

## A Conserved Domain in the Coronavirus Membrane Protein Tail Is Important for Virus Assembly<sup>∇</sup>

Ariel L. Arndt,<sup>1,2,3</sup> Blake J. Larson,<sup>2,3</sup> and Brenda G. Hogue<sup>2,3\*</sup>

*Microbiology Graduate Program,<sup>1</sup> School of Life Sciences,<sup>2</sup> and The Biodesign Institute, Center for Infectious Diseases and Vaccinology,<sup>3</sup> Arizona State University, Tempe, Arizona*

Received 26 May 2010/Accepted 9 August 2010

**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. The 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<sub>121</sub> 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<sub>121</sub>Q and E<sub>121</sub>L M proteins were capable of forming virus-like particles (VLPs) when coexpressed with E, whereas E<sub>121</sub>R and E<sub>121</sub>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.**

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 CoV (SARS-CoV), HCoV-NL63, and HCoV-HKU1, were recently identified (40, 47). 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 (CDs) 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 (13). The M glycoprotein is the most abundant component of the viral envelope and plays required, key roles in virus assembly (9, 20, 31, 33, 41). 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 (2, 5, 41).

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 (14) (Fig. 1A). M localizes in the Golgi region when expressed alone (20, 22). M molecules interact with each other and also with the spike and nucleocapsid during virus assembly (8–10, 23, 31, 33). 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 as being important for coronavirus assembly (6, 8, 9, 17, 43). Coronaviruses assemble and bud at intracellular membranes in the region of the endoplasmic reticulum (ER) Golgi intermediate compartment (ERGIC) (22, 39). Coexpression of only the M and the E proteins is sufficient for virus-like particle (VLP) assembly for most coronaviruses (2, 41).

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. 1A). The amphipathic domain appears to be closely associated with the membrane (34). At the amino terminus of the amphipathic domain, there is a highly conserved 12-amino-acid domain (SWWSFNPETNNL), consisting of residues 114 to 125 in the mouse hepatitis virus (MHV) A59 M protein (Fig. 1B) (19). These residues are almost identically conserved across the entire *Coronaviridae* family. Because of the crucial role that M plays in virus assembly and the high conservation of this domain, we hypothesized that it is functionally important for virus assembly. To test this, a series of changes were introduced in

\* Corresponding author. Mailing address: The Biodesign Institute, P.O. Box 875401, Arizona State University, Tempe, AZ 85287-5401. Phone: (480) 965-9478. Fax: (480) 727-7615. E-mail: Brenda.Hogue@asu.edu.

<sup>∇</sup> Published ahead of print on 18 August 2010.

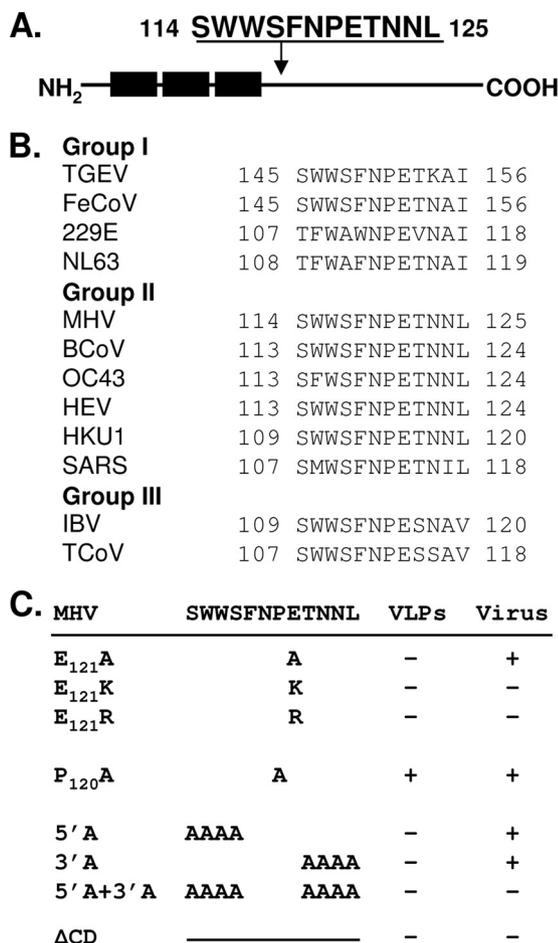


FIG. 1. M protein conserved domain and mutants. (A) A linear schematic of the M protein illustrating the relative positions of the three TM domains (black boxes) and the position of the CD in the tail. (B) Alignment of CDs from representative coronaviruses. Full-length amino acid sequences from transmissible 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 using CLUSTAL W (25). (C) Mutations introduced into the MHV CD, with + and - symbols used to indicate VLP production and virus recovery for each mutant.

the CD. The functional impact of the changes was studied in the context of the virus by genetic analysis and the 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 clone 1 (17C11) 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) (48) 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 (48). Human 293T cells (ATCC) were grown on 0.25% gelatin-coated flasks in DMEM containing 10% FCS, antibiotics, and glutamine. Stocks of wild-type (WT) MHV A59 and cloned viruses were grown in mouse 17C11 or L2 cells. Virus titers were determined in L2 cells.

