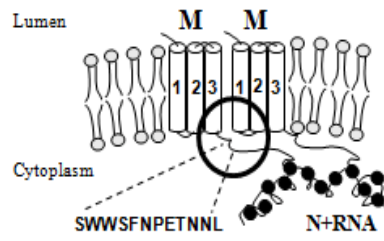
The M protein in Coronavirus Assembly

The M glycoprotein is the most abundant component of the viral envelope that plays key roles in virus assembly through interactions with itself as well as the other viral structural proteins (47, 91, 139, 143, 194). The results on the M protein CD and charged residues within the hydrophilic tail provide valuable insight into regions of M that play important roles in virus assembly as well as infectivity.

Although it has been known for quite some time that M interacts with itself to form multimers in the envelope, the regions involved in these interactions have only been very roughly mapped (47). Recent cyro-tomography and EM data has confirmed such hypotheses, as striations observed in the envelope were attributed to M-M interactions (7). The results on the CD of M provide valuable information on a particular region of M that helps mediate M lattice formation. Additionally, the results suggest that amphipathic region of M is intimately associated with the inside of the viral envelope as changes in the TM1 partially compensate for changes in the CD (Fig.28A). Such detailed mapping of the association between M molecules has not been described previously. Moreover, the results suggesting the N protein stabilizes assembly complexes, highlights a previously unrecognized role of N in virus assembly.

Fig.28. Model of coronavirus assembly protein complexes. The M protein is depicted with the three TM domains (1-3). A) The position and sequence of the CD and the N protein associated with the genomic RNA are shown. The CD of one M protein may interact with the TM1 of an adjacent M molecule (circled) and the N protein may stabilize M-M complexes. B) The position and sequence of the hydrophilic tail is shown. The endodomain of the S protein may interact with charges in the hydrophilic tail of M (circled)

A

Lumen

M   M

1 2 3   1 2 3

Cytoplasm

SWWSFNPETNNL   N+RNA

S

B

Lumen

M   M

1 2 3   1 2 3

Cytoplasm

VKSKVGNYRLPSNKPSGADTALLRT
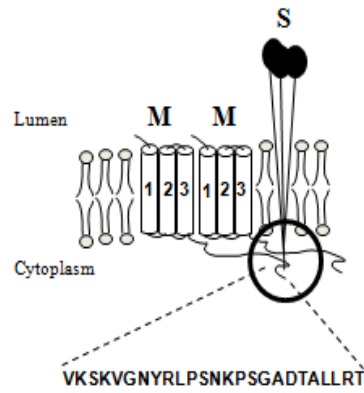
The M protein amphipathic region, containing the CD at the amino end has an unknown structure. Past studies suggest this region of M to be tightly associated with the membrane (155) and it is assumed to be globular. The amphipathic region of M may adopt an α-helix structure. Amphipathic α-helices are found in many proteins and play roles in anchoring proteins to the membrane, destabilizing membranes and promoting membrane curvature (4, 38, 53, 60, 63, 85, 108). Due to the way the hydrophobic and polar residues are on the two opposite sides of the helix, it actually lies parallel to the membrane with hydrophobic residues against the membrane and polar residues towards the lipid heads (79). A previous study found that several unrelated proteins, a Golgi-associated protein (ArfGAP1), a human coiled-coil protein (GMAP-210), a yeast sterol-binding protein (Kes1p) and a human nuclear pore complex protein (Nup133) have amphipathic helices that all contain serine/threonine rich regions on the polar face of the helix (51). It was found all these amphipathic helices were sensors of membrane curvature, suggesting the polar residues are important for their activity. It was suggested that asparagines (N) and glutamine (Q) residues may also maintain the proper composition of the helix for sensing curvature. It is possible the S, T and N residues, at positions S_{114}, S_{117}, T_{122}, N_{123} and N_{124}, in the CD are playing a role in membrane curvature sensing (Fig.8). Indeed when these residues were changed to alanine in the 5'A and 3'A mutants, it resulted in less virus production (Fig.10). Interestingly, when positive charges were introduced into the ArfGAP1 amphipathic helix, it lost its sensitivity to detect curved

133

membranes (51). Perhaps when positive charges were attempted to be introduced into the CD, it resulted in a similar decrease in sensing curved membranes.

What role may sensing membrane curvature have for the virus? Membrane curvature is required during the assembly process. What induces membrane curvature is not known, although it is thought the E protein might be involved (194). It is possible that as the E protein induces membrane curvature during early assembly steps, the M protein is able to sense this via its amphipathic region, which would lead to a clustering of M proteins at the site of assembly where budding is most efficient. Further studies could look at the CD mutant's ability to sense membrane curvature by analyzing binding to liposomes of different sizes. Such studies would potentially broaden our understanding of role of the M protein in assembly beyond protein-protein interactions.

Our work on the hydrophilic charged tail of M identifies three positive charges that are important for assembly of infectious particles. Our results suggest that the S protein is incorporated into the viral envelope less efficiently when mutations are present in the extreme carboxy tail of M. This data, along with the previous genetic work on the regions of M involved in M-S interactions, maps the interaction to a small highly charged region in the MHV M tail, from aa205-212. This includes Y_{211} that was identified previously to be important (46). Presumably the endodomains of M and S interact, which is mediated by charged regions in M and S (Fig.28B) (211). However we know that there does appear to be multiple contact points between the proteins as the amphipathic region of M plays a role in

M-S interactions as well (46). Most attention is given to the role of the M protein's tail in interactions with the nucleocapsid. However, these results emphasize the importance of the extreme carboxy tail of the M protein for incorporation of S.

REFERENCES

1. **Abdul-Rasool, S., and B. C. Fielding.** 2010. Understanding Human Coronavirus HCoV-NL63. Open Virol. J. **4:**76-84.

2. **Almazan, F., C. Galan, and L. Enjuanes.** 2004. The nucleoprotein is required for efficient coronavirus genome replication. J. Virol. **78:**12683-12688.

3. **Almazan, F., J. M. Gonzalez, Z. Penzes, A. Izeta, E. Calvo, J. Plana-Duran, and L. Enjuanes.** 2000. Engineering the largest RNA virus genome as an infectious bacterial artificial chromosome. Proc. Natl. Acad. Sci. U. S. A. **97:**5516-5521.

4. **Antonny, B., S. Beraud-Dufour, P. Chardin, and M. Chabre.** 1997. N-terminal hydrophobic residues of the G-protein ADP-ribosylation factor-1 insert into membrane phospholipids upon GDP to GTP exchange. Biochemistry. **36:**4675-4684.

5. **Arndt, A. L., B. J. Larson, and B. G. Hogue.** 2010. A conserved domain in the coronavirus membrane protein tail is important for virus assembly. J. Virol. **84:**11418-11428.

6. **Asanaka, M., and M. M. Lai.** 1993. Cell fusion studies identified multiple cellular factors involved in mouse hepatitis virus entry. Virology. **197:**732-741.

7. **Barcena, M., G. T. Oostergetel, W. Bartelink, F. G. Faas, A. Verkleij, P. J. Rottier, A. J. Koster, and B. J. Bosch.** 2009. Cryo-electron tomography of mouse hepatitis virus: Insights into the structure of the coronavirion. Proc. Natl. Acad. Sci. U. S. A. **106:**582-587.

8. **Baric, R. S., G. W. Nelson, J. O. Fleming, R. J. Deans, J. G. Keck, N. Casteel, and S. A. Stohlman.** 1988. Interactions between coronavirus nucleocapsid protein and viral RNAs: implications for viral transcription. J. Virol. **62:**4280-4287.

9. **Baric, R. S., S. A. Stohlman, and M. M. Lai.** 1983. Characterization of replicative intermediate RNA of mouse hepatitis virus: presence of leader RNA sequences on nascent chains. J. Virol. **48:**633-640.

10. **Beaudette F.R., C. B. H.** 1937. Cultivation of the virus of infectious bronchitis. J Am Vet Med Assoc. **90:**51-60.

11. **Benbacer, L., E. Kut, L. Besnardeau, H. Laude, and B. Delmas.** 1997. Interspecies aminopeptidase-N chimeras reveal species-specific receptor

recognition by canine coronavirus, feline infectious peritonitis virus, and transmissible gastroenteritis virus. J. Virol. **71:**734-737.

12. **Bisht, H., A. Roberts, L. Vogel, A. Bukreyev, P. L. Collins, B. R. Murphy, K. Subbarao, and B. Moss.** 2004. Severe acute respiratory syndrome coronavirus spike protein expressed by attenuated vaccinia virus protectively immunizes mice. Proc. Natl. Acad. Sci. U. S. A. **101:**6641-6646.

13. **Bos, E. C., L. Heijnen, W. Luytjes, and W. J. Spaan.** 1995. Mutational analysis of the murine coronavirus spike protein: effect on cell-to-cell fusion. Virology. **214:**453-463.

14. **Bos, E. C., W. Luytjes, H. V. van der Meulen, H. K. Koerten, and W. J. Spaan.** 1996. The production of recombinant infectious DI-particles of a murine coronavirus in the absence of helper virus. Virology. **218:**52-60.

15. **Bosch, B. J., R. van der Zee, C. A. de Haan, and P. J. Rottier.** 2003. The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex. J. Virol. **77:**8801-8811.

16. **Bost, A. G., R. H. Carnahan, X. T. Lu, and M. R. Denison.** 2000. Four proteins processed from the replicase gene polyprotein of mouse hepatitis virus colocalize in the cell periphery and adjacent to sites of virion assembly. J. Virol. **74:**3379-3387.

17. **Brian D.A., B.G. Hogue, and T.E. Kienzle.** 1995. The coronavirus hemagglutinin esterase glycoprotein, p. 165-180. *In* S.G. Siddell (ed.), The Coronaviridae. Plenum Press, New York.

18. **Brierley, I., M. E. Boursnell, M. M. Binns, B. Bilimoria, V. C. Blok, T. D. Brown, and S. C. Inglis.** 1987. An efficient ribosomal frame-shifting signal in the polymerase-encoding region of the coronavirus IBV. EMBO J. **6:**3779-3785.

19. **Brockway, S. M., X. T. Lu, T. R. Peters, T. S. Dermody, and M. R. Denison.** 2004. Intracellular localization and protein interactions of the gene 1 protein p28 during mouse hepatitis virus replication. J. Virol. **78:**11551-11562.

20. **Calvo, E., D. Escors, J. A. Lopez, J. M. Gonzalez, A. Alvarez, E. Arza, and L. Enjuanes.** 2005. Phosphorylation and subcellular localization of transmissible gastroenteritis virus nucleocapsid protein in infected cells. J. Gen. Virol. **86:**2255-2267.

21. **Casais, R., V. Thiel, S. G. Siddell, D. Cavanagh, and P. Britton.** 2001. Reverse genetics system for the avian coronavirus infectious bronchitis virus. J. Virol. **75:**12359-12369.

