Interaction of SARS-CoV spike protein with lipid rafts of host cells during viral entry

Lu, Yanning

2007

Lu, Y. N. (2007). Interaction of SARS-CoV spike protein with lipid rafts of host cells during viral entry. Doctoral thesis, Nanyang Technological University, Singapore.

https://hdl.handle.net/10356/6585

https://doi.org/10.32657/10356/6585

Interaction of SARS-CoV Spike Protein with Lipid Rafts of Host Cells during Viral Entry

Lu Yanning

School of Biological Sciences

A thesis submitted to the Nanyang Technological University
in fulfilment of the requirement for the degree of
Doctor of Philosophy

2007

Acknowledgements

This project was performed at the School of Biological Sciences, Nanyang Technological University between October 2002 and September 2006 to fulfill the requirement for the PhD degree. I would like to express my deep gratitude to Prof. James P. Tam, my thesis supervisor who never got tired of explanations and discussions, and helped me to develop and to realize my goals. His positive spirit and guidance has been inspirational to me, not only in academic research but also in life. A special thank goes to Prof. Ding Xiang Liu from whom I benefited extensively from his expertise in virology. He could always be consulted when I ran into experimental troubles. I would also like to thank all the present and former staff of Professor Tam's laboratory, all the students at School of Biological Sciences and my friends for their help and support over these four years, and for creating an exciting atmosphere to carry out this research work.

Finally, I am deeply grateful to my family for their understanding, encouragement and support through good and not-so-good periods. Without their supports, this work would not have been possible.

Table of Contents

List of Tables

List of Figures

Abbreviations

ACE2: angiotensin-converting enzyme 2

AcMNPV: Autographa Californica multiple nucleopolyhedrovirus

AIDS: acquired immuno-deficiency syndrome

Ala: alanine

APN: aminopeptidase N

APS: ammonium persulfate

ATCC: American type culture collection

BCV: bovine coronavirus

BEVS: baculovirus expression vector system

BSA: bovine serum albumin

CEACAMs: carcinoembryonic antigen-cell adhesion molecules

CCV: canine coronavirus

CD4: cluster differentiation marker 4

CD13: cluster differentiation marker 13

CD71: cluster differentiation marker 71

CD209L: cluster differentiation marker 209L

3CLp: chymotrypsin-picornavirus 3C-like protease

CMV: human cyto-megalovirus

CP: cytoplasmic domain

cryoEM: cryo-electron microscopy

CT-B: cholera toxin B subunit

Cys: cysteine

DIGs: Detergent-insoluble glycosphingolipidenriched microdomains

DMEM: Dulbecco's Modified Eagle Media

DNA: deoxyribonucleic acid

dpi: days post-infection

DRMs: detergent-resistant membranes

DTT: dithiothreitol

EboV: Ebola virus

EBV: Epstein-Barr virus

ECL: enhanced chemiluminescence solution

ELISA: enzyme-linked immunosorbent assay

EM: electron microscopy

Env: Envelope protein

E protein: small envelope protein

ER: endoplasmic reticulum

FCV: feline coronavirus

FIV: feline immunodeficiency virus

FIPV: feline infectious peritonitis virus

FP: fusion peptide

gp160: glycoprotein 160

gp120: glycoprotein 120

gp41: glycoprotein 41

GPs: glycoproteins

GPI: glycosylphosphotidylinositol

GPLs: glycerophospholipids

GSLs: glycosphingolipids

HA: influenza hemagglutin

HBV: hepatitis B virus

HCoV-229E: human coronavirus 229E

HCoV-OC43: human coronavirus OC43

HE: hemagglutinin-esterase

HeBS: Hepes buffered saline

HEPES: 4-(2-hydroxyethyl)-1-piperazineethane-sulphonic acid

His: Histidine

HIV-1: human immunodeficiency virus type 1

HPV: human papillomavirus

HR: heptad repeat

HRP: horse radish peroxidase

IBV: infectious bronchitis virus

Ile: isoleucine

IMCB: Institute of Molecular and Cellular Biology of Singapore

IPTG: isopropyl-beta-D-thiogalactopyranoside

kDa: kilo Dalton

lc: liquid-crystalline

lo: liquid-order

L-SIGN: L-specific ICAM-grabbing non-integrin

MAb: monoclonal antibody

MBP: maltose-binding protein

MβCD: β-methyl-cyclodextrin

MHV: mouse hepatitis virus

MMLV: Moloney murine leukemia virus

moi: multiplicity of infection

MPER: membrane-proximal external region

M protein: membrane protein

MSD: membrane-spanning domain

N protein: nucleocapsid protein

NS: non-structural protein

2OHpβCD: 2-hydroxypropyl-β-cyclodextrin

ORFs: open reading frames

PBS: phosphate-buffered saline

PCR: polymerase chain reaction

PDZ: post synaptic density protein (PSD95), drosophila disc large tumor suppressor (DlgA), and zo-1 protein

