Molecular characterization of coronavirus envelope and membrane proteins

Yuan, Quan

2008

Yuan, Q. (2008). Molecular characterization of coronavirus envelope and membrane proteins. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/6568>

<https://doi.org/10.32657/10356/6568>

Nanyang Technological University

Downloaded on 26 Apr 2023 16:24:10 SGT

MOLECULAR CHARACTERIZATION OF CORONAVIRUS ENVELOPE AND MEMBRANE PROTEINS

Yuan Quan

**SCHOOL OF BIOLOGICAL SCIENCES
NANYANG TECHNOLOGICAL UNIVERSITY**

2007

Molecular Characterization of Coronavirus Envelope and Membrane Proteins

Yuan Quan

School of Biological Sciences

A thesis submitted to the Nanyang Technological University

in fulfillment of the requirement for the degree of

Doctor of Philosophy

2007

Abstract

Coronaviruses are enveloped viruses with a single strand, positive-sense RNA genome of 27-32 kb in length. In virus-infected cells, a 3'-coterminally nested set of six to nine mRNA species, including the genome-length mRNA (mRNA1) and five to eight subgenomic mRNA species (mRNA2-9), is produced. The four major structural proteins, spike (S), envelope (E), membrane (M) and nucleocapsid (N), are encoded by different subgenomic mRNAs.

Coronavirus E protein is a small integral membrane protein with multi-functions in virion assembly, morphogenesis and virus – host interaction. Different coronavirus E proteins share striking similarities in biochemical properties and biological functions, but seem to adopt distinct membrane topology. In this report, we study the membrane topology of the SARS-CoV E protein by immunofluorescent staining of cells differentially permeabilized with detergents and proteinase K protection assay. It was revealed that both the N- and C-termini of the SARS-CoV E protein are exposed to the cytoplasmic side of the membranes (N_{cyto}C_{cyto}). In contrast, parallel experiments showed that the E protein from IBV spanned the membranes once, with the N-terminus exposed lumenally and the C-terminus exposed cytoplasmically (N_{exo(lum)}-C_{cyto}). Intriguingly, a minor proportion of the SARS-CoV E protein was found to be modified by N-linked glycosylation on Asn 66 and inserted into the membranes once with the C-terminus exposed to the luminal side. The presence of two distinct membrane topologies of the SARS-CoV E protein may provide a useful

clue to the pathogenesis of SARS-CoV. Immunofluorescence staining and cell fractionation studies demonstrate that SARS-CoV E protein and IBV E protein are partially associated with lipid rafts. Snapin was identified as an interaction partner of the SARS-CoV E protein in the yeast two-hybrid system. The interaction of SARS-CoV E protein with snapin was confirmed by co-immunoprecipitation and GST pull-down assay.

Coronavirus M protein is the key player in virion assembly. In this study, the N-linked glycosylation sites of IBV M protein were defined. Site-directed mutations in the two glycosylation sites were incorporated into the IBV genome to get the N3D/6D mutant virus. The M protein glycosylation state did not influence the tissue culture growth characteristics of viruses, virion morphology and heterogeneity. The effects of mutations in the N-terminal domain of IBV M protein were analyzed systematically. Mutations in the luminal domain had no effect on the formation of VLPs. However, some of these mutations in this domain had severe effects on the infectivity of the viral RNA to Vero cells. These observations suggested that N-linked glycosylation is not important for the virus assembly, but some other mutations in the luminal domain had effects on the virus assembly.

Acknowledgements

I would like to express my deepest gratitude to my supervisor Dr. Liu Ding Xiang for his full support, scholastic guidance and enthusiastic encouragement through my graduate career.

My deepest gratitude also goes to Dr. James P. Tam, my supervisor and Dean of SBS. Thank you for your guidance and supervision. Over the years, I have witnessed the rapid development of our school, and I am proud to have grown with our school together.

I am deeply in debt to my family, my dad Yuan Qi Min, my mom Lu Jian, my sister Yuan Yan and brother in law Yang Wen Hai. I want to thank you for all the love and support that you have given me throughout my life. Without your encouragement, I would not accomplish this work.

I want to thank my colleagues in Dr. Liu and Dr. Tam's labs. The success of my work would be impossible without help of all these people. I am also grateful to my friends Lu Yi Wei, Xu Ting and Shi Min Long. I will always remember our friendship.

This work was supported by the Agency for Science, Technology and Research, Singapore, and grants from the Biomedical Research Council, Agency for Science, Technology and Research, Singapore.

Table of Contents

Abstract	ii
Acknowledgements	iv
Table of Contents	v
List of Figures and Tables	xiii
List of Abbreviations	xvi
Chapter 1	
Introduction	1
1.1. Classification	2
1.2. Virion Morphology	4
1.3. Genome Structure and mRNAs of Coronavirus	7
1.4. Structural Proteins	11
1.4.1. The Spike (S) glycoprotein	11
1.4.2. The membrane (M) protein	14
1.4.3. The envelope (E) protein	15
1.4.4. The nucleocapsid (N) protein	18
1.5. The Replicase Polyprotein	20
1.6. Accessory Proteins	24
1.7. Coronavirus Lifecycle.....	24
1.7.1. Attachment, penetration and uncoating	24
1.7.2. Transcription of viral mRNA	29
1.7.3. Replication of viral genomic RNA	32

1.7.4. Translation of viral proteins -----	33
1.7.4.1. Ribosomal frameshifting within the polymerase gene -----	33
1.7.4.2. Internal initiation of translation of the IBV and MHV E protein mRNA -----	34
1.7.4.3. Translation of accessory proteins by leaky ribosomal scanning -----	35
1.7.5. Assembly and release of coronavirus virions -----	35
1.8. Coronavirus Genetics and Reverse Genetics -----	38
1.8.1. Virus mutants -----	39
1.8.2. Defective interfering RNA -----	39
1.8.3. Targeted recombination -----	40
1.8.4. Full-length infectious cDNA -----	41
1.9. Thesis Objectives -----	44
Chapter 2	
Materials and Methods -----	45
2.1. General Reagents -----	46
2.1.1. Enzymes -----	46
2.1.2. Commercially available kits -----	46
2.1.3. DNA vectors -----	46
2.1.4. Antibodies -----	51
2.2. Solutions, Buffers and Media -----	51

2.2.1. Laboratory stocks -----	51
2.2.2. Media for bacterial culture -----	52
2.2.3. Solutions for making competent bacteria -----	53
2.2.4. Solutions for tissue culture -----	53
2.2.5. Solutions for the FITC labeling of cells -----	53
2.2.6. Buffers for SDS-PAGE -----	53
2.2.7. Western blot buffers -----	54
2.2.8. Northern blot buffers -----	54
2.3. Methods -----	55
2.3.1. Methods for DNA manipulation and subcloning -----	55
2.3.1.1. Polymerase chain reaction and site-directed mutagenesis -----	55
2.3.1.2. Purification of plasmid DNA -----	55
2.3.1.3. Quantitation of DNA -----	55
2.3.1.4. Purification of DNA fragments by agarose gel electrophoresis -----	56
2.3.1.5. DNA ligation -----	56
2.3.1.6. Construction of plasmids -----	56
2.3.1.7. Preparation of <i>E.coli</i> competent cells -----	61
2.3.1.8. Transformation of <i>E.coli</i> with plasmid DNA -----	61
2.3.2. Cell culture methods -----	61
2.3.2.1. Cells -----	62
2.3.2.2. Cell storage in liquid nitrogen -----	62
2.3.2.3. Cells recovery from liquid nitrogen -----	62

2.3.2.4. Transient expression in mammalian cells -----	62
2.3.3. <i>Experiment Assays</i> -----	63
2.3.3.1. Polyacrylamide gel electrophoresis in sodium dodecyl sulphate -----	63
2.3.3.2. Western blot analysis -----	63
2.3.3.3. Indirect immunofluorescence -----	64
2.3.3.4. Glycosylation study of E protein -----	65
2.3.3.5. Proteinase K protection assay -----	65
2.3.3.6. Isolation of low-density detergent-insoluble membrane fraction on flotation gradients -----	66
2.3.3.7. Lipid rafts aggregation and confocal microscopy -----	66
2.3.3.8. RNA extraction from mammalian cells -----	67
2.3.3.9. Northern blot analysis -----	67
2.3.3.10. Reverse transcription reaction -----	68
2.3.3.11. Analysis of VLP -----	68
2.3.3.12. Yeast two-hybrid screen -----	69
2.3.3.13. In vitro translation -----	70
2.3.3.14. Preparation of GST (glutathione S-transferase) and snapin fusion proteins -----	70
2.3.3.15. GST pull-down assay -----	70
2.3.3.16. Coimmunoprecipitation -----	70
2.3.4. <i>Generation of recombinant IBV and virus assay methods</i> -----	71

2.3.4.1. Virus -----	71
2.3.4.2. In vitro assembly of full-length cDNA clones -----	71
2.3.4.3. In vitro transcription and electroporation -----	74
2.3.4.4. Introduction of in vitro synthesized transcripts into Vero cells by electroporation -----	74
2.3.4.5. IBV plaque infectivity assay -----	74
2.3.4.6. Growth curve of the recombinant viruses on Vero cells -----	75
2.3.4.7. Virus purification -----	75

Chapter 3

Mixed Membrane Topologies, Lipid Rafts Association and Host Cell Interaction of the Severe Acute Respiratory Syndrome Coronavirus Envelope Protein -----	77
---	----

3.1. Introduction -----	78
-------------------------	----

3.2. Results -----	78
--------------------	----

3.2.1. <i>Prediction of the hydrophobicity and membrane topology of SARS-CoV E protein -----</i>	78
--	----

3.2.2. <i>Analysis of the membrane topology of SARS-CoV E protein by immunofluorescence microscopy -----</i>	81
--	----

3.2.3. <i>Analysis of the membrane topology of SARS-CoV E protein by differential permeabilization of the plasma membrane and limited proteinase K treatment -----</i>	84
--	----

3.2.4. <i>Glycosylation of SARS-CoV E protein -----</i>	88
---	----

3.2.5. <i>Cell surface expression of the SARS-CoV E protein -----</i>	92
---	----

3.2.6. <i>Partial association of SARS-CoV E protein and IBV E protein with lipid rafts</i> -----	95
3.2.7. <i>Identification of snapin as a SARS-CoV E-interacting protein</i> -----	99
3.3. Summary -----	103
Chapter 4	
Systematic Analysis of the N-Terminal Domain of the Infectious Bronchitis Virus Membrane Protein -----	
4.1. Introduction -----	104
4.2. Results -----	105
4.2.1. <i>N-linked glycosylation of IBV M protein</i> -----	105
4.2.2. <i>Functional analysis of the N-linked glycosylation of IBV M protein</i> –	109
4.2.2.1. Construction of mutant IBV virus with mutations of the two N-linked glycosylation sites in the M protein -----	109
4.2.2.2. Characterization of the N3D/6D mutant virus -----	111
4.2.2.3. Genetic stability and growth kinetic of N3D/6D mutant virus -----	114
4.2.2.4. Effect of mutations of N-linked glycosylation sites on the subcellular localization of M protein -----	117
4.2.2.5. Analysis of the N3D/6D mutant virus by electron microscopy -----	119
4.2.2.6. Sedimentation analysis of particles in sucrose gradients -----	121

4.2.3. Systematic analysis of the N-terminal domain of the IBV M protein -	123
4.2.3.1. Construction of full-length IBV clones with mutations in the N-terminal domain and recovery of infectious mutant viruses -----	123
4.2.3.2. Growth properties of the mutant viruses -----	127
4.2.3.3. Effect of the N-terminal domain mutations on VLP assembly -----	130
4.2.3.4. Analysis of negative-strand RNA replication and subgenomic RNA transcription in cells transfected with wild-type and mutant full-length transcripts -----	133
4.3. Summary -----	137
Chapter 5	
Discussion and Future Direction -----	138
5.1. Characterization of SARS-CoV and IBV E protein -----	139
5.1.1. Mixed membrane topologies of the SARS-CoV small envelope protein -----	139
5.1.1.1. An $N_{\text{cyto}}C_{\text{cyto}}$ topology adopted by majority of the SARS-CoV E -----	139
5.1.1.2. A small proportion of the SARS-CoV E protein with the C-terminus exposed to the luminal side -----	142
5.1.1.3. Cell surface expression of the SARS-CoV E protein -----	143

5.1.2. <i>N-linked glycosylation of SARS-CoV E protein</i> -----	144
5.1.3. <i>Partial association of SARS-CoV E and IBV E proteins with lipid rafts</i> -----	145
5.1.4. <i>Interaction of SARS-CoV E protein with snapin</i> -----	146
5.2. <i>Characterization of the N-terminal domain of IBV M protein</i> -----	147
5.2.1. <i>N-linked glycosylation of IBV M protein is not required for virus assembly in cultured cells</i> -----	147
5.2.2. <i>Mutations in the N-terminal domain of IBV M protein do not affect the formation of VLP</i> -----	148
5.2.3. <i>Mutations in the N-terminal domain of IBV M protein had effects on the recovery of infectious virus</i> -----	150
5.3. <i>Future Direction</i> -----	151
5.3.1. <i>Is the IBV E protein essential for virus replication</i> -----	151
5.3.2. <i>Is palmitoylation of E protein necessary for its association with lipid rafts</i> -----	152
5.3.3. <i>The involvement of N-linked glycosylation of IBV M in ER stress</i> -----	152
Reference List -----	154
Curriculum Vitae -----	178

List of Figures and Tables

Chapter 1

Table 1 Coronavirus species and groups -----	3
Table 2 Coronavirus receptors -----	27
Fig. 1-1. Morphology of coronavirus -----	6
Fig. 1-2. Comparative genome structure of the different coronaviruses -----	9
Fig. 1-3. IBV genomic and subgenomic RNA -----	10
Fig. 1-4. Schematic fusion model of the coronavirus S fusion protein -----	13
Fig. 1-5. Scheme of coronavirus gene 1 organization -----	23
Fig. 1-6. The coronavirus life cycle -----	28
Fig. 1-7. Discontinuous transcription model of subgenomic RNA synthesis -----	31

Chapter 2

Table 3. Primers used to amplify SARS E, IBV E and snapin -----	59
Table 4. Primers used in IBV M protein mutagenesis -----	60
Fig. 2-1. pKT0 vector circle map and MCS sequence -----	49
Fig. 2-2. pFlag vector circle map and MCS sequence -----	50
Fig. 2-3. Strategy for the assembly of an IBV infectious cDNA clone -----	73

Chapter 3

Fig. 3-1. Diagram of wild-type and mutant SARS-CoV E constructs, and hydrophathicity score of SARS-CoV E protein -----	80
---	----

Fig. 3-2. Cytoplasmic exposure of the Flag epitope tagged at the N- and C-terminus of the SARS-CoV E protein and the C-terminus of the IBV E protein	83
Fig. 3-3. Limited proteinase K digestion of microsomal membranes prepared from cells expressing the SARS-CoV and IBV E protein tagged with the Flag epitope at the N- and C-termini, respectively	87
Fig. 3-4. N-linked glycosylation of SARS-CoV E protein	91
Fig. 3-5. Surface staining of HeLa cells expressing the SARS-CoV E protein with Flag tagged at the N- and C-terminus, respectively, and IBV E protein with the Flag tagged at the N-terminus	94
Fig. 3-6. Association of SARS-CoV and IBV E protein with lipid rafts	97
Fig. 3-7. IBV E protein colocalizes with GM1	98
Fig. 3-8. Interaction of SARS-CoV E protein and snapin	101

Chapter 4

Fig. 4-1. Diagram of wild-type and mutant IBV M constructs	107
Fig. 4-2. Analysis of glycosylation of IBV M	108
Fig. 4-3. In vitro assembly of the cDNA for the full-length IBV genome	110
Fig. 4-4. Characterization of mutant virus mutation site, proteins and mRNAs	112
Fig. 4-5. Plaque morphologies and growth kinetics of the rIBV and N3D/6D viruses	115
Fig. 4-6. Genetic stability of N3D/6D mutant virus	116

Fig. 4-7. Intracellular distribution of M in rIBV infected cells and N3D/6D infected cells -----	118
Fig. 4-8. Electron microscopic analysis -----	120
Fig. 4-9. Sedimental analysis of rIBV and N3D/6D -----	122
Fig. 4-10. Alignment of N-terminal sequences of four membrane proteins encoded by various coronaviruses -----	125
Fig. 4-11. Fusion foci of Vero cells electroporated by transcripts of GFPL15 and N -----	126
Fig. 4-12. Plaque morphologies and growth kinetics of rIBV, C7A, L9A, Q16A, V15A/Q16A and F18A viruses -----	128
Fig. 4-13. Effects of mutations in N-terminal domain of M protein on VLP assembly -----	131
Fig. 4-14. Co-expression of the M gene with S, E and N genes in various combinations and analysis of structure proteins in VLPs -----	132
Fig. 4-15. Analysis of negative strand RNA replication and subgenomic RNA transcription in cells transfected with wild-type and mutant full-length transcripts -----	135

Chapter 5

Fig. 5-1. Proposed membrane topologies of the unglycosylated form of the SARS-CoV E (a), the IBV E (a), and the glycosylated form of the SARS CoV E protein (b) -----	141
---	-----

List of Abbreviations

ACE 2	angiotensin converting enzyme 2
APN	aminopeptidase N
BAC	bacterial artificial chromosome
BCoV	bovine coronavirus
CEA	carcinoembryonic antigen
CECoV	canine enteric coronavirus
CPE	cytopathic effect
CRCoV	canine respiratory coronavirus
CS	conserved sequence
DCoV	duck coronavirus
DI	defective interfering
DMEM	.Dulbecco's Modified Eagle Medium
E	Envelope
ERGIC	ER-Golgi intermediate compartment
ExoN	Exonuclease
FBS	fetal bovine serum
FCoV	feline coronavirus
FIPV	feline infectious peritonitis virus
GCoV	goose coronavirus

GFP	green fluorescent protein
HCoV 229E	human coronavirus 229E
HE	hemagglutinin-esterase
HEV	haemagglutinating encephalomyelitis coronavirus
hnRNA	heterogeneous nuclear ribonucleoprotein
IBV	infectious bronchitis virus
Ig	Immunoglobulin
IRES	internal ribosome entry site
M	membrane protein
MHV	mouse hepatitis virus
MOI	multiplicity of infection
NendoU	Endoribonuclease
PEDV	porcine epidemic diarrhea virus
PFU	plaque-forming unit
PiCoV	pigeon coronavirus
PMSF	phenylmethylsulfonyl fluoride
PTB	polypyrimidine tract-binding protein
rIBV	recombinant infectious bronchitis virus
RbCV	rabbit coronavirus
RER	rough endoplasmic reticulum
RF	replicative form

RdRp	RNA-dependant RNA polymerase
S	spike protein
SARS	severe acute respiratory syndrome
SARS-CoV	severe acute respiratory syndrome coronavirus
TCID ₅₀	50% tissue culture infection dose
TCoV	turkey coronavirus
TGEV	transmissible gastroenteritis virus
TRS	transcriptional regulatory sequence
ts	temperature-sensitive
UTR	untranslated region
UV	Unltraviolet
VLP	virus-like-particle
vTF7-3	vaccinia virus-T7
W/V	weight/volume

Chapter 1

Introduction

1.1. Classification

Nidovirales, from the Latin *nidus*, has been designated for the order, as all members produce mRNAs in an extensive nested-set arrangement. The order *Nidovirales* is comprised of *Arteriviridae* and *Coronaviridae* families (Cavanagh & Horzinek, 1993; Cavanagh, 1997). The *Coronaviridae* includes both the coronavirus and torovirus genus (Cavanagh & Horzinek, 1993; Cavanagh, 1997). Based on the sequence comparisons of entire viral genomes, coronaviruses have been classified into three distinct groups (Table 1.1). Group 1 coronaviruses include transmissible gastroenteritis virus (TGEV), porcine epidemic diarrhea virus (PEDV), feline infectious peritonitis virus (FIPV), rabbit coronavirus (RbCV) and human coronavirus 229E (HCoV 229E). Group 2 coronaviruses include mouse hepatitis virus (MHV), bovine coronavirus (BCoV) and human coronavirus OC43. Infectious bronchitis virus (IBV) belongs to Group 3. In year 2003, it comes as a shock that the causative agent of severe acute respiratory syndrome (SARS) was a coronavirus (Rota *et al.*, 2003). According to phylogenetic analysis, Rota *et al.* suggest that SARS-CoV might form a subgroup within Group 2 (Rota *et al.*, 2003). Following the SARS epidemic, renewed efforts to detect unknown viruses led to the discovery of two more human coronaviruses and three distinct bat coronaviruses (van der *et al.*, 2004; Woo *et al.*, 2005; Lau *et al.*, 2005; Li *et al.*, 2005; Poon *et al.*, 2005).

Table 1 Coronavirus Species and Groups

Group I

Subgroup 1a

Porcine transmissible gastroenteritis coronavirus (TGEV)

Canine enteric coronavirus (CECoV)

Feline coronavirus (FCoV)

Subgroup 1b

Porcine epidemic diarrhea coronavirus (PEDV)

Human coronavirus 229E (HCoV)

Bat coronavirus

Group II

Subgroup 2a

Murine hepatitis coronavirus (MHV)

Human coronavirus OC43 (HCoV OC43)

Bovine coronavirus (BCoV)

Canine respiratory coronavirus (CRCoV)

Porcine haemagglutinating encephalomyelitis coronavirus (HEV)

Subgroup 2b

SARS-coronavirus (SARS-CoV)

Group III

Infectious bronchitis coronavirus (IBV)

Turkey coronavirus (TCoV)

Goose coronavirus (GCoV)

Duck coronavirus (DCoV)

Pigeon coronavirus (PiCoV)

1.2. Virion Morphology

Coronaviruses are spherical to pleomorphic enveloped viruses with diameters ranging from 60 to 220 nm (Fig. 1-1a). The virus core is enclosed in a host cell derived envelope (Tooze *et al.*, 1984; Griffiths & Rottier, 1992). The large surface projections of the coronaviruses give them the appearance of a crown, hence the name coronavirus. Inside the virion, the genomic RNA is enclosed with a helical nucleocapsid (Macneughton & Davies, 1978; Sturman *et al.*, 1980). All coronaviruses contain four major structural proteins (Fig. 1-1b). The spike (S) glycoprotein is the outmost component, which gives the virion its corona-like appearance, and is responsible for the attachment of the virus to cells (Collins *et al.*, 1982; Godet *et al.*, 1994; Kubo *et al.*, 1994). The second structural protein is membrane (M) glycoprotein, which spans the membrane three times (Machamer & Rose, 1987; Machamer *et al.*, 1990; Machamer *et al.*, 1993). The third protein is a small envelope (E) protein, which plays an essential role in virion assembly and is present in much smaller amounts (Liu *et al.*, 1991; Godet *et al.*, 1992; Yu *et al.*, 1994). The fourth protein, nucleocapsid (N), is a phosphoprotein, which interacts with viral genomic RNA and provides the structural basis for the helical nucleocapsid. The ratios of the S/E/M/N proteins for TGEV virions have been estimated to be 20:1:300:100 respectively. In addition, the virions of some antigenic group 2 coronavirus (MHV, BCoV) have a hemagglutinin-esterase (HE) glycoprotein attached to the surface of the virion (Stern & Sefton, 1982a; King *et al.*, 1985; Hogue *et al.*, 1989; Shieh *et al.*, 1989; Yokomori *et al.*, 1989). Recently, the supermolecular architecture of SARS-CoV was revealed

by electron cryomicroscopy (Neuman et al., 2006). It was shown that proteins near the viral membrane were arranged in overlapping lattices surrounding a disordered core, the trimeric glycoprotein spikes were in register with four underlying ribonucleoprotein densities and the ribonucleoprotein particles displayed coiled shapes when released from the viral membrane.

a.

b.

Fig. 1-1. Morphology of coronavirus. (a) Single virus particle negatively stained with PTA (Proc Natl Acad Sci 57: 933–940). Magnification 192,000 $\times$. (b) A schematic diagram showing the structure proteins: spike (S), envelope (E), membrane (M), nucleocapsid (N) and hemagglutinin-esterase (HE) proteins and the RNA genome.

1.3. Genome Structure and mRNAs of Coronavirus

Coronaviruses are positive-sense, single-stranded RNA viruses. The viral genome consists of a positive-sense RNA ranging from 27.6-31 kb, which is the largest viral RNA genome known. The virion RNA functions as an mRNA and is infectious. It contains approximately 7-10 functional genes, 4 or 5 of which encode structural proteins. The genes are arranged in the order 5'-polymerase-(HE)-S-E-M-N-3', with a variable number of mostly nonstructural and largely nonessential genes interspersed among them (Fig. 1-2). However, rearrangement of the gene order of MHV was found to be completely tolerated by the virus (de Haan et al., 2002). At the termini of the genome are a 5' untranslated region (UTR), ranging from 210 to 530 nucleotides, and a 3' UTR, ranging from 270 to 500 nucleotides. Almost two-thirds of the entire RNA is occupied by the polymerase gene, which comprises two overlapping ORFs, 1a and 1b. At the overlapping region is a specific seven-nucleotide slippery sequence and a pseudoknot structure, characteristic of the ribosomal frameshifting signal (Brierley *et al.*, 1987; Brierley *et al.*, 1989; Lee *et al.*, 1991; Herold & Siddell, 1993).

Coronavirus RNA synthesis occurs via an RNA-dependent RNA transcription process. Coronavirus has six to nine species of mRNAs depending on the coronavirus species and strains. The largest mRNA is equivalent to the genomic RNA, and the remainders are subgenomic in size. They have a nested-set structure, and all of them contain sequences starting at the 3' terminus and extending to various distances toward the 5' end (Fig. 1-3) (Stern & Kennedy, 1980; Lai *et al.*, 1981; Leibowitz *et*

al., 1981). Another unique structural feature of coronavirus mRNAs is that their 5' ends have a leader sequence of approximately 60-90 nucleotides which is derived from the 5' end of the genomic RNA (Lai *et al.*, 1982; Spaan *et al.*, 1983; Lai *et al.*, 1983; Lai *et al.*, 1984).

Fig. 1-2. Comparative genome structures of different coronaviruses. Genome organizations of coronavirus representatives of group 1 (HCoV-229E), group 2 (MHV; SARS-CoV) and group 3 (IBV) are shown. Gene 1 sequences are interrupted and shortened to highlight the remaining genes. Red boxes represent the accessory genes. The positions of the leader sequence (L) and poly (A) tract are indicated.

Fig. 1-3. IBV genomic and subgenomic RNAs. RNAs 1-6 form a 3'co-terminal nested set. RNA 1 encodes proteins of the replicase complex, 2 encodes the spike protein, 3 encodes the 3a, 3b and E proteins, 4 encodes the M protein, 5 encodes the 5a and 5b proteins and 6 encodes the nucleocapsid protein.

1.4. Structural Proteins

1.4.1. The Spike (S) glycoprotein

The spike glycoprotein makes up the surface projections which form the distinctive corona. It is a class I fusion protein that is co-translationally glycosylated to a 128-160 kDa glycoprotein, which is processed to a 150-200 kDa form during transport from the endoplasmic reticulum through the Golgi complex. The S protein has an N-terminal signal sequence and a membrane-anchoring sequence near the C-terminus. The ectodomain makes up most of the molecule, with only a small carboxy-terminal segment about 71 or fewer residues constituting the transmembrane domain and endodomain.

The S protein is responsible for the attachment of the virus to cells and for instigating the fusion of the virus envelope with cell membranes (Collins *et al.*, 1982; Godet *et al.*, 1994; Kubo *et al.*, 1994). The S protein functions to define viral tropism by their receptor-binding specificity. Another important biological activity of S protein is the induction of membrane fusion. This activity may be required for viral entry into cells or for cytopathic effects. Expression of the recombinant S gene alone is sufficient to cause membrane fusion, as shown by syncytium formation (de Groot *et al.*, 1989; Pflleiderer *et al.*, 1990; Yoo *et al.*, 1991; Taguchi, 1993). The S protein is the primary target for host's immune response, and neutralizing antibodies are mainly induced by S protein (Collins *et al.*, 1982).

The S protein is heavily glycosylated (Luytjes *et al.*, 1987). A mass spectrometric analysis of the SARS S protein showed that at least 12 out of the 23

candidate sites are glycosylated (Mandal *et al.*, 2003). Disulfide bonds are required for correct folding and the S protein is found in trimer on the viral envelope (Delmas *et al.*, 1990; Opstelten *et al.*, 1993). Some coronaviruses, such as IBV, BCV, TCV, PEDV and MHV, have a cleavage site for a trypsin-like protease near the middle of the S protein, but there is no cleavage of S protein in others (TGEV, FCoV, CCV and SARS-CoV) (Garwes & Pocock, 1975; Cavanagh, 1983; Xiao *et al.*, 2003). When cleaved, S protein yields two proteins called S1 (N-terminal) and S2 (C-terminal) that are roughly equal in size (Sturman & Holmes, 1977; Frana *et al.*, 1985). Two heptad repeats present in the S2 portion are presumably important for the trimerization of S and may be important for fusion of the viral envelope with the host membrane (Frana *et al.*, 1985; de Groot *et al.*, 1987).

Fig. 1-4. Schematic fusion model of the coronavirus S fusion protein. State 1, the S protein is denoted S1 and S2 for its N- and C-terminal domains; state 2, S1 is dissociated to expose the fusion peptide (FP) region, and state 3, the collapsed S2 domain brings the viral and cellular membranes together causing fusion. HR-N and HR-C represent coiled coils at the N terminus and C terminus of the S2 domain. X represents the host cell surface receptor.

1.4.2. The membrane (M) protein

The M protein is associated with the viral membrane, the major portion of which appears to be embedded in the virion membrane. The sequence of the M protein reveals that the M polypeptide comprises 221-262 amino acids. The N-terminal 20 or so residues are hydrophilic, exposed at the virion surface, and have a number of glycosylation sites. For most group 2 coronaviruses, glycosylation is O-linked except MHV strain 2 and SARS-CoV (Yamada *et al.*, 2000; Oostra *et al.*, 2006). Group 1 and group 3 coronavirus M proteins are N-glycosylated exclusively (Stern & Sefton, 1982b; Cavanagh & Davis, 1988). Although the roles of M protein glycosylation are not fully understood, the glycosylation status of M can influence the organ tropism *in vivo* and the alpha interferon inductivity *in vitro* (Charley & Laude, 1988; Laude *et al.*, 1992; de Haan *et al.*, 2003a). The remainder of the N-terminal half of M protein forms three helical membrane-spanning domains. The structure of the C-terminal half is uncertain, but it is believed to be situated on the inside of the viral envelope (Rottier *et al.*, 1984; Cavanagh *et al.*, 1986). However, some M molecule of TGEV virions have the C terminus exposed at the virion surface (Risco *et al.*, 1995). Multiple higher-molecular-mass glycosylated forms of M protein are often observed by SDS-PAGE (Sturman, 1977; Locker *et al.*, 1992).