22. **Cavanagh D., D. A. Brien, M. Brinton, L. Enjuanes and K. V. Holmes, et al.** 1994. Revision of the taxonomy of the Coronavirus, Torovirus and Arterivirus genera. Arch. Virol. **135:**227-237.

23. **Cavanagh, D.** 1995. The coronavirus surface glycoprotein, p. p. 73. *In* S. G. Siddell (ed.), The Coronaviridae. Plenum Press, New York, NY.

24. **Cavanagh, D.** 1997. Nidovirales: a new order comprising Coronaviridae and Arteriviridae. Arch. Virol. **142:**629-633.

25. **Chang, K. W., Y. Sheng, and J. L. Gombold.** 2000. Coronavirus-induced membrane fusion requires the cysteine-rich domain in the spike protein. Virology. **269:**212-224.

26. **Chang, R. Y., and D. A. Brian.** 1996. cis Requirement for N-specific protein sequence in bovine coronavirus defective interfering RNA replication. J. Virol. **70:**2201-2207.

27. **Cheever, F. S., and J. B. Daniels.** 1949. A murine virus (JHM) causing disseminated encephalomyelitis with extensive destruction of myelin. J. Exp. Med. **90:**181-210.

28. **Chen, C. Y., C. K. Chang, Y. W. Chang, S. C. Sue, H. I. Bai, L. Riang, C. D. Hsiao, and T. H. Huang.** 2007. Structure of the SARS coronavirus nucleocapsid protein RNA-binding dimerization domain suggests a mechanism for helical packaging of viral RNA. J. Mol. Biol. **368:**1075-1086.

29. **Chen, H., A. Gill, B. K. Dove, S. R. Emmett, C. F. Kemp, M. A. Ritchie, M. Dee, and J. A. Hiscox.** 2005. Mass spectroscopic characterization of the coronavirus infectious bronchitis virus nucleoprotein and elucidation of the role of phosphorylation in RNA binding by using surface plasmon resonance. J. Virol. **79:**1164-1179.

30. **Chen, S. C., E. van den Born, S. H. van den Worm, C. W. Pleij, E. J. Snijder, and R. C. Olsthoorn.** 2007. New structure model for the packaging signal in the genome of group IIa coronaviruses. J. Virol. **81:**6771-6774.

31. **Cheng, R. H., R. J. Kuhn, N. H. Olson, M. G. Rossmann, H. K. Choi, T. J. Smith, and T. S. Baker.** 1995. Nucleocapsid and glycoprotein organization in an enveloped virus. Cell. **80:**621-630.

32. **Coley, S. E., E. Lavi, S. G. Sawicki, L. Fu, B. Schelle, N. Karl, S. G. Siddell, and V. Thiel.** 2005. Recombinant mouse hepatitis virus strain A59 from cloned, full-length cDNA replicates to high titers in vitro and is fully pathogenic in vivo. J. Virol. **79:**3097-3106.

33. **Cologna, R., J. F. Spagnolo, and B. G. Hogue.** 2000. Identification of nucleocapsid binding sites within coronavirus-defective genomes. Virology. **277:**235-249.

34. **Cook J.K.A. and A.P.A. Mockett.** Epidemiology of infectious bronchitis virus, p. 317-336. *In* S.G. Siddell (ed.), The Coronaviridae. Plenum Press, New York.

35. **Corse, E., and C. E. Machamer.** 2000. Infectious bronchitis virus E protein is targeted to the Golgi complex and directs release of virus-like particles. J. Virol. **74:**4319-4326.

36. **Corse, E., and C. E. Machamer.** 2002. The cytoplasmic tail of infectious bronchitis virus E protein directs Golgi targeting. J. Virol. **76:**1273-1284.

37. **Curtis, K. M., B. Yount, and R. S. Baric.** 2002. Heterologous gene expression from transmissible gastroenteritis virus replicon particles. J. Virol. **76:**1422-1434.

38. **Dathe, M., and T. Wieprecht.** 1999. Structural features of helical antimicrobial peptides: their potential to modulate activity on model membranes and biological cells. Biochim. Biophys. Acta. **1462:**71-87.

39. **Davies, H. A., R. R. Dourmashkin, and M. R. Macnaughton.** 1981. Ribonucleoprotein of avian infectious bronchitis virus. J. Gen. Virol. **53:**67-74.

40. **de Groot, R. J., W. Luytjes, M. C. Horzinek, B. A. van der Zeijst, W. J. Spaan, and J. A. Lenstra.** 1987. Evidence for a coiled-coil structure in the spike proteins of coronaviruses. J. Mol. Biol. **196:**963-966.

41. **De Groot, R. J., R. W. Van Leen, M. J. Dalderup, H. Vennema, M. C. Horzinek, and W. J. Spaan.** 1989. Stably expressed FIPV peplomer protein induces cell fusion and elicits neutralizing antibodies in mice. Virology. **171:**493-502.

42. **de Haan, C. A., M. de Wit, L. Kuo, C. Montalto-Morrison, B. L. Haagmans, S. R. Weiss, P. S. Masters, and P. J. Rottier.** 2003. The glycosylation status of the murine hepatitis coronavirus M protein affects the

interferogenic capacity of the virus in vitro and its ability to replicate in the liver but not the brain. Virology. **312:**395-406.

43. **de Haan, C. A., L. Kuo, P. S. Masters, H. Vennema, and P. J. Rottier.** 1998. Coronavirus particle assembly: primary structure requirements of the membrane protein. J. Virol. **72:**6838-6850.

44. **de Haan, C. A., P. S. Masters, X. Shen, S. Weiss, and P. J. Rottier.** 2002. The group-specific murine coronavirus genes are not essential, but their deletion, by reverse genetics, is attenuating in the natural host. Virology. **296:**177-189.

45. **de Haan, C. A., P. Roestenberg, M. de Wit, A. A. de Vries, T. Nilsson, H. Vennema, and P. J. Rottier.** 1998. Structural requirements for O-glycosylation of the mouse hepatitis virus membrane protein. J. Biol. Chem. **273:**29905-29914.

46. **de Haan, C. A., M. Smeets, F. Vernooij, H. Vennema, and P. J. Rottier.** 1999. Mapping of the coronavirus membrane protein domains involved in interaction with the spike protein. J. Virol. **73:**7441-7452.

47. **de Haan, C. A., H. Vennema, and P. J. Rottier.** 2000. Assembly of the coronavirus envelope: homotypic interactions between the M proteins. J. Virol. **74:**4967-4978.

48. **Dea, S., A. J. Verbeek, and P. Tijssen.** 1990. Antigenic and genomic relationships among turkey and bovine enteric coronaviruses. J. Virol. **64:**3112-3118.

49. **Delmas, B., and H. Laude.** 1990. Assembly of coronavirus spike protein into trimers and its role in epitope expression. J. Virol. **64:**5367-5375.

50. **Doyle L.P., L. M. H.** 1946. A transmissible gastroenteritis in pigs. **108:**257-259.

51. **Drin, G., J. F. Casella, R. Gautier, T. Boehmer, T. U. Schwartz, and B. Antonny.** 2007. A general amphipathic alpha-helical motif for sensing membrane curvature. Nat. Struct. Mol. Biol. **14:**138-146.

52. **Drosten, C., S. Gunther, W. Preiser, S. van der Werf, H. R. Brodt, S. Becker, H. Rabenau, M. Panning, L. Kolesnikova, R. A. Fouchier, A. Berger, A. M. Burguiere, J. Cinatl, M. Eickmann, N. Escriou, K. Grywna, S. Kramme, J. C. Manuguerra, S. Muller, V. Rickerts, M. Sturmer, S. Vieth, H. D. Klenk, A. D. Osterhaus, H. Schmitz, and H. W. Doerr.** 2003. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. N. Engl. J. Med. **348:**1967-1976.

53. **Dunne, S. J., R. B. Cornell, J. E. Johnson, N. R. Glover, and A. S. Tracey.** 1996. Structure of the membrane binding domain of CTP:phosphocholine cytidylyltransferase. Biochemistry. **35:**11975-11984.

54. **Dveksler, G. S., M. N. Pensiero, C. B. Cardellichio, R. K. Williams, G. S. Jiang, K. V. Holmes, and C. W. Dieffenbach.** 1991. Cloning of the mouse hepatitis virus (MHV) receptor: expression in human and hamster cell lines confers susceptibility to MHV. J. Virol. **65:**6881-6891.

55. **Egloff, M. P., F. Ferron, V. Campanacci, S. Longhi, C. Rancurel, H. Dutartre, E. J. Snijder, A. E. Gorbalenya, C. Cambillau, and B. Canard.** 2004. The severe acute respiratory syndrome-coronavirus replicative protein nsp9 is a single-stranded RNA-binding subunit unique in the RNA virus world. Proc. Natl. Acad. Sci. U. S. A. **101:**3792-3796.

56. **Ekstrom, M., P. Liljestrom, and H. Garoff.** 1994. Membrane protein lateral interactions control Semliki Forest virus budding. EMBO J. **13:**1058-1064.

57. **El-Sahly, H. M., R. L. Atmar, W. P. Glezen, and S. B. Greenberg.** 2000. Spectrum of clinical illness in hospitalized patients with "common cold" virus infections. Clin. Infect. Dis. **31:**96-100.

58. **Escors, D., J. Ortego, H. Laude, and L. Enjuanes.** 2001. The membrane M protein carboxy terminus binds to transmissible gastroenteritis coronavirus core and contributes to core stability. J. Virol. **75:**1312-1324.

59. **Fan, H., A. Ooi, Y. W. Tan, S. Wang, S. Fang, D. X. Liu, and J. Lescar.** 2005. The nucleocapsid protein of coronavirus infectious bronchitis virus: crystal structure of its N-terminal domain and multimerization properties. Structure. **13:**1859-1868.

60. **Farsad, K., N. Ringstad, K. Takei, S. R. Floyd, K. Rose, and P. De Camilli.** 2001. Generation of high curvature membranes mediated by direct endophilin bilayer interactions. J. Cell Biol. **155:**193-200.

61. **Fischer, F., C. F. Stegen, P. S. Masters, and W. A. Samsonoff.** 1998. Analysis of constructed E gene mutants of mouse hepatitis virus confirms a pivotal role for E protein in coronavirus assembly. J. Virol. **72:**7885-7894.

62. **Fleming, J. O., S. A. Stohlman, R. C. Harmon, M. M. Lai, J. A. Frelinger, and L. P. Weiner.** 1983. Antigenic relationships of murine coronaviruses: analysis using monoclonal antibodies to JHM (MHV-4) virus. Virology. **131:**296-307.

63. **Ford, M. G., I. G. Mills, B. J. Peter, Y. Vallis, G. J. Praefcke, P. R. Evans, and H. T. McMahon.** 2002. Curvature of clathrin-coated pits driven by epsin. Nature. **419:**361-366.