**Construction of amino acid substitution and deletion mutants.** pGEM-5Zi(+) M-N, a pGEM5Zi(+) vector (Promega) containing the M and N genes (EcoRV-SacI fragment), was used for mutagenesis (42). 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 BssHII sites (48).

**Reverse genetics.** All mutant viruses were made by reverse genetics using an MHV A59 infectious clone (48). Full-length cDNA clones were assembled, transcribed, and electroporated into BHK-MHVR cells as described previously (42). Following electroporation, BHK-MHVR cells were seeded alone or in some cases were seeded concurrently with L2 cells. At 24 to 46 h after electroporation, media were harvested and an aliquot was used to infect 17C11 or L2 cells. The medium was removed from infected cells at 24 to 72 h postinfection (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 reverse transcription-PCR (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 17C11 cells five times for all viable viruses, except for the 3'A mutant, which 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 of 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 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 17C11 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 the 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 to 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, the 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. The 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 (32). 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 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 gene, the 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 1 mM phenylmethylsulfonyl fluoride (PMSF). The medium was clarified at 14,000 × 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 SW55 Ti 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

(4, 27). 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 (6). 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  $\mu$ Ci/ml of Expre<sup>35</sup>S<sup>35</sup>S labeling mixture (Perkin Elmer). At 24 h after transfection, cells were washed with phosphate-buffered saline (PBS) and lysed on ice with radio-immunoprecipitation 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 2 $\times$  RIPA buffer on ice and sonication for 1.5 min at 30-s 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 immunoprecipitated with anti-MHV M monoclonal J1.3, which recognizes the amino terminus of the protein (12), or with anti-MHV serum F88 (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 five times with RIPA buffer, followed by 1 wash with RIPA buffer without 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.** Colocalization 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-Tek chamber slides (Nunc) at 1 day prior to transfection with MirusTrans-293 reagent (Mirus). At 12 h after transfection, cells were washed two times with PBS and fixed in methanol for 15 min at  $-20^{\circ}\text{C}$ . 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 (12) 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 and washed with 2% gelatin in PBS before incubation with Alexa-labeled secondary antibodies. Cells were washed with 2% gelatin in PBS and then given a final wash with PBS alone. Slides were mounted in ProLong Gold antifade reagent (Molecular Probes) containing 4,6-diamino-2-phenylindole (DAPI) to stain nuclei. Images were collected by using an epifluorescence Nikon inverted microscope (Nikon, Inc., Melville, NY) with MetaMorph imaging software (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 tails 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. 1C). Two nucleotides were introduced for all codon changes to assist in identification of compensatory changes. An 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 coexpression 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<sub>121</sub>) was changed to neutrally charged alanine (A), as well as to positively charged arginine (R) or lysine (K) (Fig. 1C). The negative charge was of interest for its potential to mediate protein-protein interactions through electrostatic means. Sequence analysis of the E<sub>121</sub>A virus from cells infected with medium taken directly from electroporated cells showed that the introduced mutations were retained and that no other changes were present in the structural genes. In contrast, the positive charges in both the E<sub>121</sub>K and E<sub>121</sub>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, threonine (T) to A at position 428 in the E<sub>121</sub>Q mutant and glycine (G) to A at position 94 in the E<sub>121</sub>L mutant (Table 1).

The recovered E<sub>121</sub>A, E<sub>121</sub>Q, and E<sub>121</sub>L 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<sub>121</sub>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 isoleucine, in the N gene (Table 1). Parallel growth analysis of the E<sub>121</sub>A virus with and without the change in N indicated that both grew comparably (data not shown). Therefore, the E<sub>121</sub>A mutant virus with no other changes was used for further analysis.

Four out of six E<sub>121</sub>Q plaque-purified virus isolates maintained the T<sub>428</sub>A change in the N gene that was seen in the population analysis following electroporation, and two had only the E<sub>121</sub>Q change (Table 1). To determine if the change in the N gene was providing a growth advantage, one of the plaque-purified E<sub>121</sub>Q viruses without the T<sub>428</sub>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<sub>121</sub>Q virus with the T<sub>428</sub>A change in N was used for subsequent finer-detailed growth analysis.