The M protein is an essential element for the formation of the viral membrane and budding of virions (Stern & Sefton, 1982a). When the M protein was expressed alone, it was localized in the Golgi complex, near the location where virus particles bud (Tooze *et al.*, 1984; Tooze & Tooze, 1985). However, studies showed that the

site of M protein retention in the Golgi was slightly different from that for viral particle budding (Klumperman *et al.*, 1994), suggesting that additional factors are involved in virus particle assembly. The M and E proteins are essential for viral envelope formation and the release of virus particles. Virus-like-particles (VLPs) which are indistinguishable from authentic coronavirus virions in size and shape are released from cells that express both M and E proteins (Corse & Machamer, 2000). In addition, the M protein directs the incorporation of the S protein and the N protein into the budding particle (Opstelten *et al.*, 1995; Nguyen & Hogue, 1997; Narayanan *et al.*, 2000).

1.4.3. The envelope (E) protein

Coronavirus E protein is a small envelope protein present in virions at low levels, and ranges in size from 76 to 109 amino acids (Liu *et al.*, 1991; Godet *et al.*, 1992; Yu *et al.*, 1994). The E proteins from avian coronavirus IBV and MHV are translated from the third and second ORFs of mRNA 3 and 5 of the respective viruses by a cap-independent, internal ribosomal entry mechanism (Bournsnell *et al.*, 1985; Skinner *et al.*, 1985; Budzilowicz & Weiss, 1987; Leibowitz *et al.*, 1988; Liu *et al.*, 1991; Liu & Inglis, 1992b; Thiel & Siddell, 1994). In some coronaviruses, for example SARS-CoV, the E protein is derived from a monocistronic mRNA. There is a low degree of sequence identity among E proteins from three coronavirus groups, and in some cases, among members of the same group. A number of general features can be highlighted in all coronavirus E proteins, i.e., a short hydrophilic N terminal region,

followed by a long hydrophobic stretch of 21-29 amino acid residues, two to four cysteine residues immediately downstream of the hydrophobic stretch, and a hydrophilic C-terminal tail.

In cells infected with some coronaviruses, such as IBV, MHV and BCoV, the E protein exhibits a granular or punctuate pattern by immunofluorescent staining (Smith *et al.*, 1990; Abraham *et al.*, 1990; Yu *et al.*, 1994). A portion of the E protein was transported to the cell surface (Smith *et al.*, 1990; Godet *et al.*, 1992; Yu *et al.*, 1994). The IBV E protein was shown to be localized to the Golgi complex when expressed on its own, co-expressed with M protein and in IBV-infected cells (Corse & Machamer, 2000). The C-terminal tail of the IBV E protein directs the Golgi targeting (Corse & Machamer, 2002). The SARS-CoV and IBV E proteins may be also transiently localized to the ER at early stages of the infection cycles (Lim & Liu, 2001; Nal *et al.*, 2005; Liao *et al.*, 2006). The MHV E protein was found to accumulate in the ER-Golgi intermediate compartment as judged by their co-labeling with antibodies against E and Rab-1 (Raamsman *et al.*, 2000).

Biochemical characterization of coronavirus E protein showed that the protein may undergo post-translational modifications. For example, the IBV and SARS-CoV E proteins are modified by palmitoylation on one or more cysteine residues (Corse & Machamer, 2002; Liao *et al.*, 2006). However, attempts to label the MHV and TGEV E proteins with ³H-labeled palmitic acid were unsuccessful (Godet *et al.*, 1992; Raamsman *et al.*, 2000). It is still unknown whether this modification is a general characteristic of the coronavirus E protein. The coronavirus E protein can form

homooligomers. Molecular dynamic simulations predict that SARS E forms pentamers (Torres *et al.*, 2005).

As mentioned earlier, different coronavirus E proteins share significant similarities in their biochemical properties as well as biological functions. However, they seem to adopt distinct membrane topology. Early studies showed that cell surface staining was detected in cells infected with TGEV using specific antibody against C-terminal of the TGEV E protein, suggesting that the C-terminal region of the protein exposes to the external surface of the cells (Godet *et al.*, 1992). Immunofluorescent staining and biochemical analysis of cells expressing MHV E protein revealed that the C-terminal region is exposed to the cytoplasm and the N-terminal region is probably buried in the membrane near the cytoplasmic side ($N_{\text{endo}}C_{\text{endo}}$) (Raamsman *et al.*, 2000; Maeda *et al.*, 2001). This topology is supported by the observation that treatment of MHV particles with proteinase K did not alter the molecular mass of the E protein (Raamsman *et al.*, 2000; Maeda *et al.*, 2001). In the case of IBV E protein, immunofluorescence assay showed that C-terminally specific antibodies could detect the protein in cells treated with either digitonin or Triton X-100, but the N-terminally specific antibodies can detect the protein only in cells treated with Triton X-100, demonstrating that the IBV E protein adopts an $N_{\text{exo}}C_{\text{endo}}$ topology (Corse & Machamer, 2000). Recently, the biophysical properties and membrane topology of the SARS-CoV E protein were studied by several groups (Arbely *et al.*, 2004; Khattari *et al.*, 2006; Torres *et al.*, 2006). Based on *in vitro* biophysical studies, Arbely *et al.* proposed an unusual short, palindromic

transmembrane helical hairpin for the putative transmembrane domain of the protein (Arbely *et al.*, 2004; Khattari *et al.*, 2006). However, a regular α -helical structure with the N-terminal and C-terminal regions in opposite sides of the lipid bilayer was proposed based on similar *in vitro* biophysical studies (Torres *et al.*, 2006). The reason for this discrepancy is not clear.

Coronavirus E protein is a multi-functional protein. Functionally, coronavirus E protein plays a pivotal role in viral morphogenesis, assembly and budding. Co-expression of MHV E and M was shown to result in the production of VLPs, roughly the same size and shape as virions (Vennema *et al.*, 1996; Maeda *et al.*, 1999). Other coronaviruses such as SARS-CoV and IBV showed the same phenomena (Corse & Machamer, 2003; Ho *et al.*, 2004). When the E protein was expressed in mammalian cells on its own, the E-containing vesicles can be produced and released to the culture medium (Maeda *et al.*, 1999; Corse & Machamer, 2003). Furthermore, MHV and SARS-CoV E proteins were shown to induce apoptosis (An *et al.*, 1999; Yang *et al.*, 2005). However, induction of apoptosis does not appear to be a general feature of the coronavirus E protein and is cell-type specific. More recently, the SARS-CoV E protein was found to induce membrane permeability changes as well as to form pores in artificial membranes (Wilson *et al.*, 2004; Liao *et al.*, 2004b; Liao *et al.*, 2006). It was also reported that MHV E protein could modify the membrane permeability in both *E. coli* and mammalian cells (Madan *et al.*, 2005).

1.4.4. The nucleocapsid (N) protein

The N protein associates with the viral genomic RNA to form a long helical nucleocapsid. Unlike the S and M proteins, which are glycosylated, the N protein is phosphorylated. The N protein varies from 377 to 455 amino acids and has a high (7-11%) serine content, potential targets for phosphorylation. Based on sequence comparison, three domains in the N protein have been identified (Parker & Masters, 1990). The first and second domains are rich in arginines and lysines. By contrast, the third domain is negatively charged due to an excess of acidic over basic residues. The main activity of N protein is to bind the viral RNA. Both sequence specific and nonspecific modes of RNA binding are reported (Stohlman *et al.*, 1988b; Masters, 1992; Nelson & Stohlman, 1993; Molenkamp & Spaan, 1997; Cologna *et al.*, 2000; Chen *et al.*, 2005). Specific RNA substrates for N protein include the transcription regulating sequence, the 3'UTR, the N gene and the genomic RNA package signal (Stohlman *et al.*, 1988b; Masters, 1992; Nelson & Stohlman, 1993; Zhou *et al.*, 1996; Molenkamp & Spaan, 1997; Cologna *et al.*, 2000; Chen *et al.*, 2005). The RNA binding domain was mapped to domain 2 of MHV N protein (Masters, 1992; Nelson & Stohlman, 1993). However, for IBV, two separate RNA-binding sites were mapped to N-terminal and C-terminal regions of N protein (Zhou & Collisson, 2000). For SARS-CoV, the RNA binding activity was found in domains 1 and 2 (Huang *et al.*, 2004a). According to the X-ray structure of the N-terminal domain of IBV N protein and the NMR structure of the N-terminal domain of SARS-CoV N protein, the N-terminal RNA-binding domain of N protein from both IBV and SARS-CoV contains a flexible and positively charged hairpin loop that extends much beyond the

protein core and that could grasp a nucleic acid substrate by neutralizing its phosphate groups (Huang *et al.*, 2004a; Fan *et al.*, 2005; Jayaram *et al.*, 2006). Based on the X-ray structure of the N-terminal domain of IBV N protein, the Arg-76 and Tyr-94 residues were identified to be essential for the RNA-binding activity of the IBV N protein and viral infectivity by site-directed mutagenesis (Tan *et al.*, 2006).

Besides the structural role in the virion, the N protein has a number of potential activities. The N protein plays an important role in the replication of the genomic RNA (Chang & Brian, 1996), and in the transcription of subgenomic RNAs (Baric *et al.*, 1988; Stohlman *et al.*, 1988a). Finally, it was shown that, in addition to its presence in the cytoplasm, IBV N protein localizes to the nucleolus in the absence of other viral proteins as well as in the virus-infected cell (Hiscox *et al.*, 2001). TGEV and MHV N proteins also localize to the nucleolus in the absence of other viral proteins, though not in the virus-infected cells as 90% of the infected cells do not have detectable N protein in the nucleolus (Wurm *et al.*, 2001). Both MHV and IBV N proteins were found to bind to two nucleolar proteins, fibrillarin and nucleolin (Chen *et al.*, 2002). Correlated with such localization, N protein might inhibit host cell proliferation or delay cell growth (Baric *et al.*, 1988; Chen *et al.*, 2002).

1.5. The Replicase Polypeptide

Upon penetration of the virus into the host cell cytoplasm, the RNA genome is uncoated and translated through a -1 ribosomal frameshift to generate two polypeptides, pp1a and pp1ab ranging from 440 to 500 kDa and from 740 to 810 kDa,

respectively. These polyproteins are proteolytically processed into 16 cleavage products, designated nsp1-nsp16 (Fig. 1-5). IBV lacks the counterpart of nsp1, so there are 15 cleavage products. A number of functional domains within replicase have been predicted by computer-based motif analysis (Bournsnel *et al.*, 1987; Hodgman, 1988; Gorbalenya *et al.*, 1989; Lee *et al.*, 1991; Snijder *et al.*, 2003).

The cleavage products in ORF 1a contain three viral protease domains, two cysteine papain-like protease domains (PLP-1 and PLP-2) in nsp3 and a picornavirus 3C-like protease domain (3CL^{pro}) in nsp5. MHV, HCoV-229E and TGEV have two PLP domains, with PLP2 corresponding to the single PLP domain of IBV. Unlike all other coronaviruses, SARS-CoV does not have an ortholog of PLP-1. In addition to the papain-like proteinases, nsp3 contains a domain with ADP-ribose-1''-monophosphatase activity (Putics *et al.*, 2005). The construction of active-site mutants has shown that this activity is dispensable for replication of HCoV-229E in tissue culture. The coronavirus main proteinase constitutes the nsp5. This enzyme is also called the 3C-like proteinase (3CL^{pro}), because of its resemblance to the 3C proteinase of picornaviruses. Crystal structures reveal that the 3C-like proteinase is a dimer, each monomer of which has a three-domain structure, with an active site located in a cleft between the first and second domains (Anand *et al.*, 2002; Anand *et al.*, 2003; Yang *et al.*, 2003). Multiple structures determined for the 3C-like proteinase of SARS-CoV suggested that it undergoes pH-dependent conformational changes (Tan *et al.*, 2005). Nsp3, nsp4 and nsp6 contain transmembrane domains which anchor the replication complex to the intracellular membranes. At the carboxy-terminus of pp1a is a cluster

of small proteins, nsp7-nsp10. Based on the co-crystal structure of SARS-CoV nsp7 and nsp8, this complex was proposed to be able to encircle an RNA template, possibly acting as a processivity factor for the RNA polymerase (Zhai et al., 2005). The crystal structure of nsp9 revealed that it has nonspecific RNA-binding activity (Egloff *et al.*, 2004; Sutton *et al.*, 2004). Recently, the crystal structure of nsp10 was resolved by two different groups (Joseph *et al.*, 2006; Su *et al.*, 2006). Gel shift assays indicated that nsp10 binds single- and double-stranded RNA and DNA with high-micromolar affinity and without obvious sequence specificity. It is possible that nsp10 functions within a larger RNA-binding protein complex.

The functional domains for RNA synthesis are located in ORF 1b which encodes nsp12-nsp16. Nsp12 is the coronavirus RNA-dependant RNA polymerase (RdRp). It has the fingers, palm and thumb domains common to a number of viral RdRps and reverse transcriptases. Nsp13 is a helicase with a highly processive unwinding activity for both DNA and RNA substrates. It also has RNA-dependent NTPase and dNTPase activities which provide energy for its translocation along RNA template. Nsp14 has been predicted to be an exonuclease (ExoN). Nsp15 is an endoribonuclease (NendoU), which was shown to hydrolyze both single- and double-stranded RNA, with a specificity for cleavage immediately upstream and downstream of uridylate residus (Bhardwaj et al., 2004). Nsp16 has been predicted to contain 2'-O-methyltransferase activity, which takes a role in RNA capping.

Fig. 1-5. Scheme of coronavirus gene 1 organization. Cleavage sites and processed products of pp1a and of pp1ab are shown. Predicted or experimentally demonstrated activities are indicated.

1.6. Accessory Proteins

The coronavirus accessory proteins were originally labeled nonstructural proteins, but some of them, such as the group 2 HE proteins and the SARS 3a, 7a, 7b proteins, were shown to be incorporated into the virions (Ito *et al.*, 2005; Huang *et al.*, 2006; Schaecher *et al.*, 2006). Coronaviruses exhibit great heterogeneity with respect to the number and genome location of nonstructural proteins. For IBV, four accessory proteins (3a, 3b, 5a and 5b) are produced by subgenomic mRNAs 3 and 5. The 3a, 3b, 5a and 5b proteins of IBV have been identified in virus infected cells (Liu *et al.*, 1991; Liu & Inglis, 1992a). For SARS-CoV, there are eight nonstructural proteins (3a, 3b, 6, 7a, 7b, 8a, 8b and 9b). The functions of these accessory proteins are largely unknown. Recently, the SARS-CoV 3a protein was found to form an ion channel that may promote virus release (Lu *et al.*, 2006). Expression of SARS-CoV 3a and 7a proteins in A549 cells can enhance the production of inflammatory chemokines (Kanzawa *et al.*, 2006). Expression of SARS-CoV protein 6 in mammalian cells stimulated cellular DNA synthesis (Geng *et al.*, 2005).

In all cases examined, through natural or engineered mutants, accessory proteins have been found nonessential for viral replication in tissue culture. The dispensability has been determined for 3a, 3b, 5a and 5b proteins of IBV and 3a, 3b, 6, 7a and 7b proteins of SARS-CoV (Casais *et al.*, 2005; Yount *et al.*, 2005; Hodgson *et al.*, 2006).

1.7. Coronavirus Life Cycle

1.7.1. Attachment, penetration and uncoating

Coronavirus lifecycle begins with specific binding of the virion to the plasma membranes of target cells (Fig. 1-6). The main determinant of cell tropism is the binding of the S protein to a specific receptor on the cell surface. Receptors have been identified for several coronaviruses. The known cellular receptors for coronaviruses are listed in Table 2. Aminopeptidase N (APN) serves as a receptor for antigenic group 1 coronaviruses including TGEV, HCoV 229E, CCoV and FIPV (Delmas *et al.*, 1992; Yeager *et al.*, 1992; Tresnan *et al.*, 1996). APN is a cell-surface, zinc-binding protease that contributes to the digestion of small peptides in respiratory and enteric epithelia; it is also found in human neural tissue (Lachance *et al.*, 1998). The MHV receptor (mCEACAM1) is a member of the carcinoembryonic antigen (CEA) family, a group of proteins within the immunoglobulin (Ig) superfamily (Williams *et al.*, 1991; Dveksler *et al.*, 1993). Other group 2 coronaviruses use different receptors. The 9-*O*-acetyl sialic acid was identified as a cell attachment factor for BCoV (Schultze *et al.*, 1991). The angiotensin converting enzyme 2 (ACE2) was identified as the receptor of SARS-CoV (Li *et al.*, 2003). ACE2 is a zinc-binding carboxypeptidase that is involved in the regulation of heart function. Human CD209L (also called L-SIGN or DC-DIGNR), a lectin family member, was found to act as a second receptor for SARS-CoV, but it has much lower efficiency than does ACE2 (Jeffers *et al.*, 2004). Recently, IBV was shown to preferentially recognize α 2,3-linked sialic acid and this binding may be used by IBV for primary attachment to the cell surface (Winter *et al.*, 2006).

After the virus binds to a receptor, it enters the cell through the fusion of the viral envelope with the plasma or endosomal membranes. IBV, MHV and BCoV exhibit a neutral or slightly alkaline optimal pH for the induction of cell to cell fusion (Payne & Storz, 1988; Weismiller *et al.*, 1990; Sturman *et al.*, 1990; Li & Cavanagh, 1992). These viruses probably enter the cell by virus-cell fusion at the plasma membrane. MHV may also enter cells through the endosomal pathway (Krzystyniak & Dupuy, 1984; Mizzen *et al.*, 1985). The transmembrane domain of the S protein has been shown to induce conformation changes after binding to the receptor (Weismiller *et al.*, 1990). The conformational change that separates S1 from the rest of the molecule, in turn, transmits a major change to S2. The coronavirus S2 domain contains two separated heptad repeats, HR1 and HR2, with a fusion peptide upstream of HR1 and the transmembrane domain immediately after HR2. The receptor-mediated conformational change in S1 and the dissociation of S1 from S2 initiate a rearrangement in the remaining S2 trimer. This rearrangement exposes a fusion peptide that interacts with host cellular membrane and brings together the two heptad repeats to form an antiparallel six-helix bundle. The result is the sufficient proximity of the viral and cellular membrane which allows the mixing of their lipid bilayers. The mechanism by which the virus uncoats and releases its genomic RNA into the cytoplasm is not clear yet.

Table 2 Coronavirus Receptors

Group	Virus species	Receptor
1	TGEV	Porcine aminopeptidase N (pAPN)
	PRCoV	Porcine aminopeptidase N (pAPN)
	FIPV	Feline aminopeptidase N (fAPN)
	FCoV	Feline aminopeptidase N (fAPN)
	CCoV	Canine aminopeptidase N (cAPN)
	HCoV-229E	Human aminopeptidase N (hAPN)
	HCoV-NL63	Angiotensin-converting enzyme 2 (ACE2)
2	MHV	Murine carcinoembryonic antigen-related adhesion molecules 1 (mCEACAM1)
	BCoV	9- <i>O</i> -acetyl sialic acid
	SARS-CoV	Angiotensin-converting enzyme 2 (ACE2) CD209L (L-SIGN)
3	IBV	α 2,3-linked sialic acid

Fig. 1-6. The coronavirus life cycle.

1.7.2. Transcription of viral mRNA

The virus-specific RNA-dependent RNA polymerase (RdRp) is responsible for the synthesis of negative-strand RNA from the genomic RNA and subsequent transcription of viral mRNA from the negative-strand template.

The coronavirus genomic RNA is transcribed into negative strand RNA, which directs the synthesis of full-length and subgenomic length mRNAs. In addition to the viral mRNAs, both full-length and subgenomic length negative strand RNAs and replicative form (RF) RNAs have been detected in TGEV, BCoV and MHV infected cells (Sethna *et al.*, 1989; Sawicki & Sawicki, 1990; Sethna *et al.*, 1991b; Schaad & Baric, 1994; den Boon *et al.*, 1996; Baric & Yount, 2000). The negative strand RNAs are complementary to the mRNAs, all containing a 5' poly(U) tract and 3' anti-leader sequence (Hofmann & Brain, 1991; Sethna *et al.*, 1991a). During viral infection, synthesis of the negative-sense RNA peaks earlier and drops more rapidly than the positive-sense RNA (Sawicki & Sawicki, 1986). Coronavirus-infected cells contain 10-100 fold more positive-sense viral RNA than negative-sense RNA (Sawicki & Sawicki, 1986; Perlman *et al.*, 1986).

Each of the 5' most ORF in the mRNA is preceded by a transcriptional regulatory sequence (TRS) which contains a highly conserved sequence (CS) that functions in the synthesis of the subgenomic mRNAs (Budzilowicz *et al.*, 1985; Makino *et al.*, 1991; Sethna *et al.*, 1991b; Joo & Makino, 1992; Makino & Joo, 1993). The length and sequence of the CS vary depending on the particular coronavirus (5'-CUAAAC-3' for TGEV; 5'-UCUAAAC-3' for MHV; 5'-CUUAACAA-3' for

IBV). Acquisition of leader sequence during mRNA synthesis appears to be TRS-dependent, and the leader to body joining is guided by a base-pairing interaction involving the leader and body TRS.

A discontinuous transcription model has been proposed as the mechanism for subgenomic mRNAs generation (Makino *et al.*, 1988; Sawicki & Sawicki, 1995; Sawicki & Sawicki, 1998) (Fig. 1-7). According to this model, a positive-sense, genome-length RNA is used as template for negative-sense RNA transcription. However, the polymerase complex appears to pause at the TRS, and then “jump” to transcribe the positive-sense leader sequence from the 5’ end of the genome. Since transcription of all negative-sense RNA species originates at the 3’ end of the positive-sense template, this “polymerase jumping” produces subgenomic RNAs that are identical at both the 5’ and 3’ termini. The negative-sense subgenomic RNAs are subsequently used as templates for positive-sense subgenomic RNA transcription, which serve as mRNAs for the expression of structural and accessory proteins.

Fig. 1-7. Discontinuous transcription model of subgenomic RNA synthesis. 1. Genome RNA is used as template for nascent negative-sense sgRNA. The polymerase complex may pause at any TRS, then “jump” to the leader sequence. 2. Polymerase jumping results in fusion of an antileader sequence to the body of the negative-sense sgRNA. 3. Negative-sense sgRNA is used as template for synthesis of positive-sense sgRNA.

1.7.3. Replication of viral genomic RNA

Coronavirus genome replication is mediated through the synthesis of a negative-strand RNA, which in turn is the template for the synthesis of the progeny virus genomes. Deletion analysis of various cloned MHV defective interfering (DI) RNAs indicated that 470 nucleotides from the 5' end and 436 nucleotides from the 3' end are required to support replication (Kim *et al.*, 1993; Lin & Lai, 1993). Both ends of the genome are necessary for positive-strand synthesis, whereas only the last 55 nucleotides from the 3' end and the poly(A) tail are required for the synthesis of negative-strand (Lin *et al.*, 1994). For IBV, the minimal 5' and 3' *cis*-acting replication signals have been limited to 544 and 338 nucleotides (Dalton *et al.*, 2001).

A number of host factors have been proposed to participate in coronavirus RNA synthesis. Heterogeneous nuclear ribonucleoprotein (hnRNA) A1 has been identified as a major protein species binding to nucleotides 90 to 170 and to nucleotides 260 to 350 from the 3' end of MHV RNA (Furuya & Lai, 1993; Zhang & Lai, 1995; Li *et al.*, 1997). These binding sites are complementary to the sites on the negative-strand RNA that bind another cellular protein, polypyrimidine tract-binding protein (PTB). PTB and hnRNP A1 also bind to the complementary strands at the 5' end of MHV RNA (Li *et al.*, 1997; Li *et al.*, 1999). Overexpression of hnRNP A1 facilitates MHV replication, whereas expression of dominant-negative mutants of hnRNP A1 reduced replication (Shi *et al.*, 2000). Furthermore, partial deletion of the major PTB-binding site in the 3'-UTR resulted in the complete abolition of transcriptional activity,

suggesting a functional role in the replication and transcription for PTB (Huang & Lai, 1999).

1.7.4. Translation of viral proteins

Coronavirus mRNAs are structurally polycistronic, containing multiple ORFs, except for the smallest mRNA. Only the 5' most ORF in the mRNA is translated by a cap-dependent ribosomal scanning mechanism. Thus, most of these mRNAs are functionally monocistronic. Most of the structural proteins including S, M and N are translated from separate mRNAs by this mechanism. Many accessory proteins, however, are translated from truly polycistronic mRNAs. For these mRNAs, the first ORF, e.g., 3a of IBV and 5a of MHV, is probably also translated by the same mechanism as the structural protein genes. For the translation of internal ORFs, several different mechanisms are used by coronaviruses.

1.7.4.1. Ribosomal frameshifting within the polymerase gene

All of the coronavirus polymerase genes sequenced so far contain two large overlapping ORFs. It was demonstrated that translation of ORF 1b involved ribosomal frameshifting from ORF 1a, thus synthesizing a large polyprotein containing both 1a and 1b sequences. This mechanism has been shown to operate in gene 1 of MHV, HCoV-229E, IBV and TGEV (Brierley *et al.*, 1987; Brierley *et al.*, 1989; Brierley *et al.*, 1990; Bredenbeek *et al.*, 1990; Herold & Siddell, 1993; Somogyi *et al.*, 1993; Eleouet *et al.*, 1995). In all cases, the mechanism involves two

essential elements: a heptanucleotide slippery sequence and a downstream, hairpin-type pseudoknot. The slippery site of IBV has the sequence UUUAAAC. The pseudoknots of IBV comprise two base-paired regions stacked coaxially in a quasi-continuous manner and connected by two single-strand loop regions. It is the overall shape and stability of the pseudoknot that are important, not the nucleotide sequence per se. In addition, the space between these elements is critical. It is thought that the pseudoknot impedes the elongation of ribosome. The delay for the ribosome to melt out the secondary structural element allows the simultaneous slippage of the P and A site tRNAs by one base in the -1 direction. Normal translation then resumes. Why coronaviruses employ ribosomal frameshifting strategy is less well known. It is speculated that the frame shifting mechanism provides a fixed ratio of translation products for assembly into a macromolecular complex.

1.7.4.2. Internal initiation of translation of the IBV and MHV E protein mRNA

E protein of both IBV and MHV is translated by a cap-independent, internal ribosomal entry mechanism (Liu & Inglis, 1992b; Thiel & Siddell, 1994). If the 3a and 3b ORFs were eliminated from the IBV mRNA, translation of the E protein did not occur, this suggested that the 3a/3b region contains an internal ribosome entry site (IRES) for the E protein ORF (Liu & Inglis, 1992b). Secondary structures in the 3a/3b region of IBV which resemble the IRES elements of picornaviruses have been predicted by Le and colleagues (Le et al., 1994).

1.7.4.3. Translation of accessory proteins by leaky ribosomal scanning

The genomic RNA of BCoV and MHV contains an additional internal ORF within the N protein gene. This ORF is in a different frame from the ORF of N protein and encodes a hydrophobic protein (Senanayake *et al.*, 1992; Fischer *et al.*, 1997a). This protein is translated in virus infected cells by a leaky ribosomal scanning mechanism from the bicistronic mRNA of N gene (Senanayake *et al.*, 1992).

1.7.5. Assembly and release of coronavirus virions

Coronaviruses typically contain four structural proteins, *i.e.* a type I S glycoprotein required for entry, a phosphorylated N protein that interacts with the viral genome to form a helical core, a major integral M protein and a minor E protein (Holmes *et al.*, 1981; Hogue & Brian, 1986; Delmas & Laude, 1990; Liu & Inglis, 1991; Locker *et al.*, 1992). Some coronaviruses also contain HE as a small spike in their envelope (King & Brian, 1982; Dea & Tijssen, 1988; Spaan *et al.*, 1988). Unlike many other enveloped viruses, coronaviruses acquire their membrane envelope by budding into the lumen of Golgi and pre-Golgi compartments. Early in infection, virus budding occurs at the smooth membranes of the intermediate compartment, between rough endoplasmic reticulum (RER) and Golgi complex (Tooze *et al.*, 1984; Narayanan & Makino, 2001; Narayanan *et al.*, 2003). Later in the infection cycle, budding virions can also be detected in the RER and nuclear membrane. Subsequently, virions are thought to travel in vesicles through the secretory pathway and exit the cell through the fusion of these vesicles with the plasma membrane.

Virus assembly begins with the formation of a helical nucleocapsid which associates with smooth membranes belonging to the ER-Golgi intermediate compartment (ERGIC) (Krijnse-Locker *et al.*, 1994). The coronavirus N protein has been shown to bind viral RNAs in both sequence-specific and non-sequence-specific manners (Robbins *et al.*, 1986; Stohlman *et al.*, 1988b; Cologna *et al.*, 2000). A specific RNA element, usually referred to as a packaging signal or an encapsidation signal, determines the selective and specific binding of viral N protein to the genomic RNA. For MHV, N protein binds tightly to RNA packaging signal sequence located at a position 1,381 to 1,441 nucleotides from the 3' end of MHV gene 1 (Fosmire *et al.*, 1992). Studies showed that when RNA transcripts consisted of a non-MHV sequence and the packaging signal were expressed in MHV-infected cells, they were packaged into MHV particles (Woo *et al.*, 1997). The exact details of N protein-RNA binding and RNA packaging are not well understood. In fact, studies suggested that MHV genomic RNA packaging may occur without the N protein, rather the M protein is responsible for the packaging of viral RNA (Narayanan & Makino, 2001; Narayanan *et al.*, 2003).