64. **Fosmire, J. A., K. Hwang, and S. Makino.** 1992. Identification and characterization of a coronavirus packaging signal. J. Virol. **66:**3522-3530.

65. **Gagneten, S., O. Gout, M. Dubois-Dalcq, P. Rottier, J. Rossen, and K. V. Holmes.** 1995. Interaction of mouse hepatitis virus (MHV) spike glycoprotein with receptor glycoprotein MHVR is required for infection with an MHV strain that expresses the hemagglutinin-esterase glycoprotein. J. Virol. **69:**889-895.

66. **Gallagher, T. M., and M. J. Buchmeier.** 2001. Coronavirus spike proteins in viral entry and pathogenesis. Virology. **279:**371-374.

67. **Gallagher, T. M., C. Escarmis, and M. J. Buchmeier.** 1991. Alteration of the pH dependence of coronavirus-induced cell fusion: effect of mutations in the spike glycoprotein. J. Virol. **65:**1916-1928.

68. **Godet, M., R. L'Haridon, J. F. Vautherot, and H. Laude.** 1992. TGEV corona virus ORF4 encodes a membrane protein that is incorporated into virions. Virology. **188:**666-675.

69. **Goldsmith, C. S., K. M. Tatti, T. G. Ksiazek, P. E. Rollin, J. A. Comer, W. W. Lee, P. A. Rota, B. Bankamp, W. J. Bellini, and S. R. Zaki.** 2004. Ultrastructural characterization of SARS coronavirus. Emerg. Infect. Dis. **10:**320-326.

70. **Gombold, J. L., S. T. Hingley, and S. R. Weiss.** 1993. Fusion-defective mutants of mouse hepatitis virus A59 contain a mutation in the spike protein cleavage signal. J. Virol. **67:**4504-4512.

71. **Gosert, R., A. Kanjanahaluethai, D. Egger, K. Bienz, and S. C. Baker.** 2002. RNA replication of mouse hepatitis virus takes place at double-membrane vesicles. J. Virol. **76:**3697-3708.

72. **Groseth, A., S. Wolff, T. Strecker, T. Hoenen, and S. Becker.** 2010. Efficient budding of the tacaribe virus matrix protein z requires the nucleoprotein. J. Virol. **84:**3603-3611.

73. **Hamming, I., W. Timens, M. L. Bulthuis, A. T. Lely, G. J. Navis, and H. van Goor.** 2004. Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first step in understanding SARS pathogenesis. J. Pathol. **203:**631-637.

74. **Hamre, D., and J. J. Procknow.** 1966. A new virus isolated from the human respiratory tract. Proc. Soc. Exp. Biol. Med. **121:**190-193.

75. **Hofmann, H., K. Pyrc, L. van der Hoek, M. Geier, B. Berkhout, and S. Pohlmann.** 2005. Human coronavirus NL63 employs the severe acute respiratory syndrome coronavirus receptor for cellular entry. Proc. Natl. Acad. Sci. U. S. A. **102:**7988-7993.

76. **Hogue, B. G., T. E. Kienzle, and D. A. Brian.** 1989. Synthesis and processing of the bovine enteric coronavirus haemagglutinin protein. J. Gen. Virol. **70 (Pt 2):**345-352.

77. **Hogue, B. G., and C. E. Machamer.** 2008. Coronavirus structural proteins and virus assembly, p. 179. *In* Perlman S., Gallagher T., and Snijder E. (ed.), Nidoviruses. ASM Press, Washington, DC.

78. **Holmes, K. V., R. M. Welsh, and M. V. Haspel.** 1986. Natural cytotoxicity against mouse hepatitis virus-infected target cells. I. Correlation of cytotoxicity with virus binding to leukocytes. J. Immunol. **136:**1446-1453.

79. **Hristova, K., W. C. Wimley, V. K. Mishra, G. M. Anantharamiah, J. P. Segrest, and S. H. White.** 1999. An amphipathic alpha-helix at a membrane interface: a structural study using a novel X-ray diffraction method. J. Mol. Biol. **290:**99-117.

80. **Hsue, B., T. Hartshorne, and P. S. Masters.** 2000. Characterization of an essential RNA secondary structure in the 3' untranslated region of the murine coronavirus genome. J. Virol. **74:**6911-6921.

81. **Huang, Q., L. Yu, A. M. Petros, A. Gunasekera, Z. Liu, N. Xu, P. Hajduk, J. Mack, S. W. Fesik, and E. T. Olejniczak.** 2004. Structure of the N-terminal RNA-binding domain of the SARS CoV nucleocapsid protein. Biochemistry. **43:**6059-6063.

82. **Hurst, K. R., C. A. Koetzner, and P. S. Masters.** 2009. Identification of in vivo-interacting domains of the murine coronavirus nucleocapsid protein. J. Virol. **83:**7221-7234.

83. **Hurst, K. R., L. Kuo, C. A. Koetzner, R. Ye, B. Hsue, and P. S. Masters.** 2005. A major determinant for membrane protein interaction localizes to the carboxy-terminal domain of the mouse coronavirus nucleocapsid protein. J. Virol. **79:**13285-13297.

84. **Ivanov, K. A., T. Hertzig, M. Rozanov, S. Bayer, V. Thiel, A. E. Gorbalenya, and J. Ziebuhr.** 2004. Major genetic marker of nidoviruses encodes a replicative endoribonuclease. Proc. Natl. Acad. Sci. U. S. A. **101:**12694-12699.

85. **Jao, C. C., A. Der-Sarkissian, J. Chen, and R. Langen.** 2004. Structure of membrane-bound alpha-synuclein studied by site-directed spin labeling. Proc. Natl. Acad. Sci. U. S. A. **101:**8331-8336.

86. **Jayaram, H., H. Fan, B. R. Bowman, A. Ooi, J. Jayaram, E. W. Collisson, J. Lescar, and B. V. Prasad.** 2006. X-ray structures of the N- and C-terminal domains of a coronavirus nucleocapsid protein: implications for nucleocapsid formation. J. Virol. **80:**6612-6620.

87. **Jones, D. T.** 1999. Protein secondary structure prediction based on position-specific scoring matrices. J. Mol. Biol. **292:**195-202.

88. **Kang, H., K. Bhardwaj, Y. Li, S. Palaninathan, J. Sacchettini, L. Guarino, J. L. Leibowitz, and C. C. Kao.** 2007. Biochemical and genetic analyses of murine hepatitis virus Nsp15 endoribonuclease. J. Virol. **81:**13587-13597.

89. **Kapke, P. A., F. Y. Tung, B. G. Hogue, D. A. Brian, R. D. Woods, and R. Wesley.** 1988. The amino-terminal signal peptide on the porcine transmissible gastroenteritis coronavirus matrix protein is not an absolute requirement for membrane translocation and glycosylation. Virology. **165:**367-376.

90. **Kazi, L., A. Lissenberg, R. Watson, R. J. de Groot, and S. R. Weiss.** 2005. Expression of hemagglutinin esterase protein from recombinant mouse hepatitis virus enhances neurovirulence. J. Virol. **79:**15064-15073.

91. **Klumperman, J., J. K. Locker, A. Meijer, M. C. Horzinek, H. J. Geuze, and P. J. Rottier.** 1994. Coronavirus M proteins accumulate in the Golgi complex beyond the site of virion budding. J. Virol. **68:**6523-6534.

92. **Komla-Soukha, I., and C. Sureau.** 2006. A tryptophan-rich motif in the carboxyl terminus of the small envelope protein of hepatitis B virus is central to the assembly of hepatitis delta virus particles. J. Virol. **80:**4648-4655.

93. **Krijnse-Locker, J., M. Ericsson, P. J. Rottier, and G. Griffiths.** 1994. Characterization of the budding compartment of mouse hepatitis virus: evidence that transport from the RER to the Golgi complex requires only one vesicular transport step. J. Cell Biol. **124:**55-70.

94. **Kunkel, F., and G. Herrler.** 1993. Structural and functional analysis of the surface protein of human coronavirus OC43. Virology. **195:**195-202.

95. **Kuo, L., and P. S. Masters.** 2002. 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.

96. **Kuo, L., and P. S. Masters.** 2003. The small envelope protein E is not essential for murine coronavirus replication. J. Virol. **77:**4597-4608.

97. **Lai M.M. and K.V. Holmes.** 2001. Coronaviridae: The viruses and their replication, p. 641-663. *In* D.M. Knipe and P.M. Howley (ed.), Fundamental Virology. Lippincott Williams and Wilkins, Philadelphia.

98. **Lai M.M., S. Perlman, and L.J. Anderson.** 2007. The Coronaviridae, p. 1305-1335. *In* D.M. Knipe and P.M. Howley (ed.), Fields Virology, Fifth Edition ed., . Lippincott Williams and Wilkins, Philadelphia.

99. **Lai, M. M.** 1986. Coronavirus leader-RNA-primed transcription: an alternative mechanism to RNA splicing. Bioessays. **5:**257-260.

100. **Lai, M. M., R. S. Baric, P. R. Brayton, and S. A. Stohlman.** 1984. Characterization of leader RNA sequences on the virion and mRNAs of mouse hepatitis virus, a cytoplasmic RNA virus. Proc. Natl. Acad. Sci. U. S. A. **81:**3626-3630.

101. **Lai, M. M., and D. Cavanagh.** 1997. The molecular biology of coronaviruses. Adv. Virus Res. **48:**1-100.

102. **Landolt-Marticorena, C., K. A. Williams, C. M. Deber, and R. A. Reithmeier.** 1993. Non-random distribution of amino acids in the transmembrane segments of human type I single span membrane proteins. J. Mol. Biol. **229:**602-608.

103. **Larkin, M. A., G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson, T. J. Gibson, and D. G. Higgins.** 2007. Clustal W and Clustal X version 2.0. Bioinformatics. **23:**2947-2948.

104. **Lau, S. K., P. C. Woo, K. S. Li, Y. Huang, H. W. Tsoi, B. H. Wong, S. S. Wong, S. Y. Leung, K. H. Chan, and K. Y. Yuen.** 2005. Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. Proc. Natl. Acad. Sci. U. S. A. **102:**14040-14045.

105. **Lau, S. K., P. C. Woo, C. C. Yip, H. Tse, H. W. Tsoi, V. C. Cheng, P. Lee, B. S. Tang, C. H. Cheung, R. A. Lee, L. Y. So, Y. L. Lau, K. H. Chan, and K. Y. Yuen.** 2006. Coronavirus HKU1 and other coronavirus infections in Hong Kong. J. Clin. Microbiol. **44:**2063-2071.

106. **Laude H. and P.S. Masters.** The coronavirus nucleocapsid protein, p. 141-164. *In* S.G. Siddell (ed.), The Coronaviridae. Plenum Press, New York.

107. **Lee, H., and D. T. Brown.** 1994. Mutations in an exposed domain of Sindbis virus capsid protein result in the production of noninfectious virions and morphological variants. Virology. **202:**390-400.