Sequence analysis of the E<sub>121</sub>L plaque-purified viruses showed that all contained the G<sub>94</sub>A change seen immediately following electroporation in the N gene through P5. To determine the impact of the G<sub>94</sub>A second-site change, an independent clone that contained only the E<sub>121</sub>L change was constructed. Recovered virus was plaque purified. Seven individual plaque isolates had only the E<sub>121</sub>L change. Interestingly however, an equal number had additional changes scattered across the N gene (G<sub>23</sub>E, E<sub>236</sub>K plus K<sub>395</sub>E, N<sub>153</sub>D plus E<sub>173</sub>K, E<sub>374</sub>K, S<sub>417</sub>N, R<sub>13</sub>G, and D<sub>346</sub>N) (Table 1). The virus isolate that contained the N<sub>153</sub>D plus E<sub>173</sub>K changes in the N gene also had a silent change in the E<sub>1321</sub> 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<sub>1321</sub>K). Since growth kinetic analysis indicated that the G<sub>94</sub>A change in the N gene was not providing additional compensation over the E<sub>121</sub>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

TABLE 1. Summary of recovered M conserved domain (SWWSFNPETNNL) mutant viruses

Mutant virus <sup>a</sup>	No. of plaques analyzed	No. of passages	Retention of mutation	Change(s) in <sup>d</sup> :	
				M	N
E <sub>121</sub> A	4	5	Yes		
E <sub>121</sub> Q (K)	1	5	Yes		V <sub>410</sub> I
	4	5	No	K <sub>121</sub> Q	T <sub>428</sub> A
E <sub>121</sub> L (R) no. 1	2	5	No	K <sub>121</sub> Q	
	4	5	No	R <sub>121</sub> L	G <sub>94</sub> A
E <sub>121</sub> L no. 2 <sup>b</sup>	7	2-5	Yes		G <sub>23</sub> E; E <sub>236</sub> K; K <sub>395</sub> E; N <sub>153</sub> D; E <sub>173</sub> K <sup>c</sup> ; E <sub>374</sub> K; S <sub>417</sub> N; R <sub>13</sub> G; D <sub>346</sub> N
P <sub>120</sub> A	7	2-5	Yes		
SWWS → AAAA no. 1	5	5	Yes		
	2	5	Yes	G <sub>31</sub> R	
SWWS → AAAA no. 2	Blind passage	5	Yes	Q <sub>42</sub> R	
	Electroporation	1	Yes	T <sub>38</sub> N	
	6	2-5	Yes	Q <sub>42</sub> R (3) Q <sub>42</sub> R (2) L <sub>35</sub> P (1)	N <sub>409</sub> S
TNNL → AAAA	Blind passage	5	Yes		

<sup>a</sup> No. 1 and 2 refer to independent constructions described in the text.

<sup>b</sup> Two additional plaque-purified viruses were isolated. One had a change in E (E<sub>71</sub>G), and another had a change in S (E<sub>1321</sub>K).

<sup>c</sup> N<sub>153</sub>D, E<sub>173</sub>K also contained a silent change in S in the E<sub>1321</sub>K codon.

<sup>d</sup> Numbers in parentheses refer to the number of plaque isolates with indicated changes.

parallel, along with the WT virus, for their growth properties (Fig. 2). Viruses E<sub>121</sub>A, E<sub>121</sub>Q with T<sub>428</sub>A, and E<sub>121</sub>L with G<sub>94</sub>A all produced plaques and exhibited growth properties similar to those of the WT virus. These results indicate that a negative charge at position 121 within the MHV M protein conserved domain is not absolutely 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<sub>120</sub>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 those of the WT virus (Fig. 2). These 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. 1C). 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. 3A). Only 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. 3B). 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. 3B).

We then constructed a second independent clone with the SWWS-to-AAAA change. When the medium 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<sub>38</sub>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. Five of the viruses had the Q<sub>42</sub>R change observed earlier. One of the viruses, surprisingly, had an L<sub>35</sub>P second-site change. Two of the isolates (no. 5 and 6) with the additional Q<sub>42</sub>R changes also had a change in the N gene, N<sub>409</sub>S.

All of the recovered 5'A viruses with the TM1 changes except T<sub>38</sub>N, which was recovered only during blind passage, were compared for their growth characteristics in parallel with the WT virus at an MOI of 0.01 PFU/cell (Fig. 3C, left). All of the recovered viruses exhibited significant growth impairment, with a ~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<sub>35</sub>P change actually grew even less well, producing at least 1,000-fold less virus at 20 to 30 h p.i. While we expected that the N<sub>409</sub>S second-site change in the N gene might be a compensatory change, 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