Unlike most other enveloped RNA viruses, coronaviruses employ a nucleocapsid-independent strategy to drive the assembly and budding. In studies of MHV and TGEV particle formation, co-expression of the E and M proteins resulted in the release of VLPs from cells, which are similar in size and morphology to authentic MHV and TGEV virions (Vennema *et al.*, 1996; Baudoux *et al.*, 1998). Additionally, expression of E protein on its own results in the release of E-containing

membrane vesicles (Maeda *et al.*, 1999). However, SARS-CoV M and N proteins were reported to be necessary and sufficient for the formation of VLP (Huang *et al.*, 2004b), though it is not consistent with a report showing that SARS-CoV M and E are sufficient for VLP formation (Ho *et al.*, 2004). It either indicates that SARS-CoV might adopt a unique mechanism for virion assembly, or simply suggests that formation of VLP may vary with the expression systems used.

The precise function of E protein in virion assembly is not fully elucidated, but its low abundance in virions and in VLPs implied that it might serve to induce membrane curvature or act to pinch off the neck of the viral particles at the final stage of the budding process. However, a mutant MHV with deletion of the E gene was recovered later, although the mutant virus produced tiny plaques and showed low growth rate and titer (Kuo & Masters, 2003). Alanine scanning insertion mutagenesis of the hydrophobic domain of the MHV E protein suggested that positioning of polar hydrophilic residues within the predicted transmembrane domain are important for virus production (Ye & Hogue, 2007). Substitution of the MHV E gene with heterologous E genes from viruses spanning all three groups of coronavirus family showed that the E proteins of group 2 and 3 coronaviruses could almost fully replace the MHV E protein (Kuo *et al.*, 2007). However, E protein of the group 1 coronavirus, TGEV, could functionally replace the MHV E protein only after acquisition of particular mutations (Kuo *et al.*, 2007). As these E proteins share a low degree of sequence identity, it indicates that sequence-specific contacts with other viral components may not be essential. More recently, a recombinant SARS-CoV that lacks

the E gene was rescued in Vero E6 cells, and the recovered deletion mutant was attenuated *in vitro* and in the hamster model (DeDiego *et al.*, 2007). In general, coronavirus E protein is not essential for viral replication, but seems to be a non-obligate budding enhancer.

How the budding-site is selected for coronaviruses is not well understood. It is generally thought that E protein may be responsible for driving the budding process by selecting the site of budding, envelope formation and virion assembly. In previous reports, the coronavirus M protein was shown to be capable of interacting with itself, S, E, HE and N proteins (Nguyen & Hogue, 1997; de Haan *et al.*, 1999; Narayanan *et al.*, 2000; Lim & Liu, 2001). The interactions of the M protein with these structural proteins may be the means for their incorporation into assembled virion particles.

1.8. Coronavirus Genetics and Reverse Genetics

Coronavirus genetics analysis has been limited to analysis of temperature-sensitive (ts) mutants (Fu & Baric, 1992; Schaad & Baric, 1994; Fu & Baric, 1994; Lai & Cavanagh, 1997; Stalcup *et al.*, 1998), DI RNAs which rely on helper viruses for replication (Repass & Makino, 1998; Williams *et al.*, 1999; Izeta *et al.*, 1999; Narayanan & Makino, 2001) and recombinant viruses generated by targeted recombination (Fischer *et al.*, 1997b; Hsue & Masters, 1999; Kuo *et al.*, 2000). Systematic assembly of full-length coronavirus infectious clone has been hampered by the large size of the genome, regions of instability and the inability to synthesize

full-length transcripts (Masters, 1999). In recent years, several methods have been developed to assemble full-length coronavirus infectious clone.

1.8.1. Virus mutants

Coronaviruses, like other RNA viruses, have high frequency of mutation. This characteristic results from a combination of the large genome size and high error rate of RdRp. The high mutation rate has led to the isolation of a number of temperature-sensitive mutants. A number of these groups are RNA-negative mutants, defective in negative strand synthesis, that do not replicate RNA at the nonpermissive temperature. RNA-positive mutants have mutations within the S or N structural genes and exhibit altered cytopathic effect (CPE) or fail to produce infectious virions (Leibowitz *et al.*, 1982; Sturman *et al.*, 1987; Baric *et al.*, 1990). Besides naturally occurring viral variants, a lot of mutants were generated through chemical mutagenesis or neutralization selection by monoclonal antibodies (Robb *et al.*, 1979; Koolen *et al.*, 1983; Sturman *et al.*, 1987; Martin *et al.*, 1988; Schaad *et al.*, 1990). Virus mutants have provided some insights into coronavirus replication. However, the use of these mutants is severely restricted by the inability to manipulate the viral genome.

1.8.2. Defective interfering RNA

Upon passage of MHV at high multiplicity of infection (MOI), subgenomic sized RNA species that interfere with genomic RNA replication were identified that

contained large deletions as compared with the wild-type genomic RNA (Makino *et al.*, 1984). DI RNAs are composed of discontinuous regions of the viral genome and contain the 5' and 3' termini of genomic RNA. Generation of coronavirus DI RNAs occurs by recombination and/or polymerase jumping (Makino *et al.*, 1985; Furuya *et al.*, 1993). For the first time, DI RNA systems allowed genetic analysis through manipulation of the DI RNA sequence. Many of the major observations concerning TRS sequences and discontinuous transcription have been made using DI RNAs. Additionally, MHV and TGEV DI RNAs have been engineered as expression vectors and used to express hemagglutinin-esterase and reporter gene encoding beta-glucuronidase (Liao & Lai, 1995; Zhang *et al.*, 1997; Izeta *et al.*, 1999). Despite the utility of DI RNA system, these RNAs require a helper virus for their replication and likely replicate at a higher rate than full-length genomic RNAs. Additionally, they are not authentic with respect to an intact genomic RNA and it is unclear how findings from DI studies correlate to full-length viral RNA genomes.

1.8.3. Targeted recombination

The observations of high coronavirus recombination frequencies led to the development of the targeted recombination approach for constructing recombinant coronaviruses (Koetzner *et al.*, 1992; Masters, 1999). This strategy initially made use of an N gene mutant Alb4, which is temperature sensitive and thermolabile. Upon co-transfection of Alb4 genomic RNA and *in vitro* synthesized mRNA7 containing a marker within the 3' UTR, recombinant viruses that result from a single crossover

event were selected based on the loss of temperature sensitive phenotype. The recombination frequency with this system has been significantly increased through the use of helper virus and replication defective RNAs (van der Most *et al.*, 1992). This approach has been used for genetic analysis of a number of coronavirus genes. The pathogenicity and tropism of MHV have been modified by the introduction of mutations into the S gene (Leparc-Goffart *et al.*, 1998; Navas *et al.*, 2001). Targeted recombination was also used to elucidate the important role of E and M proteins in coronavirus assembly (de Haan *et al.*, 1998; Fischer *et al.*, 1998). In addition, the expression of foreign gene from the MHV genome was demonstrated using target recombination to insert the gene encoding green fluorescent protein (GFP) between the S and E genes (Fischer *et al.*, 1997b). However, at present the targeted recombination approach is limited to the analysis of the 3' end of the MHV genome and the viral replicase.

1.8.4. Full-length infectious cDNA

Analysis of recombinant virus with specific mutations has proved to be a powerful method to understand the molecular biology of the RNA viruses and study the role of individual genes. Such reverse genetics systems have been achieved for a number of positive-strand RNA viruses such as piconaviruses, arteriviruses and Citrus tristeza virus (Racaniello & Baltimore, 1981; Satyanarayana *et al.*, 1999; Satyanarayana *et al.*, 2001). Despite these successes, the reverse genetics methods used for positive-strand RNA viruses are still not entirely routine procedures.

Coronaviruses have the largest RNA genome of any known RNA virus and some cDNAs in the regions of the replicase gene are unstable in bacteria. All these hampered the generation of a full-length infectious cDNA clone. In these years, several different methods were described for the assembly of coronavirus full-length cDNAs.

The first method involves the assembly of a TGEV full-length cDNA in a low copy-number bacterial artificial chromosome (BAC), downstream of a viral RNA polymerase II promoter. Transfection of this construct into susceptible cells resulted in the recovery of recombinant virus (Almazan et al., 2000). Further stabilization of the full-length BAC clone in bacteria was achieved through the insertion of a eukaryotic intron into either of two positions in the toxic region of TGEV cDNA (Gonzalez et al., 2002). This allowed stable propagation of BAC for over 200 generations.

The second system involves the *in vitro* assembly of the TGEV full-length cDNA by using a series of contiguous cDNAs with engineered unique restriction sites. T7 RNA polymerase derived RNA transcripts of full-length cDNA were used to generate recombinant virus (Yount et al., 2000). The boundaries of the subcloned genomic cDNA allowed easy manipulation for site-directed mutagenesis. Most importantly, some fragment boundaries were arranged in such a way as to interrupt unstable regions of the cDNA. This approach was facilitated by restriction enzymes that cut at a distance from their recognition sequences. This ensured that no rearranged extra sequence was generated during the ligation. Now this *in vitro*

assembly technique has been successfully used to engineer the genomes of MHV, SARS-CoV and IBV (Yount *et al.*, 2002; Yount *et al.*, 2003; Youn *et al.*, 2005; Fang *et al.*, 2006).

The third system involves the *in vitro* assembly of the HCoV full-length cDNA followed by cloning into the genome of vaccinia virus (Thiel *et al.*, 2001). Infectious HCoV RNA synthesized *in vitro* from the vaccinia virus DNA was transfected into the cells and resulted in the recovery of recombinant virus. The use of vaccinia as a vector has allowed manipulation of the resulting cloned cDNA by methods that have been developed for poxvirus reverse genetics. Originally applied to HCoV-229E, it has been used to engineer the genome of IBV and MHV (Casais *et al.*, 2001; Coley *et al.*, 2005).

The availability of reverse genetics system allows to produce defined and genetically modified viruses, which will be important for the analysis of the molecular biology and pathogenesis of coronaviruses. This approach has been used for genetic analysis of a number of coronavirus genes. Mutant viruses generated by full-length infectious IBV cDNA provided the evidence that neither gene 3 nor gene 5 of IBV is essential for replication (Casais *et al.*, 2005; Hodgson *et al.*, 2006). In addition, using a full-length infectious construct of TGEV, it was demonstrated that subgenomic transcription is heavily influenced by upstream flanking sequences (Curtis *et al.*, 2004). Coronavirus reverse genetic investigations also led to the development for vaccines, expression systems and gene delivery vectors (Alonso *et*

al., 2002; Thiel *et al.*, 2003; de Haan *et al.*, 2003b; Haijema *et al.*, 2004; de Haan *et al.*, 2005).

1.9. Thesis Objectives

Coronavirus E protein is a small membrane protein and minor component of the virus particles. It plays important roles in virion assembly and morphogenesis, alteration of the membrane permeability of host cells and virus-host cell interaction.

The objectives of the first part of this study are:

- A). To determine the membrane topology of SARS-CoV E protein.
- B). To characterize the posttranslational modification of SARS-CoV E protein.
- C). To determine the lipid rafts association of IBV and SARS-CoV E proteins.
- D). To discover SARS-CoV E protein-interacting partners.

Coronavirus M protein is a major component of the virus particles. It plays an important role in virus assembly. The IBV M protein was choosed for the study because the M protein mutants will be introduced into the virus genome to see the effect on the virus replication, however our lab only has the reverse genetic system for IBV and the SARS-CoV cannot be cultured in the lab. The objectives of the second part of this study are:

- E). To characterize the N-linked glycosylation of IBV M protein.
- F). To determine the effects of the N-terminal domain mutants of IBV M protein on virus assembly and VLP formation.

Chapter 2

Materials and Methods

2.1. General Reagents

2.1.1. Enzymes

Restriction Endonuclease	New England biolabs Inc.
Taq DNA Polymerase	New England biolabs Inc.
PNGase F	New England biolabs Inc.
Pfu DNA Polymerase	Stratagene Inc.
Turbo Pfu DNA Polymerase	Stratagene Inc.
Proteinase K	KPL

2.1.2. Commercially available kits

QIAprep® Spin Miniprep Kit	QIAGEN Inc.
QIAprep® Spin Midiprep Kit	QIAGEN Inc.
QIAprep® Spin Maxiprep Kit	QIAGEN Inc.
QIAQuick® Gel Extraction Kit	QIAGEN Inc.
QIAQuick® PCR Purification Kit	QIAGEN Inc.
Effectene® Transfection Reagent	QIAGEN Inc.
Lipofectamine 2000	Invitrogen
ECL™ Western Blotting Analysis System	Amersham Pharmacia Biotech
Protein A-Agarose Beads	KPL
QuikChange® Site-Directed Mutagenesis Kit	Stratagene
Vybrant® Lipid Raft Labeling Kits	Molecular Probes

2.1.3. DNA vectors

pKT0: A modified form of plasmid pING14 (Liu *et al.*, 1991), in which a 23-nucleotide fragment representing the T7 RNA polymerase promoter sequence was

inserted between the SP6 promoter and the 5' globin noncoding region by site-directed mutagenesis (Fig. 2-1).

pFLAG: A hydrophilic FLAG epitope tag (D-Y-K-D-D-D-D-K) was inserted between *NcoI* and *EcoRV* sites of pKT0 vector to obtain the pFlag plasmid (Fig. 2-2).

pGEX-5X-1 (GE Healthcare): A bacterial vector that permits inducible protein expression with IPTG is used to express GST-fused proteins. A multiple cloning site (MCS) is present downstream of the GST coding region. The vector carries the Amp^r for selection in *Escherichia coli* (*E. coli*).

pGBKT7 (Clontech): The pGBKT7 vector expresses proteins fused to amino acids 1–147 of the GAL4 DNA binding domain (DNA-BD). In yeast, fusion proteins are expressed at high levels from the constitutive ADH1 promoter (PADH1); transcription is terminated by the T7 and ADH1 transcription termination signals (TT7 & ADH1). pGBKT7 also contains the T7 promoter, a c-Myc epitope tag, and a MCS. pGBKT7 replicates autonomously in both *E. coli* and *Saccharomyces cerevisiae* (*S. cerevisiae*) from the pUC and 2 μ ori, respectively. The vector carries the Kan^r for selection in *E. coli* and the *TRP1* nutritional marker for selection in yeast.

pACT2 (Clontech): pACT2 generates a fusion of the GAL4 AD (amino acids 768–881), an HA epitope tag, and a protein of interest (or protein encoded by a cDNA in a fusion library) cloned into the MCS in the correct orientation and reading frame. pACT2 contains a unique *EcoR I* site in the MCS. The hybrid protein is expressed at high levels in yeast host cells from the constitutive ADH1 promoter; transcription is terminated at the ADH1 transcription termination signal. The protein

is targeted to the yeast nucleus by the nuclear localization sequence from SV40 T-antigen which has been cloned into the 5' end of the GAL4 AD sequence. pACT2 is a shuttle vector that replicates autonomously in both *E. coli* and *S. cerevisiae* and carries the *bla* gene, which confers ampicillin resistance in *E. coli*. pACT2 also contains the *LEU2* nutritional gene that allows yeast auxotrophs to grow on limiting synthetic media. Transformants with AD/library plasmids can be selected by complementation by the *LEU2* gene by using an *E. coli* strain that carries a *leuB* mutation.

pcDNA3.1(+) (Invitrogen): pcDNA3.1(+) is a 5.4 kb vector derived from pcDNA3 and designed for high-level stable and transient expression in mammalian cells. The vectors contain the following elements: Human cytomegalovirus immediate-early (CMV) promoter for high-level expression in a wide range of mammalian cells; multiple cloning sites in the forward (+) and reverse (-) orientations to facilitate cloning; neomycin resistance gene for selection of stable cell lines; episomal replication in cells lines that are latently infected with SV40 or that express the SV40 large T antigen.

Fig. 2-1. pkT0 vector circle map and MCS sequence.

Fig. 2-2. pFlag vector circle map and MCS sequence.

2.1.4. Antibodies

Mouse Anti-Flag	From Stratagene
Rabbit Anti-HA	From Santa Cruz Biotechnology
Rabbit Anti- β -tubulin	From Santa Cruz Biotechnology
Rabbit Anti-PDI	From Santa Cruz Biotechnology
Rabbit Anti-SARS CoV E	Raised in Rabbit
Rabbit Anti-IBV E	Raised in Rabbit
Rabbit Anti-IBV M	Raised in Rabbit
Rabbit Anti-IBV N	Raised in Rabbit
Rabbit Anti-IBV S	Raised in Rabbit
Goat Anti-Rabbit-HRP	From DAKO
Goat Anti-Mouse-HRP	From DAKO
Goat Anti-Rabbit-FITC	From DAKO
Goat Anti-mouse-TRITC	From DAKO

2.2. Solutions, Buffers and Media

2.2.1. Laboratory stocks

1 M DTT	Stored at -20 °C
0.5 M EDTA, pH 7.5	Stored at RT
0.1% ethidium bromide	Stored at 4 °C
Glycerol loading dyes	30% glycerol, 5 mM EDTA pH 7.5, 0.1% bromophenol
3 M NaAc, pH 5.2	Stored at RT

Phenol chloroform	Equal volumes of Biophenol-Tris and chloroform were mixed and stored at 4 °C
10% (w/v) SDS	Stored at RT
10x TBE	0.89 M Tris pH 8.0, 0.89 M Boric acid, 25 mM EDTA
TE buffer, pH 8.0	10 mM Tris pH 8.0, 0.1 mM EDTA
1 M Tris-HCl, pH 8.0	Stored at RT
1.5 M Tris-HCl, pH 8.8	Stored at RT
1 M Tris-HCl, pH 6.8	Stored at RT
1X PBS	100 mM NaCl, 80 mM Na ₂ HPO ₄ , 20 mM NaH ₂ PO ₄ , stored at RT

2.2.2. Media for bacterial culture

LB medium	1% (w/v) Bacto-tryptone (BD), 0.5% (w/v) Yeast (BD), 1% (w/v) NaCl (Merck)
LB agar	LB medium plus 1.5% (w/v) Bacto-agar(BD)
LB plates (amp ^r)	LB agar plates of 100µg/ml Ampicillin
LB plates (kana ^r)	LB agar plates of 50 µg/ml

kanamycin

2.2.3. Solutions for making competent bacteria

0.1 M CaCl ₂	0.1 M, stored at 4 °C
0.1 M MgCl ₂ -CaCl ₂	80 mM MgCl ₂ , 20 mM CaCl ₂ , stored at 4 °C

2.2.4. Solutions for tissue culture

Dulbecco's Modified Eagle Medium (DMEM)	Gibco, Invitrogen Corp.
Fetal Bovine Serum (FBS)	Hyclone
Cell freezing mixture	10% DMSO in FBS
Trypsin-EDTA	Gibco, Invitrogen Corp.

2.2.5. Solutions for the immunofluorescence labeling of cells

Wash buffer	PBS
Fixative	4% paraformaldehyde in PBS
Antibody solution	5% FBS in PBS

2.2.6. Buffers for SDS-PAGE

10x Running buffer	0.25 M Tris, 1.9 M Glycine, 1% SDS
SDS loading buffer	100 mM Tris pH 6.8, 200 mM DTT, 4% SDS, 20% glycerol and 0.2%

bromphenol blue

2.2.7. Western blot buffers

Blotting buffer	12.5 mM Tris pH 8.0, 96 mM glycine, 10% (w/v) ethanol
Blocking buffer	5% fat free milk, 0.1% Tween [®] 20 in PBS
Washing buffer	0.1% (v/v) Tween [®] 20 in PBS

2.2.8. Northern blot buffers

10 × MOPS	200 mM MOPS, 50 mM NaAc, 20 mM EDTA
20 × SSC	3 M NaCl, 300 mM sodium citrate, pH 7.0
RNA loading buffer	250 µl of 100% Formamide, 83 µl of 37% Formaldehyde, 50 µl 10 × MOPS, 50 µl 100% Glycerol, 10 µl 2.5% Bromphenolblue, 57 µl DEPC treated water
Running buffer	1 × MOPS
Stripping buffer	0.1% SDS

2.3. Methods

2.3.1. Methods for DNA manipulation and subcloning

2.3.1.1. Polymerase chain reaction and site-directed mutagenesis

Amplification of respective template DNAs with appropriate primers was performed with Pfu DNA polymerase (Stratagene). The PCR conditions were 30 cycles of 95 °C for 45 seconds, 45~60 °C for 45 seconds, and 72 °C for different extension time. The annealing temperature and extension time were subjected to adjustments according to the melting temperatures of the primers used and the lengths of the PCR fragments synthesized.

Mutations were introduced into the corresponding sites by using QuickChange site-directed mutagenesis kit (Stratagene) and confirmed by sequencing of the plasmids.

2.3.1.2. Purification of plasmid DNA

For small scale purification of plasmid DNA, 5 ml LB broth was inoculated with a single colony from an agar plate and incubated at 37 °C with constant shaking for ~16 h. For sequencing, plasmid DNA was prepared using a Plasmid Mini Kit (QIAGEN Inc.) according to the manufacturer's instructions. For large scale purification of plasmid DNA, a Plasmid Midi Kit or Plasmid Maxi Kit (QIAGEN Inc.) was used according to the manufacturer's instructions.

2.3.1.3. Quantitation of DNA

The concentration of DNA was determined by its absorbance at a wavelength of

260 nm (using a SHIMAZHU ultraviolet spectrophotometer) based on the calculation:

50 µg/ml double-stranded DNA gives an OD₂₆₀ of 1.

2.3.1.4. Purification of DNA fragments by agarose gel electrophoresis

Gel slices containing the DNA fragments to be purified were cut from the gel using a scalpel blade and avoided long time exposure to the UV light. DNA was extracted using a QIAquick Gel Extraction Kit (QIAGEN Inc.)

2.3.1.5. DNA ligation

DNA fragments with complementary ends to be ligated were prepared by restriction enzyme digestion and purified using agarose gel electrophoresis. Vector DNA (~100 ng) and insert DNA (200~400 ng) were ligated using 1 unit of T4 DNA ligase in a 10 µl reaction mixture containing 1 mM ATP, 50 mM Tris-HCl pH 8.0, 10 mM MgCl₂, 20 mM DTT, 50 µg/ml BSA and incubated at 16 °C overnight. As controls, a reaction without insert DNA and a reaction without both insert DNA and T4 DNA ligase were included in the experiment. The ligation product was used directly to transform competent cells or stored at -20 °C.

2.3.1.6. Construction of plasmids

Plasmid Flag-SARS E, which covers the SARS-CoV E sequence, was constructed by cloning an *EcoRV*/*EcoRI*-digested PCR fragment into *EcoRV*/*EcoRI*-digested pFLAG vector. The PCR fragment was generated using

primers Flag-SARS E 5' and Flag-SARS E 3' (Table 3). Plasmid SARS E-Flag was created by cloning a *BglIII*/*EcoRI*-digested PCR fragment into *BglIII*/*EcoRI*-digested pKT0 vector. The PCR fragment was generated using primers SARS E-Flag 5' and SARS E-Flag 3' (Table 3). Plasmid Flag-IBV E was generated by cloning an *EcoRV*/*EcoRI*-digested PCR fragment into *EcoRV*/*EcoRI*-digested pFLAG vector. The two primers used to generate the PCR fragment are Flag-IBV E 5' and Flag-IBV E 3' (Table 3). Plasmid IBV E-Flag was created by cloning a *BglIII*/*EcoRI*-digested PCR fragment into *BglIII*/*EcoRI*-digested pKT0. The PCR fragment was generated using primers IBV E-Flag 5' and IBV E-Flag 5' (Table 3). Plasmid pKT0-SARS E was constructed by cloning a *BglIII*/*EcoRI*-digested PCR fragment into *BglIII*/*EcoRI*-digested pKT0. The PCR fragment was generated using primers PKT0-SARS E 5' and PKT0-SARS E 3' (Table 3). Plasmid pKT0-IBV E was constructed by cloning a *BglIII*/*EcoRI*-digested PCR fragment into *BglIII*/*EcoRI*-digested pKT0. The PCR fragment was generated using primers PKT0-IBV E 5' and PKT0-IBV E 3' (Table 3).

Mutations were introduced into the SARS-CoV E gene by two rounds of PCR. The PCR amplified fragments were cloned into *EcoRV*- and *EcoRI*-digested pFlag.

Plasmid pGEX-snapin was constructed by cloning a *BamHI*/*XhoI*-digested PCR fragment into *BamHI*/*XhoI*-digested pGEX-5X-1. The PCR fragment was generated using primers pGEX-snapin 5' and pGEX-snapin 3' (Table 3). Plasmid pcDNA-HA snapin was constructed by cloning a *HindIII*/*XhoI*-digested PCR fragment into *HindIII*/*XhoI*-digested pcDNA-HA. The PCR fragment was generated using primers

pcDNA-HA snapin 5' and pcDNA-HA snapin 3' (Table 3).

Plasmid pKT0-IBV M was constructed by cloning a *BglIII*-/*EcoRI*-digested PCR fragment into *BglIII*-/*EcoRI*-digested pKT0. The PCR fragment was generated using primers PKT0-M 5 and PKT0-M 3 (Table 4). The IBV M mutants PKT0-N3D, PKT0-N6D and PKT0-N3D/6D were made by cloning a *BglIII*-/*EcoRI*-digested PCR fragment into *BglIII*-/*EcoRI*-digested pKT0. The PCR fragments were generated using following sets of primers PKT0 N3D-5 and PKT0-M 3, PKT0 N6D-5 and PKT0-M 3 and PKT0 N36D-5 and PKT0 PKT0-M 3 respectively (Table 4).

To make IBV M protein mutant virus N3D/6D, C7A, L9A, Q16A, V15A/Q16A, F18A and F18A/K19A mutations were first introduced into an infectious clone E of IBV by using QuickChange site-directed mutagenesis kit (Stratagene) with primers in table 4. Substitution mutations of GFP48, GFPL15, GFPM15 and GFPR18 were introduced into an infectious clone E of IBV by two rounds of PCR with primers in table 4.

All plasmids and the introduced mutations were confirmed by automated DNA sequencing.

Table 3. Primers used to amplify SARS E, IBV E and snapin

Name	Sequence
Flag-SARS E 5'	GCAAGATATCCTACTCATTTCGTTTCGGAA
Flag-SARS E 3'	CCGGAATTCTTAGACCAGAAGATCAG
Flag-IBV E 5'	GCAAGATATCCAATTTATTGAATAAGTCG
Flag-IBV E 3'	CCGGAATTCTCA AGAG TACAATTTGTC
SARS E-Flag 5'	TGGAAGATCTCCACCATGTACTCATTTCGTTTCGGAA
SARS E-Flag 3'	CCGGAATTCTTACTTGTCATCGTCGTCCTTGTAATCG ACCAGAAG ATCAGGAA
IBV E-Flag 5'	TGGAAGATCTCCACCATGAATTTATTGAATAA
IBV E-Flag 3'	CCGGAATTCTCACTTGTC TCGTCGTCCTTGTAATC AGAGTACAATTTGTCTCG
PKT0-SARS E 5'	TGGAAGATCTCCACCAT GTACTCATTTCGTTTCGGAA
PKT0-SARS E 3'	CCGGAATTCTTAGACCAGAAGATCAG
PKT0-IBV E 5'	TGGAAGAT CTCCACCATGAATTTATTGAATAA
PKT0-IBV E 3'	CCGGAATTCTCAAGAGTACAATTTGTC
pcDNA-HA snapin 5'	TTGTAAGCTTACCATGGCGGGGGCTGGTTCC
pcDNA-HA snapin 3'	CCGTCTCGAGTTTGCCTGGGGAGCCAGG
pGEX-snapin 5'	CGCGGATCCCGATGGCGGGGGCTGGTTCC
pGEX-snapin 3'	CCGTCTCGAGTTATTTGCCTGGGGAGCCA

a. Underlined letters are restriction enzyme recognition sequence.

Table 4. Primers used in IBV M protein mutagenesis

Name	5'→3' sequence	Genomic Location	Purpose
A7-5	CCCAACGAGACAAATGCTACTCTTGACTTTG	24508-24538	C7 to A mutation
A7-3	CAAAGTCAAGAGTAGCATTGTCTCGTTGGG	24508-24538	C7 to A mutation
A9-5	GAGACAAATTGTACTGCTGACTTTGAACAGTC	24514-24545	L9 to A mutation
A9-3	GACTGTTCAAAGTCAGCAGTACAATTGTCTC	24514-24545	L9 to A mutation
A15A16-5	GACTTTGAACAGTCAGCTGCGCTTTTAAAGAGTAT	24532-24567	V15Q16 to AA mutation
A15A16-3	ATACTCTTTAAAAAGCGCAGCTGACTGTTCAAAGTC	24532-24567	V15Q16 to AA mutation
A18A19-5	CAGTCAGTTCAGCTTGCTGCAGAGTATAATTATTT	24541-24576	F18K19 to AA mutation
A18A19-3	AAATAAATTATACTCTGCAGCAAGCTGAACTGACTG	24541-24576	F18K19 to AA mutation
A16-5	GACTTTGAACAGTCAGTTGCGCTTTTAAAGAGTAT	24532-24567	Q16 to A mutation
A16-3	ATACTCTTTAAAAAGCGCAACTGACTGTTCAAAGTC	24532-24567	Q16 to A mutation
A18-5	CAGTCAGTTCAGCTTGCTAAAGAGTATAATTATTT	24541-24576	F18 to A mutation
A18-3	AAATAAATTATACTCTTTAGCAAGCTGAACTGACTG	24541-24576	F18 to A mutation
N3-D-5	CAAATTTTCAAGATGCCCCGACGAGACAAATTTGACTC	24493-24529	N3 to D mutation
N3-D-3	GAGTACAATTGTCTCGTCGGGCATCTTGAAAATTTG	24493-24529	N3 to D mutation
N3, 6-D-5	CAAGATGCCCCGACGAGACAGATTGTACTCTTGACTTTG	24501-24538	N3N6 to D3D6 mutation
N3, 6-D-3	CAAAGTCAAGAGTACAATCTGTCTCGTCGGGCATCTTG	24501-24538	N3N6 to D3D6 mutation
GFPL15-5	CCAGTAAAGGAGAAGAACTCTTGACTTTGAACAGT		N3 to C7 replacement
GFPL15-3	GTTTCTTCTCCTTTACTGGGCATCTTGAAAATTTGC		N3 to C7 replacement
GFPM15-5	GTCTTTTCACTGGAGTTCAGTCAGTTCAGCTTTTAA		T8 to E12 replacement
GFPM15-3	TGAACTCCAGTGAAAAGACAATTTGTCTCGTTGGGC		T8 to E12 replacement
GFPR18-5	GTCCCAATTCTTGTGAAAAAGAGTATAATTTATTTAACTGC		Q13 to F18 replacement
GFPR18-3	TTTCAACAAGAATTGGGACTTCAAAGTCAAGAGTAC		Q13 to F18 replacement
GFP48-5	GAACTTTTCACTGGAGTTGTCCCAATCTTGTGAAAAAGAGT		N3 to F18 replacement
	ATAATTTATTTATAACTGC		
GFP48-3	TGGGACAACTCCAGTGAAAAGTTCTTCTCCTTTACTGGGCATC		N3 to F18 replacement
	TTGAAAATTTGC		
BLPI 23016	TGTTTTAGCATCTGCTAAGCA	23001-23021	
BSMBI 26086	AAACCTGGCTTGCGTCTCCA	26083-26103	
PKT0 M-5	TGGAAGATCTCCACCATGCCCAACGAGACAAATTG		
PKT0 M-3	CCGGAATTCTTATGTGTAAAGACTACTTC		
PKT0 N3D-5	TGGAAGATCTCCACCATGCCCGACGAGACAAATTGTACT		PKT0-N3D
PKT0 N6D-5	TGGAAGATCTCCACCATGCCCAACGAGACAGATTGTACTCTTG		PKT0-N6D
PKT0 N36D-5	TGGAAGATCTCCACCATGCCCGACGAGACAGATTGTACTCTTG		PKT0-N36D

a. Underlined letters are restriction enzyme recognition sequence.

b. Bold letter is mutation which is introduced into plasmid.