PEDV: porcine epidemic diarrhea virus

Phe: phenylalanine

PHEV: haemagglutinating encephalomyelitis virus of swine

PLP: papain-like cysteine protease

Pol: polymerase

preTM: pretransmembrane region

PVDF: polyvinylidene fluoride

RBD: receptor-binding domain

RNA: ribonucleic acid

RT: reverse transcriptase

SARS: severe acute respiratory syndrome

SARS-CoV: SARS associated coronavirus

SDAV: syaloacryoadenitis virus of rats

SDS: sodium dodecyl sulfate

SDS-PAGE: SDS polyacrylamide gel electrophoresis

Ser: serine

SP: signal peptide

S protein: spike glycoprotein

SV40: simian virus 40

SV5: simian virus 5

TAE: Tris-acetate-EDTA

TCR: T-cell receptor

TCV: turkey coronavirus

TEMED: N,N,N',N'-Tetramethylethylenediamine

TGEV: porcine transmissible gastroenteritis virus

TM: transmembrane domain

Trp: tryptophan

TRSs: transcription-regulatory sequences

Tyr: tyrosine

VSV: vesicular stomatitis virus

VSV-G: vesicular stomatitis virus glycoprotein

Summary

SARS-CoV, a novel virus belonging to group IV coronavirus was discovered in association with cases of severe acute respiratory syndrome (SARS) in 2003. The spike glycoprotein (S) of coronavirus mediates viral entry. S protein contains distinct functional domains near the amino (S1) and carboxyl (S2) termini. The S1 subunit binds to cellular receptor(s) that triggers the integral membrane S2 subunit to mediate fusion and viral entry.

Importantly, the interactions between a virus and cell surface receptors must be specific and strong in successful infection. Lipid rafts are special membrane domains that are rich in sphingolipids, glycosphingolipids and cholesterol. In viruses such as SIV, HIV and MHV, lipid rafts are used as entry sites, either as signaling platforms or as concentration devices. The receptor for SARS-CoV has been identified as angiotensin-converting enzyme 2 (ACE2), the receptor-binding domain of S protein has been defined between residues 318-510 of S1 and the crystal structure of this domain complexed with ACE2 solved. However, there are still unknowns concerning whether SARS-CoV would require lipid rafts to cluster receptors and S protein as a specific and concentration device for its entry.

My study focuses on the interaction of SARS-CoV S protein with host cell membrane lipid rafts during viral entry. I have found that lipid rafts are involved in SARS-CoV entry into Vero E6 cells, a permissive cell line for SARS-CoV infectivity. The integrity of lipid rafts is required for productive infection of Vero E6 cells with pseudotyped SARS-CoV as determined by single-cycle infectivity assay. Depletion

of cholesterol with methyl-β-cyclodextrin (MβCD) relocalizes raft-markers and ACE2 to a nonraft environment, but does not affect the translocation of ACE2 onto the cell surface. MβCD-treatment inhibits productive infection of pseudotyped SARS-CoV by 90%. In contrast, infection with vesicular stomatitis virus envelope glycoprotein (VSV-G) pseudotyped virus, which enters the cells through a lipid rafts-independent pathway, is not suppressed by MβCD. Biochemical fractionation and confocal microscopy imaging of SARS-CoV receptor distribution in live cells confirm that ACE2 colocalizes with raft-markers, caveolin-1 and ganglioside GM1. S protein ectodomain could associate with lipid rafts after binding to its receptor, and colocalized with GM1. These results strongly support that lipid rafts serve as a port for SARS-CoV entry and bear similarity to the entry mechanism of HIV and MHV.