108. **Lee, M. C., L. Orci, S. Hamamoto, E. Futai, M. Ravazzola, and R. Schekman.** 2005. Sar1p N-terminal helix initiates membrane curvature and completes the fission of a COPII vesicle. Cell. **122:**605-617.

109. **Lee, S., K. E. Owen, H. K. Choi, H. Lee, G. Lu, G. Wengler, D. T. Brown, M. G. Rossmann, and R. J. Kuhn.** 1996. Identification of a protein binding site on the surface of the alphavirus nucleocapsid and its implication in virus assembly. Structure. **4:**531-541.

110. **Lewicki, D. N., and T. M. Gallagher.** 2002. Quaternary structure of coronavirus spikes in complex with carcinoembryonic antigen-related cell adhesion molecule cellular receptors. J. Biol. Chem. **277:**19727-19734.

111. **Li, D., and D. Cavanagh.** 1992. Coronavirus IBV-induced membrane fusion occurs at near-neutral pH. Arch. Virol. **122:**307-316.

112. **Li, M., P. T. Schmitt, Z. Li, T. S. McCrory, B. He, and A. P. Schmitt.** 2009. Mumps virus matrix, fusion, and nucleocapsid proteins cooperate for efficient production of virus-like particles. J. Virol. **83:**7261-7272.

113. **Li, W., M. J. Moore, N. Vasilieva, J. Sui, S. K. Wong, M. A. Berne, M. Somasundaran, J. L. Sullivan, K. Luzuriaga, T. C. Greenough, H. Choe, and M. Farzan.** 2003. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. Nature. **426:**450-454.

114. **Li, W., Z. Shi, M. Yu, W. Ren, C. Smith, J. H. Epstein, H. Wang, G. Crameri, Z. Hu, H. Zhang, J. Zhang, J. McEachern, H. Field, P. Daszak, B. T. Eaton, S. Zhang, and L. F. Wang.** 2005. Bats are natural reservoirs of SARS-like coronaviruses. Science. **310:**676-679.

115. **Licata, J. M., R. F. Johnson, Z. Han, and R. N. Harty.** 2004. Contribution of ebola virus glycoprotein, nucleoprotein, and VP24 to budding of VP40 virus-like particles. J. Virol. **78:**7344-7351.

116. **Liu, D. X., and S. C. Inglis.** 1991. Association of the infectious bronchitis virus 3c protein with the virion envelope. Virology. **185:**911-917.

117. **Locker, J. K., D. J. Opstelten, M. Ericsson, M. C. Horzinek, and P. J. Rottier.** 1995. Oligomerization of a trans-Golgi/trans-Golgi network retained protein occurs in the Golgi complex and may be part of its retention. J. Biol. Chem. **270:**8815-8821.

118. **Lopez, L. A., A. J. Riffle, S. L. Pike, D. Gardner, and B. G. Hogue.** 2008. Importance of conserved cysteine residues in the coronavirus envelope protein. J. Virol. **82:**3000-3010.

119. **Lopez, S., J. S. Yao, R. J. Kuhn, E. G. Strauss, and J. H. Strauss.** 1994. Nucleocapsid-glycoprotein interactions required for assembly of alphaviruses. J. Virol. **68:**1316-1323.

120. **Luo, H., D. Wu, C. Shen, K. Chen, X. Shen, and H. Jiang.** 2006. Severe acute respiratory syndrome coronavirus membrane protein interacts with nucleocapsid protein mostly through their carboxyl termini by electrostatic attraction. Int. J. Biochem. Cell Biol. **38:**589-599.

121. **Luo, Z., and S. R. Weiss.** 1998. Roles in cell-to-cell fusion of two conserved hydrophobic regions in the murine coronavirus spike protein. Virology. **244:**483-494.

122. **Luytjes, W., P. J. Bredenbeek, A. F. Noten, M. C. Horzinek, and W. J. Spaan.** 1988. Sequence of mouse hepatitis virus A59 mRNA 2: indications for RNA recombination between coronaviruses and influenza C virus. Virology. **166:**415-422.

123. **Macneughton, M. R., and H. A. Davies.** 1978. Ribonucleoprotein-like structures from coronavirus particles. J. Gen. Virol. **39:**545-549.

124. **Madan, V., J. Garcia Mde, M. A. Sanz, and L. Carrasco.** 2005. Viroporin activity of murine hepatitis virus E protein. FEBS Lett. **579:**3607-3612.

125. **Maeda, J., A. Maeda, and S. Makino.** 1999. Release of coronavirus E protein in membrane vesicles from virus-infected cells and E protein-expressing cells. Virology. **263:**265-272.

126. **Maeda, J., J. F. Repass, A. Maeda, and S. Makino.** 2001. Membrane topology of coronavirus E protein. Virology. **281:**163-169.

127. **Makino, S., and M. M. Lai.** 1989. High-frequency leader sequence switching during coronavirus defective interfering RNA replication. J. Virol. **63:**5285-5292.

128. **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.

129. **McBride, C. E., and C. E. Machamer.** 2010. A single tyrosine in the severe acute respiratory syndrome coronavirus membrane protein cytoplasmic tail is important for efficient interaction with spike protein. J. Virol. **84:**1891-1901.

130. **McIntosh, K., J. H. Dees, W. B. Becker, A. Z. Kapikian, and R. M. Chanock.** 1967. Recovery in tracheal organ cultures of novel viruses from patients with respiratory disease. Proc. Natl. Acad. Sci. U. S. A. **57:**933-940.

131. **Mohandas, D. V., and S. Dales.** 1991. Endosomal association of a protein phosphatase with high dephosphorylating activity against a coronavirus nucleocapsid protein. FEBS Lett. **282:**419-424.

132. **Molenkamp, R., and W. J. Spaan.** 1997. Identification of a specific interaction between the coronavirus mouse hepatitis virus A59 nucleocapsid protein and packaging signal. Virology. **239:**78-86.

133. **Nakajima, A., H. Iijima, M. F. Neurath, T. Nagaishi, E. E. Nieuwenhuis, R. Raychowdhury, J. Glickman, D. M. Blau, S. Russell, K. V. Holmes, and R. S. Blumberg.** 2002. Activation-induced expression of carcinoembryonic antigen-cell adhesion molecule 1 regulates mouse T lymphocyte function. J. Immunol. **168:**1028-1035.

134. **Narayanan, K., A. Maeda, J. Maeda, and S. Makino.** 2000. Characterization of the coronavirus M protein and nucleocapsid interaction in infected cells. J. Virol. **74:**8127-8134.

135. **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.

136. **Nelson, G. W., and S. A. Stohlman.** 1993. Localization of the RNA-binding domain of mouse hepatitis virus nucleocapsid protein. J. Gen. Virol. **74 (Pt 9):**1975-1979.

148

137. **Nelson, G. W., S. A. Stohlman, and S. M. Tahara.** 2000. High affinity interaction between nucleocapsid protein and leader/intergenic sequence of mouse hepatitis virus RNA. J. Gen. Virol. **81:**181-188.

138. **Neuman, B. W., B. D. Adair, C. Yoshioka, J. D. Quispe, G. Orca, P. Kuhn, R. A. Milligan, M. Yeager, and M. J. Buchmeier.** 2006. Supramolecular architecture of severe acute respiratory syndrome coronavirus revealed by electron cryomicroscopy. J. Virol. **80:**7918-7928.

139. **Nguyen, V. P., and B. G. Hogue.** 1997. Protein interactions during coronavirus assembly. J. Virol. **71:**9278-9284.

140. **Niwa, H., K. Yamamura, and J. Miyazaki.** 1991. Efficient selection for high-expression transfectants with a novel eukaryotic vector. Gene. **108:**193-199.

141. **Oleszak, E. L., S. Perlman, and J. L. Leibowitz.** 1992. MHV S peplomer protein expressed by a recombinant vaccinia virus vector exhibits IgG Fc-receptor activity. Virology. **186:**122-132.

142. **Oostra, M., C. A. de Haan, R. J. de Groot, and P. J. Rottier.** 2006. Glycosylation of the severe acute respiratory syndrome coronavirus triple-spanning membrane proteins 3a and M. J. Virol. **80:**2326-2336.

143. **Opstelten, D. J., M. J. Raamsman, K. Wolfs, M. C. Horzinek, and P. J. Rottier.** 1995. Envelope glycoprotein interactions in coronavirus assembly. J. Cell Biol. **131:**339-349.

144. **Ortego, J., D. Escors, H. Laude, and L. Enjuanes.** 2002. Generation of a replication-competent, propagation-deficient virus vector based on the transmissible gastroenteritis coronavirus genome. J. Virol. **76:**11518-11529.

145. **Owen, K. E., and R. J. Kuhn.** 1997. Alphavirus budding is dependent on the interaction between the nucleocapsid and hydrophobic amino acids on the cytoplasmic domain of the E2 envelope glycoprotein. Virology. **230:**187-196.

146. **Parker, M. M., and P. S. Masters.** 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.

147. **Payne, H. R., and J. Storz.** 1988. Analysis of cell fusion induced by bovine coronavirus infection. Arch. Virol. **103:**27-33.

148. **Pensaert M.B.** 1999. Porcine epidemic diarrhea, p. 179-185. *In* Straw B.D., S. D'Allaire, W.L. Mengeling, D.I. Taylor (ed.), Diseases of Swine, 8th ed., . Iowa State University Press.

149. **Raamsman, M. J., J. K. Locker, A. de Hooge, A. A. de Vries, G. Griffiths, H. Vennema, and P. J. Rottier.** 2000. Characterization of the coronavirus mouse hepatitis virus strain A59 small membrane protein E. J. Virol. **74:**2333-2342.

150. **Raman, S., P. Bouma, G. D. Williams, and D. A. Brian.** 2003. Stem-loop III in the 5' untranslated region is a cis-acting element in bovine coronavirus defective interfering RNA replication. J. Virol. **77:**6720-6730.

151. **Ricagno, S., M. P. Egloff, R. Ulferts, B. Coutard, D. Nurizzo, V. Campanacci, C. Cambillau, J. Ziebuhr, and B. Canard.** 2006. Crystal structure and mechanistic determinants of SARS coronavirus nonstructural protein 15 define an endoribonuclease family. Proc. Natl. Acad. Sci. U. S. A. **103:**11892-11897.

152. **Risco, C., I. M. Anton, L. Enjuanes, and J. L. Carrascosa.** 1996. The transmissible gastroenteritis coronavirus contains a spherical core shell consisting of M and N proteins. J. Virol. **70:**4773-4777.

153. **Risco, C., I. M. Anton, C. Sune, A. M. Pedregosa, J. M. Martin-Alonso, F. Parra, J. L. Carrascosa, and L. Enjuanes.** 1995. Membrane protein molecules of transmissible gastroenteritis coronavirus also expose the carboxy-terminal region on the external surface of the virion. J. Virol. **69:**5269-5277.