2.3.1.7. Preparation of *E.coli* competent cells

Competent cells of *E.coli* strain DH-5 α were routinely used for plasmid transformation. The bacteria were streaked out a LB agar plate and incubated at 37 °C for ~18 h. A single colony was used to inoculate 20 ml LB broth in a 250 ml flask and incubated at 37 °C with constant shaking for 3 h (OD₆₀₀ of 0.35). The culture was transferred into a 50 ml polypropylene tube (Falcon) and cooled down to 0 °C by leaving on ice for 10 min. The bacteria were sedimented by centrifugation at 4000 rpm for 10 min at 4 °C and resuspended in 150 ml pre-chilled 0.1 M CaCl₂-MgCl₂ (80 mM MgCl₂, 20 mM CaCl₂) solution. The bacteria were pelleted by centrifugation at 4000 rpm for 10 min at 4 °C and resuspended in 10 ml pre-chilled 0.1M CaCl₂ solution. The competent bacteria were divided into 600 μ l aliquots in pre-chilled microfuge tubes, flash-frozen in liquid nitrogen and stored at -70 °C.

2.3.1.8. Transformation of *E.coli* with plasmid DNA

Competent *E.coli* cells were thawed on ice. 100 μ l competent cells were mixed with 5 μ l of a ligation mix, or with 5 ng plasmid DNA, in a pre-chilled 1.5 ml microfuge tube and incubated on ice for 30 min and then heated at 42 °C for 1 min. 200 μ l LB was added and the cells were allowed to recover at 37 °C for 1 h. The cells were pelleted by centrifugation and resuspended in 50 μ l LB. The suspension was plated on LB argar plate. Plate was incubated at 37 °C for 18 h.

2.3.2. Cell culture and transfection methods

2.3.2.1. Cells

Vero, HeLa, Huh-7 and COS-7 cells were cultured at 37 °C in DMEM supplemented with 10% FBS, penicillin (100 units/ml), and streptomycin (100 µg/ml).

2.3.2.2. Cell storage in liquid nitrogen

Cells were sedimented at 800 × g, resuspended in freezing mixture and aliquotted into Cryo Vials (Greiner Lab.). Cells were frozen for 24 h at -70 °C. After that the vials were transferred to liquid nitrogen.

2.3.2.3. Cells recovery from liquid nitrogen

Cells were removed from the liquid nitrogen, quickly brought to 37 °C, and washed in 10 ml of culture media. The cells were sedimented at 800 × g, resuspended in 10 ml of growth media, and transferred to flask for growth.

2.3.2.4. Transient expression in mammalian cells

HeLa cells were grown at 37 °C in 5% CO₂ and maintained in DMEM supplemented with 10% FBS. IBV M and mutants, SARS-CoV E and mutants were placed under the control of a T7 promoter and transiently expressed in mammalian cells using a vaccinia/T7 expression system described by Fuerst et al. (Fuerst *et al.*, 1986). Briefly, 60-80% confluent monolayers of HeLa cells grown on dishes were infected with a recombinant vaccinia virus (vTF7-3) that expresses T7 RNA

polymerase at a MOI of 10. Two hours later, cells were transfected with plasmid DNA mixed with Effectene according to the instructions of the manufacturer (QIAGEN Inc.). Cells were harvested at 12 to 24 h post-transfection.

2.3.3. Experiment Assays

2.3.3.1. Polyacrylamide gel electrophoresis in sodium dodecyl sulphate

Gels containing 5% polyacrylamide were made as stacking gel and gels containing 8-15% polyacrylamide were made as separating gel according to the method of Laemmli (Laemmli, 1970). Proteins were reduced prior to electrophoresis by incubating in an equal volume of sample and 2 × SDS loading buffer containing a final concentration of 20 mM DTT for 5 min at 100 °C. Electrophoresis was carried out at a constant voltage of 100 V in stacking gel and 200 V in separating gel.

2.3.3.2. Western blot analysis

Total proteins extracted from cells were lysed with 2 × SDS loading buffer in the presence of 20 mM DTT and subjected to SDS-PAGE. Proteins were transferred to PVDF membrane (Bio-Rad) and blocked overnight at 4 °C in blocking buffer containing 5% fat free milk powder in PBST buffer pH7.5 (80 mM Na₂HPO₄, 20 mM NaH₂PO₄, 100 mM NaCl, 0.1% Tween 20). The membrane was incubated with 1:1000~1:5000 diluted primary antibodies in blocking buffer for 1 h at RT. After washing three times with PBST, the membrane was incubated with 1:1000 diluted anti-mouse or anti-rabbit IgG antibodies conjugated with horseradish peroxidase

(DAKO) in blocking buffer for 1 h at RT. After washing for three times with PBST, the polypeptides were detected with a chemiluminescence detection kit (ECL, Amersham Pharmacia Biotech) according to the instructions of the manufacturer.

2.3.3.3. Indirect immunofluorescence

HeLa cells expressing Flag-tagged SARS-CoV or IBV E protein were fixed with 4% paraformaldehyde for 10 min at 16 h post-transfection, washed three times with 1x PBS, permeabilized with 0.2% Triton X-100 for 10 min at RT, and washed three times with 1x PBS. Monoclonal anti-Flag antibody (Stratagene) was used to detect E protein and mutants. Cells were then subjected to three washes with PBS and incubated with anti-mouse IgG conjugated to TRITC (DAKO) diluted in PBS (1:30~1:50) for 2 h at 4 °C. After three times washes with PBS, cells were mounted with glass coverslips using a fluorescence mounting medium (DAKO). To selectively permeabilize the plasma membrane, HeLa cells were fixed with 4% paraformaldehyde for 10 min and permeabilized with 5 µg of digitonin per ml for 5 min at RT. Staining was as described above. To detect subcellular localization of IBV M protein, Vero cells were grown on coverslips, fixed at 8 h post-infection in 4% paraformaldehyde for 10 min. Incubation with 0.2% Triton X-100/PBS at RT for 10 min was used to penetrate cells before they were rinsed three times by PBS. Polyclonal rabbit anti-M antisera were diluted to 1:200 and incubated with cells for at least 1 h at RT. Cells were then subjected to three washes with PBS and incubated with anti-rabbit IgG conjugated to FITC (DAKO) diluted in PBS (1:30~1:50) for 2 h

at 4 °C. After three times washes with PBS, cells were mounted with glass coverslips using a fluorescence mounting medium (DAKO). Images were collected with a META 510 confocal laser-scanning microscope (Zeiss).

2.3.3.4. Glycosylation study of E protein

HeLa cells expressing E protein were treated with glycoprotein denaturing buffer (0.5% SDS and 1% β -mercaptoethanol) at 100 °C for 10 min. The denatured proteins were incubated with 1 μ l glycosidase PNGase F (Research Biolabs) in G7 reaction buffer (50 mM sodium phosphate, pH7.5) (Research Biolabs) supplemented with 1% NP40 at 37 °C for 1 h. The deglycosylated proteins were analyzed by Western blot.

2.3.3.5. Proteinase K protection assay

HeLa cells grown in 60 mm dishes were transfected with SARS-CoV E and IBV E tagged with the Flag epitope at either N- or C-terminus, respectively, by using the vTF7-3 system. After incubation for 18 h, the cells were washed twice with ice-cold PBS, scraped off the dish, and homogenized with 20 strokes in a tight-fitting Dounce homogenizer. Nuclei were removed by centrifugation at 1,200 rpm for 15 min at 4 °C. Individual samples with the tagged SARS-CoV E or IBV E proteins were split into three microcentrifuge tubes. One tube was taken as control, the remaining two tubes in the absence or presence of 1% Triton X-100 were subjected to 20 μ g/ml proteinase K digestion for 40 min on ice. The reaction was stopped by adding 4 mM phenylmethylsulfonyl fluoride (PMSF). The samples were incubated at 100 °C for 15 min in the SDS loading buffer and then analyzed on SDS-PAGE.

2.3.3.6. Isolation of low-density detergent-insoluble membrane fraction on flotation gradients

HeLa cells expressing the Flag-tagged SARS-CoV E or Flag-tagged IBV E protein from two 10 cm dishes were washed twice with ice-cold PBS and lysed on ice for 30 min in 1 ml of 1% Triton X-100 TNE lysis buffer (25 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM EDTA) supplemented with protease inhibitor cocktail (Roche). The cell lysates were homogenized with 25 strokes using a Dounce homogenizer and centrifuged at $3000 \times g$ at 4 °C for 5 min to remove insoluble materials and nuclei. The supernatants were mixed with 1 ml of 80% sucrose in lysis buffer, placed at the bottom of a ultracentrifuge tube, overlaid with 6 ml of 30% and 3 ml of 5% sucrose in TNE lysis buffer, and ultracentrifuged at 38,000 rpm at 4 °C in a SW41 rotor (Beckman) for 18 h. After centrifugation, 11 fractions (1 ml each) were collected from the top to the bottom and analyzed immediately by Western blot or stored at -80 °C.

2.3.3.7. Lipid rafts aggregation and confocal microscopy

HeLa cells were seeded on four-chambered slides (IWAKI). HeLa cells grown in these four-chambered slides were transfected with Flag tagged IBV E by using the vTF7-3 system. Twelve hours post-transfection, the cells were incubated for 10 min at 4 °C. The working solution was prepared by adding 2 µl of the 1mg/ml CT-B conjugate stock solution (Molecular Probes) and 10 µl of anti-flag monoclonal

antibody (Stratagene) to 2 ml chilled DMEM medium. HeLa cells were incubated in working solution for 30 min at 4 °C. After gently washing several times with chilled 1 × PBS, the cells were incubated at 4 °C for 30 min in 0.5 ml chilled DMEM medium containing 2.5 µl of the anti-CT-B rabbit serum (Molecular Probes) and 10 µl of TRITC labeled anti-mouse secondary antibody (DAKO). After this incubation, the cells were gently washed several times with chilled 1 × PBS. The cells were then fixed in 4% paraformaldehyde for 15 min. After fixation, cells were mounted with glass coverslips using a fluorescence mounting medium (DAKO). Images were collected with a META 510 confocal laser-scanning microscope (Zeiss).

2.3.3.8. RNA extraction from mammalian cells

Media in the culture flask was poured out and Trizol reagent (Invitrogen) was directly added onto the cells. Cell lysate was used to extract RNA by using Trizol reagent (Invitrogen) according to the manufacturer's instruction. RNA was stored at -80 °C.

2.3.3.9. Northern blot analysis

Huh-7 cells were infected with wild-type or N3D/6D mutant viruses at a MOI of 1. Total RNA was extracted from cells infected with wild-type or mutant viruses. Ten micrograms of RNA were added to a mixture of 1 × MOPS, 37% formaldehyde and formamide and incubated at 65 °C for 20 min before subjected to gel electrophoresis. The segregated RNA were transferred onto a Hybond N+ membrane (Amersham

Pharmacia Biotech) via capillary action overnight and fixed by ultraviolet crosslinking. A DNA probe, corresponding to the last 222 nucleotides of the IBV 3' untranslated region (UTR), generated with primers 27387 5'-GGAAACGAACGGTA GACCCTTAGA-3' and 27608 5'-TGCTCTAACTCTATACTAGCCTAT-3', was used to detect all virus-specific RNA species. Hybridization of DIG-labeled DNA probes was carried out at 50 °C in hybridization buffer (0.5 M sodium phosphate pH 7.2, 7% sodium dodecyl sulfate, and 1 mM EDTA) overnight. After hybridization, the membrane was washed three times in wash buffer (0.2 × SSC, 0.1% SDS) at 68 °C, before proceeding to detection with CDP-Star (Roche) according to the manufacturer's instructions.

2.3.3.10. Reverse transcription reaction

The first cDNA reaction was done using the Expand Reverse Transcriptase (Roche) according to the manufacture's protocol. Briefly, 1 µg of RNA in 4.5 µl of nuclease free water with 5 µl of 10 pmol of reverse primer were incubated at 65 °C for 10 min. Reaction were chilled on ice for 5 min before adding 4 µl of 5 × Expand reverse transcriptase buffer, 2 µl of 100 mM DTT, 0.5 µl of RNase inhibitor (Promega), 2 µl of 10 mM dNTP and 1 µl of Expand reverse transcriptase (50 unit/µl) and then incubated at 43 °C for 1 h. The reaction was stopped by placing on ice.

2.3.3.11. Analysis of VLP

COS-7 cells were seeded one day prior to achieve 90% confluency for infection at a MOI of 10 with vTF7-3 (Fuerst *et al.*, 1986). Cells were transfected with PKT0

plasmids containing either wild-type or mutated IBV M genes in combination with the wild-type IBV E gene by using lipofectamine 2000 (Invitrogen). Cells were incubated in DMEM medium for 15 h. The media was clarified at $1,500 \times g$ for 10 min at 4 °C. VLPs were collected by pelleting the clarified media through a 20% sucrose gradient cushion for 2 h in a Beckman SW41Ti rotor at 38,000 rpm at 4 °C. Pellets were resuspended directly in 1 × SDS loading buffer. Intracellular and VLP samples were analyzed by SDS-PAGE. Protein samples were transferred to PVDF membranes and incubated with anti-IBV M and anti-IBV E antibody. Following incubation with anti-rabbit secondary antibodies, blots were visualized by ECL chemiluminescence (Amersham Pharmacia Biotech).

2.3.3.12. Yeast two-hybrid screen

Screening for interacting proteins for SAR-CoV E was carried out with the yeast two-hybrid system. SARS-CoV E, cloned into pGBKT7 vector, was used as the bait. The yeast strain AH109 was transformed by lithium acetate method and expression of the bait protein was confirmed by Western blot. The cDNA library was sequentially transformed into yeast expressing the bait protein, and transformants expressing both the bait and interacting prey proteins were selected on Trp-, Leu-, His-, and Ade-minus plates by incubation at 30 °C for 5 days and confirmed by testing for β -galactosidase activity using the filter lift assay. Subsequently, constructs containing the interacting prey proteins were isolated from positive yeast colonies, transformed into *E. coli* strain by electroporation, and the positive inserts were determined by sequencing.

2.3.3.13. In vitro translation

PKT0 vector containing cDNA encoding SARS-CoV E protein was *in vitro* translated with a TNT[®] T7 Coupled Wheat Germ Extract System (Promega).

2.3.3.14. Preparation of GST (glutathione S-transferase) and snapin fusion proteins

Full-length snapin was fused to the C-terminus of GST in frame by subcloning snapin cDNA fragment into the GST vector pGEX-5X-1 (GE Healthcare). The GST and GST fusion proteins were purified according to the manufacturer's instructions.

2.3.3.15. GST pull-down assay

2 µg of GST alone or GST-snapin fusion protein were prebound to glutathione sepharose beads by incubation with agitation for 1 h at RT in 0.5 ml of binding buffer (20 mM HEPES pH7.8, 150 mM KCl, 0.1% Gelatin, 0.1% Triton X-100, 0.1% NP40, 5 mM MgCl₂, 2 mM DTT). 15 µl of ³⁵S-labeled SARS-CoV E was then added, and incubation was continued for at least 2 h. The beads were washed five times with washing buffer (40 mM HEPES pH7.5, 0.1% NP40, 1 mM EDTA, and 100 mM KCl). Labeled protein bound on the beads was recovered by heating at 100 °C in 10 µl of SDS loading buffer and was analyzed by SDS-PAGE. Radiolabeled bands were visualized by autoradiography.

2.3.3.16. Coimmunoprecipitation

Transiently transfected HeLa cells in 100-mm dishes were lysed in 1 ml of lysis buffer (150 mM NaCl, 1% NP-40, and 50 mM Tris-HCl, pH 8.0) with protease inhibitor cocktail (Sigma). The lysates were centrifuged at 12,000 rpm for 20 min at 4 °C. The supernatants were added with anti-HA antibody at 4 °C for 2 h. 40 µl of Protein-A agarose beads were added to the lysates and incubated with shaking for 1 h at 4 °C. The beads were collected by centrifugation and washed for three times with RIPA buffer (150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.05% SDS, and 50 mM Tris-HCl, pH 8.0). Proteins binding to the beads were eluted by adding SDS loading buffer and analyzed by Western blotting with anti-Flag antibody.

2.3.4. Generation of recombinant IBV and virus assay methods

2.3.4.1. Virus

The Beaudette strain of IBV (ATCC VR-22) was obtained from the American Type Culture Collection (ATCC) and was adapted to Vero and H1299 cells as previously described (Liu *et al.*, 1998).

2.3.4.2. In vitro assembly of full-length cDNA clones

To construct a full-length IBV clone, five plasmids which contain five fragments (A to E) spanning the entire IBV genome were constructed in our lab (Fang *et al.*, 2006). Briefly, five fragments spanning the entire IBV genome were obtained by RT-PCR from Vero cells infected with Vero cell-adapted IBV P65. To facilitate the assembly of the full-length cDNA *in vitro*, restriction sites for either BsmBI or BsaI

were introduced into both the 5' and 3' ends of the fragments (Fig. 2-3). The PCR products were purified from agarose gels and cloned into pCR-XL-TOPO(Invitrogen) or pEGM-T Easy (Promega) vectors. Subsequently, fragment A was removed from pCR-XL-TOPO by digestion with NheI and EcoRI, and subcloned into pKT0 vector. In fragment A, a 19-nucleotide sequence corresponding to the T7 promotor was inserted into the 5' end of the IBV genome to facilitate *in vitro* transcription by T7 polymerase (Fig. 2-3). Plasmids were digested with either BsmBI (fragment A) or BsaI (fragments B, C, D and E). The digested plasmids were separated on 0.8% agarose gels containing crystal violet. Bands corresponding to each of the fragments were cut from the gels and purified with QIAquick gel extraction kit (QIAGEN Inc.).

Fragments A, B, C, D and E were ligated with T4 DNA ligase at 16 °C overnight. The ligation products were extracted with phenol/chloroform, precipitated with ethanol, and detected by electrophoresis on 0.4% agarose gels.

Fig. 2-3. Strategy for the assembly of an IBV infectious cDNA clone. The regions coding for the replicase polyproteins, the structural proteins S, E, M, and N, the accessory proteins 3a, 3b, 5a, and 5b, and the 5'- and 3'-UTR are shown. Also shown are the regions of the five RT-PCR fragments, the T7 promoter at the 5'-end of fragment A, and the 30 As at the 3'-end of fragment E.

2.3.4.3. In vitro transcription and electroporation

Full-length transcripts were generated *in vitro* using the mMessage mMachine T7 kit (Ambion) according to the manufacturer's instructions with certain modifications. Briefly, 30 µl of transcription reaction with a 1:1 ratio of GTP to cap analog was sequentially incubated at 40.5 °C for 25 min, 37.5 °C for 50 min, 40.5 °C 25 min and 37.5 °C for 20 min.

The N transcripts were generated by using a linearized pKTO-IBV N containing IBV N gene and the 3'-UTR region as templates. A 1:2 ratio of GTP to cap analog was used for the transcription of IBV N gene.

2.3.4.4. Introduction of in vitro synthesized transcripts into Vero cells by electroporation

The *in vitro* synthesized full-length RNA and N transcripts were treated with DNase I and purified with phenol/chloroform. Vero cells were grown to 90% confluency, trypsinized, washed twice with cold PBS, and resuspended in PBS. RNA transcripts were added to 400 µl of Vero cell suspension in a 4 mm electroporation cuvette (Bio-Rad), and electroporated with one pulse at 450 V, 50 µF with a Bio-Rad Gene Pulser II electroporator. The electroporated Vero cells were cultured overnight in 1% FBS containing MEM in a 60 mm dish or a six-well plate and further cultured in MEM without FBS.

2.3.4.5. IBV plaque infectivity assay

Confluent Vero cells in 6 well plates (IWAKI) were infected with 100 μ l of 100-fold diluted virus stock. After 1 h of incubation, cells were washed twice with PBS and cultured in 2 ml DMEM containing 1% carboxymethylcellulose (Sigma) and 1% FBS. After 4 days of incubation, the cells were fixed with 4% paraformaldehyde for 15 min, and visualized by staining the intact cells with 0.1% toluidine blue (Sigma) for 30 min. The yield of virus was calculated in terms of plaque-forming unit (PFU) per ml.

2.3.4.6. Growth curve of the recombinant viruses on Vero cells

Confluent monolayers of Vero cells on 6-well plates were infected with wild-type and recombinant IBV, and harvested at different times post-infection. Viral stocks were prepared by freezing/thaw of the cells three times. The 50% tissue culture infection dose (TCID₅₀) of each sample was determined by infecting five wells of Vero cells on 96-well plates with 10-fold serial dilution of each viral stock.

2.3.4.7. Virus purification

Vero cells in T-75 Flask were washed once with PBS and infected with IBV at a MOI of 10. After incubation for 90 min, the inoculum was replaced with 12 ml of DMEM. At 18 h post-infection the supernatant was collected and clarified by centrifugation at $450 \times g$ for 15 min. IBV particles were partially purified by ultracentrifugation on a 20% sucrose cushion in a Beckman SW28 rotor at $75,000 \times g$ for 3 h. The pellet was resuspended in 1 ml TNE (25 mM Tris-HCl pH 7.5, 150 mM

NaCl, 5 mM EDTA) and further applied on a discontinuous sucrose gradient of 20 to 60% sucrose. The samples were centrifuged at 28,000 rpm for 18 h. Subsequently, 10 fractions were collected. Each fraction was subjected to Western blot analysis with anti-M protein antibody. The fraction 8 and 9 were diluted in TNE and pelleted through a 20% sucrose cushion at 38,000 rpm for 2 h with a Beckman SW41 rotor. Pellet was stored at -80 °C until use.

Chapter 3

Mixed Membrane Topologies, Lipid Rafts Association and Host Cell Interaction of the Severe Acute Respiratory Syndrome Coronavirus Envelope Protein

3.1. Introduction

The SARS-CoV is an enveloped virus with a single strand, positive-sense RNA genome of 29.7 kb in length. Similar to other coronaviruses, its envelope typically contains three major structural proteins: S, M and E, with functions including recognition of target cells and fusion, interaction with other viral components, and involvement in virion assembly and budding.

In this report, the membrane topology of SARS-CoV E protein was determined by immunofluorescent staining of cells differentially permeabilized with detergents and by limited proteinase K digestion of microsomal membranes. These studies have revealed that both the N- and C-terminal regions of the protein are located in the cytoplasm ($N_{\text{cyto}}C_{\text{cyto}}$). However, a minor proportion of the protein is found to be posttranslationally modified by N-linked glycosylation on the asparagine 66 residue. It suggests that a certain proportion of the protein may also adopt either an $N_{\text{cyto}}C_{\text{exo}}$ or $N_{\text{exo}}C_{\text{exo}}$ topology. This is the first coronavirus E protein with two distinct membrane topologies. Furthermore, the lipid rafts association of SARS-CoV or IBV E protein was determined. To address host factors that might interact with SARS-CoV E protein, a yeast two-hybrid screen of a HeLa cDNA library was performed. We identified snapin, which was found to interact specifically with SARS-CoV E protein both in vitro and in vivo.

3.2. Results

3.2.1. Prediction of the hydrophobicity and membrane topology of SARS-CoV E

protein

The hydrophobicity of the SARS-CoV E protein is shown as a Kyte-Doolittle hydropathy plot in Fig. 3-1b. The protein is largely hydrophobic in the region from amino acids 9 to 59. In both N- and C-terminal regions, some hydrophilic amino acid stretches were found.

Three computer programs, *i.e.*, TMHMM (Krogh et al., 2001), HMMTOP (Tusnady & Simon, 1998), and MEMSAT (Jones et al., 1994), were used to predict the membrane topology of the SARS-CoV E protein. All three programs predicted that SARS-CoV E protein contains one transmembrane domain. Both TMHMM and MEMSAT predicted that SARS-CoV E protein may assume an $N_{\text{cyto}}C_{\text{exo}}$ topology, whereas HMMTOP indicated that SARS-CoV E protein may adopt a $C_{\text{cyto}}N_{\text{exo}}$ topology.

a

	1	10	20	30	40	50	60	70	
E	MYSFVSEET <u>GTLIVNSVLLFLAEVVFLLVTLAILTALRLCAYCCNI</u> V NVSLVKPTVYVYSRVKNL N SSEGVPDLLV								
N48-D	-----D-----								
N66-D	-----D-----								

b

Fig. 3-1. Diagram of wild-type and mutant SARS-CoV E constructs, and hydrophobicity score of SARS-CoV E protein. (a) Amino acid sequences of wild-type and mutant SARS-CoV E protein. The amino acid sequence of the putative transmembrane domain is underlined, the three cysteine residues and the two potential glycosylation sites are indicated in bold. Also shown are the mutations introduced into each of the mutant constructs. (b) The hydropathy profile of SARS-CoV E protein determined by Kyte and Doolittle with a 7-residue window. It displays the highly hydrophobic character of this protein.

3.2.2. Analysis of the membrane topology of SARS-CoV E protein by immunofluorescence microscopy

To test these predictions and to experimentally determine the membrane topology of the SARS-CoV E protein, both SARS-CoV and IBV E proteins were tagged with Flag at N- and C-termini, respectively, and expressed in HeLa cells. Cells were then permeabilized with either 0.5% Triton X-100 or 5 µg/ml digitonin, and the expression of the Flag-tagged E protein was detected by indirect immunofluorescent staining using anti-Flag monoclonal antibody. Treatment of cells with digitonin at low concentrations selectively permeabilizes the plasma membrane but leaves the intracellular membranes intact, while Triton X-100 treatment permeabilizes both plasma and intracellular membranes (Plutner et al., 1992). If the Flag tag is exposed to the cytoplasmic side, the epitope can be recognized by anti-Flag antibody after digitonin treatment. On the other hand, if the Flag tag is exposed luminally, the epitope cannot be recognized by the anti-Flag antibody after digitonin treatment. It can only be recognized after Triton X-100 treatment.

As shown in Fig. 3-2, clear detection of the N- or C-terminally tagged SARS-CoV E protein at perinuclear regions was obtained in HeLa cells permeabilized with either 0.5% Triton X-100 or 5 µg/ml digitonin (Fig. 3-2), suggesting that both N- and C-termini of the protein may be located in the cytoplasmic side. However, in cells expressing the N-terminally tagged IBV E protein, expression of the protein was observed in cells permeabilized with 0.5% Triton X-100 only (Fig. 3-2). No detection of the protein expression in cells permeabilized with 5

$\mu\text{g/ml}$ digitonin was obtained (Fig. 3-2). It was also noted that no significant difference in the staining intensity was observed between cells treated with digitonin and Triton X-100. In cells expressing the C-terminally tagged IBV E protein, the protein was detected in cells permeabilized under either condition (Fig. 3-2). These results are consistent with the topology of IBV E protein established before (Corse & Machamer, 2000), and justify the experimental conditions used. As a control, PDI, a host protein residing in the ER lumen, was detected in cells permeabilized with 0.5% Triton X-100, but was not detected in cells permeabilized with 5 $\mu\text{g/ml}$ digitonin (Fig. 3-2), confirming that the ER membrane was still intact after treatment with digitonin (Fig. 3-2). Taken together, these results indicate that both the N- and C-termini of the SARS-CoV E protein may be located in the cytoplasmic side of the cell, whereas IBV E protein adopts an $N_{\text{exo}}C_{\text{cyto}}$ topology.

Fig. 3-2. Cytoplasmic exposure of the Flag epitope tagged at the N- or C-terminus of the SARS-CoV E protein and the C-terminus of the IBV E protein. HeLa cells expressing Flag-SARS E, SARS E-Flag, Flag-IBV E or IBV E-Flag, respectively, were permeabilized using either 5 μ g/ml digitonin or 0.5% Triton X-100 and immunostained with anti-Flag antibody as the primary antibody and TRITC-conjugated anti-mouse IgG antibody as the secondary antibody. Untransfected HeLa cells were treated with digitonin or Triton X-100 and immunostained with anti-PDI antibody and TRITC-labeled anti-rabbit IgG secondary antibody.

3.2.3. Analysis of the membrane topology of SARS-CoV E protein by differential permeabilization of the plasma membrane and limited proteinase K treatment

The topology of SARS-CoV E protein on cellular membranes was further analyzed by limited proteinase K digestion of the membrane fraction prepared from cells expressing the N- or C-terminally tagged E protein. Digestion with the nonspecific proteinase K would degrade proteins (or a portion of the protein) protruding from the exterior face of the microsomal membranes, while proteins (or a portion of the protein) orientated towards the lumen are protected. For this purpose, HeLa cells expressing the Flag-tagged SARS-CoV E protein at either N- or C-terminus were broken by homogenization, and the membrane fraction was collected. The membrane fraction was then divided into three aliquots: one was treated with both 1% Triton X-100 and 20 µg/ml proteinase K, one treated with 20 µg/ml proteinase K only, and one without any treatment. Total protein was then separated by SDS-PAGE and analyzed by Western blot with anti-Flag antibody. As shown in Fig. 3-3, both N- or C-terminally tagged SARS-CoV E protein was efficiently detected in the untreated membrane fraction (Fig. 3-3, lanes 1 and 4). Upon treatment of the membrane fraction with 20 µg/ml proteinase K in the presence (Fig. 3-3, lanes 3 and 6) or absence (Fig. 3-3, lanes 2 and 5) of 1% Triton X-100, no detection of the SARS-CoV E protein was observed from cells expressing either the N- (lanes 2 and 3) or C- (lanes 5 and 6) terminally tagged E protein, demonstrating that treatment with proteinase K led to the removal of the Flag tag from either terminus of the SARS-CoV E protein.