The Trp-rich region is absolutely conserved in members of coronaviruses and highly conserved in other RNA viruses such as HIV, FIV and EboV. The role of the Trp-rich region in SARS-CoV S protein in viral entry was further investigated. Mutations of Trp with Ala or Phe dramatically decrease the infectivity of pseudotyped SARS-CoV, suggesting the importance of this region in viral entry.

Finally, peptide library was prepared with overlapping peptides representing almost the entire length of S protein and screened for binding of S protein ectodomain. It was found that peptides capable of binding S protein ectodomain are largely limited to the exposed protein surfaces with ordered structure of α-helices and β-sheets that are located in receptor-binding domain and HR1 regions. These findings provide useful clues for understanding the entry mechanism and developing potential

therapeutics for SARS-CoV.

Chapter 1. Introduction

In 2003, a novel virus was discovered in association with cases of severe acute respiratory syndrome (SARS) (Drosten *et al.*, 2003; Ksiazek *et al.*, 2003; Peiris, Guan, and Yuen, 2004). This virus belongs to the coronavirus family, hence called SARS-CoV (Marra *et al.*, 2003; Rota *et al.*, 2003).

A. Virus

In 1898, virus was discovered by Friedrich Loeffler and Paul Frosch as the cause of foot-and-mouth disease in livestock (Loeffler F., 1897; Sobrino *et al.*, 2001). Virus was an infectious particle smaller than bacteria. They are metabolically inert outside of host cells and exist as a protein coat or capsid, sometimes enclosed within a membrane (Bamford, Grimes, and Stuart, 2005). After inserting its genetic material into host and taking over the host's functions, a virus becomes metabolically active.

Viruses cause a number of diseases in eukaryotes. In humans, the common cold (Wat, 2004), chickenpox (Hambleton, 2005), influenza (Adungo *et al.*, 2005), polio (Minor, 2004), ebola (Pourrut *et al.*, 2005) and AIDS (Mullins and Jensen, 2006) are examples of viral diseases. Even some types of cancer have been linked to viruses, such as hepatitis B virus (HBV), human papillomavirus (HPV) and Epstein-Barr virus (EBV) (Markowska *et al.*, 2005; Puscas, 2005; Serraino *et al.*, 2005).

More than 5,450 viruses in more than 2,000 species, 287 genera, 73 families and 3 orders are recognized (Van Regenmortel, 1999) and could be classified into seven groups (Table 1) based on the Baltimore classification system (Baltimore, 1971).

Table 1. The Baltimore Classification of Viruses

Group	Characteristics	Example
I Double-stranded DNA	Some replicate in the nucleus e.g. adenoviruses using cellular proteins. Poxviruses replicate in the cytoplasm and make their own enzymes for nucleic acid replication.	Adenoviruses Herpesviruses Poxviruses
II Single-stranded (+) sense DNA	Replication occurs in the nucleus, involving the formation of a (-) sense strand, which serves as a template for (+) strand RNA and DNA synthesis.	Parvoviruses
III Double-stranded RNA	These viruses have segmented genomes. Each genome segment is transcribed separately to produce monocistronic mRNAs.	Reoviruses Birnaviruses
IV Single-stranded (+) sense RNA	a) Polycistronic mRNA e.g. Picornaviruses; Hepatitis A. Genome RNA = mRNA. Means naked RNA is infectious, no virion particle associated polymerase. Translation results in the formation of a polyprotein product, which is subsequently cleaved to form the mature proteins. b) Complex Transcription e.g. Togaviruses. Two or more rounds of translation are necessary to produce the genomic RNA.	Picorna-viruses Togaviruses
V Single-stranded (-) sense RNA	Must have a virion particle RNA directed RNA polymerase. a) Segmented e.g. Orthomyxoviruses. First step in replication is transcription of the (-) sense RNA genome by the virion RNA-dependent RNA polymerase to produce monocistronic mRNAs, which also serve as the template for genome replication. b) Non-segmented e.g. Rhabdoviruses. Replication occurs as above and monocistronic mRNAs are produced.	Orthomyxo-viruses, Rhabdo-viruses
VI Single-stranded (+)sense RNA with DNA intermediate in life-cycle	Genome is (+) sense but unique among viruses in that it is DIPLOID, and does not serve as mRNA, but as a template for reverse transcription.	Retroviruses
VII Double-stranded DNA with RNA intermediate	This group of viruses also relies on reverse transcription, but unlike the Retroviruses, this occurs inside the virus particle on maturation. On infection of a new cell, the first event to occur is repair of the gapped genome, followed by transcription.	Hepadna-viruses