154. **Rota, P. A., M. S. Oberste, S. S. Monroe, W. A. Nix, R. Campagnoli, J. P. Icenogle, S. Penaranda, B. Bankamp, K. Maher, M. H. Chen, S. Tong, A. Tamin, L. Lowe, M. Frace, J. L. DeRisi, Q. Chen, D. Wang, D. D. Erdman, T. C. Peret, C. Burns, T. G. Ksiazek, P. E. Rollin, A. Sanchez, S. Liffick, B. Holloway, J. Limor, K. McCaustland, M. Olsen-Rasmussen, R. Fouchier, S. Gunther, A. D. Osterhaus, C. Drosten, M. A. Pallansch, L. J. Anderson, and W. J. Bellini.** 2003. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. Science. **300:**1394-1399.

155. **Rottier, P., D. Brandenburg, J. Armstrong, B. van der Zeijst, and G. Warren.** 1984. Assembly in vitro of a spanning membrane protein of the endoplasmic reticulum: the E1 glycoprotein of coronavirus mouse hepatitis virus A59. Proc. Natl. Acad. Sci. U. S. A. **81:**1421-1425.

156. **Rottier, P. J., G. W. Welling, S. Welling-Wester, H. G. Niesters, J. A. Lenstra, and B. A. Van der Zeijst.** 1986. Predicted membrane topology of the coronavirus protein E1. Biochemistry. **25:**1335-1339.

157. **Routledge, E., R. Stauber, M. Pfleiderer, and S. G. Siddell.** 1991. Analysis of murine coronavirus surface glycoprotein functions by using monoclonal antibodies. J. Virol. **65:**254-262.

158. **Saif L.J and R.D. Wesley.** Transmissible gastroenteritis and porcine respiratory coronavirus, p. 295-325. *In* Straw B.E., S. D'Allaire, W.L. Mengeling, D.I. Taylor (ed.), Diseases of Swine. Iowa State University Press.

159. **Saikatendu, K. S., J. S. Joseph, V. Subramanian, T. Clayton, M. Griffith, K. Moy, J. Velasquez, B. W. Neuman, M. J. Buchmeier, R. C. Stevens, and P. Kuhn.** 2005. Structural basis of severe acute respiratory syndrome coronavirus ADP-ribose-1"-phosphate dephosphorylation by a conserved domain of nsP3. Structure. **13:**1665-1675.

160. **Saikatendu, K. S., J. S. Joseph, V. Subramanian, B. W. Neuman, M. J. Buchmeier, R. C. Stevens, and P. Kuhn.** 2007. Ribonucleocapsid formation of severe acute respiratory syndrome coronavirus through molecular action of the N-terminal domain of N protein. J. Virol. **81:**3913-3921.

161. **Sawicki, S. G., and D. L. Sawicki.** 2005. Coronavirus transcription: a perspective. Curr. Top. Microbiol. Immunol. **287:**31-55.

162. **Schalk A.F., M. C. H.** 1931. An apparently new respiratory disease of baby chicks. J Am Vet Med Assoc. **78:**413-422.

163. **Schelle, B., N. Karl, B. Ludewig, S. G. Siddell, and V. Thiel.** 2005. Selective replication of coronavirus genomes that express nucleocapsid protein. J. Virol. **79:**6620-6630.

164. **Schmitt, A. P., G. P. Leser, D. L. Waning, and R. A. Lamb.** 2002. Requirements for budding of paramyxovirus simian virus 5 virus-like particles. J. Virol. **76:**3952-3964.

165. **Schultze, B., and G. Herrler.** 1992. Bovine coronavirus uses N-acetyl-9-O-acetylneuraminic acid as a receptor determinant to initiate the infection of cultured cells. J. Gen. Virol. **73 (Pt 4):**901-906.

166. **Seybert, A., A. Hegyi, S. G. Siddell, and J. Ziebuhr.** 2000. The human coronavirus 229E superfamily 1 helicase has RNA and DNA duplex-unwinding activities with 5'-to-3' polarity. RNA. **6:**1056-1068.

167. **Shapiro, L. H., R. A. Ashmun, W. M. Roberts, and A. T. Look.** 1991. Separate promoters control transcription of the human aminopeptidase N gene in myeloid and intestinal epithelial cells. J. Biol. Chem. **266:**11999-12007.

168. **Siddel S.G.** 1995. The small-membrane protein, p. 181-190. *In* S.G. Siddell (ed.), The Coronaviridae. Plenum Press, New York.

169. **Siddell S.G. (ed.),** 1995. The Coronaviridae. Plenum Press, New York.

170. **Simmons, G., D. N. Gosalia, A. J. Rennekamp, J. D. Reeves, S. L. Diamond, and P. Bates.** 2005. Inhibitors of cathepsin L prevent severe acute respiratory syndrome coronavirus entry. Proc. Natl. Acad. Sci. U. S. A. **102:**11876-11881.

171. **Sims, A. C., J. Ostermann, and M. R. Denison.** 2000. Mouse hepatitis virus replicase proteins associate with two distinct populations of intracellular membranes. J. Virol. **74:**5647-5654.

172. **Siu, Y. L., K. T. Teoh, J. Lo, C. M. Chan, F. Kien, N. Escriou, S. W. Tsao, J. M. Nicholls, R. Altmeyer, J. S. Peiris, R. Bruzzone, and B. Nal.** 2008. The M, E, and N structural proteins of the severe acute respiratory syndrome coronavirus are required for efficient assembly, trafficking, and release of virus-like particles. J. Virol. **82:**11318-11330.

173. **Sloots, T. P., P. McErlean, D. J. Speicher, K. E. Arden, M. D. Nissen, and I. M. Mackay.** 2006. Evidence of human coronavirus HKU1 and human bocavirus in Australian children. J. Clin. Virol. **35:**99-102.

174. **Snijder, E. J., P. J. Bredenbeek, J. C. Dobbe, V. Thiel, J. Ziebuhr, L. L. Poon, Y. Guan, M. Rozanov, W. J. Spaan, and A. E. Gorbalenya.** 2003. Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. J. Mol. Biol. **331:**991-1004.

175. **Snijder, E. J., Y. van der Meer, J. Zevenhoven-Dobbe, J. J. Onderwater, J. van der Meulen, H. K. Koerten, and A. M. Mommaas.** 2006. Ultrastructure and origin of membrane vesicles associated with the severe acute respiratory syndrome coronavirus replication complex. J. Virol. **80:**5927-5940.

176. **Spaan W.J., D. Cavanagh, M.C. Horzinek.** 1990. Coronaviruses, p. 359-379. *In* A. R. N. M.H.V. Regenmortel (ed.), Immunochemistry of Viruses, The Basis of Serodiagnosis and Vaccinesvol. 2. Elsevier, Amsterdam.

177. **Spaan, W. J., D. Cavanagh, and R.J. de Groot, L. Enjuanes, A.E. Gorbalenya, E.J. Snijder, P.J. Walker.** 2005. Order Nidovirales, p. 937-945. *In*

C.M. Fauquet, M.A. Mayo, J. Maniloff, U. Desselberger, L.A. Ball (ed.), Virus Taxonomy, Eighth Report of the International Committee on Taxonomy of Viruses. Elsevier, Academic Press, Amsterdam.

178. **Spaan, W., H. Delius, M. Skinner, J. Armstrong, P. Rottier, S. Smeekens, B. A. van der Zeijst, and S. G. Siddell.** 1983. Coronavirus mRNA synthesis involves fusion of non-contiguous sequences. EMBO J. **2:**1839-1844.

179. **Stohlman, S. A., J. O. Fleming, C. D. Patton, and M. M. Lai.** 1983. Synthesis and subcellular localization of the murine coronavirus nucleocapsid protein. Virology. **130:**527-532.

180. **Sturman, L. S., and K. V. Holmes.** 1977. Characterization of coronavirus II. Glycoproteins of the viral envelope: tryptic peptide analysis. Virology. **77:**650-660.

181. **Sturman, L. S., C. S. Ricard, and K. V. Holmes.** 1990. Conformational change of the coronavirus peplomer glycoprotein at pH 8.0 and 37 degrees C correlates with virus aggregation and virus-induced cell fusion. J. Virol. **64:**3042-3050.

182. **Surjit, M., R. Kumar, R. N. Mishra, M. K. Reddy, V. T. Chow, and S. K. Lal.** 2005. The severe acute respiratory syndrome coronavirus nucleocapsid protein is phosphorylated and localizes in the cytoplasm by 14-3-3-mediated translocation. J. Virol. **79:**11476-11486.

183. **Takeda, M., C. K. Chang, T. Ikeya, P. Guntert, Y. H. Chang, Y. L. Hsu, T. H. Huang, and M. Kainosho.** 2008. Solution structure of the c-terminal dimerization domain of SARS coronavirus nucleocapsid protein solved by the SAIL-NMR method. J. Mol. Biol. **380:**608-622.

184. **Thiel, V., J. Herold, B. Schelle, and S. G. Siddell.** 2001. Infectious RNA transcribed in vitro from a cDNA copy of the human coronavirus genome cloned in vaccinia virus. J. Gen. Virol. **82:**1273-1281.

185. **Thiel, V., and S. G. Siddell.** 1994. Internal ribosome entry in the coding region of murine hepatitis virus mRNA 5. J. Gen. Virol. **75 (Pt 11):**3041-3046.

186. **Thorp, E. B., J. A. Boscarino, H. L. Logan, J. T. Goletz, and T. M. Gallagher.** 2006. Palmitoylations on murine coronavirus spike proteins are essential for virion assembly and infectivity. J. Virol. **80:**1280-1289.

187. **Tooze, J., S. Tooze, and G. Warren.** 1984. Replication of coronavirus MHV-A59 in sac- cells: determination of the first site of budding of progeny virions. Eur. J. Cell Biol. **33:**281-293.

188. **Tooze, J., and S. A. Tooze.** 1985. Infection of AtT20 murine pituitary tumour cells by mouse hepatitis virus strain A59: virus budding is restricted to the Golgi region. Eur. J. Cell Biol. **37:**203-212.

189. **Turner, B. C., E. M. Hemmila, N. Beauchemin, and K. V. Holmes.** 2004. Receptor-dependent coronavirus infection of dendritic cells. J. Virol. **78:**5486-5490.

190. **Tyrrell D.A.J, J.D. Almeida, D.M. Berry, et al.** 1968. Coronaviruses. Nature. **220:**650.

191. **Tyrrell, D. A., and M. L. Bynoe.** 1965. Cultivation of a Novel Type of Common-Cold Virus in Organ Cultures. Br. Med. J. **1:**1467-1470.