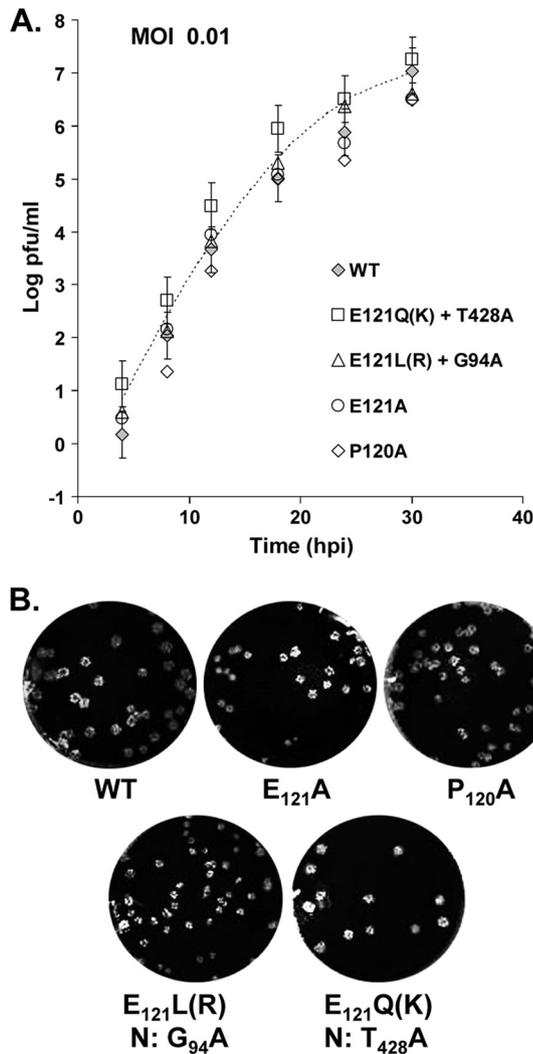


FIG. 2. Growth properties of single-substitution mutant viruses. (A) Mouse 17C11 cells were infected with mutant viruses at an MOI of 0.01 PFU/cell. Titers were determined by plaque assay on L2 cells at the indicated times. Data points are shown for all viruses, but the growth curve is included for only the WT virus. Error bars and exponential growth curves are as described in the Materials and Methods. The estimated doubling time was  $\sim 0.62$  h, and the saturation parameter was 1.5% per h. (B) Plaque characteristics were determined for the indicated viruses in L2 cells. Changes in the recovered viruses with positive-charge (R and K) substitutions and corresponding noncompensating changes in the N gene are indicated.

similar to that of the amino end (Fig. 1B). Alanine substitutions were introduced in place of the TNNL residues to also assess the importance of the region. Like for 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 that of 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 medium from the electroporated cells, but no second-site changes were present in the recovered virus population. The

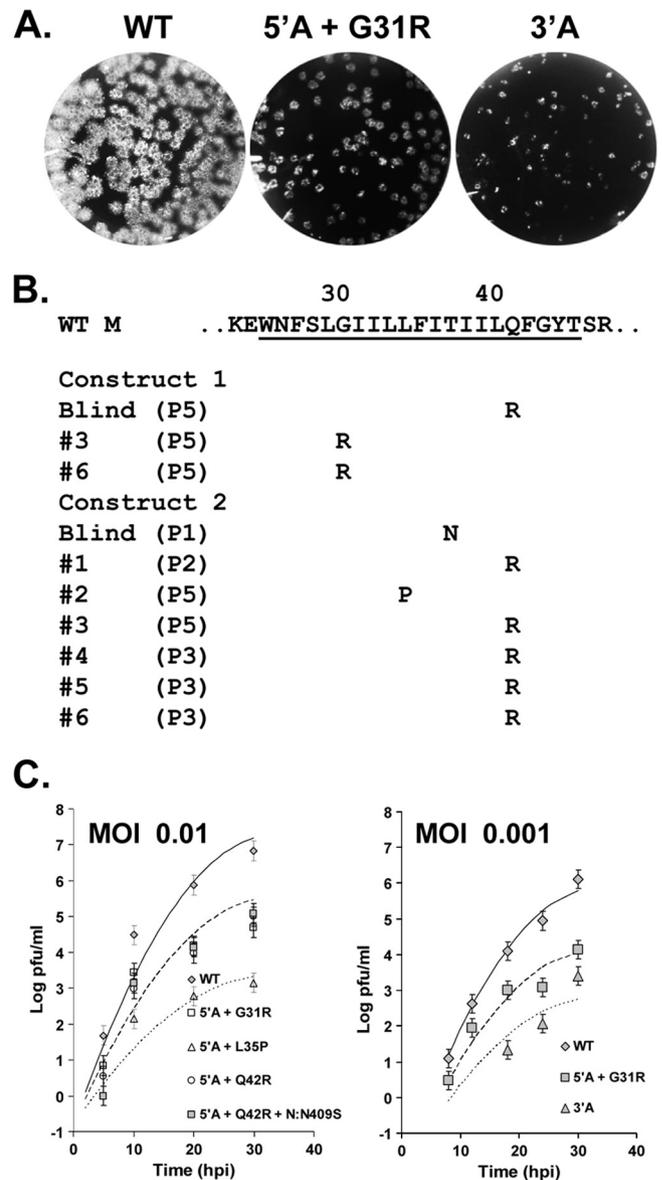


FIG. 3. Growth properties of 5'A and 3'A mutant viruses. (A) Plaque characteristics of WT, 5'A+G<sub>31</sub>R, and 3'A viruses were analyzed in mouse L2 cells. (B) Summary of second-site changes in the TM1 of recovered viruses from two independent virus constructions, designated 1 and 2, are shown under 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 with mouse 17C11 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 Materials and Methods. Estimated doubling times were  $\sim 0.62$  h for WT virus,  $\sim 0.79$  h for all 5'A viruses except L<sub>35</sub>P ( $\sim 1.2$  h), and 1 h for the 3'A virus.

virus stock still produced very tiny plaques, and growth kinetic analysis at an MOI of 0.001 PFU/cell in parallel with the WT and 5'A+G<sub>31</sub>R viruses showed that the virus was very crippled, reaching peak titers about 1,000 times lower than those of the WT virus (Fig. 3A and C, right).