B. Coronavirus

First isolated from chickens in 1937, coronavirus is a diverse group of large enveloped viruses that exhibit a marked tropism for epithelial cells of the respiratory and enteric tracts in humans and animals, such as cattle, pigs, rodents, cats, dogs and birds (Weiss and Navas-Martin, 2005). They belong to group IV single-strand (+) sense RNA virus (Baltimore classification). In 1965, Tyrrell and Bynoe (Tyrrell and Bynoe, 1965) used cultures of human ciliated embryonic trachea to propagate the first human coronavirus *in vitro*. There are now more than 15 species in this family. Some are serious veterinary pathogens, especially for chickens.

Coronaviruses belong to *nidovirales* order which consists of two families: Coronaviridae and Arteriviridae. Coronaviridae contains two genera - *coronavirus* and *torovirus* - and a possible third, *arterivirus* (Mackie, 2003). The distinctive characteristic of *nidovirales* is that they produce a nested set of mRNA with a common 3' end (Gorbalenya *et al.*, 2006). Coronavirus family has been divided into four groups on the basis of genomic sequence homology. Group I coronaviruses (canine, feline coronavirus, feline infectious peritonitis, porcine transmissible gastroenteritis and porcine respiratory viruses, human coronavirus 229E), and group II (bovine coronavirus, murine hepatitis virus, rat sialodacryoadenitis viruses, human coronavirus OC43) contain mammalian viruses, while group III contain only avian viruses (avian infectious bronchitis, turkey coronavirus). The recently discovered SARS-CoV forms group IV, whereas phylogenetic analyses (Stadler *et al.*, 2003) indicated that SARS-CoV and group II coronaviruses are closely related and may

share a common ancestor (Figure 1).

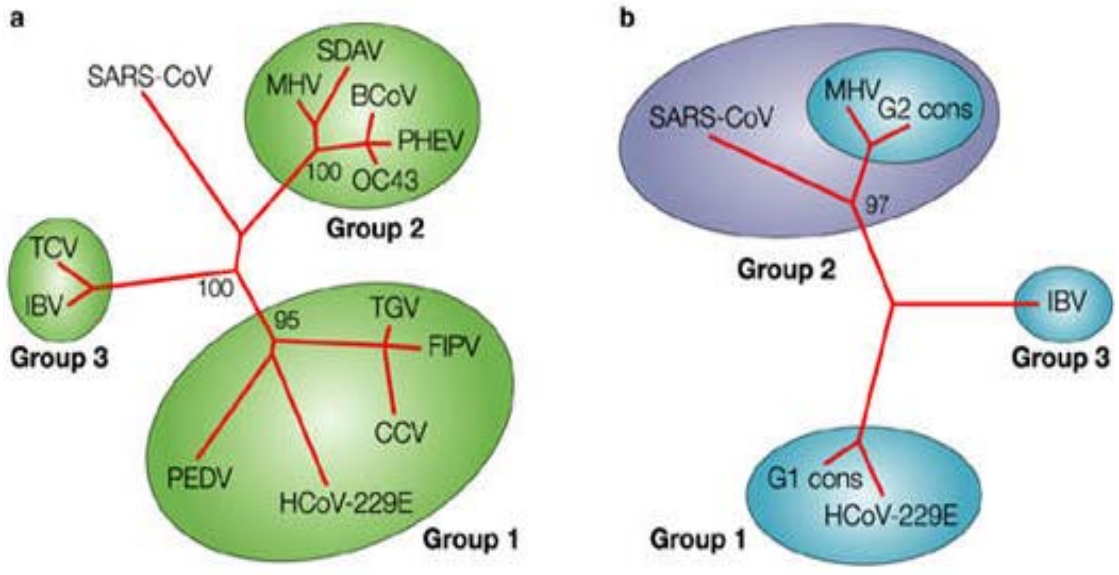

Figure 1. Relationship between SARS-CoV and other coronaviruses using different phylogenetic strategies (Taken from Stadler *et al*., 2003).