192. **Urata, S., T. Noda, Y. Kawaoka, S. Morikawa, H. Yokosawa, and J. Yasuda.** 2007. Interaction of Tsg101 with Marburg virus VP40 depends on the PPPY motif, but not the PT/SAP motif as in the case of Ebola virus, and Tsg101 plays a critical role in the budding of Marburg virus-like particles induced by VP40, NP, and GP. J. Virol. **81:**4895-4899.

193. **van der Hoek, L., K. Pyrc, M. F. Jebbink, W. Vermeulen-Oost, R. J. Berkhout, K. C. Wolthers, P. M. Wertheim-van Dillen, J. Kaandorp, J. Spaargaren, and B. Berkhout.** 2004. Identification of a new human coronavirus. Nat. Med. **10:**368-373.

194. **Vennema, H., G. J. Godeke, J. W. Rossen, W. F. Voorhout, M. C. Horzinek, D. J. Opstelten, and P. J. Rottier.** 1996. Nucleocapsid-independent assembly of coronavirus-like particles by co-expression of viral envelope protein genes. EMBO J. **15:**2020-2028.

195. **Verma, S., V. Bednar, A. Blount, and B. G. Hogue.** 2006. Identification of functionally important negatively charged residues in the carboxy end of mouse hepatitis coronavirus A59 nucleocapsid protein. J. Virol. **80:**4344-4355.

196. **Verma, S., L. A. Lopez, V. Bednar, and B. G. Hogue.** 2007. Importance of the penultimate positive charge in mouse hepatitis coronavirus A59 membrane protein. J. Virol. **81:**5339-5348.

197. **Voss, D., S. Pfefferle, C. Drosten, L. Stevermann, E. Traggiai, A. Lanzavecchia, and S. Becker.** 2009. Studies on membrane topology, N-glycosylation and functionality of SARS-CoV membrane protein. Virol. J. **6:**79.

198. **Wang, C. J., P. J. Chen, J. C. Wu, D. Patel, and D. S. Chen.** 1991. Small-form hepatitis B surface antigen is sufficient to help in the assembly of hepatitis delta virus-like particles. J. Virol. **65:**6630-6636.

199. **Wang, J., S. Fang, H. Xiao, B. Chen, J. P. Tam, and D. X. Liu.** 2009. Interaction of the coronavirus infectious bronchitis virus membrane protein with beta-actin and its implication in virion assembly and budding. PLoS One. **4:**e4908.

200. **Wang, S. M., Y. F. Chang, Y. M. Chen, and C. T. Wang.** 2008. Severe acute respiratory syndrome coronavirus nucleocapsid protein confers ability to efficiently produce virus-like particles when substituted for the human immunodeficiency virus nucleocapsid domain. J. Biomed. Sci. **15:**719-729.

201. **West, J., and D. T. Brown.** 2006. Role of a conserved tripeptide in the endodomain of Sindbis virus glycoprotein E2 in virus assembly and function. J. Gen. Virol. **87:**657-664.

202. **White, T. C., Z. Yi, and B. G. Hogue.** 2007. Identification of mouse hepatitis coronavirus A59 nucleocapsid protein phosphorylation sites. Virus Res. **126:**139-148.

203. **Wilkinson, T. A., T. L. Tellinghuisen, R. J. Kuhn, and C. B. Post.** 2005. Association of sindbis virus capsid protein with phospholipid membranes and the E2 glycoprotein: implications for alphavirus assembly. Biochemistry. **44:**2800-2810.

204. **Williams, R. K., G. S. Jiang, and K. V. Holmes.** 1991. Receptor for mouse hepatitis virus is a member of the carcinoembryonic antigen family of glycoproteins. Proc. Natl. Acad. Sci. U. S. A. **88:**5533-5536.

205. **Wilson, L., P. Gage, and G. Ewart.** 2006. Hexamethylene amiloride blocks E protein ion channels and inhibits coronavirus replication. Virology. **353:**294-306.

206. **Wilson, L., C. McKinlay, P. Gage, and G. Ewart.** 2004. SARS coronavirus E protein forms cation-selective ion channels. Virology. **330:**322-331.

207. **Woo, P. C., S. K. Lau, C. M. Chu, K. H. Chan, H. W. Tsoi, Y. Huang, B. H. Wong, R. W. Poon, J. J. Cai, W. K. Luk, L. L. Poon, S. S. Wong, Y. Guan, J. S. Peiris, and K. Y. Yuen.** 2005. Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. J. Virol. **79:**884-895.

208. **Xu, H. Y., K. P. Lim, S. Shen, and D. X. Liu.** 2001. Further identification and characterization of novel intermediate and mature cleavage products released from the ORF 1b region of the avian coronavirus infectious bronchitis virus 1a/1b polyprotein. Virology. **288:**212-222.

209. **Xu, X., Y. Zhai, F. Sun, Z. Lou, D. Su, Y. Xu, R. Zhang, A. Joachimiak, X. C. Zhang, M. Bartlam, and Z. Rao.** 2006. New antiviral target revealed by the hexameric structure of mouse hepatitis virus nonstructural protein nsp15. J. Virol. **80:**7909-7917.

210. **Ye Y.** 2007. Multifunctional Facets of Coronavirus Proteins: Viroporins to Inate Immunity. Ph.D. Arizona State University.

211. **Ye, R., C. Montalto-Morrison, and P. S. Masters.** 2004. Genetic analysis of determinants for spike glycoprotein assembly into murine coronavirus virions: distinct roles for charge-rich and cysteine-rich regions of the endodomain. J. Virol. **78:**9904-9917.

212. **Yokomori, K., L. R. Banner, and M. M. Lai.** 1991. Heterogeneity of gene expression of the hemagglutinin-esterase (HE) protein of murine coronaviruses. Virology. **183:**647-657.

213. **Youn, S., J. L. Leibowitz, and E. W. Collisson.** 2005. In vitro assembled, recombinant infectious bronchitis viruses demonstrate that the 5a open reading frame is not essential for replication. Virology. **332:**206-215.

214. **Yount, B., K. M. Curtis, and R. S. Baric.** 2000. Strategy for systematic assembly of large RNA and DNA genomes: transmissible gastroenteritis virus model. J. Virol. **74:**10600-10611.

215. **Yount, B., K. M. Curtis, E. A. Fritz, L. E. Hensley, P. B. Jahrling, E. Prentice, M. R. Denison, T. W. Geisbert, and R. S. Baric.** 2003. Reverse genetics with a full-length infectious cDNA of severe acute respiratory syndrome coronavirus. Proc. Natl. Acad. Sci. U. S. A. **100:**12995-13000.

216. **Yount, B., M. R. Denison, S. R. Weiss, and R. S. Baric.** 2002. Systematic assembly of a full-length infectious cDNA of mouse hepatitis virus strain A59. J. Virol. **76:**11065-11078.

217. **Yount, B., R. S. Roberts, A. C. Sims, D. Deming, M. B. Frieman, J. Sparks, M. R. Denison, N. Davis, and R. S. Baric.** 2005. Severe acute respiratory syndrome coronavirus group-specific open reading frames encode nonessential functions for replication in cell cultures and mice. J. Virol. **79:**14909-14922.

218. **Yu, I. M., M. L. Oldham, J. Zhang, and J. Chen.** 2006. Crystal structure of the severe acute respiratory syndrome (SARS) coronavirus nucleocapsid protein dimerization domain reveals evolutionary linkage between corona- and arteriviridae. J. Biol. Chem. **281:**17134-17139.

219. **Yu, X., W. Bi, S. R. Weiss, and J. L. Leibowitz.** 1994. Mouse hepatitis virus gene 5b protein is a new virion envelope protein. Virology. **202:**1018-1023.

220. **Yuan, Q., Y. Liao, J. Torres, J. P. Tam, and D. X. Liu.** 2006. Biochemical evidence for the presence of mixed membrane topologies of the severe acute respiratory syndrome coronavirus envelope protein expressed in mammalian cells. FEBS Lett. **580:**3192-3200.

221. **Zelus, B. D., J. H. Schickli, D. M. Blau, S. R. Weiss, and K. V. Holmes.** 2003. Conformational changes in the spike glycoprotein of murine coronavirus are induced at 37 degrees C either by soluble murine CEACAM1 receptors or by pH 8. J. Virol. **77:**830-840.

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

223. **Ziebuhr, J.** 2005. The coronavirus replicase. Curr. Top. Microbiol. Immunol. **287:**57-94.

224. **Ziebuhr, J., and S. G. Siddell.** 1999. Processing of the human coronavirus 229E replicase polyproteins by the virus-encoded 3C-like proteinase: identification of proteolytic products and cleavage sites common to pp1a and pp1ab. J. Virol. **73:**177-185.

225. **Ziebuhr, J., E. J. Snijder, and A. E. Gorbalenya.** 2000. Virus-encoded proteinases and proteolytic processing in the Nidovirales. J. Gen. Virol. **81:**853-879.

226. **Zuker, M.** 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. **31:**3406-3415.

227. **Zuniga, S., I. Sola, J. L. Moreno, P. Sabella, J. Plana-Duran, and L. Enjuanes.** 2007. Coronavirus nucleocapsid protein is an RNA chaperone. Virology. **357:**215-227.

APPENDIX A

CONSTRUCTION OF SARS EXPRESSION PLASMIDS

DATA COLLECTED 2004-2005

The pcDNA3.1Zeo(-):SARS S plasmid contains the entire SARS S gene. The S gene fragment was obtained by PCR from a PCR-XL-TOPO:SARS S clone generated by Louisa Ruetz using primers SARS S NheI For (5' GCC GGC TAG CAT GTT TAT TTT CTT A 3') and SARS S XhoI Rev (5' GCG CCT CGA GTT ATG TGT AAT GTA A 3'). The PCR product was digested with NheI and XhoI enzymes and cloned into the pcDNA3.1Zeo(-) vector.

The pCAGGS:SARS S plasmid contains the entire SARS S gene. The S gene fragment was obtained by PCR from a pcDNA3.1Zeo(-):SARS S clone using primers SARS S SmaI For (5' GCC CCG GGA TGT TTA TTT TCT TA 3') and SARS S XmaI Rev (5' TCC CCC GGG TTA TGT GTA ATG TAA A 3'). The PCR product was subjected to digestion was SmaI and XmaI enzymes and cloned into the pCAGGS vector.

The pCAGGS:SARS 3a plasmid contains the entire SARS 3a gene. The 3a gene fragment was obtained from an incorrect pCAGGS:SARS 3a clone by digesting with KpnI and XhoI. The SARS 3a gene was then cloned into the pCAGGS vector. The pCAGGS:SARS HA-3a plasmid contains the entire SARS 3a gene with a 16 amino acid HA tag on the N-terminus. The 3a gene with the HA tag was obtained by PCR from a pCRUZ:SARS HA-3a clone using primers SARS 3a KpnI For (5' GCG GGT ACC ATG GAT TTG TTT ATG AG 3') and SARS 3A XhoI Rev (5' GTA TCT CGA GTT ACA AAG GCA CGC TAG 3'). The PCR product was subsequently digested with KpnI and XhoI and cloned into the pCAGGS vector.