Altogether, the results demonstrate that the conserved res-

idues 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. 1C). 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 medium of electroporated cells was also attempted, but virus could not be recovered. These results strongly indicate that the 5'A+3'A and  $\Delta$ CD mutations were lethal to the virus.

Since significant changes were introduced with the 5'A+3'A and  $\Delta$ CD mutations, the localization of the proteins were analyzed by immunofluorescence. Wild-type M colocalized with the Golgi marker giantin, as expected (20, 22). 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 apparatus 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 (2, 5, 41). Thus, the mutant proteins were coexpressed with the wild-type E protein in 293T cells. In addition, the M genes with the E<sub>121</sub>Q and E<sub>121</sub>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. 4). The only M mutant to support VLP production was P<sub>120</sub>A. Interestingly, even though the E<sub>121</sub>A change supported virion assembly in the context of the virus, the protein was not competent for VLP assembly. The E<sub>121</sub>-to R or -K mutant M proteins did not support VLP assembly; however, the two recovered changes, E<sub>121</sub>Q and E<sub>121</sub>L, were competent for VLP production. As expected, the 5'A+3'A and  $\Delta$ CD 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 (17, 23, 29, 43). Recent studies indicate that N contributes to efficient assembly of SARS-CoV VLPs (37, 46). Thus, we asked if the presence of N during VLP formation would affect VLP output of our mutant M proteins. We tested the three mutants that were capable of forming VLPs, P<sub>120</sub>A, E<sub>121</sub>Q, and E<sub>121</sub>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 a WT N clone (Fig. 5). All of the mutant M proteins yielded fewer 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 an ~25 to

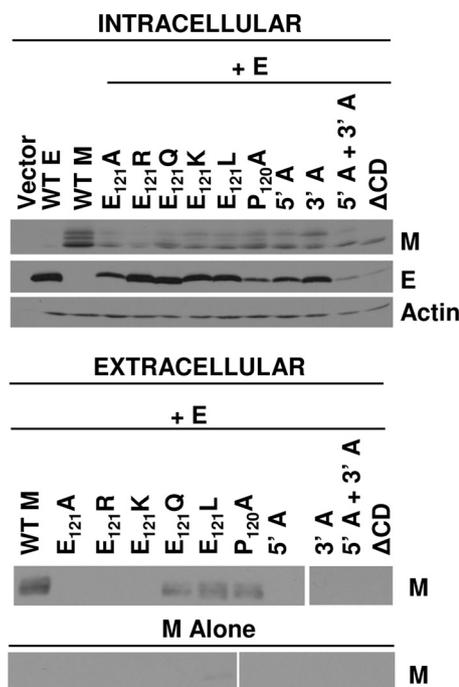


FIG. 4. Effects of CD mutations on VLP production. 293T cells were transfected with pCAGGS vectors containing WT or mutant M genes singly or in combination with pCAGGS containing the WT E gene. Control empty vector (vector) was analyzed in parallel. Intracellular cell lysates and pelleted extracellular 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.

45% increase in VLP output, compared with an ~10% increase for WT M. This correlated well with the growth phenotypes of the E<sub>121</sub>Q, E<sub>121</sub>L, and P<sub>120</sub>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<sub>31</sub>R, Q<sub>42</sub>R, T<sub>38</sub>N, and L<sub>35</sub>P) (Fig. 3B). 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 plus TM1 changes were coexpressed with E, in both the absence and presence of the N protein (Fig. 6). The 5'A and 5'A+TM1 proteins produced no or at most a very small amount of VLPs in the absence of N coexpression. Interestingly, however, the 5'A+TM1 mutant M proteins produced ~20 to 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<sub>31</sub>R M, the increases represented changes from undetectable levels in the absence of N. These data suggest that the TM1 changes are