In cells expressing the N- or C-terminally tagged IBV E protein, the protein was clearly detected in the membrane fraction without treatment with Triton X-100 and proteinase K (Fig. 3-3, lanes 7 and 10). Treatment of the membrane fraction with 20 µg/ml proteinase K resulted in the detection of a fragment of approximately 10 kDa representing the N-terminal region of the IBV E protein (Fig. 3-3, lane 8). Upon treatment of the membrane fraction with 1% Triton X-100 and 20 µg/ml proteinase K, the IBV E protein was no longer detected (Fig. 3-3, lane 9). In cells expressing the C-terminally tagged IBV E protein, the protein was not detected after treatment of the membrane fraction with proteinase K in the presence or absence of Triton X-100 (Fig. 3-3, lanes 10-12). Consistent with the immunofluorescence data, these results reinforce the conclusion that the N-terminus of the IBV E protein was located in the lumen of the ER and the Golgi apparatus.

As an internal control for the integrity of microsomal membranes after limited proteolytic digestion, the detection of PDI was included in the experiment. As shown in Fig. 3-3, the full-length PDI was detected in the membrane fraction before and after treatment with 20 µg/ml proteinase K in the absence of 1% Triton X-100 (Fig. 3-3, lanes 1, 2, 4, 5, 7, 8, 10 and 11). In addition, a shorter form of the protein was also detected in the membrane fraction after treatment with 20 µg/ml proteinase K in the absence of 1% Triton X-100 (Fig. 3-3, lanes 2, 5, 8 and 11). In the presence of 1% Triton X-100, treatment of the membrane fraction with proteinase K showed that no full-length PDI was detected (Fig. 3-3, lanes 3, 6, 9 and 12). However, the shorter form of PDI was still detected (Fig. 3-3, lanes 3, 6, 9 and 12), suggesting that it may

represent a proteinase K-resistant fragment of PDI. The fact that the shorter form of PDI was detected in the membrane fraction after treatment with 20 $\mu\text{g/ml}$ proteinase K in the absence of 1% Triton X-100 suggests that the procedure used to prepare the membrane fraction may result in partial disruption of the microsomes. In the same experiment, β -tubulin was also included as an internal control for the degradation of cytosolic proteins. The protein was degraded by the treatment with 20 $\mu\text{g/ml}$ proteinase K in the presence or absence of Triton X-100 (Fig. 3-3). Together with the immunofluorescence studies, these results confirm that both the N-terminus and C-terminus of SARS-CoV E are cytosolic with an orientation of $N_{\text{cyto}}C_{\text{cyto}}$.

Fig. 3-3. Limited proteinase K digestion of microsomal membranes prepared from cells expressing the SARS-CoV or IBV E protein tagged with the Flag epitope at the N- and C-termini, respectively. Intact microsomes were isolated from cells expressing Flag-SARS E, SARS E-Flag, Flag-IBV E and IBV E-Flag, and subjected to digestion by proteinase K in the absence (lanes 2, 5, 8, 11) or presence (lanes 3, 6, 9, 12) of Triton X-100. The proteinase K-treated samples were separated on SDS - 17.5% polyacrylamide gel and analyzed by Western blot. PDI (an ER luminal protein) was detected by anti-PDI antibody (Santa Cruz), and β -tubulin (a cytosolic protein) was detected by anti-tubulin antibody (Santa Cruz). Numbers on the left indicate molecular masses in kilodaltons

3.2.4. Glycosylation of SARS-CoV E protein

Examination of the SARS-CoV E protein sequence showed the presence of two potential N-linked glycosylation sites on asparagines 48 (N48) and 66 (N66) (Fig. 3-1a). If the protein adopted a sole N_{cyto}C_{cyto} topology, glycosylation at the two positions would not occur. However, multiple bands were usually detected when the protein was expressed in different systems, suggesting that it may undergo posttranslational modifications. To address the possibility that the protein may be modified by N-linked glycosylation, mutations were introduced into the E protein to change the predicted N-linked glycosylation sites from asparagine to aspartic acid. Two mutants, N48-D and N66-D, were generated and expressed (Fig. 3-1a). Western blot analysis of cells expressing wild-type and the N48-D mutant showed the detection of three bands on an SDS-17.5% polyacrylamide gel (Fig. 3-4a, lanes 1 and 2). These may represent three isoforms of the E protein. In cells expressing the N66-D mutant, only two bands were detected; the most slowly migrating species of the three isoforms was not observed (Fig. 3-4a, lane 3).

To analyze further the N-linked glycosylation of the E protein, cells expressing wild-type (Fig. 3-4b, lanes 1 and 2) and mutant E (Fig. 3-4b, lanes 3-6) were lysed. Total cell lysates were first treated with the N-linked glycosidase PNGase F, and analyzed by Western blot with anti-E polyclonal antibody. In cells expressing wild-type and the N48-D mutant E protein, the protein was separated into two major bands in the gel system used (Fig. 3-4b, lanes 1 and 3). Treatment of the same total cell lysates led to the removal of the upper band (Fig. 3-4b, lanes 2 and 4), confirming that

it represents the glycosylated form of the protein. In this gel system, the two unglycosylated isoforms of wild-type SARS-CoV E protein were not well separated, so three major bands were detected with Flag-SARS E but only two major bands with wild-type SARS-CoV E protein. In addition, a minor species of approximately 20 kDa, representing the dimeric form of the protein, was observed in cells expressing the two constructs (Fig. 3-4b, lanes 1-4). In cells expressing the N66-D mutant, the glycosylated form was not detected either before or after treatment with PNGase F (Fig. 3-4b, lanes 5 and 6). Interestingly, in cells expressing this mutant E protein, more dimeric form and a species representing trimer of the E protein were detected (Fig. 3-4b, lanes 5 and 6). It suggests that this mutant E protein tends to form multimers or aggregates. As a control, cells expressing IBV E protein were treated with the same glycosidase. The protein was detected as a single band in total cell lysates with or without PNGase F treatment (Fig. 3-4b, lanes 7 and 8). In cells expressing the Flag-tagged E protein, this treatment resulted in the disappearance of the most slowly migrating species of the three isoforms (Fig. 3-4b, lanes 9 and 10), further confirming that it is the glycosylated form of the E protein. Because minor amount of Flag-tagged E protein is glycosylated, this form cannot be detected in protease K protection assay (Fig. 3-3, lanes 1 and 4). These results demonstrate that a minor proportion of the SARS-CoV E protein is modified by N-linked glycosylation and the N66 residue is the site for this modification. More importantly, it suggests that this portion of the SARS-CoV E protein would assume a different membrane

topology from the majority of the E protein, i.e., the C-terminal region of the protein would be located in the lumens of the ER and the Golgi apparatus.

a

b

Fig. 3-4. N-linked glycosylation of SARS-CoV E protein. (a) HeLa cells were transfected with the Flag-tagged wild-type (lane 1) and three mutant constructs containing mutations of the N48 (lane 2) and N66 (lane 3). Cell lysates were prepared 24 h post-transfection, polypeptides were separated by SDS - PAGE and analyzed by Western blot using the anti-Flag antibody. Numbers on the left indicate molecular masses in kilodaltons. (b) Total cell lysates prepared from HeLa cells expressing wild-type SARS-CoV E (lanes 1 and 2), N48-D (lanes 3 and 4), N66-D (lanes 5 and 6), wild-type IBV E (lanes 7 and 8) or the Flag-tagged SARS-CoV E (lanes 9 and 10) were treated either with (lanes 2, 4, 6, 8 and 10) or without (lanes 1, 3, 5, 7, and 9) PNGase F. Polypeptides were separated by SDS - PAGE and analyzed by Western blot using anti-SARS-CoV E antibodies (lanes 1-6), anti-IBV E (lanes 7 and 8) or anti-Flag (lanes 9 and 10). Numbers on the left indicate molecular masses in kilodaltons.

3.2.5. Cell surface expression of the SARS-CoV E protein

Immunofluorescent staining of cells expressing the Flag-tagged SARS-CoV E protein at either N- or C-terminus was then carried out to test if the E protein translocated to the cell surface could be accessed by the antibody. As shown in Fig 3-5, immunofluorescent staining of HeLa cells expressing the N-terminally Flag-tagged IBV E using anti-Flag antibody exhibited typical cell surface staining. In cells expressing SARS-CoV E protein with the Flag epitope tagged at either N- or C-terminus, no cell with obvious positive staining was detected (Fig. 3-5). However, cells with a few fluorescent dots on the surface were consistently observed (Fig. 3-5).

Fig. 3-5. Surface staining of HeLa cells expressing the SARS-CoV E protein with Flag tagged at the N- or C-terminus, respectively, and IBV E protein with the Flag tagged at the N-terminus. HeLa cells expressing Flag-SARS E, SARS E-Flag and Flag-IBV E, respectively, were immunostained with anti-Flag antibody as the primary antibody and TRITC-conjugated anti-mouse IgG antibody as the secondary antibody.

3.2.6. Partial association of SARS-CoV E protein and IBV E protein with lipid rafts

To test if the E protein may be associated with lipid rafts, the low-density, detergent-insoluble membrane fraction was isolated from HeLa cells overexpressing the SARS-CoV E protein and IBV E protein. As shown in Fig. 3-6, the majority of the SARS-CoV E protein was detected at the bottom fractions (lanes 9-11). However, a certain proportion of SARS-CoV E protein associated with lipid rafts was detected in fraction 3, 4, and 5 (Fig. 3-6a). The GM1 was detected in fraction 4 and 5 (Fig. 3-6a). GM1 is a glycosphingolipid that localizes to plasma membrane lipid rafts and is used as a raft marker (Parton, 1994). The IBV E protein also showed the similar pattern as SARS-CoV E protein after isolation of low-density detergent-insoluble membrane fraction on flotation gradients (Fig. 3-6b).

To visualize the IBV E protein associated with lipid rafts on the plasma membrane, confocal immunofluorescence microscopy was performed with HeLa cells. HeLa cells grown on chamber slides were transfected with Flag-IBV E. At 14 h post-transfection, cells were labeled simultaneously at 4 °C with fluorescently labeled CT-B and anti-FLAG monoclonal antibodies. The fluorescently conjugated CT-B can specifically label the GM1 (Merritt et al., 1994). After washing with PBS, cells were incubated with the anti-CT-B rabbit serum (Molecular Probes) and TRITC labeled anti-mouse secondary antibody (DAKO) in chilled DMEM medium. The anti-CT-B rabbit serum can cross-link the raft to promote visualization of raft domain. It was shown that the IBV E protein colocalized with GM1 (Fig. 3-7). The SARS E protein

did not exhibit well surface staining as shown in Fig. 3-5, so this experiment was not performed.

Fig. 3-6. Association of SARS-CoV and IBV E protein with lipid rafts. HeLa cells expressing the Flag-tagged SARS E protein (a) or IBV E protein (b) were lysed with 1% Triton X-100, and centrifuged to remove insoluble material and nuclei. The supernatants were fractionated by ultracentrifugation with a sucrose gradient, and 11 fractions were collected. The presence of the SARS-CoV E protein (a) and IBV E protein (b) in each fraction was analyzed by Western blot using anti-Flag antibody. The presence of GM1 was determined by dot blot. Numbers on the left indicate molecular masses in kilodaltons.

Fig. 3-7. IBV E protein colocalizes with GM1. HeLa cells were transfected with the Flag-IBV E and incubated on ice with an anti-FLAG antibody and a fluorescently labeled CT-B. After wash, cells were incubated with the anti-CT-B rabbit serum and TRITC labeled anti-mouse secondary antibody. Samples were analyzed by confocal microscopy, with colocalization shown in yellow in the overlay panel.

3.2.7. Identification of snapin as a SARS-CoV E-interacting protein

To identify cellular proteins that may interact with SARS-CoV E protein, a yeast two-hybrid screen of a HeLa cDNA library was performed. Transformed yeast cells were selected on semi-stringent SD/-Trp/-Leu/-His plates, followed by replating positive colonies onto SD/-Trp/-Leu/-His/-Ade plates for a more stringent selection. The cDNA inserts from four positive clones were found to encode snapin (Fig. 3-8d).

GST pull-down assay was performed to confirm the interaction between SARS-CoV E and snapin in a context other than in yeast. As shown in Fig. 3-8a, equal amounts and similar purities of purified GST and GST-snapin were used in the GST pull-down assay. It was found that GST-snapin was able to precipitate the ³⁵S-labeled SARS-CoV E protein (Fig. 3-8b). In contrast, the SARS-CoV E protein could not be precipitated by using equal amount of GST protein (Fig. 3-8b). These results confirmed that SARS-CoV E protein interacted with snapin *in vitro*.

Subsequently, the interaction of SARS-CoV E and snapin in intact cells was verified by coimmunoprecipitation experiments. For this purpose, HeLa cells were transfected with Flag-tagged SARS-CoV E, HA-tagged snapin, or cotransfected with both plasmids (Fig. 3-8c). At 16 h post-transfection, cell lysates were prepared and subjected to immunoprecipitation with anti-HA antibody. As shown in Fig. 3-8c lane 7, a 14-kDa band which represented the precipitated SARS-CoV E could only be detected when the two proteins were co-expressed.

To map the interacting domain on snapin, three fragments covering snapin sequences from amino acids 1-79, 79-136 and 36-136 were cloned into pACT2, and

co-transformed into yeast cells with the bait construct pGBKT7-E. As shown in Fig. 3-8d, normal growth of yeast was observed only when the fragment covering 36-136 was used. These results demonstrated that the interacting domain is located in amino acids 36-136 in snapin and that the C-terminal region from amino acids 79-136 is not sufficient for the interaction.

a

b

c

d

Fig. 3-8. Interaction of SARS-CoV E protein and snapin (a) Verification of purified GST and GST fusion proteins with Coomassie Blue staining. (b) GST pull-down assays. *In vitro* translated and ^{35}S -labelled SARS-CoV E protein was incubated with GST alone or GST-snapin fusion proteins (lanes 1 and 2) bound to glutathione-sepharose beads. After extensive washing, the ^{35}S -labelled protein bound to the beads was eluted and analyzed by autoradiography. (c) Coimmunoprecipitation of SARS-CoV E protein and snapin from cellular extracts. HeLa cells expressing Flag-tagged SARS-CoV E (lane 2), HA-tagged snapin (lane 3), or co-expressing both proteins (lane 1) were lysed and analyzed by Western blotting using anti-HA antibody (Lanes 1-3) or anti-Flag antibody (lanes 4-6). The cell lysis was subjected to immunoprecipitation with anti-HA antibody before analysis on SDS-15% polyacrylamide gel (Lanes 7-9). Numbers on the left indicate molecular masses in kilodaltons. (d) Mapping the region in snapin responsible for interaction with SARS-CoV E protein. Schematic map indicates the positions of the N-terminal hydrophobic domain (amino acids 1-20) and two predicted helical regions (amino acids 37-65 and 81-126, respectively). Four fragments covering snapin sequences from amino acid 1-79, 79-136 and 36-136 were cloned into plasmid pACT2 and co-transformed into yeast cells with bait construct pGBKT7-E. The growth of the yeast cells on SD/-Trp/-Leu/-His/-Ade plate is shown.

3.3. Summary

Coronavirus envelope (E) protein is a small integral membrane protein with multi-functions in virion assembly, morphogenesis and virus-host interaction. Different coronavirus E proteins share striking similarities in biochemical properties and biological functions, but seem to adopt distinct membrane topology. In this report, the membrane topology of the SARS-CoV E protein was studied by immunofluorescent staining of cells differentially permeabilized with detergents and proteinase K protection assay. It was revealed that both the N- and C-termini of the SARS-CoV E protein are exposed to the cytoplasmic side of the membranes ($N_{\text{cyto}}C_{\text{cyto}}$). In contrast, parallel experiments showed that the E protein from IBV spanned the membranes once, with the N-terminus exposed luminally and the C-terminus exposed cytoplasmically ($N_{\text{exo(lum)}}C_{\text{cyto}}$). Intriguingly, a minor proportion of the SARS-CoV E protein was found to be modified by N-linked glycosylation on Asn 66 and inserted into the membranes with the C-terminus exposed to the luminal side. The presence of two distinct membrane topologies of the SARS-CoV E protein may provide a useful clue to the pathogenesis of SARS-CoV. Further immunofluorescence staining and cell fractionation studies demonstrate that SARS-CoV E protein and IBV E protein are partially associated with lipid rafts. Snapin was identified as an interaction partner of the SARS-CoV E protein in the yeast two-hybrid system. The interaction of SARS-CoV E protein with snapin was confirmed by co-immunoprecipitation and GST pull-down assay.

Chapter 4

Systematic Analysis of the N-Terminal Domain of the Infectious Bronchitis

Virus Membrane Protein

4.1. Introduction

M protein is a multispanning membrane protein. It consists of a short amino terminus exposed on the exterior of the virion, three hydrophobic transmembrane domains and a large carboxy terminus situated in the interior of the virion. The ectodomain, which is the least conserved region of M protein, is glycosylated. Many different functions have been assigned to oligosaccharide side chains. The carbohydrates have been shown to be important for folding, structure stability and intracellular sorting of proteins and to play a role in the generation of immune response (Drickamer & Taylor, 1998; Helenius & Aebi, 2001).

In the present study, the effects of N-linked glycosylation and other N-terminal domain mutations of IBV M protein on virus assembly were investigated. For this purpose a set of mutations was introduced into the IBV genome by full-length infectious cDNA clone of IBV. Meanwhile the effects of these IBV M mutants on the VLP formation were investigated. The results demonstrated that all these IBV M mutants can support the VLP formation without showing much difference, but some substitutions in this domain had severe effects on the recovery of infectious virus with full-length infectious cDNA clone system.

4.2. Results

4.2.1. N-linked glycosylation of IBV M protein

IBV M contains two consensus sites for N-linked glycosylation near its N terminus. Given that the N terminus is luminal and M protein is a glycoprotein,

mutagenesis studies were done to investigate if both sites are used (Cavanagh *et al.*, 1986; Cavanagh & Davis, 1988). Mutations were introduced into the M protein to change the predicted N-linked glycosylation sites from asparagine to aspartic acid. Three mutants, N3D, N6D and N3D/6D were generated and expressed (Fig. 4-1). Western blot analysis of cells expressing wild-type M protein detected three bands (Fig. 4-2, lane 1). In cells expressing the N3D and N6D mutants, two bands were detected (Fig. 4-2, lane 3 and lane 5). For double mutant N3D/6D, only the lowest band representing the unglycosylated form of M protein was detected (Fig. 4-2, lane 7). To further analyze the N-linked glycosylation status of M protein, cells expressing wild-type and mutant M were treated with PNGase F and analyzed by Western blot with anti-M polyclonal antibodies. This treatment resulted in the disappearance of the two upper bands in wild-type M protein and one upper band in each single mutant M protein (Fig. 4-2, lane 2, 4 and 6). No difference was observed for the N3D/6D mutant after PNGase F treatment (Fig. 4-2, lane 7 and 8). These results confirmed that the IBV M protein is N-link glycosylated at both the N3 and N6 residues.

	1	10	20	virus recovery
wt	MPNETNCTLDFEQSVQLFKE			+
N3D/6D	MPDET D CTLDFEQSVQLFKE			+
N3D	MPDET N CTLDFEQSVQLFKE			+
N6D	MPNET D CTLDFEQSVQLFKE			+
GFP48	MP SKGEELFTGVVPIL VEKE			-
GFPL15	MP SKGEET LDFEQSVQLFKE			-
GFPM15	MPNETN CLFTGVQ SVQLFKE			-
GFPR18	MPNETNCTLD FEVPIL VEKE			-
C7A	MPNETN A TLDFEQSVQLFKE			+
L9A	MPNETNCT A DFEQSVQLFKE			+
Q16A	MPNETNCTLD FEQSV ALFKE			+
V15A/Q16A	MPNETNCTLD FEQSA ALFKE			+
F18A	MPNETNCTLD FEQSVQL AKE			+
F18A/K19A	MPNETNCTLD FEQSVQL AAE			-

Fig. 4-1. Diagram of wild-type and mutant IBV M constructs. Amino acid substitutions in the N-terminal region of IBV M are indicated in red letters. The recovery ability of mutant viruses was indicated by + and -, respectively.

Fig. 4-2. Analysis of glycosylation of IBV M. Total cell lysates prepared from HeLa cells expressing wild-type IBV M (lanes 1 and 2), N3D (lanes 3 and 4), N6D (lanes 5 and 6) and N3D/6D (lanes 7 and 8) were treated either with (lanes 2, 4, 6 and 8) or without (lanes 1, 3, 5 and 7) PNGase F. Polypeptides were separated by SDS - PAGE and analyzed by Western blot using anti-IBV M antibody. Numbers on the left indicate molecular masses in kilodaltons.

4.2.2. Functional analysis of the N-linked glycosylation of IBV M protein

4.2.2.1. Construction of mutant IBV virus with mutations of the two N-linked glycosylation sites in the M protein

To characterize the role of glycosylation of the ectodomain of the M protein, two mutations from asparagine to aspartic acid at N3 and N6 were introduced into the genome of IBV by a full-length infectious cDNA clone. Plasmids were digested with either BsmBI (fragment A) or BsaI (fragment B, C, D, E). Bands corresponding to each of the fragments were purified with QIAquick gel extraction kit (QIAGEN Inc.). The full-length clone was made by ligation of the purified fragments *in vitro* (Fig. 4-3). Using this ligation product as template, the full-length *in vitro* synthesized transcripts were generated using the mMessage mMachine T7 kit (Ambion) (Fig. 4-3). As coronavirus N gene transcripts were shown to enhance the recovery of the rescued virus from the *in vitro* synthesized full-length transcripts, the N transcripts were generated from a linearized pKT0-IBV N construct. The full-length transcripts together with the N transcripts were electroporated into Vero cells. Two mutations from asparagine to aspartic acid at N3 and N6 were introduced into fragment E by using the QuickChange site-directed mutagenesis kit (Stratagen). Typical CPE were observed in cells transfected with either wild-type RNA transcripts or N3D/6D mutant transcripts at 4 days post-electroporation and recombinant viruses were recovered.

Fig. 4-3. In vitro assembly of the cDNA for the full-length IBV genome. The six cDNA fragments covering IBV sequences from nucleotides 1 – 5751 (lane 1), 5752 – 8693 (lane 2), 8694 – 15,520 (lane 3), 15,521 – 20,900 (lane 4), and 20,901 – 27,608 (lane 5 and 6) respectively, were obtained by digestion of corresponding plasmid DNA with either BsmBI or BsaI, purified from agarose gel, and analyzed on a 0.8% agarose gel. Equal amounts of the purified fragments were ligated using T4 DNA ligase (lane 9 and 10) and analyzed on a 0.4% agarose gel. The *in vitro* assembled ligation products of rIBV (lane 9) and N3D/6D (lane 10) were used as templates for generation of the full-length *in vitro* transcripts, which were analyzed on a 0.8% agarose gel (lane 12 and 13). Lanes 7, 8, and 11 show DNA markers.

4.2.2.2. Characterization of the N3D/6D mutant virus

Sequencing of the RT-PCR fragment covering regions containing the amino acid mutations was performed to confirm that the mutations were introduced into the viral genome. To verify whether the mutations resulted in the desired phenotype, the glycosylation status of M protein was analysed by PNGase F treatment. Cells infected by rIBV or N3D/6D mutant virus were treated with PNGase F and analyzed by Western blot with anti-M polyclonal antibody. The electrophoretic mobility of the M protein in rIBV-infected cells was changed by PNGase F digestion (Fig. 4-4b, lane 1 and 2). However, in N3D/6D-infected cells the M protein has the same mobility with or without the treatment of PNGase F (Fig. 4-4b, lane 3 and 4). A nonspecific band was detected by anti-M antibody and marked by an asterisk (Fig. 4-4a and b). These results demonstrated the N3D/6D mutant virus had the desired phenotype with the removal of N-linked glycosylation.

Further characterization of N3D/6D was carried out by analyzing the structural proteins and viral RNA. Northern blot detected similar amount and ratio of viral RNAs in cells infected with rIBV or N3D/6D mutant virus (Fig. 4-4c). Western blot analysis also showed similar pattern of S and N proteins (Fig. 4-4a). The S protein appears as the mature 180 kDa glycoprotein. A fraction of the S protein had been cleaved into two subunits, both with a molecular weight of approximately 90 kDa. The M proteins of rIBV and N3D/6D mutant showed clear difference in their electrophoretic mobilities (Fig. 4-4a).

Fig. 4-4. Characterization of mutant virus proteins and mRNAs. Vero cells were infected with N3D/6D or rIBV. (a) Western blot analysis of viral protein expression in cells infected with N3D/6D and rIBV. Vero cells infected with N3D/6D (lane 1) and rIBV (lane 2) were harvested, lysed and separated on SDS-10% polyacrylamide gel. The expression of S, N, and M proteins was analyzed by Western blot with polyclonal anti-S, anti-N, and anti-M antibodies respectively. (b) PNGase F treatment of the cell lysates from N3D/6D and rIBV infected cells. Total cell lysates prepared from N3D/6D infected Vero cells (lane 1 and 2) and rIBV infected Vero cells (lane 3 and 4) were treated either with (lanes 2 and 4) or without (lanes 1 and 3) PNGase F. Polypeptides were separated by SDS-PAGE and analyzed by Western blot using anti-IBV M antibody. Numbers on the left indicate molecular masses in kilodaltons. (c) Northern blot analysis of the genomic and subgenomic RNAs in cells infected with N3D/6D and rIBV. Ten micrograms of total RNA extracted from Vero cells infected with N3D/6D and rIBV, respectively, was separated on 1% agarose gel and transferred to a Hybond N+ membrane. Viral RNAs were probed with a DIG-labeled DNA probe corresponding to the 3' end 680 nucleotides of the IBV genome. Numbers on the right indicate the genomic and subgenomic RNA species of IBV.

4.2.2.3. Genetic stability and growth kinetics of N3D/6D mutant virus

The growth properties of the N3D/6D mutant virus on Vero cells were tested by analysis of plaque sizes and growth curves. Compared to cells infected with rIBV, the N3D/6D mutant virus displayed no obvious difference with respect to plaque size or growth property (Fig. 4-5). These results indicated that the glycosylation status of M does not influence the virus growth property in tissue culture.

The genetic stability of N3D/6D mutant viruses was tested by propagation on Vero cells for 8 passages. Western blot analysis and sequencing of the M protein of N3D/6D showed that the mutations were stable (Fig 4-6).

Fig. 4-5. Plaque morphologies and growth kinetics of the rIBV and N3D/6D viruses.

Plaque size and morphologies of rIBV and N3D/6D viruses were analysed in Vero cells. Monolayers of Vero cells on a 6-well plate were infected with 100 μ l of 1000-fold diluted virus stocks and cultured in the presence of 0.5% carboxymethyl cellulose at 37 °C for 3 days. Cells were then fixed and stained with 0.1% toluidine. To determine the growth curves of rIBV and N3D/6D viruses, Vero cells were infected with the viruses and harvested at 0, 8, 16, 24, and 32 h post-inoculation respectively. Viral stocks were prepared by freezing/thawing of the cells three times, and TCID₅₀ of each viral stock was determined by infecting five wells of Vero cells on 48-well plates with 10-fold serial dilution of each viral stock.

Fig. 4-6. Genetic stability of N3D/6D mutant virus. (a) Western blot analysis of the M protein expression in cells infected with N3D/6D and rIBV. Vero cells infected with N3D/6D (lane 1) and rIBV (lane 2) were harvested, lysates prepared and separated on SDS-10% polyacrylamide gel. The expression of M protein was analyzed by Western blot with polyclonal anti-M antibody. (b) Nucleotide sequencing of the region covering the N-linked glycosylation site mutations. Total RNA was prepared from Vero cells infected with the N3D/6D mutant virus 24 h post-infection. The 5' 400-nucleotide region covering the amino acid mutations was amplified by RT-PCR and sequenced by automated nucleotide sequencing.

4.2.2.4. Effect of mutations of N-linked glycosylation sites on the subcellular localization of M protein

As carbohydrate is important for folding, structure stability and intracellular sorting of proteins, the intracellular localization of M protein in rIBV- and N3D/6D-infected cells was examined to check whether the glycosylation status of M protein can affect the localization of M protein. It is well established that IBV M localizes to the Golgi (Swift & Machamer, 1991). Immunofluorescent microscopy studies of rIBV or N3D/6D mutant infected Vero cells stained with anti-M antibodies showed a perinuclear, dot-like staining pattern which was indicative for Golgi localization (Fig. 4-7). This indicated that the glycosylation state of M protein did not affect the normal transport and localization of the protein.

Fig. 4-7. Intracellular distribution of M in rIBV infected cells and N3D/6D infected cells. rIBV and N3D/6D infected Vero cells were fixed, permeabilized and incubated with anti-M antibody. Bound antibody was detected using a FITC-labelled anti-rabbit antibody.

4.2.2.5. Analysis of the N3D/6D mutant virus by electron microscopy

The oligosaccharides on the coronavirus M protein constitute a hydrophilic cover on the surface of the virion. In order to determine whether removal of the oligosaccharides has any effect on virion morphology, viral particles in the extracellular media were pelleted through 30% sucrose cushion and analysed by electron microscopy. As shown in Fig. 4-8a and b, the N3D/6D particles had the same appearance as rIBV particles. Both N3D/6D and rIBV particles showed heterogeneity in virion morphology. This might be attributed to the distorting effects of negative-staining procedures or the existence of particles lacking the nucleocapsid.

Fig. 4-8. Electron microscopic analysis. Viruses released from rIBV (a) and N3D/6D (b) were purified by pelleting through a 30% sucrose cushion, negatively stained and viewed by electron microscopy.

4.2.2.6. Sedimentation analysis of particles in sucrose gradients

IBV particles are heterogeneous in sucrose gradient and they sediment in a broad band of density (Collins *et al.*, 1976; Macnaughton & Davies, 1980). N3D/6D and rIBV were compared by ultracentrifugation in 20-60% discontinuous sucrose gradients. Fractions were collected and analysed by Western blot with anti-M antibody. The sedimentation profiles were similar. Most of N3D/6D and rIBV particles were detected in fractions 8 and 9 (Fig. 4-9). As shown in Fig. 4-9, most of the M proteins in rIBV displayed slow migration which represented the glycosylated form.