a. Unrooted tree obtained by comparing the well-conserved polymerase protein sequence. According to this approach, SARS-CoV belongs to a new group. The tree has been constructed using the protein sequences of the RNA-dependent RNA polymerase of the following coronaviruses: porcine epidemic diarrhea virus (PEDV), human coronavirus 229E (HCoV-229E), canine coronavirus (CCV), feline infectious peritonitis virus (FIPV), transmissible gastroenteritis virus (TGV), mouse hepatitis virus (MHV), bovine coronavirus (BCV), syaloacryoadenitis virus of rats (SDAV), human coronavirus OC43 (HCoV-OC43), haemagglutinating encephalomyelitis virus of swine (PHEV), turkey coronavirus (TCV), avian infectious bronchitis virus (IBV) and SARS-CoV.

b. Tree obtained using the sequence of the S1 domain of the spike protein. The multiple sequence alignment was constructed using consensus sequences generated from group 1 and group 2 coronaviruses (G1 cons and G2 cons), the sequence of IBV (group 3) and of SARS-CoV. The neighbor-joining algorithm was used to build the tree (Saitou and Nei, 1987).

Numbers represent the result of a bootstrap analysis performed with 100 replicates.

i. Coronavirus morphology

The coronavirus family is named for its crown-like (corona = crown) appearance of the surface projections due to a large spike glycoprotein when viewed by electron microscopy (Figure 2). Coronaviruses are the largest RNA viruses with the virion size ranging from 75 nm to 160 nm in diameter. Their RNA genome is associated with the nucleocapsid protein N to form a long, flexible, helical nucleocapsid. The nucleocapsid lies within a lipoprotein envelope which is formed during budding from intracellular membranes but not the plasma membrane (Griffiths and Rottier, 1992).

The envelopes of all coronaviruses contain three different viral-specified structural glycoproteins, spike protein (S), membrane protein (M) and small envelope protein (E). S protein (180-220 kDa) is a trimeric membrane protein and gives rise to the typical club-shaped projections (approximately 20 nm in length) of the virion membrane (Lai and Cavanagh, 1997). Such 20-40 nm surface projections also surround the periphery of SARS-CoV particles (Ksiazek *et al.*, 2003).

ii. Coronavirus genome and genome map of SARS-CoV

The genome of coronaviruses is about 27 to 32 kb non-segmented RNA (Lee *et al.*, 1991). It is capped, polyadenylated and coupled with RNA polymerases. These polymerases lack proof reading capacities and lead to the high mutation frequency in coronaviruses (Lai, 1992). Several coronaviruses, including the SARS-CoV (Marra *et al*., 2003), have been sequenced and conservation at genomic level was found. At the 5' end of coronaviruses genome is the polymerase (pol) followed by four structural proteins: S, E, M and N proteins, and at the 3' end is a poly A tract.