160

All expression constructs were transiently expressed. The pcDNA3.1Zeo(-): SARS S clone was expressed in BHK-21 cells using vaccinia T7 polymerase, vTF7-3. The pCAGGS:SARS S clone was expressed in BHK-21 cells. The pCAGGS:SARS 3a and pCAGGS:SARS HA-3a clones were expressed in 293T cells. Protein expression was confirmed for all constructs using SDS-PAGE and Western blotting analysis (Fig.29).

To determine the cellular localization of the SARS 3a protein, indirect immunofluorescence was carried out. Previous attempts to determine the localization of 3a resulted in inconclusive results so the MHV M protein was used as a marker for Golgi localization. BHK-21 cells were transfected with the pCAGGS:MHV M and pCAGGS:SARS 3a clones and were fixed at 12 hpt. The MHV M protein colocalizes with the Golgi marker, Giantin (Fig.30A). The SARS 3a protein colocalizes with the MHV M protein (Fig.30B). Therefore, the SARS 3a was determined to localize in the Golgi region.
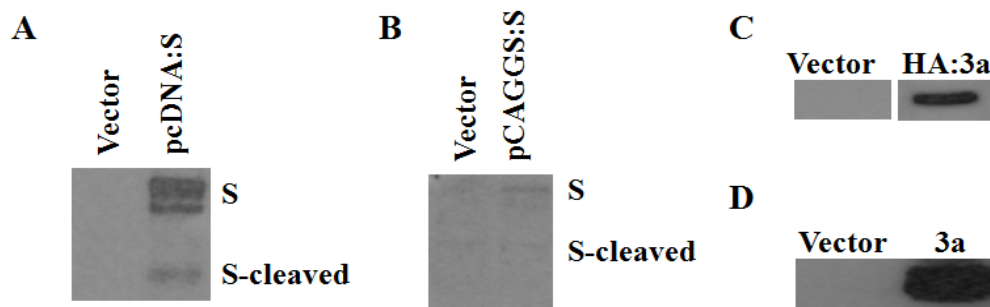
Fig.29. Transient expression of SARS-CoV S and 3a proteins. pcDNA3.1:S (A), pCAGGS:S (B), pCAGGS:3a (C) and pCAGGS:HA:3a (D) proteins were expressed. Both pcDNA3.1:S and pCAGGS:S were expressed in BHK-21 cells. pcDNA3.1:S was expressed using vaccina T7 polymerase vTF7.3. Both pCAGGS:3a clones were expressed in 293T cells. Proteins were analyzed by SDS-PAGE and Western blotting. Transfection of empty pCAGGS vector was used as the vector control. The positions of full length and cleaved S are indicated in panel A and B.
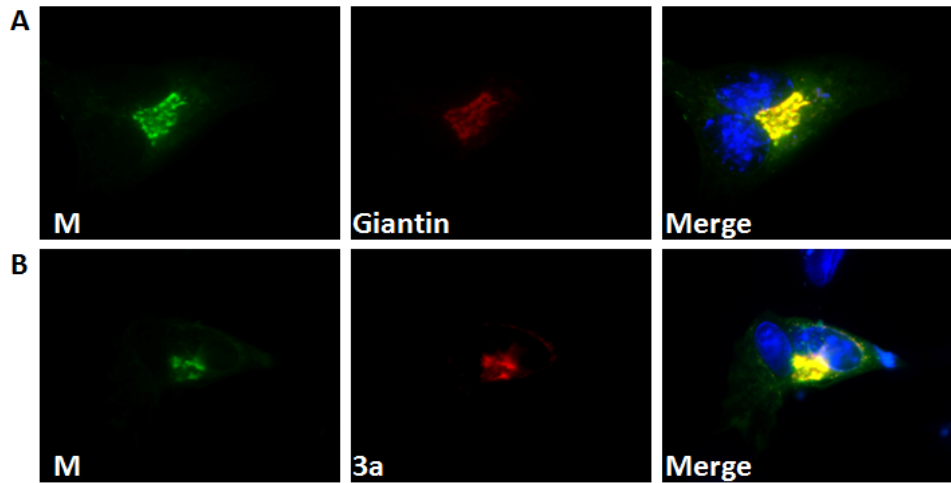
FIG.30. Localization of SARS 3a. BHK-21 cells were transfected with pCAGGS vectors containing WT MHV M or SARS 3a genes. Cells were fixed at 12 h after transfection and analyzed by immunofluorescence using mouse and rabbit antibodies against the M and 3a proteins, respectively. The Golgi was visualized using a rabbit anti-Giantin antibody. Alexa Fluor 488-conjugated mouse and Alexa Fluor 594-conjugated rabbit secondary antibodies were used to visualize the M and 3a or Golgi, respectively. Nuclei were stained with DAPI. A) Colocalization of M and Giantin shown in the merged image on the right B) Colocalization of M and 3a proteins in the merged image on the right

APPENDIX B

CONSTRUCTION OF MVA VIRUSES

DATA COLLECTED 2004-2006

Five Modified vaccinia virus Ankara (MVA) recombinant viruses were constructed that express SARS-CoV proteins. MVA is a highly attenuated recombinant vaccinia virus that was created after being serially passaged more than 500 times in chicken embryo fibroblast (CEF) cells (12). SARS-CoV S protein expressed by attenuated vaccinia virus protectively immunizes mice (12). MVA has been used to express large quantities of foreign proteins off of the early-late promoters. The SARS-CoV genes that were introduced into the MVA genome were S, E, M, 3a and HA-3a.

The full length genes for the SARS-CoV structural proteins are cloned into the pLW44 transfer vector. The pLW44 vector contains two MVA flanking regions that contain a DNA sequence that is similar to wild-type MVA. These flanking regions will allow for homologous recombination to occur between the transfer vector and the Del III region of the MVA genome. Located between the two flanking regions are two vaccinia virus early-late promoters. One of these promoters, the p11 promoter, drives the transcription of green fluorescent protein (GFP). GFP expression is used for selection of the recombinant viruses that underwent homologous recombination. The other promoter is a H5 promoter that drives the transcription of the gene of interest. First, any poxvirus transcription termination motifs (TTTTTNT) that were present in the SARS genes had to be mutated to ensure full length mRNA transcripts are produced. The S gene contained two such motifs at positions 22569 and 24718 in the SARS Urbani

genome. The E gene contained one motif at position 26172. The M gene contained one motif at 26527. The 3a gene contained one motif at position 25580. The poxvirus transcription termination motifs were changed using whole plasmid PCR with primers listed in Table 2. Once mutation of the motifs were confirmed by sequence analysis, PCR was carried out using primers listed in Table 3 that contained appropriate restriction sites to clone into the pLW44 transfer vector. The PCR products were digested with the appropriate enzymes and the SARS-CoV genes were cloned into the pLW44 vector. For the MVA HA tagged 3a recombinant, the 3a gene was first cloned into the pLW44 vector and then whole plasmid PCR was done.

Once the genes were successfully cloned into the pLW44 transfer vector, the generation of SARS MVA recombinants was carried out. First, BHK-21 cells were infected with WT MVA virus at an MOI of 0.01-0.1. Next, the infected cells were transfected with the appropriate pLW44 SARS clone. After two days the cells were harvested, freeze/thawed, and sonicated. The virus is then plated onto BHK-21 cells and overlaid with low melting point agar. After two days, GFP positive plaques are picked. After freeze/thaw and sonication cycles, the plaques are plated onto BHK-21 cells and allowed to incubate for two days. A total of six rounds of plaque purification were done to ensure the stability of the gene insertion. Each purified plaque was checked by SDS-PAGE and Western blotting to determine if the SARS protein expressed in Vero cells (Fig.31). The best expressing plaque was used to grow stocks of the SARS MVA recombinant.

166

Note: The MVA SARS HA:3a virus could never be detected by Western blot using the anti-HA antibody only with the anti-3a antibody (Hogue Lab).

Table 2. Primers used for Whole Plasmid PCR

Primer Name	Primer sequence (5'-3')
SARS S 22571 T to C For	GCT CTA CAA CTC AAC ATT CTT TTC AAC C
SARS S 22571 A to G Rev	GCA CTT AAA GGT TGA AAA GAA TGT TGA G
SARS S 24722 For T to C	CCC TCG TGA AGG TGT TTT CGT GTT AAT GG
SARS S 24722 Rev A to G	CCA AGA AGT GCC ATT AAA CAC GAA AAC ACC
SARS E DoubFor	CGT TAA TAG TTA ATA GCG TAC TTC TCT TCC TTG CTT TCG TGG
SARS E DoubRev	CCA CGA AAG CAA GGA AGA GAA GTA CGC TAT TAA CTA TTA ACG
SARS M DoubFor	CTA ATC GGA ACA GGT TCC TGT ACA TAA TAA AGC TTG TTT TCC TCT GG
SARS M Rev AA	CTT TAT TAT GTA CAG GAA CCT GTT CCG ATT AGA ATA GGC AAA TTG
SARS 3a For T25	CGC TGC AGG TAT GGA GGC GCA ATT CTT GTA CCT CTA TG
SARS 3a Rev A25	GCA TTG TAG AAA ATA TAT CAA GGC ATA GAG GTA CAA GAA TTG CGC CTC C
SARS 3a For T25	CGC TGC AGG TAT GGA GGC GCA ATT CTT GTA CCT CTA TG
SARS 3a Rev A25	GCA TTG TAG AAA ATA TAT CAA GGC ATA GAG GTA CAA GAA TTG CGC CTC C