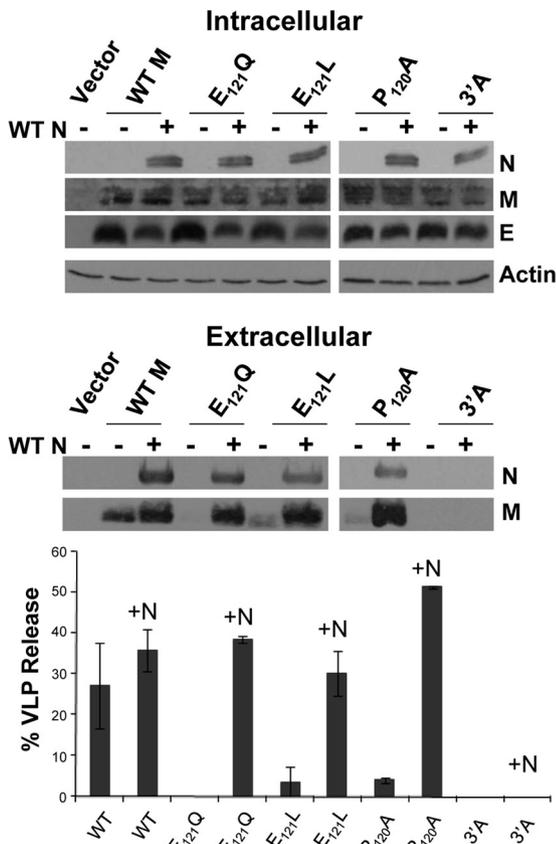


FIG. 5. N protein enhancement of VLP production. 293T cells were transfected as indicated with pCAGGS vectors containing WT or mutant M genes in combination with pCAGGS-E and pCAGGS-N where indicated. Intracellular and extracellular VLP fractions were analyzed by SDS-PAGE and Western blotting as indicated in Fig. 4. 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 compared to total M (intracellular plus extracellular) protein. The data represent averages and standard deviations from two experiments.

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. 5).

**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 an M protein called A2A3, which was previously described (6). 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 (6, 7).

Our mutant M proteins were coexpressed with A2A3 M and

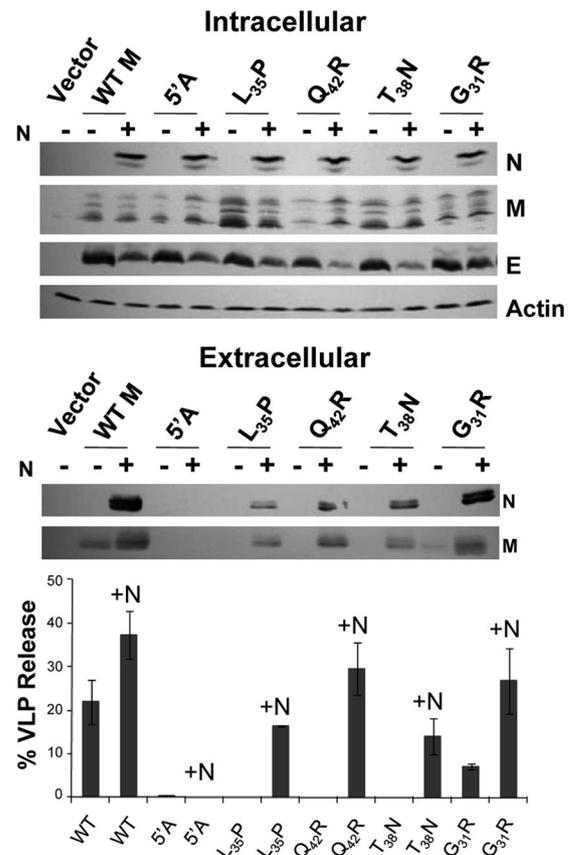


FIG. 6. VLP analysis of 5'A M mutants with second-site changes in TM1 coexpressed with the N protein. 293T cells were transfected with WT and mutant genes in pCAGGS vectors. M genes were expressed with E in the absence and presence of the N gene. Intracellular and extracellular fractions were analyzed 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. 5. Error bars represent the standard deviations from two independent experiments.

wild-type E. A lower concentration of mutant plasmid DNA was transfected to ensure efficient VLP formation, as previously determined (6). Cells were metabolically labeled with Expre<sup>35</sup>S<sup>35</sup>S labeling mixture at 4 to 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. 7, lanes 5 and 6). When the 5'A or 3'A mutant M protein was 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. 7, lanes 7 to 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. 7, lanes 11 and 12). The M mutants P<sub>120</sub>A, E<sub>121</sub>Q, and E<sub>121</sub>L, which supported VLP production when expressed alone with E, were all also

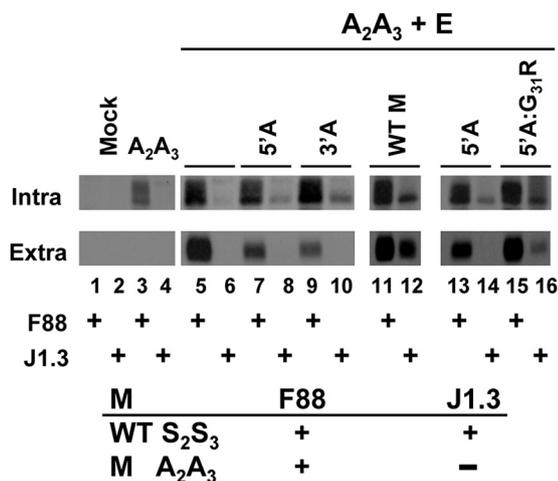


FIG. 7. Rescue of CD mutants into VLPs. 293T cells were transfected with the pCAGGS vector containing WT or mutant M proteins as indicated. Cells were labeled for 20 h at 4 h after transfection with Expre<sup>35</sup>S<sup>35</sup>S labeling mixture. Intracellular lysates (Intra) and extracellular media (Extra) were divided in half and immunoprecipitated with monoclonal J1.3 or polyclonal F88 antibodies, which recognize only WT and both WT and A<sub>2</sub>A<sub>3</sub> proteins, respectively. Samples were analyzed by SDS-PAGE and autoradiography.