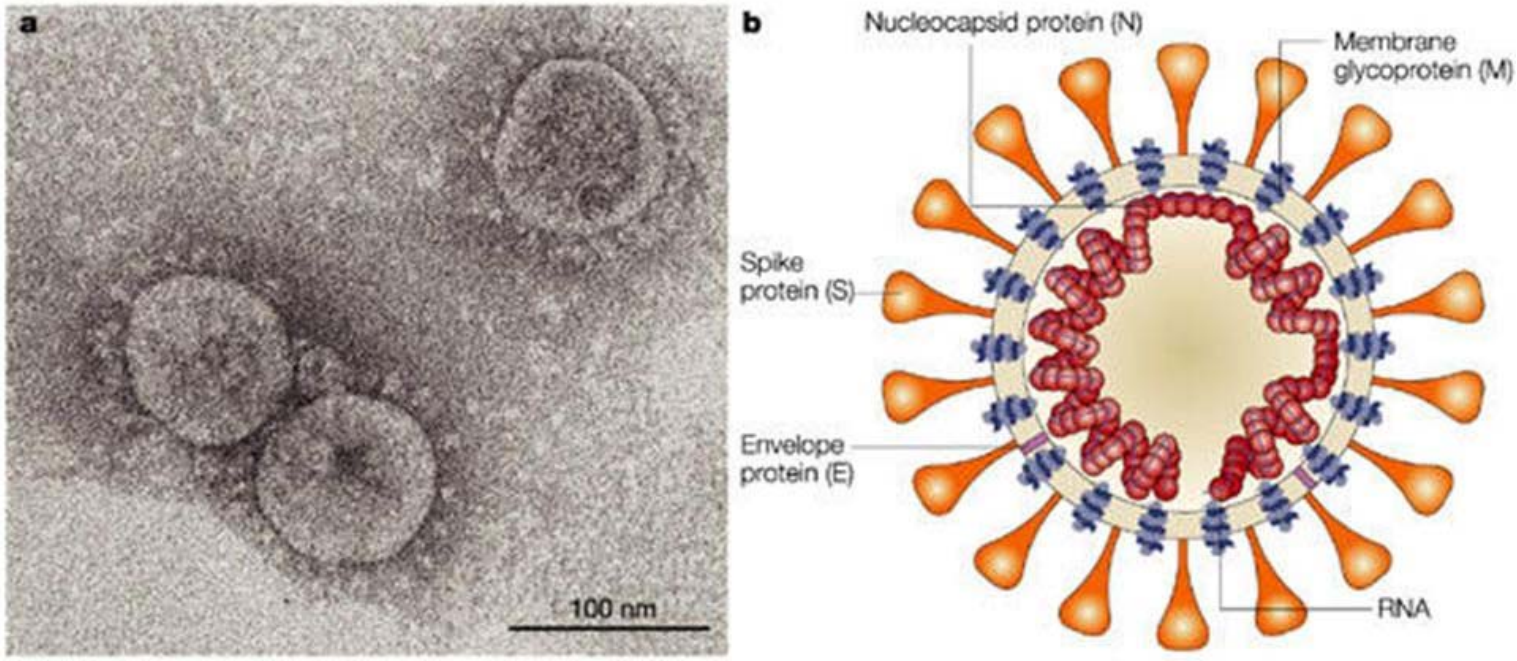

Figure 2. Morphology of the coronavirus (Taken from Stadler *et al.*, 2003).

a. Electron micrograph of SARS-CoV virus that was cultivated in Vero cells (Image courtesy of Dr L. Kolesnikova, Institute of Virology, Marburg, Germany). Large, club-shaped protrusions consisting of spike protein form a crown-like corona that gives the virus its name. **b.** Schematic representation of the coronavirus. A lipid bilayer comprising the spike protein, the membrane glycoprotein and the envelope protein cloaks the helical nucleocapsid, which consists of the nucleocapsid protein that is associated with the viral RNA. In the case of coronaviruses, the lipid envelope is derived from intracellular membranes.

The first three structural proteins are incorporated into intracellular membranes of host cells (particularly the Golgi body). N protein is the fourth major structural protein which interacts with M protein as a core to hold virion together (Narayanan *et al.*, 2000). Group II coronaviruses including human respiratory coronavirus (HCoV-OC43), also have a hemagglutinin-esterase (HE) gene between the pol and S gene (Klausegger *et al.*, 1999). HE is a membrane protein which may be involved in virion entry or release (Brian, 1995).

All coronaviruses also produce additional small non-structural proteins (NS). Their number, location, size, sequence, and function vary considerably between viruses. In addition to the protein-coding genes in the genomic RNA, all coronaviruses have 7-base sequences called intergenic sequences that are at the 5' end of each gene (Zuniga *et al.*, 2004). If the intergenic sequence is altered, the sub-genomic mRNA that starts at this point is not made.

SARS-CoV genome (Figure 3) contains five major open reading frames (ORFs) that encode the replicase, S, E, M and N proteins (Marra *et al.*, 2003). ORF 1a/b polymerase is predicted to be a 700-800 kDa precursor protein which is produced from two ORFs, 1a and b (Marra *et al.*, 2003). They are joined by a frameshift codon, and function as viral RNA-dependent RNA polymerase. The primary gene product is autocatalytically cleaved co/post-translationally to give a number of smaller products that contain polymerase, helicase domains and at least one PLP (papain-like cysteine protease) domain (two in some viruses), and a 3CLp (chymotrypsin-picornavirus 3C-like protease) domain.