Fig. 4-9. Sedimental analysis of rIBV and N3D/6D. Vero cells were infected with rIBV and N3D/6D. Viral particles in the media were pelleted through 30% sucrose cushion and then sedimented into a discontinuous sucrose gradient consisting of 20, 30, 40, 50 and 60% sucrose 2 ml each in TNE buffer. Fractions, 1 ml each, were collected from the top to the bottom for analysis by Western blot with anti-M antibody.

4.2.3. Systematic analysis of the N-terminal domain of the IBV M protein

4.2.3.1. Construction of full-length IBV clones with mutations in the N-terminal domain and recovery of infectious mutant viruses

To understand the roles of M protein N-terminal mutations in the virus assembly and replication, mutations were introduced into the genome of IBV using full-length infectious clone of IBV. For this purpose a set of mutants was constructed with various mutations in this domain. Mutant GFP48 has a large internal substitution, replacing residues N₃ through F₁₈ with the same length of the N-terminal sequence of GFP protein (Fig. 4-1). Mutants GFP15L, GFP15M and GFP18R have smaller substitutions ranging from N₃ to C₇, T₈ to E₁₂ and Q₁₃ to F₁₈ with sequences corresponding to the GFP48 respectively (Fig. 4-1). Based on the alignment of four different coronavirus M proteins, the conserved amino acids were chosen for mutation (Fig. 4-10). In mutant C7A, the cysteine at position 7 was replaced by alanine, while in mutant L9A, the leucine at position 9 was replaced by alanine (Fig. 4-1). In mutant Q16A, the glutamine at position 16 was replaced by alanine (Fig. 4-1). Mutant V15A/Q16A has replacement of valine and glutamine at position 15 and 16 with alanine (Fig. 4-1). Mutant F18A has replacement of phenylalanine at position 18 with alanine, while mutant F18A/K19A has replacement of phenylalanine and lysine at position 18 and 19 with alanine (Fig. 4-1).

These mutations were first introduced into an infectious clone of IBV and *in vitro* transcribed into full-length RNA molecules using the T7 RNA polymerase in the presence of a cap analog. The RNA transcripts were electroporated into Vero cells

together with RNA transcripts of N protein. After three to four days post-electroporation, the mutant viruses C7A, L9A, Q16A, V15A/Q16A and F18A were recovered and confirmed by RT-PCR sequencing. In contrast, some substitutions in this domain had detrimental effect on the recovery of mutant viruses F18A/K19A, GFP15L, GFP15M, GFP18R and GFP48. For GFP15L mutant, a large fusion foci was observed at 48 h post-electroporation, however this virus failed to infect neighboring cells at 60 h post-electroporation and could not be recovered (Fig. 4-11).

Fig. 4-10. Alignment of the N-terminal sequences of four membrane proteins encoded by various coronaviruses. The proposed first transmembrane domains are underlined. Identical residues are highlighted in yellow. Conserved residues are highlighted in blue. Similar residues are highlighted in green.

Fig. 4-11. Fusion foci of Vero cells electroporated by transcripts of GFPL15 and N. Vero cells were electroporated with the *in vitro* synthesized transcripts derived from an *in vitro* assembled full-length clone (GFPL15+N). One large fusion foci was observed at 48 h post-electroporation. At 60 h post-electroporation, the cells which formed this fusion foci detached from the dish. The mutant virus failed to propagate to neighboring cells and could not be rescued.

4.2.3.2. Growth properties of the mutant viruses

The growth properties of the C7A, L9A, Q16A, V15A/Q16A and F18A mutant viruses on Vero cells were tested by analysis of plaque size and growth curves. Compared to cells infected with rIBV, the mutant virus C7A, L9A, Q16A, V15A/Q16A displayed no obvious difference with respect to plaque sizes and growth properties (Fig. 4-12). However, in L9A and F18A infected-cells smaller-sized plaques were observed (Fig. 4-12a). Analysis of growth curves of F18A showed that it replicated to titers about one log lower than rIBV, while L9A showed a slight delay of growth (Fig. 4-12b).

Fig. 4-12. Plaque morphologies and growth kinetics of rIBV, C7A, L9A, Q16A, V15A/Q16A and F18A viruses. (a) Plaque size and morphologies of rIBV, C7A, L9A, Q16A, V15A/Q16A and F18A viruses were analysed in Vero cells. Monolayers of Vero cells on a 6-well plate were infected with 100 μ l of 1000-fold diluted virus stock and cultured in the presence of 0.5% carboxymethyl cellulose at 37 °C for 3 days. The cells were fixed and stained with 0.1% toluidine. (b) To determine the growth curves of rIBV, C7A, L9A, Q16A, V15A/Q16A and F18A viruses, Vero cells were infected with the viruses and harvested at 0, 8, 16, 24, and 32 h postinoculation, respectively. Viral stocks were prepared by freezing/thawing of the cells three times, and TCID₅₀ of each viral stock was determined by infecting five wells of Vero cells on 48-well plates with 10-fold serial dilution of each viral stock.

4.2.3.3. Effect of the N-terminal domain mutations on VLP assembly

Next, the question is whether VLP formation is sensitive to the N-terminal domain mutations of M protein. Each mutant protein was co-expressed with WT E protein. The proteins were expressed under the control of T7 promoter using the vaccinia recombinant vTF7-3 that expresses the T7 polymerase (Fuerst *et al.*, 1986). At 12 h post-transfection the media was removed from cells and clarified by centrifugation. VLPs were isolated from clarified media by centrifugation through a sucrose cushion. Cell lysates and extracellular VLPs were analyzed by SDS-PAGE and Western blotting (Fig. 4-13). All of the mutants could form the VLPs as shown by the presence of extracellular M. This indicated that the N-terminal domain of M is not important for the VLP formation. When expressed alone, a small amount of M could be detected in the media. This could result from the unclarified cell debris. Interestingly, only the unglycosylated M and mutants could be detected in the VLPs (Fig. 4-13).

To determine whether the N or S can enhance the incorporation of glycosylated M into the VLPs, VLPs which were produced by co-expression of S, M, E and N in four different combination were pelleted and analyzed by Western blot (Fig. 4-14). As shown in Fig. 4-14, the expression level of M protein in cell lysates were similar. It is clear from Fig. 4-14 that neither S nor N could enhance the incorporation of glycosylated M into VLPs. When co-expressing all these four proteins S, M, E and N, it was still found that most of the M proteins assembled into the VLPs were unglycosylated form (Fig. 4-14).

Fig. 4-13. Effects of mutations in the N-terminal domain of M protein on VLP assembly. VLPs were produced in COS-7 cells using vaccinia recombinant vTF7-3 that expresses T7 RNA polymerase. Following infection cells were transfected with plasmids containing WT, C7A, L9A, Q16A, V15A/Q16A, F18A, F18A/K19A, N3D/6D, GFP48, GFP15L, GFP15M and GFP18R M genes in combination with WT E gene. Single expression of M protein was taken as a negative control. Intracellular cell lysates and pelleted extracellular VLPs were analyzed by Western blot with anti-M antibody.

Fig. 4-14. Co-expression of the M gene with S, E and N genes in various combinations and analysis of structure proteins in VLPs. Recombinant vaccinia virus vTF7-3-infected COS-7 cells were transfected with various combinations of plasmids containing IBV genes S, E, M or N. At 12 h post-transfection, the cells were lysed by SDS loading buffer. The cell lysates were analyzed by Western blot with polyclonal anti-S, anti-N and anti-M antibodies. The medium was removed from the cells, clarified by centrifugation. The VLPs were pelleted by ultracentrifugation through sucrose gradient. The sedimented materials were resuspended in SDS loading buffer and then analyzed by Western blot with polyclonal anti-S, anti-N and anti-M antibodies.

4.2.3.4. Analysis of negative-strand RNA replication and subgenomic RNA transcription in cells transfected with wild-type and mutant full-length transcripts

As no infectious virus was recovered from cells transfected with F18A/K19A, GFP15L, GFP15M, GFP18R and GFP48 mutant transcripts, total RNA was extracted from cells electroporated with wild-type and mutant full-length transcripts and RT-PCR amplification of subgenomic mRNAs was carried out to check whether a low level of RNA replication and subgenomic mRNA transcription occurred in cells transfected with the mutant transcripts. The forward primer used in this reaction corresponds to the leader sequence from nucleotides 26–46 in the genomic RNA and the downstream primers covers IBV sequences from nucleotides 24 784 to 24 803. If transcription of subgenomic mRNAs did occur, a 415 bp PCR product corresponding to the 5'-terminal region of the subgenomic mRNA4 would be expected. As shown in Fig. 4-15, a 415 bp band was observed in cells electroporated with wild-type full-length transcripts at 2 days post-electroporation. The same PCR products were not detected in cells electroporated with the GFP48, GFPR18 and F18A/K19A mutant transcripts, respectively, at 2 days post-electroporation (Fig. 4-15). In cells electroporated with GFPL15 and GFPM15 mutant transcripts at 2 days post-electroporation, RT-PCR amplification of the subgenomic mRNA 4 detected the 415 bp bands (Fig. 4-15).

RT-PCR amplification of the negative strand RNA was performed to check if RNA replication occurred in these transfected cells. The primer pair was chosen so

that the IBV sequence from nucleotides 14 931 to 15 600 would be amplified by the RT-PCR. If replication of viral RNA occurred, a 670 bp PCR fragment would be expected. As shown in Fig. 4-15, RT-PCR fragments amplified from negative strand RNA templates were obtained from cells transfected with the mutant transcripts. As a positive control, the same RT-PCR fragment for the negative strand RNA was observed in cells transfected with wild-type transcripts (Fig. 4-15). These results confirmed that transcription of the negative strand RNA has taken place in cells transfected with the mutant transcripts.

Fig. 4-15. Analysis of negative strand RNA replication and subgenomic RNA transcription in cells transfected with wild-type and mutant full-length transcripts. To detect subgenomic RNA synthesis in cells electroporated with rIBV, GFP48, GFPL15, GFPM15, GFPR18 and F18A/K19A transcripts, total RNA was prepared from Vero cells electroporated with *in vitro* synthesized full-length transcripts 2 days post-electroporation. Region corresponding to the 5'-terminal 415 nucleotides of the subgenomic mRNA4 was amplified by RT-PCR and analyzed on 1.5% agarose gel. Lane 7 shows DNA markers and numbers on the right indicate the length of DNA in bases. To detect negative strand RNA synthesis in cells electroporated with rIBV, GFP48, GFPL15, GFPM15, GFPR18 and F18A/K19A transcripts, total RNA was prepared from Vero cells electroporated with *in vitro* synthesized full-length transcripts 2 days post-electroporation. Regions corresponding to nucleotides 14 931–15 600 of the negative sense IBV RNA were amplified by RT-PCR and analyzed on 1.5% agarose gel. Lane 1 shows DNA markers. Numbers on the left indicate nucleotides in bases.

4.3. Summary

Coronavirus M protein, the most abundant structural protein, plays a predominant role in the formation of virus particles. The N-linked glycosylation is a conserved characteristic for group 1 and group 3 coronaviruses. In this report, it was revealed that N-linked glycosylation of IBV M protein is not required for virus assembly. Mutations in the two glycosylation sites were incorporated into the IBV genome to get the N3D/6D mutant virus. Comparison between N3D/6D and recombinant wild type virus (rIBV) showed no difference in tissue culture growth characteristics, virion morphology and heterogeneity. Mutations in the luminal domain had no effect on the formation of VLPs. However, some of these mutations in this domain had severe effects on the infectivity of viral RNA to Vero cells. These observations suggested that N-linked glycosylation is not important for virus assembly, but some other mutations in the luminal domain had effects on it.

Chapter 5

Discussion and Future Direction

5.1. Characterization of the E protein from SARS-CoV and IBV

5.1.1. Mixed membrane topologies of the SARS-CoV E protein

Coronaviruses encode a small integral membrane protein that is associated with the viral envelope and plays important functions in virion assembly and morphogenesis (Vennema *et al.*, 1996; Liao *et al.*, 2004a). Recently, the E protein from SARS-CoV and MHV was shown to enhance the membrane permeability of bacterial and mammalian cells to small molecules (Wilson *et al.*, 2004; Liao *et al.*, 2004b; Liao *et al.*, 2006). SARS-CoV E protein contains a putative long transmembrane domain of 29 amino acid residues. Based on the distribution of the charged amino acids flanking the transmembrane domain and the recent observation that the E protein is palmitoylated on all three cysteine residues (Liao *et al.*, 2006), the protein may insert into the ER and Golgi membranes with an N_{exo(lum)}C_{cyto} orientation. *In vitro* results reported by Arbely *et al.* suggest that the transmembrane domain of SARS-CoV may adopt a highly unusual topology, consisting of a short transmembrane helical hairpin that forms an inversion about a previously unidentified pseudo-centre of symmetry (Arbely *et al.*, 2004), although this is in contrast with simulation results based on phylogenetic data (Torres *et al.*, 2005).

5.1.1.1. An N_{cyto}C_{cyto} topology adopted by majority of the SARS-CoV E

In the present study, the membrane topology of the SARS-CoV E protein is systematically studied with N-terminus and C-terminus Flag tagged SARS-CoV E protein. As SARS-CoV E protein with the Flag epitope tag retained the activities to

produce E protein vesicles and permeabilize mammalian cells (data not shown), it was assumed that the topology of tagged SARS-CoV E proteins in the membrane was the same as that of wild-type SARS-CoV E protein and I ventured into the extended experiments. Immunofluorescent staining of cells differentially permeabilized with detergents and proteinase K protection assay revealed that both the N- and C-termini of the SARS-CoV E protein were exposed to the cytoplasmic side of the membranes ($N_{\text{cyto}}C_{\text{cyto}}$) (Fig. 5-1a). Consistent with previous studies, parallel experimental data presented in this study demonstrated that the IBV E protein spans the membranes once with the N-terminus exposed lumenally and the C-terminus exposed cytoplasmically ($N_{\text{exo(lum)}}C_{\text{cyto}}$) (Fig. 5-1a).

In the limited proteinase digestion assay, a proteinase K-resistant fragment of PDI was consistently observed. The same fragment as well as some smaller fragments was observed when PDI expressed in *E. coli* was treated with proteinase K (Koivunen et al., 2005), confirming that it represents a proteinase K-resistant fragment of PDI. The smaller fragments were not detected in this study may be due to the fact that the antibody used was raised only against the middle region of PDI from amino acid 211-370.

Fig. 5-1. Proposed membrane topologies of the unglycosylated form of the SARS-CoV E (a), the IBV E (a), and the glycosylated form of the SARSCoV E protein (b). N and C represent the N- and C-termini of the proteins, respectively. The two potential forms of SARS-E when both the N- and C-terminal ends are located in the same side of the membrane are indicated by (1) and (2). The Asn-linked high mannose carbohydrate modification is shown at the C terminus of the two proposed forms by stars.

5.1.1.2. A small proportion of the SARS-CoV E protein with the C-terminus exposed to the luminal side

Interestingly, a small proportion of the SARS-CoV E protein was shown to be modified by N-linked glycosylation on the asparagine 66 residue. Glycosylation of the C-terminal region of the E protein was an unexpected finding, since no proteinase K-protected fragment was detected in the limited proteinase K assay. Nevertheless, the detection of the N-linked glycosylation at the C-terminal region and the inefficient synthesis of the glycosylated form suggest that the SARS-CoV E protein may have an alternative membrane topology. Two possible models for this alternative membrane topology of the SARS-CoV E protein were proposed (Fig. 5-1b). In the first case, the protein may span the membrane once with an $N_{\text{cyto}}C_{\text{exo(lum)}}$ orientation (Fig. 5-1b). Alternatively, both N- and C-termini may be exposed to the luminal side of the membranes ($N_{\text{exo(lum)}}C_{\text{exo(lum)}}$) (Fig. 5-1b). Once again, two potential forms are proposed (Fig. 5-1b, form (1) and (2)). At present, it has not sorted out which form is more likely represents the membrane topology for this portion of the E protein, as neither the full-length nor a shorter form of the protein was detected in the limited proteinase K assay probably due to the reason that only a very small proportion of protein assumes this topology.

It is therefore established that the majority of the SARS-CoV E protein is inserted into cellular membranes with an $N_{\text{cyto}}C_{\text{cyto}}$ topology, but a small proportion of the protein is modified by N-linked glycosylation and inserted into the membranes with the C-terminus exposed to the luminal side. There are a few examples of viral

proteins that adopt more than one membrane topology. These include the TGEV membrane protein (Escors *et al.*, 2001), the fusion protein F from the Newcastle disease virus (McGinnes *et al.*, 2003) and the adenovirus E3, a 6.7K protein (Moise *et al.*, 2004). As coronavirus E protein including SARS-CoV E protein is a multi-functional protein, it would be important to establish whether different topological forms are responsible for distinct functions. The SARS-CoV E protein was found to interact with Bcl-xL (Yang *et al.*, 2005). As this interaction was mediated by the C-terminal BH₃-like region of the E protein, it would require that the C-terminal part of the protein is exposed to the cytoplasm (Yang *et al.*, 2005). The C-terminal tail of the IBV E protein is important for its interaction with the M protein during virus budding (Corse & Machamer, 2003). It is therefore likely that the N_{cyto}C_{cyto} form of the E protein may be the form that accomplishes these functions.

5.1.1.3. Cell surface expression of the SARS-CoV E protein

Most coronavirus E proteins could be translocated to the cell surface to facilitate budding and release of progeny viruses (Smith *et al.*, 1990; Liu & Inglis, 1991; Lim & Liu, 2001; Maeda *et al.*, 2001). Recently, the E protein of SARS-CoV and MHV was found to modify membrane permeability, allowing entry of small molecules into cells and leading to cell lysis (Wilson *et al.*, 2004; Liao *et al.*, 2004b; Madan *et al.*, 2005; Liao *et al.*, 2006). As cell surface expression of E protein is essential for this function, it is quite certain that SARS-CoV must be translocated to the cell surface. In this study, however, efficient cell surface expression of the SARS-CoV E protein was not

detected in cells expressing the protein. This might be explained by the conclusion that the majority of the SARS-CoV E protein adopts an $N_{\text{cyto}}C_{\text{cyto}}$ topology in cells. Interestingly, a few fluorescent dots were consistently observed in nonpermeabilized cells expressing the protein. It is currently unclear if these represent the minor proportion of the SARS-CoV E protein that expresses on the cell surface and with both N- and C-termini located outside of the cells.

5.1.2. N-linked glycosylation of SARS-CoV E protein

Two of the three coronavirus membrane-associated structural proteins, M and S, are posttranslationally modified by either N- or O-linked glycosylation (Nal *et al.*, 2005). The fourth membrane-associated structural protein, the hemagglutinin-esterase (HE), in some coronaviruses is also a glycoprotein (Nal *et al.*, 2005). So far, the E protein is an apparent exception. In this study, it was shown that the SARS-CoV E protein is also modified by N-linked glycosylation. In N-linked glycosylation, the oligosaccharides are added to specific asparagine residues in the consensus sequence Asn-X-Ser/Thr. Once in oxidizing environment of the ER lumen, the protein would fold and allow the attachment of carbohydrate moieties where N-linked glycosylation sites are accessible to the luminal compartment. The minimum distance between a functional C-terminal glycosylation acceptor site and the luminal end of the transmembrane domain is 12 to 13 residues (Nilsson & Von, 1993), suggesting that only the most C-proximal asparagine residue out of the two potential sites in the SARS-CoV E protein can be glycosylated. Consistent with this observation,

mutagenesis studies present in this paper confirmed that only N66 is modified. Furthermore, it was also shown that a potential N-linked glycosylation site in the N-terminal region of the IBV E protein (N6), which is 6 amino acids upstream of the transmembrane domain, was not modified, although the N-terminal region of this protein is exposed to the luminal side of the membranes.

N-linked glycosylation affects most of the proteins present on the surface of the enveloped viruses. For this reason, it is likely to play a major role in the stability, antigenicity and other biological functions of the modified viral envelope proteins. In the early secretory pathway, the glycans also play a role in protein folding, quality control and certain sorting events. However, N-linked glycosylation seems not to affect the membrane-association and membrane-permeabilizing activity of the SARS-CoV E protein (Liao *et al.*, 2006). To further characterize the functional significance of this modification, it would be interesting to test if N-linked glycosylation of SARS-CoV E protein can also be detected in virus-infected cells and in purified virions. As the N66 residue is not conserved in the E protein of other coronaviruses, it is likely that this modification may render unique features to the SARS-CoV E protein. Further exploration of the functions and effects of this modification on SARS-CoV E protein using an infectious cloning system may shed new light on the molecular biology and pathogenesis of SARS-CoV.

5.1.3. Partial association of SARS-CoV E and IBV E proteins with lipid rafts

As observed in this study, the SARS-CoV E protein and IBV E protein are partially associated with lipid rafts. Lipid rafts are liquid-ordered, plasma membrane microdomains that are enriched in sphingolipids and cholesterol (Simons & Ikonen, 1997). The formation and structural integrity of these microdomains are dependent upon the presence of cholesterol. At low temperature, lipid rafts resist extraction with nonionic detergents and can be isolated on density gradients as a low-density fraction (Simons & Ikonen, 1997; Hooper, 1999). Rafts act as platforms enabling a sufficient quantity of protein to accumulate, enabling their functions or to promoting specific protein-protein interactions. Rafts also play roles in apical sorting in polarized cells (Keller & Simons, 1998), T cell activation (Langlet et al., 2000), Ras signaling (Prior et al., 2001) and the assembly of many enveloped viruses (Suomalainen, 2002). Further work will determine the requirement of palmitoylation, transmembrane domain, and cytoplasmic tail for the lipid association of SARS-CoV E protein and IBV E protein.

5.1.4. Interaction of SARS-CoV E protein with snapin

In a yeast two-hybrid screen for cellular proteins that interact with the SARS-CoV E protein, snapin was identified as a novel interaction partner of SARS-CoV E protein. The SARS-CoV E-snapin interaction was further confirmed by GST pull-down assay and coimmunoprecipitation.

Snapin was originally identified as a SNAP25 interacting protein and described to be an exclusively neuronal protein that regulates the association of the calcium

sensor synaptotagmin with the SNARE complex (Ilardi et al., 1999). More recently, snapin has been described as a non-neuronal protein interacting with the ubiquitously expressed SNAP23 (Buxton et al., 2003) and as a component of the biogenesis of lysosome-related complex-1 (BLOC-1) (Starcevic & Dell'Angelica, 2004). The interaction of snapin with SARS-CoV E protein might be involved in the virus particle transport processes. Further work will determine if this interaction is related to the secretion of VLPs.

5.2. Characterization of the N-terminal domain of IBV M protein

5.2.1. N-linked glycosylation of IBV M protein is not required for virus assembly in cultured cells

Coronavirus M proteins are invariably glycosylated by N-linked or O-linked oligosaccharides. The function of this modification has not been identified clearly. The presence of N-linked glycosylation on TGEV M protein is important for an efficient induction of IFN- α by TGEV (Laude *et al.*, 1992). A mutant of TGEV in which the N-linked glycosylation site of M protein was disrupted was isolated (Laude *et al.*, 1992). The glycosylation defect has no obvious consequences on the viral replication and infectivity. For MHV, the recombinant virus with O-linked glycosylation site mutation was recovered by target recombination (Masters, 1999). This recombinant virus appeared to be a poor interferon inducer in virus-infected cells (de Haan *et al.*, 2003a). However, the *in vivo* interferogenic capability of MHV was not affected by the glycosylation status (de Haan *et al.*, 2003a). Our results showed

that IBV M protein has N-linked glycosylation sites on asparagine 3 and 6. To identify the function of this modification, the asparagines 3 and 6 to aspartic acid mutations were introduced into the IBV genome by full-length infectious cDNA clone of IBV. The mutant virus N3D/6D was recovered in Vero cells.

When compared with rIBV, the N3D/6D mutant virus exhibited similar growth properties and plaque phenotypes. The results indicated that glycosylation of the IBV M protein is not important for virus assembly and replication. After 8 passages, the mutations were stably maintained. As glycosylation has been shown to be important for the intracellular sorting of proteins, the localization of M protein of N3D/6D was characterized. The M in N3D/6D infected cells localized like the wild type M protein in the Golgi. Meanwhile, no effects of glycosylation on the virion morphology and sedimental properties were found.

However the conservation of either the N- or O-linked glycosylation of coronavirus M protein suggested that these modifications are beneficial to their respective virus. As glycosylation is not involved in virus assembly or replication, it might play roles in the virus-host interaction. Here the function of N-linked glycosylation of IBV M protein was not ruled out by the *in vitro* experiments, characterization of their *in vivo* replication might shed light on it.

5.2.2. Mutations in the N-terminal domain of IBV M protein do not affect the formation of VLP

The N-terminal domain of the coronavirus M protein is exposed lumenally.

Substitution of the N-terminal domain from residues 3 to 7, 8 to 12, 13 to 18 and 3 to 18 with GFP sequences did not affect VLP assembly. Single or double mutations such as N3D/6D, C7A, L9A, Q16A, V15A/Q16A and F18A also showed no effect on VLP assembly. Only mutant M protein F18A/K19A showed a slightly lower secretion of the VLPs. F18A/K19A mutant contains mutations that change the phenylalanine on residue 18 to alanine and lysine on 19 to alanine. Lysine 19 is close to the first transmembrane domain of IBV M which is predicted to start from tyrosine 21 (Fig. 4-10). Translocation of F18A/K19A might not be affected because the ratio of glycosylated versus unglycosylated form of F18A/K19A in the cell lysates was similar to that of wild type IBV M. Altogether these results demonstrated that the N-terminal domain of IBV M protein may not be required for VLP formation. For MHV, deletion of the middle part of the ectodomain or some substitutions in this domain had severe effects on the VLP formation (de Haan *et al.*, 1998). It was suggested that these N-terminal MHV M mutants might impair lateral interactions causing the budding defect.

It was interesting to observe that most of the IBV M protein incorporated into the VLPs was the unglycosylated form. However the M protein in the virion of IBV is glycosylated as shown in the sedimentation analysis (Fig. 4-9). The difference of the M protein glycosylation status between the VLPs and IBV virions indicated that the mechanisms of the VLP assembly and viral assembly might not be exactly the same. The discrepancy between the VLP assembly and viral assembly was also found in the MHV M protein C-terminal mutant (de Haan *et al.*, 1998).

5.2.3. Mutations in the N-terminal domain of IBV M protein had effects on the recovery of infectious virus

The importance of the N-terminal domain of IBV M was examined by introducing mutations in this region into IBV genome. It was shown that most of the single or double mutations did not have effects on virus replication. In contrast, the inability to rescue the GFP48, GFPL15, GFPM15, GFPR18 and F18A/K19A mutants demonstrated that swap of the middle part of the ectodomain or substitution close to the first transmembrane domain had severe effects. Studies on subgenomic RNA synthesis indicated that mutations close to the first transmembrane domain had more severe effects on virus replication. Slower growth kinetics of F18A mutant virus and failure to recover the F18A/K19A mutant virus also indicated that residues close to the first transmembrane domain are important.

There are some possibilities for the detrimental effects of these middle region swap mutations on viral life cycles. First, these mutations might result in impaired interactions with other structure proteins causing the budding defect. As changes in the N-terminal domain of M do not affect the VLP formation and the lumenally exposed N-terminal domain of M protein can not interact with N protein which is localized in the cytoplasm, it is speculated that the mutations affect the ability of the M protein to associate with the S protein. Second, the N-terminal domain of M might contain *cis*-acting RNA elements that affect the transcription efficiency. M protein is produced by subgenomic mRNA 4. The TRS of subgenomic mRNA 4 is about 100 bp upstream of the N-terminal domain of M. Alteration of flanking sequences of the TRS

might suppress subgenomic mRNA 4 transcription. Third, as the ORFs of M and E overlap 29 bp, these middle region swap mutations also change an arginine into a stop codon and result in a mutant E protein short of the last six residues (RDKLYS). These six residues were identified to be an ER retention signal previously (Lim & Liu, 2001). However, deletion of the last 16 or 25 residues of the E protein resulted in no phenotypic change of the recovered viruses (unpublished data).

5.3. Future Direction

5.3.1. *Is IBV E protein essential for virus replication*

It is still unknown whether the IBV E protein is essential for virus replication. A mutant MHV with deletion of the E gene was recovered by target recombination (Kuo & Masters, 2003). This mutant virus produces tiny plaques and shows slow growth rate and low titer. Alanine scanning insertion mutagenesis of the hydrophobic domain of the MHV E protein suggested that positioning of polar hydrophilic residues within the predicted transmembrane domain are important for virus production (Ye & Hogue, 2007). More recently, a recombinant SARS-CoV that lacks the E gene was rescued in Vero E6 cells, and the recovered deletion mutant was attenuated *in vitro* and in the hamster model (DeDiego *et al.*, 2007). In contrast for group 1 TGEV, deletion of E protein is lethal as shown in two independent reverse genetic experiments (Curtis *et al.*, 2002; Ortego *et al.*, 2002).

To investigate the effects of IBV E protein on the assembly and maturation of virions, a series of mutations on IBV E protein will be introduced into the IBV genome. These mutants are designed by replacing the respective residues with a stop

codon. The ability of virus recovery and the phenotype changes of the recovered mutant viruses will be investigated.

5.3.2. Is palmitoylation of E protein necessary for its association with lipid rafts

IBV and SARS-CoV E proteins have been shown to be palmitoylated on cysteine residues (Corse & Machamer, 2002; Liao *et al.*, 2006). Palmitoylation is a common posttranslational modification that may increase partitioning into lipid rafts (Arni *et al.*, 1998). As IBV and SARS-CoV E proteins were shown to be partially associated with lipid rafts, it is reasonable to speculate that the palmitoylation of E protein might determine its lipid rafts association. Whether E protein is associated with lipid rafts in the presence of 2-bromopalmitate, a specific inhibitor of protein palmitoylation (Webb *et al.*, 2000), can be determined by isolation of low-density detergent-insoluble membrane fraction on flotation gradients. The lipid rafts association of E proteins with cysteine residue mutations can also be determined.