Table 3. Primers used for PCR to introduce restriction sites

| Primer Name	Primer sequence (5'-3')
SARS S SmaI For	GCC CCG GGA TGT TTA TTT TCT TA
SARS S XmaI Rev	TCC CCC CGG GTT ATG TGT AAT GTA A
E SmaI For	GCC CGG GAT GTA CTC ATT CGT TTC GG
SARS E Rev SalI	GAG CGT CGA CTT AGA CCA GAA GAT CAG G
M SmaI For	GCC CGG GAT GGC AGA CAA CGG TAC TAT TAC CG
SARS M Rev SalI	GCG CGT CGA CTT ACT GTA CTA GCA AAG C
SARS 3a untag SmaI For	CGG CCC GGG ATG GAT TTG TTT ATG AGA
SARS 3a SalI Rev	GAG CGT CGA CTT ACA AAG GCA CGC TAG
SARS LHA SmaI For	GCC CGG GAT GGG ATC CTA CCC TTA C
SARS 3a SalI Rev	GAG CGT CGA CTT ACA AAG GCA CGC TAG
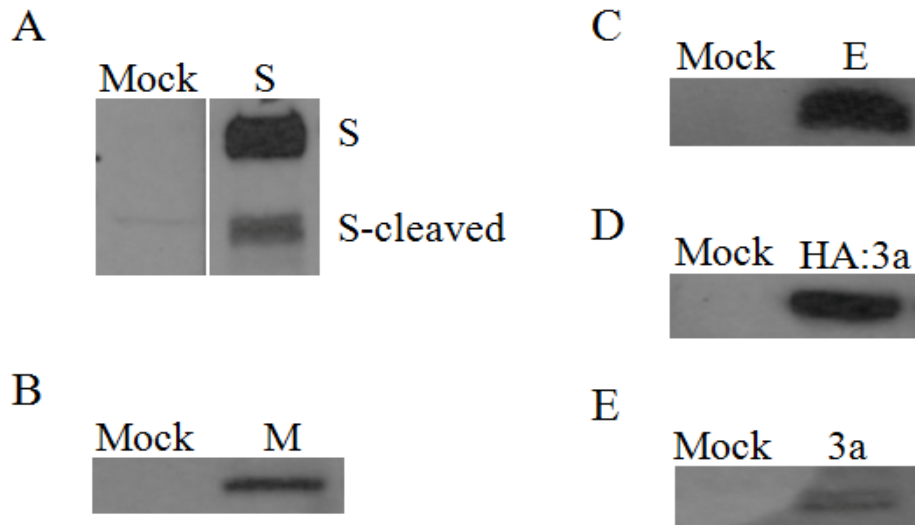
Fig.31. Expression of SARS-CoV structural proteins by MVA recombinant viruses. MVA:S, MVA:M, MVA:E, MVA:HA:3a, and MVA:3a, A-E respectively. Proteins were analyzed by SDS-PAGE and Western blotting and detected by chemiluminescence using antibodies specific for each protein. Transfection of WT MVA virus was used as the mock. The positions of full length and cleaved S are indicated in panel A.

APPENDIX C

PRIMER TABLES

Table 4. Primers used in CD Studies

Name	5' →3' Sequence[a]	Polarity	Purpose
M:E121A	GCT GGT GGA GCT TCA ACC CC**G CC**A CAA ACA ACC TTA TG	For	Mutagenesis
M:E121A	CTA TAC ACA TAA GGT TGT TTG T**GG C**GG GGT TGA AGC	Rev	Mutagenesis
M:E121R	GG**G** AGC TGG TGG AGC TTC AAC CCC **CGA** ACA AAC AAC CTT ATG TG**C** ATA G	For	Mutagenesis
M:E121R	CAT ATC TAT **G**CA CAT AAG GTT GTT TGT **TCG** GGG GTT GAA GCT CCA CCA GCT **C**CC	Rev	Mutagenesis
M:E121K	GCT GGT GGA GCT TCA ACC CC**A AG**A CAA ACA ACC TTA TG	For	Mutagenesis
M:E121K	CTA TAC ACA TAA GGT TGT TTG TC**T TG**G GGT TGA AGC TCC	Rev	Mutagenesis
M:P120A	GCT GGT GGA GCT TCA AC**G CG**G AAA CAA ACA ACC TTA TG	For	Mutagenesis
M:P120A	CTA TAC ACA TAA GGT TGT TTG TTT C**CG C**GT TGA AGC	Rev	Mutagenesis
M:5'A	CAG GAC TGG T**GC CGC CGC CGC C**TT CAA CCC CGA AAC	For	Mutagenesis
M:5'A	CGG GGT TGA A**GG CGG CGG CGG C**AC CAG TCC TGA TAA AC	Rev	Mutagenesis
M:3'A	GGA GCT TCA ACC CCG AA**G CCG CCG CCG CC**A TGT GTA TAG	For	Mutagenesis
M:3'A	CTA TAC ACA T**GG CGG CGG CGG C**TT CGG GGT TGA AGC TCC	Rev	Mutagenesis
M:5'A+3'A	CAG GAC TGG T**GC CGC CGC CGC C**TT CAA CCC CG	For	Mutagenesis
M:5'A+3'A	CGG GGT TGA A**GG CGG CGG CGG C**AC CAG TCC TGA TAA AC	Rev	Mutagenesis
M:ΔCD	GTT TAT CAG GAC TGG TAT GTG TAT AGA TAT GAA AGG	For	Mutagenesis
M:ΔCD	CTT TCA TAT CTA TAC ACA TAC CAG TCC TGA TAA ACA ACC	Rev	Mutagenesis
MHV M(+)	GGA TGA TAT CGA ATT CAA ACA TTA TG	For	Cloning into pCAGGS
MHV M(-)	GCA TCG ATT TAG GTT CTC AAC AAT GCG GTG	Rev	Cloning into pCAGGS
MHV M(A2A3)	CCG AAT TCA AAC ATT ATG GCT GCT ACT ACT CAG G	For	Mutagenesis SS(2-3)AA
MHV M-N(+)	CCACCTCTACATGCAAGGTGTTAAGC	For	RT-PCR
MHV M-N(-)	GGTCTGCCACAACCTTCTCTATCTG	Rev	RT-PCR
MHV E-M(+)	CAGAACTGTCCAACAGGCCGTTAGCAAG	For	RT-PCR and sequencing
MHV E-M(-)	GCAACCCAGAAGACACCTTCAATGC	Rev	RT-PCR
MHV S (+)	CAGACGTCTATTGCGCCTG	For	RT-PCR
MHV 3' E (-)	GATACACAGGATCCAGCGCATACAC	Rev	RT-PCR
MHV G3 Forward	TGGTTGCCTTCCTTGCGTC	For	Sequencing
MHV G4 Reverse	AGTCTGCTTTGGCTGATTCCTTC	Rev	Sequencing
MHV G6 Reverse	TTCCTGAGCCTGTCTACG	Rev	Sequencing
MHV G7 Forward	GAACCCACCAAAGATGTGTATGAGC	For	Sequencing
MHV G8 Forward	GGCAGAAGCTCCTCTGTAAACC	For	Sequencing
S Reverse	GCCAATGCCTAGCATACATGC	Rev	Sequencing

[a] Mutagenized codons are boldface and underlined.

Table 5. Primers used in Charged Residue Studies

Name	5' →3' Sequence[a]	Polarity	Purpose
M:K205AK207A	GCT GTT TAT GTG **GCG** TCC **GCG** GTC GGA AAC TAC CG	For	Mutagenesis
M:K205AK207A	GGG CAG TCG GTA GTT TCC GAC **CGC** GGA **CGC** CAC ATA AAC	Rev	Mutagenesis
M:K205DK207D	GCG GTT TTG CTG TTT ATG TG**G AC**T CC**G AC**G TCG GAA ACT ACC G	For	Mutagenesis
M:K205DK207D	CGG TAG TTT CCG ACG TCG GAG TCC ACA TAA ACA GC	Rev	Mutagenesis
M:R212AK217A	CGG AAA TTA C**GC A**CT GCC CTC AAA C**GC A**CC GAG TGG GGC GG	For	Mutagenesis
M:R212AK217A	CCG CCC CAC TCG G**TG C**GT TTG AGG GCA G**TG C**GT AAT TTC CGA CC	Rev	Mutagenesis
M:R212DK217D	CGG AAA TTA C**GA C**CT GCC CTC AAA C**GA C**CC GAG TGG GGC GG	For	Mutagenesis
M:R212DK217D	CCG CCC CAC TCG G**GT C**GT TTG AGG GCA G**GT C**GT AAT TTC CGA CC	Rev	Mutagenesis
M:K205A	GCT GTT TAT GTG **GCG** TCC AAG GTC GGA AAC TAC CG	For	Mutagenesis
M:K205A	GCA GTC GGT AGT TTC CGA CCT TGG A**CG C**CA CAT AAA C	Rev	Mutagenesis
M:K205D	GCG GTT TTG CTG TTT ATG TG**G AC**T CCA AGG TCG GAA ACT ACC T	For	Mutagenesis
M:K205D	GGG CAG TCG GTA GTT TCC GAC CTT GGA **GTC** CAC ATA AAC AGC	Rev	Mutagenesis
M:K207A	GCT GTT TAT GTG AAG TCC **GCG** GTC GGA AAC TAC CG	For	Mutagenesis
M:K207A	GCA GTC GGT AGT TTC CGA C**CG C**GG ACT TCA CAT AAA C	Rev	Mutagenesis
M:K207D	GCG GTT TTG CTG TTT ATG TGA AGT C**CG A**CG TCG GAA ACT ACC G	For	Mutagenesis
M:K207D	GGG CAG TCG GTA GTT TCC GAC **GTC** GGA CTT CAC ATA AAC AGC	Rev	Mutagenesis
M:R212A	CGG AAA TTA C**GC A**CT GCC CTC CAA CAA ACC G	For	Mutagenesis
M:R212A	GCC ACT CGG TTT GTT GGA GGG CAG **TGC** GTA ATT TCC G	Rev	Mutagenesis
M:R212D	CGG AAA TTA C**GA C**CT GCC CTC CAA CAA ACC G	For	Mutagenesis
M:R212D	GCC ACT CGG TTT GTT GGA GGG CAG **GTC** GTA ATT TCC G	Rev	Mutagenesis
M:K217A	GCC CTC AAA C**GC A**CC GAG TGG GGC GGA CAC CGC	For	Mutagenesis
M:K217A	GCG GTG TCC GCC CCA CTC GG**T GC**G TTT GAG GGC AGT CGG	Rev	Mutagenesis
M:K217D	GCC CTC AAA C**GA C**CC GAG TGG GGC GGA CAC CGC	For	Mutagenesis
M:K217D	GCG GTG TCC GCC CCA CTC GG**G TC**G TTT GAG GGC AGT CGG	Rev	Mutagenesis

[a] Mutagenized codons are boldface and underlined.

Table 5. continued

MHV M-N(+)	CCACCTCTACATGCAAGGTGTTAAGC	For	RT-PCR
MHV M-N(-)	GGTCTGCCACAACCTTCTCTATCTG	Rev	RT-PCR
MHV E-M(+)	CAGAACTGTCCAACAGGCCGTTAGCAAG	For	RT-PCR and sequencing
MHV E-M(-)	GCAACCCAGAAGACACCTTCAATGC	Rev	RT-PCR
MHV S (+)	CAGACGTCTATTGCGCCTG	For	RT-PCR
MHV 3' E (-)	GATACACAGGATCCAGCGCATACAC	Rev	RT-PCR
MHV G3 Forward	TGGTTGCCTTCCTTGCGTC	For	Sequencing
MHV G4 Reverse	AGTCTGCTTTGGCTGATTCCTTC	Rev	Sequencing
MHV G6 Reverse	TTCCTGAGCCTGTCTACG	Rev	Sequencing
MHV G7 Forward	GAACCCACCAAAGATGTGTATGAGC	For	Sequencing
MHV G8 Forward	GGCAGAAGCTCCTCTGTAAACC	For	Sequencing
S Reverse	GCCAATGCCTAGCATACATGC	Rev	Sequencing