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 mutant M proteins with the recovered G<sub>31</sub>R change. When expressed with A<sub>2</sub>A<sub>3</sub> and E, both mutant proteins were detected in the intracellular fractions with J1.3 and F88 antibodies (Fig. 7, lanes 13 to 16). However, only the 5'A+G<sub>31</sub>R mutant could be rescued into VLP particles. These data further support the conclusion that the second-site TM1 changes are 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 (31, 33). WT M protein localizes to the Golgi apparatus 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. 8). 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. 8). S localized to a large extent with the 5'A+3'A and ΔCD proteins, which 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.

DISCUSSION

The significance of the evolutionarily conserved domain (CD) located at the amino end of the long amphipathic region

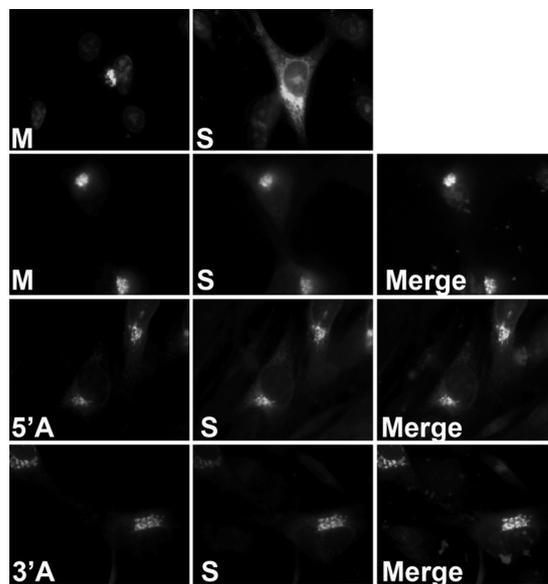


FIG. 8. S protein colocalization with WT and mutant M proteins. 293T cells were transfected with pCAGGS vectors containing WT or mutant M proteins and the S gene. Cells were fixed at 12 h after transfection and analyzed by immunofluorescence using mouse and goat antibodies against the M and S proteins, respectively. Alexa Fluor 488-conjugated mouse and Alexa Fluor 594-conjugated goat secondary antibodies were used to visualize the localizations of the M and S proteins, respectively. 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.

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 end were very sensitive to changes, since recovered viruses exhibited crippled phenotypes, with 100- to 1,000-fold reductions in virus yields. Removal of the conserved proline must not alter the domain, since assembly of VLPs 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, which 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 position. Since the E<sub>121</sub>R and E<sub>121</sub>K mutant proteins were not able to form VLPs, whereas M proteins with the recovered E<sub>121</sub>Q and E<sub>121</sub>L changes 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 L changes exhibited growth characteristics similar to those of the WT virus, suggesting that the changes in N are not providing an 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 resonances and crystal structures have been determined for parts of infectious bronchitis virus (IBV) and SARS-CoV N proteins (3, 11, 15, 18, 36, 38, 49). 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 domain (TM1) of M. The TM1 changes were not fully compensating, since the amount of virus output was significantly reduced compared with that for the WT. Nonetheless, in the contexts of both 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 (24). 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) (45). 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 (21). 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 (9). Recent cryo-electron microscopy (cryo-EM) and cryo-electron tomography analysis of single virus particles provided new insight about the three-dimensional (3D) structures of coronaviruses (1, 30). The tomography analysis revealed that the envelope is striated, presumably due to M-M interactions (1). 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 (1). 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 (34, 35). 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 (1). 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 assembly of the envelope, 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 M-N 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 (1, 30). Recent genetic analysis demonstrated that important interactions occur between the carboxy ends of the N and M proteins (17, 23, 42, 43). Tight interaction between the carboxy-terminal regions of MHV N (amino acids 380 to 454) was also recently shown to play a role in helping to mediate N-N interactions (16). 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  $\Delta$ CD M proteins were not able to support virus assembly. Both proteins localized in the ER, whereas proteins with replacement of only the amino or carboxy residues localized like WT M in the Golgi apparatus. This suggests that the structures of the 5'A+3'A and  $\Delta$ CD proteins are altered or that a localization signal was disrupted. An MHV M deletion mutant lacking residues L<sub>108</sub> through T<sub>112</sub> ( $\Delta$ LT), that are adjacent to the CD, was also retained in the ER, but when residues E<sub>121</sub> to D<sub>195</sub> were removed ( $\Delta$ C), the protein was retained in the Golgi apparatus (9). A specific localization signal has not been identified for MHV M. However, the cytoplasmic tail has been shown to be essential but not suffi-

cient for Golgi localization (26). Altogether, the results suggest that the region just beyond the third TM domain is important for Golgi retention. Interestingly, a truncated SARS-CoV M consisting of the amino-terminal 134 residues, which includes the CD (S<sub>107</sub> to L<sub>118</sub>), is retained in the Golgi apparatus (44).