5.3.3. Involvement of N-linked glycosylation of IBV M in ER stress

Viral infection induces ER stress. IBV infection also can induces ER stress as our unpublished microarray data showed that IBV infection could induce the expression of growth arrest and DNA damage-inducible protein 34 (GADD34) at the mRNA level. GADD34 is an ER stress downstream gene in the PKR-like ER kinase (PERK) pathway. Interaction of BiP with misfolded or unfolded viral proteins, most probably viral glycoproteins, triggers the ER stress response. This renders the M and S proteins two possible candidates for triggering ER stress. The N3D/6D M

glycosylation mutant virus and rIBV are intended to be compared for their inducibility of ER stress.

Reference List

- Abraham S, Kienzle TE, Lapps WE & Brian DA (1990) Sequence and expression analysis of potential nonstructural proteins of 4.9, 4.8, 12.7, and 9.5 kDa encoded between the spike and membrane protein genes of the bovine coronavirus. *Virology* **177**, 488-495.
- Almazan F, Gonzalez JM, Penzes Z, Izeta A, Calvo E, Plana-Duran J & Enjuanes L (2000) Engineering the largest RNA virus genome as an infectious bacterial artificial chromosome. *Proc Natl Acad Sci U S A* **97**, 5516-5521.
- Alonso S, Sola I, Teifke JP, Reimann I, Izeta A, Balasch M, Plana-Duran J, Moormann RJ & Enjuanes L (2002) In vitro and in vivo expression of foreign genes by transmissible gastroenteritis coronavirus-derived minigenomes. *J Gen Virol* **83**, 567-579.
- An S, Chen CJ, Yu X, Leibowitz JL & Makino S (1999) Induction of apoptosis in murine coronavirus-infected cultured cells and demonstration of E protein as an apoptosis inducer. *J Virol* **73**, 7853-7859.
- Anand K, Palm GJ, Mesters JR, Siddell SG, Ziebuhr J & Hilgenfeld R (2002) Structure of coronavirus main proteinase reveals combination of a chymotrypsin fold with an extra alpha-helical domain. *EMBO J* **21**, 3213-3224.
- Anand K, Ziebuhr J, Wadhwani P, Mesters JR & Hilgenfeld R (2003) Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS drugs. *Science* **300**, 1763-1767.
- Arbely E, Khattari Z, Brotons G, Akkawi M, Salditt T & Arkin IT (2004) A highly unusual palindromic transmembrane helical hairpin formed by SARS coronavirus E protein. *J Mol Biol* **341**, 769-779.
- Arni S, Keilbaugh SA, Ostermeyer AG & Brown DA (1998) Association of GAP-43 with detergent-resistant membranes requires two palmitoylated cysteine residues. *J Biol Chem* **273**, 28478-28485.
- Baric RS, Fu K, Schaad MC & Stohlman SA (1990) Establishing a genetic recombination map for murine coronavirus strain A59 complementation groups. *Virology* **177**, 646-656.
- Baric RS, Nelson GW, Fleming JO, Deans RJ, Keck JG, Casteel N & Stohlman SA (1988) Interactions between coronavirus nucleocapsid protein and viral RNAs: implications for viral transcription. *J Virol* **62**, 4280-4287.
- Baric RS & Yount B (2000) Subgenomic negative-strand RNA function during mouse

- hepatitis virus infection. *J Virol* **74**, 4039-4046.
- Baudoux P, Carrat C, Besnardeau L, Charley B & Laude H (1998) Coronavirus pseudoparticles formed with recombinant M and E proteins induce alpha interferon synthesis by leukocytes. *J Virol* **72**, 8636-8643.
- Bhardwaj K, Guarino L & Kao CC (2004) The severe acute respiratory syndrome coronavirus Nsp15 protein is an endoribonuclease that prefers manganese as a cofactor. *J Virol* **78**, 12218-12224.
- Bournsell ME, Binns MM & Brown TD (1985) Sequencing of coronavirus IBV genomic RNA: three open reading frames in the 5' 'unique' region of mRNA D. *J Gen Virol* **66** (Pt 10), 2253-2258.
- Bournsell ME, Brown TD, Foulds IJ, Green PF, Tomley FM & Binns MM (1987) Completion of the sequence of the genome of the coronavirus avian infectious bronchitis virus. *J Gen Virol* **68** (Pt 1), 57-77.
- Bredenbeek PJ, Pachuk CJ, Noten AF, Charite J, Luytjes W, Weiss SR & Spaan WJ (1990) The primary structure and expression of the second open reading frame of the polymerase gene of the coronavirus MHV-A59; a highly conserved polymerase is expressed by an efficient ribosomal frameshifting mechanism. *Nucleic Acids Res* **18**, 1825-1832.
- Brierley I, Bournsell ME, Binns MM, Bilimoria B, Blok VC, Brown TD & Inglis SC (1987) An efficient ribosomal frame-shifting signal in the polymerase-encoding region of the coronavirus IBV. *EMBO J* **6**, 3779-3785.
- Brierley I, Bournsell ME, Binns MM, Bilimoria B, Rolley NJ, Brown TD & Inglis SC (1990) Products of the polymerase-encoding region of the coronavirus IBV. *Adv Exp Med Biol* **276**, 275-281.
- Brierley I, Digard P & Inglis SC (1989) Characterization of an efficient coronavirus ribosomal frameshifting signal: requirement for an RNA pseudoknot. *Cell* **57**, 537-547.
- Budzilowicz CJ & Weiss SR (1987) In vitro synthesis of two polypeptides from a nonstructural gene of coronavirus mouse hepatitis virus strain A59. *Virology* **157**, 509-515.
- Budzilowicz CJ, Wilczynski SP & Weiss SR (1985) Three intergenic regions of coronavirus mouse hepatitis virus strain A59 genome RNA contain a common nucleotide sequence that is homologous to the 3' end of the viral mRNA leader sequence. *J Virol* **53**, 834-840.
- Buxton P, Zhang XM, Walsh B, Sriratana A, Schenberg I, Manickam E & Rowe T (2003) Identification and characterization of Snapin as a ubiquitously

- expressed SNARE-binding protein that interacts with SNAP23 in non-neuronal cells. *Biochem J* **375**, 433-440.
- Casais R, Davies M, Cavanagh D & Britton P (2005) Gene 5 of the avian coronavirus infectious bronchitis virus is not essential for replication. *J Virol* **79**, 8065-8078.
- Casais R, Thiel V, Siddell SG, Cavanagh D & Britton P (2001) Reverse genetics system for the avian coronavirus infectious bronchitis virus. *J Virol* **75**, 12359-12369.
- Cavanagh D (1983) Coronavirus IBV glycopolypeptides: size of their polypeptide moieties and nature of their oligosaccharides. *J Gen Virol* **64**, 1187-1191.
- Cavanagh D (1997) Nidovirales: a new order comprising Coronaviridae and Arteriviridae. *Arch Virol* **142**, 629-633.
- Cavanagh D & Davis PJ (1988) Evolution of avian coronavirus IBV: sequence of the matrix glycoprotein gene and intergenic region of several serotypes. *J Gen Virol* **69** (Pt 3), 621-629.
- Cavanagh D, Davis PJ & Pappin DJ (1986) Coronavirus IBV glycopolypeptides: locational studies using proteases and saponin, a membrane permeabilizer. *Virus Res* **4**, 145-156.
- Cavanagh D & Horzinek MC (1993) Genus Torovirus assigned to the Coronaviridae. *Arch Virol* **128**, 395-396.
- Chang RY & Brian DA (1996) cis Requirement for N-specific protein sequence in bovine coronavirus defective interfering RNA replication. *J Virol* **70**, 2201-2207.
- Charley B & Laude H (1988) Induction of alpha interferon by transmissible gastroenteritis coronavirus: role of transmembrane glycoprotein E1. *J Virol* **62**, 8-11.
- Chen H, Gill A, Dove BK, Emmett SR, Kemp CF, Ritchie MA, Dee M & Hiscox JA (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.
- Chen H, Wurm T, Britton P, Brooks G & Hiscox JA (2002) Interaction of the coronavirus nucleoprotein with nucleolar antigens and the host cell. *J Virol* **76**, 5233-5250.
- Coley SE, Lavi E, Sawicki SG, Fu L, Schelle B, Karl N, Siddell SG & Thiel V (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.
- Collins AR, Knobler RL, Powell H & Buchmeier MJ (1982) Monoclonal antibodies to murine hepatitis virus-4 (strain JHM) define the viral glycoprotein responsible for attachment and cell-cell fusion. *Virology* **119**, 358-371.
- Collins MS, Alexander DJ & Harkness JW (1976) Heterogeneity of infectious bronchitis virus grown in eggs. *Arch Virol* **50**, 55-72.
- Cologna R, Spagnolo JF & Hogue BG (2000) Identification of nucleocapsid binding sites within coronavirus-defective genomes. *Virology* **277**, 235-249.
- Corse E & Machamer CE (2000) Infectious bronchitis virus E protein is targeted to the Golgi complex and directs release of virus-like particles. *J Virol* **74**, 4319-4326.
- Corse E & Machamer CE (2002) The cytoplasmic tail of infectious bronchitis virus E protein directs Golgi targeting. *J Virol* **76**, 1273-1284.
- Corse E & Machamer CE (2003) The cytoplasmic tails of infectious bronchitis virus E and M proteins mediate their interaction. *Virology* **312**, 25-34.
- Curtis KM, Yount B & Baric RS (2002) Heterologous gene expression from transmissible gastroenteritis virus replicon particles. *J Virol* **76**, 1422-1434.
- Curtis KM, Yount B, Sims AC & Baric RS (2004) Reverse genetic analysis of the transcription regulatory sequence of the coronavirus transmissible gastroenteritis virus. *J Virol* **78**, 6061-6066.
- Dalton K, Casais R, Shaw K, Stirrups K, Evans S, Britton P, Brown TD & Cavanagh D (2001) cis-acting sequences required for coronavirus infectious bronchitis virus defective-RNA replication and packaging. *J Virol* **75**, 125-133.
- de Groot RJ, Lenstra JA, Luytjes W, Niesters HG, Horzinek MC, van der Zeijst BA & Spaan WJ (1987) Sequence and structure of the coronavirus peplomer protein. *Adv Exp Med Biol* **218**, 31-38.
- de Groot RJ, Van Leen RW, Dalderup MJ, Vennema H, Horzinek MC & Spaan WJ (1989) Stably expressed FIPV peplomer protein induces cell fusion and elicits neutralizing antibodies in mice. *Virology* **171**, 493-502.
- de Haan CA, de WM, Kuo L, Montalto-Morrison C, Haagmans BL, Weiss SR, Masters PS & Rottier PJ (2003a) 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.

- de Haan CA, Haijema BJ, Boss D, Heuts FW & Rottier PJ (2005) Coronaviruses as vectors: stability of foreign gene expression. *J Virol* **79**, 12742-12751.
- de Haan CA, Kuo L, Masters PS, Vennema H & Rottier PJ (1998) Coronavirus particle assembly: primary structure requirements of the membrane protein. *J Virol* **72**, 6838-6850.
- de Haan CA, Smeets M, Vernooij F, Vennema H & Rottier PJ (1999) Mapping of the coronavirus membrane protein domains involved in interaction with the spike protein. *J Virol* **73**, 7441-7452.
- de Haan CA, van GL, Stoop JN, Volders H & Rottier PJ (2003b) Coronaviruses as vectors: position dependence of foreign gene expression. *J Virol* **77**, 11312-11323.
- de Haan CA, Volders H, Koetzner CA, Masters PS & Rottier PJ (2002) Coronaviruses maintain viability despite dramatic rearrangements of the strictly conserved genome organization. *J Virol* **76**, 12491-12502.
- Dea S & Tijssen P (1988) Identification of the structural proteins of turkey enteric coronavirus. *Arch Virol* **99**, 173-186.
- DeDiego ML, Alvarez E, Almazan F, Rejas MT, Lamirande E, Roberts A, Shieh WJ, Zaki SR, Subbarao K & Enjuanes L (2007) A severe acute respiratory syndrome coronavirus that lacks the E gene is attenuated in vitro and in vivo. *J Virol* **81**, 1701-1713.
- Delmas B, Gelfi J, L'Haridon R, Vogel LK, Sjostrom H, Noren O & Laude H (1992) Aminopeptidase N is a major receptor for the entero-pathogenic coronavirus TGEV. *Nature* **357**, 417-420.
- Delmas B & Laude H (1990) Assembly of coronavirus spike protein into trimers and its role in epitope expression. *J Virol* **64**, 5367-5375.
- Delmas B, Rasschaert D, Godet M, Gelfi J & Laude H (1990) Four major antigenic sites of the coronavirus transmissible gastroenteritis virus are located on the amino-terminal half of spike glycoprotein S. *J Gen Virol* **71** (Pt 6), 1313-1323.
- den Boon JA, Kleijnen MF, Spaan WJ & Snijder EJ (1996) Equine arteritis virus subgenomic mRNA synthesis: analysis of leader-body junctions and replicative-form RNAs. *J Virol* **70**, 4291-4298.
- Drickamer K & Taylor ME (1998) Evolving views of protein glycosylation. *Trends Biochem Sci* **23**, 321-324.
- Dveksler GS, Pensiero MN, Dieffenbach CW, Cardellichio CB, Basile AA, Elia PE &

- Holmes KV (1993) Mouse hepatitis virus strain A59 and blocking antireceptor monoclonal antibody bind to the N-terminal domain of cellular receptor. *Proc Natl Acad Sci U S A* **90**, 1716-1720.
- Egloff MP, Ferron F, Campanacci V, Longhi S, Rancurel C, Dutartre H, Snijder EJ, Gorbalenya AE, Cambillau C & Canard B (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.
- Eleouet JF, Rasschaert D, Lambert P, Levy L, Vende P & Laude H (1995) Complete sequence (20 kilobases) of the polyprotein-encoding gene 1 of transmissible gastroenteritis virus. *Virology* **206**, 817-822.
- Escors D, Camafeita E, Ortego J, Laude H & Enjuanes L (2001) Organization of two transmissible gastroenteritis coronavirus membrane protein topologies within the virion and core. *J Virol* **75**, 12228-12240.
- Fan H, Ooi A, Tan YW, Wang S, Fang S, Liu DX & Lescar J (2005) The nucleocapsid protein of coronavirus infectious bronchitis virus: crystal structure of its N-terminal domain and multimerization properties. *Structure* **13**, 1859-1868.
- Fang S, Chen B, Tay FP, Ng BS & Liu DX (2006) An arginine-to-proline mutation in a domain with undefined functions within the helicase protein (Nsp13) is lethal to the coronavirus infectious bronchitis virus in cultured cells. *Virology*.
- Fischer F, Peng D, Hingley ST, Weiss SR & Masters PS (1997a) The internal open reading frame within the nucleocapsid gene of mouse hepatitis virus encodes a structural protein that is not essential for viral replication. *J Virol* **71**, 996-1003.
- Fischer F, Stegen CF, Koetzner CA & Masters PS (1997b) Analysis of a recombinant mouse hepatitis virus expressing a foreign gene reveals a novel aspect of coronavirus transcription. *J Virol* **71**, 5148-5160.
- Fischer F, Stegen CF, Masters PS & Samsonoff WA (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.
- Fosmire JA, Hwang K & Makino S (1992) Identification and characterization of a coronavirus packaging signal. *J Virol* **66**, 3522-3530.
- Frana MF, Behnke JN, Sturman LS & Holmes KV (1985) Proteolytic cleavage of the E2 glycoprotein of murine coronavirus: host-dependent differences in proteolytic cleavage and cell fusion. *J Virol* **56**, 912-920.

- Fu K & Baric RS (1992) Evidence for variable rates of recombination in the MHV genome. *Virology* **189**, 88-102.
- Fu K & Baric RS (1994) Map locations of mouse hepatitis virus temperature-sensitive mutants: confirmation of variable rates of recombination. *J Virol* **68**, 7458-7466.
- Fuerst TR, Niles EG, Studier FW & Moss B (1986) Eukaryotic transient-expression system based on recombinant vaccinia virus that synthesizes bacteriophage T7 RNA polymerase. *Proc Natl Acad Sci U S A* **83**, 8122-8126.
- Furuya T & Lai MM (1993) Three different cellular proteins bind to complementary sites on the 5'-end-positive and 3'-end-negative strands of mouse hepatitis virus RNA. *J Virol* **67**, 7215-7222.
- Furuya T, Macnaughton TB, La MN & Lai MM (1993) Natural evolution of coronavirus defective-interfering RNA involves RNA recombination. *Virology* **194**, 408-413.
- Garwes DJ & Pocock DH (1975) The polypeptide structure of transmissible gastroenteritis virus. *J Gen Virol* **29**, 25-34.
- Geng H, Liu YM, Chan WS, Lo AW, Au DM, Waye MM & Ho YY (2005) The putative protein 6 of the severe acute respiratory syndrome-associated coronavirus: expression and functional characterization. *FEBS Lett* **579**, 6763-6768.
- Godet M, Grosclaude J, Delmas B & Laude H (1994) Major receptor-binding and neutralization determinants are located within the same domain of the transmissible gastroenteritis virus (coronavirus) spike protein. *J Virol* **68**, 8008-8016.
- Godet M, L'Haridon R, Vautherot JF & Laude H (1992) TGEV corona virus ORF4 encodes a membrane protein that is incorporated into virions. *Virology* **188**, 666-675.
- Gonzalez JM, Penzes Z, Almazan F, Calvo E & Enjuanes L (2002) Stabilization of a full-length infectious cDNA clone of transmissible gastroenteritis coronavirus by insertion of an intron. *J Virol* **76**, 4655-4661.
- Gorbalenya AE, Koonin EV, Donchenko AP & Blinov VM (1989) Coronavirus genome: prediction of putative functional domains in the non-structural polyprotein by comparative amino acid sequence analysis. *Nucleic Acids Res* **17**, 4847-4861.
- Griffiths G & Rottier P (1992) Cell biology of viruses that assemble along the biosynthetic pathway. *Semin Cell Biol* **3**, 367-381.

- Haijema BJ, Volders H & Rottier PJ (2004) Live, attenuated coronavirus vaccines through the directed deletion of group-specific genes provide protection against feline infectious peritonitis. *J Virol* **78**, 3863-3871.
- Helenius A & Aebi M (2001) Intracellular functions of N-linked glycans. *Science* **291**, 2364-2369.
- Herold J & Siddell SG (1993) An 'elaborated' pseudoknot is required for high frequency frameshifting during translation of HCV 229E polymerase mRNA. *Nucleic Acids Res* **21**, 5838-5842.
- Hiscox JA, Wurm T, Wilson L, Britton P, Cavanagh D & Brooks G (2001) The coronavirus infectious bronchitis virus nucleoprotein localizes to the nucleolus. *J Virol* **75**, 506-512.
- Ho Y, Lin PH, Liu CY, Lee SP & Chao YC (2004) Assembly of human severe acute respiratory syndrome coronavirus-like particles. *Biochem Biophys Res Commun* **318**, 833-838.
- Hodgman TC (1988) A new superfamily of replicative proteins. *Nature* **333**, 22-23.
- Hodgson T, Britton P & Cavanagh D (2006) Neither the RNA nor the proteins of open reading frames 3a and 3b of the coronavirus infectious bronchitis virus are essential for replication. *J Virol* **80**, 296-305.
- Hofmann MA & Brain DA (1991) The 5' end of coronavirus minus-strand RNAs contains a short poly(U) tract. *J Virol* **65**, 6331-6333.
- Hogue BG & Brian DA (1986) Structural proteins of human respiratory coronavirus OC43. *Virus Res* **5**, 131-144.
- Hogue BG, Kienzle TE & Brian DA (1989) Synthesis and processing of the bovine enteric coronavirus haemagglutinin protein. *J Gen Virol* **70** (Pt 2), 345-352.
- Holmes KV, Doller EW & Sturman LS (1981) Tunicamycin resistant glycosylation of coronavirus glycoprotein: demonstration of a novel type of viral glycoprotein. *Virology* **115**, 334-344.
- Hooper NM (1999) Detergent-insoluble glycosphingolipid/cholesterol-rich membrane domains, lipid rafts and caveolae (review). *Mol Membr Biol* **16**, 145-156.
- Hsue B & Masters PS (1999) Insertion of a new transcriptional unit into the genome of mouse hepatitis virus. *J Virol* **73**, 6128-6135.
- Huang C, Ito N, Tseng CT & Makino S (2006) Severe acute respiratory syndrome coronavirus 7a accessory protein is a viral structural protein. *J Virol* **80**, 7287-7294.

- Huang P & Lai MM (1999) Polypyrimidine tract-binding protein binds to the complementary strand of the mouse hepatitis virus 3' untranslated region, thereby altering RNA conformation. *J Virol* **73**, 9110-9116.
- Huang Q, Yu L, Petros AM, Gunasekera A, Liu Z, Xu N, Hajduk P, Mack J, Fesik SW & Olejniczak ET (2004a) Structure of the N-terminal RNA-binding domain of the SARS CoV nucleocapsid protein. *Biochemistry* **43**, 6059-6063.
- Huang Y, Yang ZY, Kong WP & Nabel GJ (2004b) Generation of synthetic severe acute respiratory syndrome coronavirus pseudoparticles: implications for assembly and vaccine production. *J Virol* **78**, 12557-12565.
- Ilardi JM, Mochida S & Sheng ZH (1999) Snapin: a SNARE-associated protein implicated in synaptic transmission. *Nat Neurosci* **2**, 119-124.
- Ito N, Mossel EC, Narayanan K, Popov VL, Huang C, Inoue T, Peters CJ & Makino S (2005) Severe acute respiratory syndrome coronavirus 3a protein is a viral structural protein. *J Virol* **79**, 3182-3186.
- Izeta A, Smerdou C, Alonso S, Penzes Z, Mendez A, Plana-Duran J & Enjuanes L (1999) Replication and packaging of transmissible gastroenteritis coronavirus-derived synthetic minigenomes. *J Virol* **73**, 1535-1545.
- Jayaram H, Fan H, Bowman BR, Ooi A, Jayaram J, Collisson EW, Lescar J & Prasad BV (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.
- Jeffers SA, Tusell SM, Gillim-Ross L, *et al.* (2004) CD209L (L-SIGN) is a receptor for severe acute respiratory syndrome coronavirus. *Proc Natl Acad Sci U S A* **101**, 15748-15753.
- Jones DT, Taylor WR & Thornton JM (1994) A model recognition approach to the prediction of all-helical membrane protein structure and topology. *Biochemistry* **33**, 3038-3049.
- Joo M & Makino S (1992) Mutagenic analysis of the coronavirus intergenic consensus sequence. *J Virol* **66**, 6330-6337.
- Joseph JS, Saikatendu KS, Subramanian V, *et al.* (2006) Crystal structure of nonstructural protein 10 from the severe acute respiratory syndrome coronavirus reveals a novel fold with two zinc-binding motifs. *J Virol* **80**, 7894-7901.
- Kanzawa N, Nishigaki K, Hayashi T, *et al.* (2006) Augmentation of chemokine production by severe acute respiratory syndrome coronavirus 3a/X1 and 7a/X4 proteins through NF-kappaB activation. *FEBS Lett* **580**, 6807-6812.

- Keller P & Simons K (1998) Cholesterol is required for surface transport of influenza virus hemagglutinin. *J Cell Biol* **140**, 1357-1367.
- Khattari Z, Brotons G, Akkawi M, Arbely E, Arkin IT & Salditt T (2006) SARS coronavirus E protein in phospholipid bilayers: an x-ray study. *Biophys J* **90**, 2038-2050.
- Kim YN, Jeong YS & Makino S (1993) Analysis of cis-acting sequences essential for coronavirus defective interfering RNA replication. *Virology* **197**, 53-63.
- King B & Brian DA (1982) Bovine coronavirus structural proteins. *J Virol* **42**, 700-707.
- King B, Potts BJ & Brian DA (1985) Bovine coronavirus hemagglutinin protein. *Virus Res* **2**, 53-59.
- Klumperman J, Locker JK, Meijer A, Horzinek MC, Geuze HJ & Rottier PJ (1994) Coronavirus M proteins accumulate in the Golgi complex beyond the site of virion budding. *J Virol* **68**, 6523-6534.
- Koetzner CA, Parker MM, Ricard CS, Sturman LS & Masters PS (1992) Repair and mutagenesis of the genome of a deletion mutant of the coronavirus mouse hepatitis virus by targeted RNA recombination. *J Virol* **66**, 1841-1848.
- Koivunen P, Salo KE, Myllyharju J & Ruddock LW (2005) Three binding sites in protein-disulfide isomerase cooperate in collagen prolyl 4-hydroxylase tetramer assembly. *J Biol Chem* **280**, 5227-5235.
- Koolen MJ, Osterhaus AD, Van SG, Horzinek MC & van der Zeijst BA (1983) Temperature-sensitive mutants of mouse hepatitis virus strain A59: isolation, characterization and neuropathogenic properties. *Virology* **125**, 393-402.
- Krijnse-Locker J, Ericsson M, Rottier PJ & Griffiths G (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.
- Krogh A, Larsson B, Von HG & Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* **305**, 567-580.
- Krzystyniak K & Dupuy JM (1984) Entry of mouse hepatitis virus 3 into cells. *J Gen Virol* **65** (Pt 1), 227-231.
- Kubo H, Yamada YK & Taguchi F (1994) Localization of neutralizing epitopes and the receptor-binding site within the amino-terminal 330 amino acids of the murine coronavirus spike protein. *J Virol* **68**, 5403-5410.

- Kuo L, Godeke GJ, Raamsman MJ, Masters PS & Rottier PJ (2000) Retargeting of coronavirus by substitution of the spike glycoprotein ectodomain: crossing the host cell species barrier. *J Virol* **74**, 1393-1406.
- Kuo L, Hurst KR & Masters PS (2007) Exceptional flexibility in the sequence requirements for coronavirus small envelope protein function. *J Virol* **81**, 2249-2262.
- Kuo L & Masters PS (2003) The small envelope protein E is not essential for murine coronavirus replication. *J Virol* **77**, 4597-4608.
- Lachance C, Arbour N, Cashman NR & Talbot PJ (1998) Involvement of aminopeptidase N (CD13) in infection of human neural cells by human coronavirus 229E. *J Virol* **72**, 6511-6519.
- Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.
- Lai MM, Baric RS, Brayton PR & Stohlman SA (1984) Characterization of leader RNA sequences on the virion and mRNAs of mouse hepatitis virus, a cytoplasmic RNA virus. *Proc Natl Acad Sci U S A* **81**, 3626-3630.
- Lai MM, Brayton PR, Armen RC, Patton CD, Pugh C & Stohlman SA (1981) Mouse hepatitis virus A59: mRNA structure and genetic localization of the sequence divergence from hepatotropic strain MHV-3. *J Virol* **39**, 823-834.
- Lai MM & Cavanagh D (1997) The molecular biology of coronaviruses. *Adv Virus Res* **48**, 1-100.
- Lai MM, Patton CD, Baric RS & Stohlman SA (1983) Presence of leader sequences in the mRNA of mouse hepatitis virus. *J Virol* **46**, 1027-1033.
- Lai MM, Patton CD & Stohlman SA (1982) Further characterization of mRNA's of mouse hepatitis virus: presence of common 5'-end nucleotides. *J Virol* **41**, 557-565.
- Langlet C, Bernard AM, Drevot P & He HT (2000) Membrane rafts and signaling by the multichain immune recognition receptors. *Curr Opin Immunol* **12**, 250-255.
- Lau SK, Woo PC, Li KS, Huang Y, Tsoi HW, Wong BH, Wong SS, Leung SY, Chan KH & Yuen KY (2005) Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. *Proc Natl Acad Sci U S A* **102**, 14040-14045.
- Laude H, Gelfi J, Lavenant L & Charley B (1992) Single amino acid changes in the viral glycoprotein M affect induction of alpha interferon by the coronavirus transmissible gastroenteritis virus. *J Virol* **66**, 743-749.

- Le SY, Sonenberg N & Maizel JV, Jr. (1994) Distinct structural elements and internal entry of ribosomes in mRNA3 encoded by infectious bronchitis virus. *Virology* **198**, 405-411.
- Lee HJ, Shieh CK, Gorbalenya AE, Koonin EV, La MN, Tuler J, Bagdzhadzhyan A & Lai MM (1991) The complete sequence (22 kilobases) of murine coronavirus gene 1 encoding the putative proteases and RNA polymerase. *Virology* **180**, 567-582.
- Leibowitz JL, DeVries JR & Haspel MV (1982) Genetic analysis of murine hepatitis virus strain JHM. *J Virol* **42**, 1080-1087.
- Leibowitz JL, Perlman S, Weinstock G, DeVries JR, Budzilowicz C, Weissemann JM & Weiss SR (1988) Detection of a murine coronavirus nonstructural protein encoded in a downstream open reading frame. *Virology* **164**, 156-164.
- Leibowitz JL, Wilhelmsen KC & Bond CW (1981) The virus-specific intracellular RNA species of two murine coronaviruses: MHV-a59 and MHV-JHM. *Virology* **114**, 39-51.
- Leparc-Goffart I, Hingley ST, Chua MM, Phillips J, Lavi E & Weiss SR (1998) Targeted recombination within the spike gene of murine coronavirus mouse hepatitis virus-A59: Q159 is a determinant of hepatotropism. *J Virol* **72**, 9628-9636.
- Li D & Cavanagh D (1992) Coronavirus IBV-induced membrane fusion occurs at near-neutral pH. *Arch Virol* **122**, 307-316.
- Li HP, Huang P, Park S & Lai MM (1999) Polypyrimidine tract-binding protein binds to the leader RNA of mouse hepatitis virus and serves as a regulator of viral transcription. *J Virol* **73**, 772-777.
- Li HP, Zhang X, Duncan R, Comai L & Lai MM (1997) Heterogeneous nuclear ribonucleoprotein A1 binds to the transcription-regulatory region of mouse hepatitis virus RNA. *Proc Natl Acad Sci U S A* **94**, 9544-9549.
- Li W, Moore MJ, Vasilieva N, *et al.* (2003) Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* **426**, 450-454.
- Li W, Shi Z, Yu M, *et al.* (2005) Bats are natural reservoirs of SARS-like coronaviruses. *Science* **310**, 676-679.
- Liao CL & Lai MM (1995) A cis-acting viral protein is not required for the replication of a coronavirus defective-interfering RNA. *Virology* **209**, 428-436.
- Liao Y, Lescar J, Tam JP & Liu DX (2004b) Expression of SARS-coronavirus envelope protein in Escherichia coli cells alters membrane permeability.

Biochem Biophys Res Commun **325**, 374-380.

- Liao Y, Lescar J, Tam JP & Liu DX (2004a) Expression of SARS-coronavirus envelope protein in Escherichia coli cells alters membrane permeability. *Biochem Biophys Res Commun* **325**, 374-380.
- Liao Y, Yuan Q, Torres J, Tam JP & Liu DX (2006) Biochemical and functional characterization of the membrane association and membrane permeabilizing activity of the severe acute respiratory syndrome coronavirus envelope protein. *Virology*.
- Lim KP & Liu DX (2001) The missing link in coronavirus assembly. Retention of the avian coronavirus infectious bronchitis virus envelope protein in the pre-Golgi compartments and physical interaction between the envelope and membrane proteins. *J Biol Chem* **276**, 17515-17523.
- Lin YJ & Lai MM (1993) Deletion mapping of a mouse hepatitis virus defective interfering RNA reveals the requirement of an internal and discontinuous sequence for replication. *J Virol* **67**, 6110-6118.
- Lin YJ, Liao CL & Lai MM (1994) Identification of the cis-acting signal for minus-strand RNA synthesis of a murine coronavirus: implications for the role of minus-strand RNA in RNA replication and transcription. *J Virol* **68**, 8131-8140.
- Liu DX, Cavanagh D, Green P & Inglis SC (1991) A polycistronic mRNA specified by the coronavirus infectious bronchitis virus. *Virology* **184**, 531-544.
- Liu DX & Inglis SC (1991) Association of the infectious bronchitis virus 3c protein with the virion envelope. *Virology* **185**, 911-917.
- Liu DX & Inglis SC (1992a) Identification of two new polypeptides encoded by mRNA5 of the coronavirus infectious bronchitis virus. *Virology* **186**, 342-347.
- Liu DX & Inglis SC (1992b) Internal entry of ribosomes on a tricistronic mRNA encoded by infectious bronchitis virus. *J Virol* **66**, 6143-6154.
- Liu DX, Shen S, Xu HY & Wang SF (1998) Proteolytic mapping of the coronavirus infectious bronchitis virus 1b polyprotein: evidence for the presence of four cleavage sites of the 3C-like proteinase and identification of two novel cleavage products. *Virology* **246**, 288-297.
- Locker JK, Griffiths G, Horzinek MC & Rottier PJ (1992) O-glycosylation of the coronavirus M protein. Differential localization of sialyltransferases in N- and O-linked glycosylation. *J Biol Chem* **267**, 14094-14101.
- Lu W, Zheng BJ, Xu K, Schwarz W, Du L, Wong CK, Chen J, Duan S, Deubel V &

- Sun B (2006) Severe acute respiratory syndrome-associated coronavirus 3a protein forms an ion channel and modulates virus release. *Proc Natl Acad Sci U S A* **103**, 12540-12545.
- Luytjes W, Sturman LS, Bredenbeek PJ, Charite J, van der Zeijst BA, Horzinek MC & Spaan WJ (1987) Primary structure of the glycoprotein E2 of coronavirus MHV-A59 and identification of the trypsin cleavage site. *Virology* **161**, 479-487.
- Machamer CE, Grim MG, Esquela A, Chung SW, Rolls M, Ryan K & Swift AM (1993) Retention of a cis Golgi protein requires polar residues on one face of a predicted alpha-helix in the transmembrane domain. *Mol Biol Cell* **4**, 695-704.
- Machamer CE, Mentone SA, Rose JK & Farquhar MG (1990) The E1 glycoprotein of an avian coronavirus is targeted to the cis Golgi complex. *Proc Natl Acad Sci U S A* **87**, 6944-6948.
- Machamer CE & Rose JK (1987) A specific transmembrane domain of a coronavirus E1 glycoprotein is required for its retention in the Golgi region. *J Cell Biol* **105**, 1205-1214.
- Macnaughton MR & Davies HA (1980) Two particle types of avian infectious bronchitis virus. *J Gen Virol* **47**, 365-372.
- Macneughton MR & Davies HA (1978) Ribonucleoprotein-like structures from coronavirus particles. *J Gen Virol* **39**, 545-549.
- Madan V, Garcia MJ, Sanz MA & Carrasco L (2005) Viroporin activity of murine hepatitis virus E protein. *FEBS Lett* **579**, 3607-3612.
- Maeda J, Maeda A & Makino S (1999) Release of coronavirus E protein in membrane vesicles from virus-infected cells and E protein-expressing cells. *Virology* **263**, 265-272.
- Maeda J, Repass JF, Maeda A & Makino S (2001) Membrane topology of coronavirus E protein. *Virology* **281**, 163-169.
- Makino S, Fujioka N & Fujiwara K (1985) Structure of the intracellular defective viral RNAs of defective interfering particles of mouse hepatitis virus. *J Virol* **54**, 329-336.
- Makino S & Joo M (1993) Effect of intergenic consensus sequence flanking sequences on coronavirus transcription. *J Virol* **67**, 3304-3311.
- Makino S, Joo M & Makino JK (1991) A system for study of coronavirus mRNA synthesis: a regulated, expressed subgenomic defective interfering RNA results from intergenic site insertion. *J Virol* **65**, 6031-6041.

- Makino S, Soe LH, Shieh CK & Lai MM (1988) Discontinuous transcription generates heterogeneity at the leader fusion sites of coronavirus mRNAs. *J Virol* **62**, 3870-3873.
- Makino S, Taguchi F & Fujiwara K (1984) Defective interfering particles of mouse hepatitis virus. *Virology* **133**, 9-17.
- Mandal M, Mandal A, Das S, Chakraborti T & Chakraborti S (2003) Identification, purification and partial characterization of tissue inhibitor of matrix metalloproteinase-2 in bovine pulmonary artery smooth muscle. *Mol Cell Biochem* **254**, 275-287.
- Martin JP, Koehren F, Rannou JJ & Kirn A (1988) Temperature-sensitive mutants of mouse hepatitis virus type 3 (MHV-3): isolation, biochemical and genetic characterization. *Arch Virol* **100**, 147-160.
- Masters PS (1992) Localization of an RNA-binding domain in the nucleocapsid protein of the coronavirus mouse hepatitis virus. *Arch Virol* **125**, 141-160.
- Masters PS (1999) Reverse genetics of the largest RNA viruses. *Adv Virus Res* **53**, 245-264.
- McGinnes LW, Reitter JN, Gravel K & Morrison TG (2003) Evidence for mixed membrane topology of the newcastle disease virus fusion protein. *J Virol* **77**, 1951-1963.
- Merritt EA, Sixma TK, Kalk KH, van Zanten BA & Hol WG (1994) Galactose-binding site in Escherichia coli heat-labile enterotoxin (LT) and cholera toxin (CT). *Mol Microbiol* **13**, 745-753.
- Mizzen L, Hilton A, Cheley S & Anderson R (1985) Attenuation of murine coronavirus infection by ammonium chloride. *Virology* **142**, 378-388.
- Moise AR, Grant JR, Lippe R, Gabathuler R & Jefferies WA (2004) The adenovirus E3-6.7K protein adopts diverse membrane topologies following posttranslational translocation. *J Virol* **78**, 454-463.
- Molenkamp R & Spaan WJ (1997) Identification of a specific interaction between the coronavirus mouse hepatitis virus A59 nucleocapsid protein and packaging signal. *Virology* **239**, 78-86.
- Nal B, Chan C, Kien F, *et al.* (2005) Differential maturation and subcellular localization of severe acute respiratory syndrome coronavirus surface proteins S, M and E. *J Gen Virol* **86**, 1423-1434.
- Narayanan K, Chen CJ, Maeda J & Makino S (2003) Nucleocapsid-independent specific viral RNA packaging via viral envelope protein and viral RNA signal.

J Virol **77**, 2922-2927.

- Narayanan K, Maeda A, Maeda J & Makino S (2000) Characterization of the coronavirus M protein and nucleocapsid interaction in infected cells. *J Virol* **74**, 8127-8134.
- Narayanan K & Makino S (2001) Cooperation of an RNA packaging signal and a viral envelope protein in coronavirus RNA packaging. *J Virol* **75**, 9059-9067.
- Navas S, Seo SH, Chua MM, Sarma JD, Lavi E, Hingley ST & Weiss SR (2001) Murine coronavirus spike protein determines the ability of the virus to replicate in the liver and cause hepatitis. *J Virol* **75**, 2452-2457.
- Nelson GW & Stohlman SA (1993) Localization of the RNA-binding domain of mouse hepatitis virus nucleocapsid protein. *J Gen Virol* **74** (Pt 9), 1975-1979.
- Neuman BW, Adair BD, Yoshioka C, Quispe JD, Orca G, Kuhn P, Milligan RA, Yeager M & Buchmeier MJ (2006) Supramolecular architecture of severe acute respiratory syndrome coronavirus revealed by electron cryomicroscopy. *J Virol* **80**, 7918-7928.
- Nguyen VP & Hogue BG (1997) Protein interactions during coronavirus assembly. *J Virol* **71**, 9278-9284.
- Nilsson IM & Von HG (1993) Determination of the distance between the oligosaccharyltransferase active site and the endoplasmic reticulum membrane 34. *J Biol Chem* **268**, 5798-5801.
- Oostra M, de Haan CA, de Groot RJ & Rottier PJ (2006) Glycosylation of the severe acute respiratory syndrome coronavirus triple-spanning membrane proteins 3a and M. *J Virol* **80**, 2326-2336.
- Opstelten DJ, de GP, Horzinek MC, Vennema H & Rottier PJ (1993) Disulfide bonds in folding and transport of mouse hepatitis coronavirus glycoproteins. *J Virol* **67**, 7394-7401.
- Opstelten DJ, Raamsman MJ, Wolfs K, Horzinek MC & Rottier PJ (1995) Envelope glycoprotein interactions in coronavirus assembly. *J Cell Biol* **131**, 339-349.
- Ortego J, Escors D, Laude H & Enjuanes L (2002) Generation of a replication-competent, propagation-deficient virus vector based on the transmissible gastroenteritis coronavirus genome. *J Virol* **76**, 11518-11529.
- Parker MM & Masters PS (1990) Sequence comparison of the N genes of five strains of the coronavirus mouse hepatitis virus suggests a three domain structure for the nucleocapsid protein. *Virology* **179**, 463-468.

- Parton RG (1994) Ultrastructural localization of gangliosides; GM1 is concentrated in caveolae. *J Histochem Cytochem* **42**, 155-166.
- Payne HR & Storz J (1988) Analysis of cell fusion induced by bovine coronavirus infection. *Arch Virol* **103**, 27-33.
- Perlman S, Ries D, Bolger E, Chang LJ & Stoltzfus CM (1986) MHV nucleocapsid synthesis in the presence of cycloheximide and accumulation of negative strand MHV RNA. *Virus Res* **6**, 261-272.
- Pfleiderer M, Routledge E & Siddell SG (1990) Functional analysis of the coronavirus MHV-JHM surface glycoproteins in vaccinia virus recombinants. *Adv Exp Med Biol* **276**, 21-31.
- Plutner H, Davidson HW, Saraste J & Balch WE (1992) Morphological analysis of protein transport from the ER to Golgi membranes in digitonin-permeabilized cells: role of the P58 containing compartment. *J Cell Biol* **119**, 1097-1116.
- Poon LL, Chu DK, Chan KH, *et al.* (2005) Identification of a novel coronavirus in bats. *J Virol* **79**, 2001-2009.
- Prior IA, Harding A, Yan J, Sluimer J, Parton RG & Hancock JF (2001) GTP-dependent segregation of H-ras from lipid rafts is required for biological activity. *Nat Cell Biol* **3**, 368-375.
- Putics A, Filipowicz W, Hall J, Gorbalenya AE & Ziebuhr J (2005) ADP-ribose-1"-monophosphatase: a conserved coronavirus enzyme that is dispensable for viral replication in tissue culture. *J Virol* **79**, 12721-12731.
- Raamsman MJ, Locker JK, de HA, de Vries AA, Griffiths G, Vennema H & Rottier PJ (2000) Characterization of the coronavirus mouse hepatitis virus strain A59 small membrane protein E. *J Virol* **74**, 2333-2342.
- Racaniello VR & Baltimore D (1981) Cloned poliovirus complementary DNA is infectious in mammalian cells. *Science* **214**, 916-919.
- Repass JF & Makino S (1998) Importance of the positive-strand RNA secondary structure of a murine coronavirus defective interfering RNA internal replication signal in positive-strand RNA synthesis. *J Virol* **72**, 7926-7933.
- Risco C, Anton IM, Sune C, Pedregosa AM, Martin-Alonso JM, Parra F, Carrascosa JL & Enjuanes L (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.
- Robb JA, Bond CW & Leibowitz JL (1979) Pathogenic murine coronaviruses. III. Biological and biochemical characterization of temperature-sensitive mutants

of JHMV. *Virology* **94**, 385-399.

- Robbins SG, Frana MF, McGowan JJ, Boyle JF & Holmes KV (1986) RNA-binding proteins of coronavirus MHV: detection of monomeric and multimeric N protein with an RNA overlay-protein blot assay. *Virology* **150**, 402-410.
- Rota PA, Oberste MS, Monroe SS, *et al.* (2003) Characterization of a novel coronavirus associated with severe acute respiratory syndrome. *Science* **300**, 1394-1399.
- Rottier P, Brandenburg D, Armstrong J, van der ZB & Warren G (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.
- Satyanarayana T, Bar-Joseph M, Mawassi M, biach-Marti MR, Ayllon MA, Gowda S, Hilf ME, Moreno P, Garnsey SM & Dawson WO (2001) Amplification of Citrus tristeza virus from a cDNA clone and infection of citrus trees. *Virology* **280**, 87-96.
- Satyanarayana T, Gowda S, Boyko VP, *et al.* (1999) An engineered closterovirus RNA replicon and analysis of heterologous terminal sequences for replication. *Proc Natl Acad Sci U S A* **96**, 7433-7438.
- Sawicki SG & Sawicki DL (1986) Coronavirus minus-strand RNA synthesis and effect of cycloheximide on coronavirus RNA synthesis. *J Virol* **57**, 328-334.
- Sawicki SG & Sawicki DL (1990) Coronavirus transcription: subgenomic mouse hepatitis virus replicative intermediates function in RNA synthesis. *J Virol* **64**, 1050-1056.
- Sawicki SG & Sawicki DL (1995) Coronaviruses use discontinuous extension for synthesis of subgenome-length negative strands. *Adv Exp Med Biol* **380**, 499-506.
- Sawicki SG & Sawicki DL (1998) A new model for coronavirus transcription. *Adv Exp Med Biol* **440**, 215-219.
- Schaad MC & Baric RS (1994) Genetics of mouse hepatitis virus transcription: evidence that subgenomic negative strands are functional templates. *J Virol* **68**, 8169-8179.
- Schaad MC, Stohlman SA, Egbert J, Lum K, Fu K, Wei T, Jr. & Baric RS (1990) Genetics of mouse hepatitis virus transcription: identification of cistrons which may function in positive and negative strand RNA synthesis. *Virology* **177**, 634-645.

- Schaecher SR, Mackenzie JM & Pekosz A (2006) The ORF7b protein of SARS-CoV is expressed in virus-infected cells and incorporated into SARS-CoV particles. *J Virol*.
- Schultze B, Gross HJ, Brossmer R & Herrler G (1991) The S protein of bovine coronavirus is a hemagglutinin recognizing 9-O-acetylated sialic acid as a receptor determinant. *J Virol* **65**, 6232-6237.
- Senanayake SD, Hofmann MA, Maki JL & Brian DA (1992) The nucleocapsid protein gene of bovine coronavirus is bicistronic. *J Virol* **66**, 5277-5283.
- Sethna PB, Hofmann MA & Brian DA (1991b) Minus-strand copies of replicating coronavirus mRNAs contain antileaders. *J Virol* **65**, 320-325.
- Sethna PB, Hofmann MA & Brian DA (1991a) Minus-strand copies of replicating coronavirus mRNAs contain antileaders. *J Virol* **65**, 320-325.
- Sethna PB, Hung SL & Brian DA (1989) Coronavirus subgenomic minus-strand RNAs and the potential for mRNA replicons. *Proc Natl Acad Sci U S A* **86**, 5626-5630.
- Shi ST, Huang P, Li HP & Lai MM (2000) Heterogeneous nuclear ribonucleoprotein A1 regulates RNA synthesis of a cytoplasmic virus. *EMBO J* **19**, 4701-4711.
- Shieh CK, Lee HJ, Yokomori K, La MN, Makino S & Lai MM (1989) Identification of a new transcriptional initiation site and the corresponding functional gene 2b in the murine coronavirus RNA genome. *J Virol* **63**, 3729-3736.
- Simons K & Ikonen E (1997) Functional rafts in cell membranes. *Nature* **387**, 569-572.
- Skinner MA, Ebner D & Siddell SG (1985) Coronavirus MHV-JHM mRNA 5 has a sequence arrangement which potentially allows translation of a second, downstream open reading frame. *J Gen Virol* **66** (Pt 3), 581-592.
- Smith AR, Bournsnel ME, Binns MM, Brown TD & Inglis SC (1990) Identification of a new membrane-associated polypeptide specified by the coronavirus infectious bronchitis virus. *J Gen Virol* **71** (Pt 1), 3-11.
- Snijder EJ, Bredenbeek PJ, Dobbe JC, Thiel V, Ziebuhr J, Poon LL, Guan Y, Rozanov M, Spaan WJ & Gorbalenya AE (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.
- Somogyi P, Jenner AJ, Brierley I & Inglis SC (1993) Ribosomal pausing during translation of an RNA pseudoknot. *Mol Cell Biol* **13**, 6931-6940.

- Spaan W, Cavanagh D & Horzinek MC (1988) Coronaviruses: structure and genome expression. *J Gen Virol* **69** (Pt 12), 2939-2952.
- Spaan W, Delius H, Skinner M, Armstrong J, Rottier P, Smeekens S, van der Zeijst BA & Siddell SG (1983) Coronavirus mRNA synthesis involves fusion of non-contiguous sequences. *EMBO J* **2**, 1839-1844.
- Stalcup RP, Baric RS & Leibowitz JL (1998) Genetic complementation among three panels of mouse hepatitis virus gene 1 mutants. *Virology* **241**, 112-121.
- Starcevic M & Dell'Angelica EC (2004) Identification of snapin and three novel proteins (BLOS1, BLOS2, and BLOS3/reduced pigmentation) as subunits of biogenesis of lysosome-related organelles complex-1 (BLOC-1). *J Biol Chem* **279**, 28393-28401.
- Stern DF & Kennedy SI (1980) Coronavirus multiplication strategy. II. Mapping the avian infectious bronchitis virus intracellular RNA species to the genome. *J Virol* **36**, 440-449.
- Stern DF & Sefton BM (1982a) Coronavirus proteins: biogenesis of avian infectious bronchitis virus virion proteins. *J Virol* **44**, 794-803.
- Stern DF & Sefton BM (1982b) Coronavirus proteins: structure and function of the oligosaccharides of the avian infectious bronchitis virus glycoproteins. *J Virol* **44**, 804-812.
- Stohlman SA, Baric RS, Nelson GN, Soe LH, Welter LM & Deans RJ (1988b) Specific interaction between coronavirus leader RNA and nucleocapsid protein. *J Virol* **62**, 4288-4295.
- Stohlman SA, Baric RS, Nelson GN, Soe LH, Welter LM & Deans RJ (1988a) Specific interaction between coronavirus leader RNA and nucleocapsid protein. *J Virol* **62**, 4288-4295.
- Sturman LS (1977) I. Structural proteins: effects of preparative conditions on the migration of protein in polyacrylamide gels. *Virology* **77**, 637-649.
- Sturman LS, Eastwood C, Frana MF, Duchala C, Baker F, Ricard CS, Sawicki SG & Holmes KV (1987) Temperature-sensitive mutants of MHV-A59. *Adv Exp Med Biol* **218**, 159-168.
- Sturman LS & Holmes KV (1977) Characterization of coronavirus II. Glycoproteins of the viral envelope: tryptic peptide analysis. *Virology* **77**, 650-660.
- Sturman LS, Holmes KV & Behnke J (1980) Isolation of coronavirus envelope glycoproteins and interaction with the viral nucleocapsid. *J Virol* **33**, 449-462.

- Sturman LS, Ricard CS & Holmes KV (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.
- Su D, Lou Z, Sun F, Zhai Y, Yang H, Zhang R, Joachimiak A, Zhang XC, Bartlam M & Rao Z (2006) Dodecamer structure of severe acute respiratory syndrome coronavirus nonstructural protein nsp10. *J Virol* **80**, 7902-7908.
- Suomalainen M (2002) Lipid rafts and assembly of enveloped viruses. *Traffic* **3**, 705-709.
- Sutton G, Fry E, Carter L, *et al.* (2004) The nsp9 replicase protein of SARS-coronavirus, structure and functional insights. *Structure* **12**, 341-353.
- Swift AM & Machamer CE (1991) A Golgi retention signal in a membrane-spanning domain of coronavirus E1 protein. *J Cell Biol* **115**, 19-30.
- Taguchi F (1993) Fusion formation by the uncleaved spike protein of murine coronavirus JHMV variant cl-2. *J Virol* **67**, 1195-1202.
- Tan J, Verschueren KH, Anand K, *et al.* (2005) pH-dependent conformational flexibility of the SARS-CoV main proteinase (M(pro)) dimer: molecular dynamics simulations and multiple X-ray structure analyses. *J Mol Biol* **354**, 25-40.
- Tan YW, Fang S, Fan H, Lescar J & Liu DX (2006) Amino acid residues critical for RNA-binding in the N-terminal domain of the nucleocapsid protein are essential determinants for the infectivity of coronavirus in cultured cells. *Nucleic Acids Res* **34**, 4816-4825.
- Thiel V, Herold J, Schelle B & Siddell SG (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.
- Thiel V, Karl N, Schelle B, Disterer P, Klagge I & Siddell SG (2003) Multigene RNA vector based on coronavirus transcription. *J Virol* **77**, 9790-9798.
- Thiel V & Siddell SG (1994) Internal ribosome entry in the coding region of murine hepatitis virus mRNA 5. *J Gen Virol* **75** (Pt 11), 3041-3046.
- Tooze J, Tooze S & Warren G (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.
- Tooze J & Tooze SA (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.

- Torres J, Parthasarathy K, Lin X, Saravanan R, Kukol A & Liu DX (2006) Model of a putative pore: the pentameric alpha-helical bundle of SARS coronavirus E protein in lipid bilayers. *Biophys J* **91**, 938-947.
- Torres J, Wang J, Parthasarathy K & Liu DX (2005) The transmembrane oligomers of coronavirus protein E. *Biophys J* **88**, 1283-1290.
- Tresnan DB, Levis R & Holmes KV (1996) Feline aminopeptidase N serves as a receptor for feline, canine, porcine, and human coronaviruses in serogroup I. *J Virol* **70**, 8669-8674.
- Tusnady GE & Simon I (1998) Principles governing amino acid composition of integral membrane proteins: application to topology prediction. *J Mol Biol* **283**, 489-506.
- van der Most RG, Heijnen L, Spaan WJ & de Groot RJ (1992) Homologous RNA recombination allows efficient introduction of site-specific mutations into the genome of coronavirus MHV-A59 via synthetic co-replicating RNAs. *Nucleic Acids Res* **20**, 3375-3381.
- van der HL, Pyrc K, Jebbink MF, Vermeulen-Oost W, Berkhout RJ, Wolthers KC, Wertheim-van Dillen PM, Kaandorp J, Spaargaren J & Berkhout B (2004) Identification of a new human coronavirus. *Nat Med* **10**, 368-373.
- Vennema H, Godeke GJ, Rossen JW, Voorhout WF, Horzinek MC, Opstelten DJ & Rottier PJ (1996) Nucleocapsid-independent assembly of coronavirus-like particles by co-expression of viral envelope protein genes. *EMBO J* **15**, 2020-2028.
- Webb Y, Hermida-Matsumoto L & Resh MD (2000) Inhibition of protein palmitoylation, raft localization, and T cell signaling by 2-bromopalmitate and polyunsaturated fatty acids. *J Biol Chem* **275**, 261-270.
- Weismiller DG, Sturman LS, Buchmeier MJ, Fleming JO & Holmes KV (1990) Monoclonal antibodies to the peplomer glycoprotein of coronavirus mouse hepatitis virus identify two subunits and detect a conformational change in the subunit released under mild alkaline conditions. *J Virol* **64**, 3051-3055.
- Williams GD, Chang RY & Brian DA (1999) A phylogenetically conserved hairpin-type 3' untranslated region pseudoknot functions in coronavirus RNA replication. *J Virol* **73**, 8349-8355.
- Williams RK, Jiang GS & Holmes KV (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.
- Wilson L, McKinlay C, Gage P & Ewart G (2004) SARS coronavirus E protein forms

- cation-selective ion channels. *Virology* **330**, 322-331.
- Winter C, Schwegmann-Wessels C, Cavanagh D, Neumann U & Herrler G (2006) Sialic acid is a receptor determinant for infection of cells by avian Infectious bronchitis virus. *J Gen Virol* **87**, 1209-1216.
- Woo K, Joo M, Narayanan K, Kim KH & Makino S (1997) Murine coronavirus packaging signal confers packaging to nonviral RNA. *J Virol* **71**, 824-827.
- Woo PC, Lau SK, Chu CM, *et al.* (2005) Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. *J Virol* **79**, 884-895.
- Wurm T, Chen H, Hodgson T, Britton P, Brooks G & Hiscox JA (2001) Localization to the nucleolus is a common feature of coronavirus nucleoproteins, and the protein may disrupt host cell division. *J Virol* **75**, 9345-9356.
- Xiao X, Chakraborti S, Dimitrov AS, Gramatikoff K & Dimitrov DS (2003) The SARS-CoV S glycoprotein: expression and functional characterization. *Biochem Biophys Res Commun* **312**, 1159-1164.
- Yamada YK, Yabe M, Ohtsuki T & Taguchi F (2000) Unique N-linked glycosylation of murine coronavirus MHV-2 membrane protein at the conserved O-linked glycosylation site. *Virus Res* **66**, 149-154.
- Yang H, Yang M, Ding Y, *et al.* (2003) The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor. *Proc Natl Acad Sci U S A* **100**, 13190-13195.
- Yang Y, Xiong Z, Zhang S, *et al.* (2005) Bcl-xL inhibits T-cell apoptosis induced by expression of SARS coronavirus E protein in the absence of growth factors. *Biochem J* **392**, 135-143.
- Ye Y & Hogue BG (2007) Role of the coronavirus E viroporin protein transmembrane domain in virus assembly. *J Virol* **81**, 3597-3607.
- Yeager CL, Ashmun RA, Williams RK, Cardellicchio CB, Shapiro LH, Look AT & Holmes KV (1992) Human aminopeptidase N is a receptor for human coronavirus 229E. *Nature* **357**, 420-422.
- Yokomori K, La MN, Makino S, Shieh CK & Lai MM (1989) Biosynthesis, structure, and biological activities of envelope protein gp65 of murine coronavirus. *Virology* **173**, 683-691.
- Yoo DW, Parker MD & Babiuk LA (1991) The S2 subunit of the spike glycoprotein of bovine coronavirus mediates membrane fusion in insect cells. *Virology* **180**, 395-399.

- Youn S, Leibowitz JL & Collisson EW (2005) In vitro assembled, recombinant infectious bronchitis viruses demonstrate that the 5a open reading frame is not essential for replication. *Virology* **332**, 206-215.
- Yount B, Curtis KM & Baric RS (2000) Strategy for systematic assembly of large RNA and DNA genomes: transmissible gastroenteritis virus model. *J Virol* **74**, 10600-10611.
- Yount B, Curtis KM, Fritz EA, Hensley LE, Jahrling PB, Prentice E, Denison MR, Geisbert TW & Baric RS (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.
- Yount B, Denison MR, Weiss SR & Baric RS (2002) Systematic assembly of a full-length infectious cDNA of mouse hepatitis virus strain A59. *J Virol* **76**, 11065-11078.
- Yount B, Roberts RS, Sims AC, Deming D, Frieman MB, Sparks J, Denison MR, Davis N & Baric RS (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.
- Yu X, Bi W, Weiss SR & Leibowitz JL (1994) Mouse hepatitis virus gene 5b protein is a new virion envelope protein. *Virology* **202**, 1018-1023.
- Zhai Y, Sun F, Li X, Pang H, Xu X, Bartlam M & Rao Z (2005) Insights into SARS-CoV transcription and replication from the structure of the nsp7-nsp8 hexadecamer. *Nat Struct Mol Biol* **12**, 980-986.
- Zhang X, Hinton DR, Cua DJ, Stohlman SA & Lai MM (1997) Expression of interferon-gamma by a coronavirus defective-interfering RNA vector and its effect on viral replication, spread, and pathogenicity. *Virology* **233**, 327-338.
- Zhang X & Lai MM (1995) Interactions between the cytoplasmic proteins and the intergenic (promoter) sequence of mouse hepatitis virus RNA: correlation with the amounts of subgenomic mRNA transcribed. *J Virol* **69**, 1637-1644.
- Zhou M & Collisson EW (2000) The amino and carboxyl domains of the infectious bronchitis virus nucleocapsid protein interact with 3' genomic RNA. *Virus Res* **67**, 31-39.
- Zhou M, Williams AK, Chung SI, Wang L & Collisson EW (1996) The infectious bronchitis virus nucleocapsid protein binds RNA sequences in the 3' terminus of the genome. *Virology* **217**, 191-199.

Curriculum Vitae

Yuan Quan

Education:

- Nanyang Technological University, Singapore Aug 2002 - present
Ph.D. Candidate in virology
- Fudan University, China Sep 1996 - July 2000
B.A. in pharmaceutical chemistry

Working experience

HealthDigit Co., Ltd

Product Developer, Sep 2000 ~ July 2002

- In charge of the research of ELISA kit for 12 tumor marker antigens
- Work with other research teams on the development of protein chips

Publication:

- Yuan, Q., Liao, Y., Torres, J., Tam, J. P., Liu, D. X., 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.
- Liao, Y., Yuan, Q., Torres, J., Tam, J. P., Liu, D. X., 2006. Biochemical and functional characterization of the membrane association and membrane permeabilizing activity of the severe acute respiratory syndrome coronavirus envelope protein. Virology 349, 264-275.
- Liu, D. X., Yuan, Q., Liao, Y., 2007. Coronavirus Envelope (E) Protein: a Small Membrane Protein with Multiple Functions. Cell. Mol. Life Sci. (Epub).