A panel of MHV M mutants that provided important insight into which regions of M are involved in envelope assembly was studied previously (6, 8, 9). Three deletion mutants from the earlier studies should be noted here. Amino acids 121 to 195, which constitute a large portion of the amphipathic domain, were deleted from MHV M ( $\Delta$ C). This deletion includes the last five residues (ETNNL) of the CD (Fig. 1). Like our 3'A mutant M with alanine substitutions for TNNL residues, the  $\Delta$ C 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 (9). 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 (9). Finally, M proteins with deletion of only I<sub>110</sub> failed to participate in VLP assembly (8). 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 an impact on the overall structure of the M tail as residue deletions.

Through interaction with M, S is retained at assembly sites in the ERGIC region of the cell (31, 33). The 5'A and 3'A M proteins retained S in the Golgi apparatus, indicating that the alanine substitutions for SWWS or TNNL residues are not affecting M-S interactions. The amphipathic domain of MHV M was previously indicated to be involved in mediating interaction with S, since the  $\Delta$ C mutant described above did not retain S in the Golgi apparatus (8). A single Y<sub>211</sub>G change that is significantly distant from the CD has also been indicated to play a role in M-S interactions (8). It is interesting that the amino-terminal 134 amino acids of SARS-CoV are sufficient for retention of S in the Golgi apparatus (44). A single tyrosine (Y<sub>195</sub>) in SARS-CoV M was recently shown to be necessary to retain S in the Golgi apparatus (28). Further mapping will be required to more precisely identify requirements for MHV M-S interactions. It is clear from the present study that M proteins with changes in the CD that do not affect localization of the proteins are still competent to mediate colocalization.

Overall, our results suggest that the CD functions in formation of the viral envelope by helping to 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 is compromised. It is likely that N helps stabilize M complexes, possibly helping to 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 (10, 23, 29, 33, 42, 43). The 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.

#### ACKNOWLEDGMENTS

This work was supported by National Institutes of Health grant AI53704. A.L.A. was supported in part by National Institutes of Health predoctoral fellowship F31AI075538 and funds from the More Graduate Education at Mountain States Alliance (MGE@MSA) Alliance for Graduate Education and the Professoriate (AGEP) National Science Foundation (NSF) cooperative agreement no. HRD-9978868, by an Achievement Rewards for College Scientists (ARCS) Phoenix Chapter fellowship, and as an American Society for Microbiology Robert D. Watkins Graduate Research Fellow. B.J.L. was supported in part by funds from the Howard Hughes Medical Institute through the Undergraduate Science Education Program and from the Arizona State University School of Life Sciences.

We thank Stephanie Melbourne for construction of the pCAGGS: 3'A clone and Benjamin Neuman for helpful discussions and sharing results with us prior to publication. We also thank members of the Hogue lab for helpful discussions and comments.

#### REFERENCES

1. 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 coronavirus. *Proc. Natl. Acad. Sci. U. S. A.* **106**:582–587.
2. 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.
3. 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.
4. Cologna, R., J. F. Spagnolo, and B. G. Hogue. 2000. Identification of nucleocapsid binding sites within coronavirus-defective genomes. *Virology* **277**:235–249.
5. 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.
6. 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.
7. 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.
8. 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.
9. 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.
10. 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.
11. 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.
12. Fleming, J. O., S. A. Stohman, 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.
13. Gallagher, T. M., and M. J. Buchmeier. 2001. Coronavirus spike proteins in viral entry and pathogenesis. *Virology* **279**:371–374.
14. Hogue, B. G., and C. E. Machamer. 2008. Coronavirus structural proteins and virus assembly, p. 179. *In* S. Perlman, T. Gallagher, and E. Snijder (ed.), *Nidoviruses*. ASM Press, Washington, DC.
15. 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.
16. 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.
17. 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.
18. 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.
  19. 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.
  20. 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.
  21. 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.
  22. 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.
  23. 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.
  24. 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.
  25. 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.
  26. Locker, J. K., J. Klumperman, V. Oorschot, M. C. Horzinek, H. J. Geuze, and P. J. Rottier. 1994. The cytoplasmic tail of mouse hepatitis virus M protein is essential but not sufficient for its retention in the Golgi complex. *J. Biol. Chem.* **269**:28263–28269.
  27. 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.
  28. 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.
  29. 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.
  30. 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.
  31. Nguyen, V. P., and B. G. Hogue. 1997. Protein interactions during coronavirus assembly. *J. Virol.* **71**:9278–9284.
  32. Niwa, H., K. Yamamura, and J. Miyazaki. 1991. Efficient selection for high-expression transfectants with a novel eukaryotic vector. *Gene* **108**:193–199.
  33. 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.
  34. 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.
  35. 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.
  36. 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.
  37. 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.
  38. 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.
  39. 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.
  40. 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.
  41. 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.
  42. 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.
  43. 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.
  44. 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. *Virology* **6**:79.
  45. 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.
  46. 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.
  47. 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.
  48. 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.
  49. 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 coronavirus and arteriviridae. *J. Biol. Chem.* **281**:17134–17139.