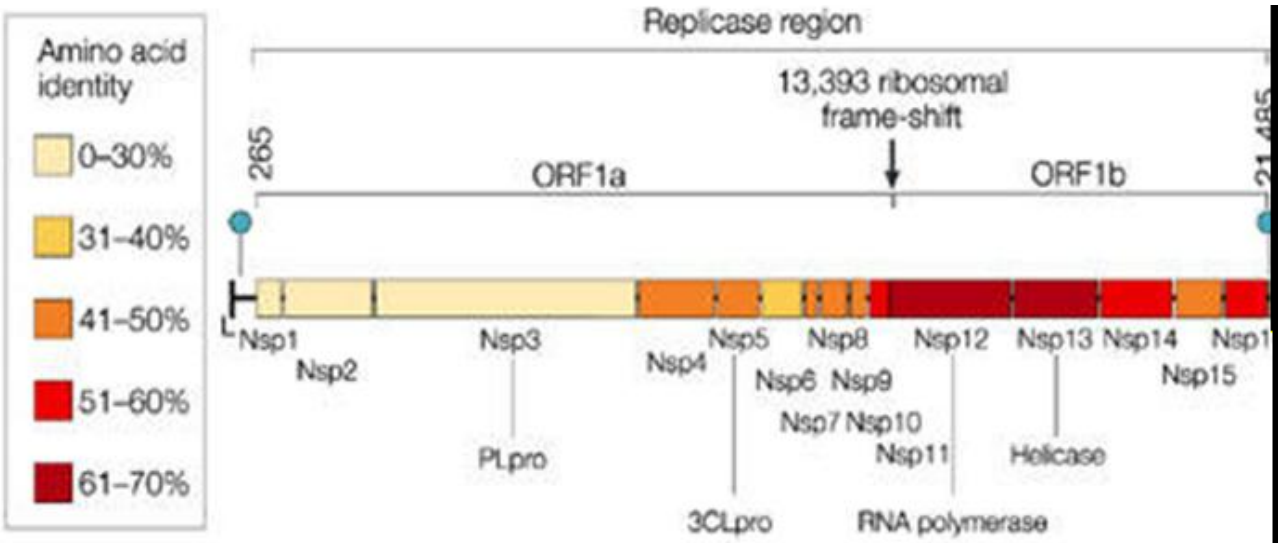

Figure 3. The genome structure of SARS-CoV (Taken from Stadler

Replicase and structural regions are shown together with the predi
ORF1a and ORF1b. The position of the leader sequence (L), the
ribosomal frameshift site between ORF1a and ORF1b are also indica
protein product (NS, non-structural protein). Colours indicate the le
with the best-matching protein of other coronaviruses. The SARS
white. Filled circles indicate the positions of the nine transcription-re
that are specific for SARS-CoV (5'ACGAAC3').

iii. Coronavirus life cycle

The life cycle of a coronavirus (Figure 4) starts when the S protein interacts with receptor(s) through its N-terminal domain (S1, see Figure 5). The entry, which is probably mediated by the C-terminal domain (S2), occurs by fusion of the viral membrane with cellular membrane (Hofmann and Pohlmann, 2004) through acidic pH-dependent endocytosis or pH-independent fusion. Uncoating then takes place. The positive single-strand RNA genome is then released into the cytoplasm where replication occurs. The virus-directed host machinery translates the overlapping open reading frames ORF 1a and ORF 1b by a ribosomal frame-shifting mechanism to produce a single polyprotein (Snijder *et al.*, 2003). Cleavage by virally encoded proteases yields the components that are necessary to assemble the viral replication complex, which synthesizes full-length negative-strand RNA as template for genomic RNA synthesis. In addition, a discontinuous transcription strategy during negative-strand synthesis produces a nested set of sub-genomic negative-sense RNAs (Antoine A. F. de Vries and Peter J. M. Rottier, 1997) as templates for the synthesis of mRNAs to produce viral proteins. N protein and genomic RNA assemble in the cytoplasm to form the helical nucleocapsid (Narayanan and Makino, 2001). There is a specific sequence near the 3' terminus of ORF 1b of the genome which facilitates assembly of the genomic RNA into virus particles (Narayanan *et al.*, 2003; Narayanan and Makino, 2001; Woo *et al.*, 1997). The assembly of MHV genomic RNA is nucleocapsid independent (Narayanan *et al.*, 2000), but whether the assembly of SARS-CoV is nucleocapsid dependent remains unclear (Hsieh *et al.*, 2005). This

core structure acquires its envelope by budding through intracellular membranes between the endoplasmic reticulum (ER) and the Golgi apparatus (Nal *et al.*, 2005) to maturate and release from the host cell.

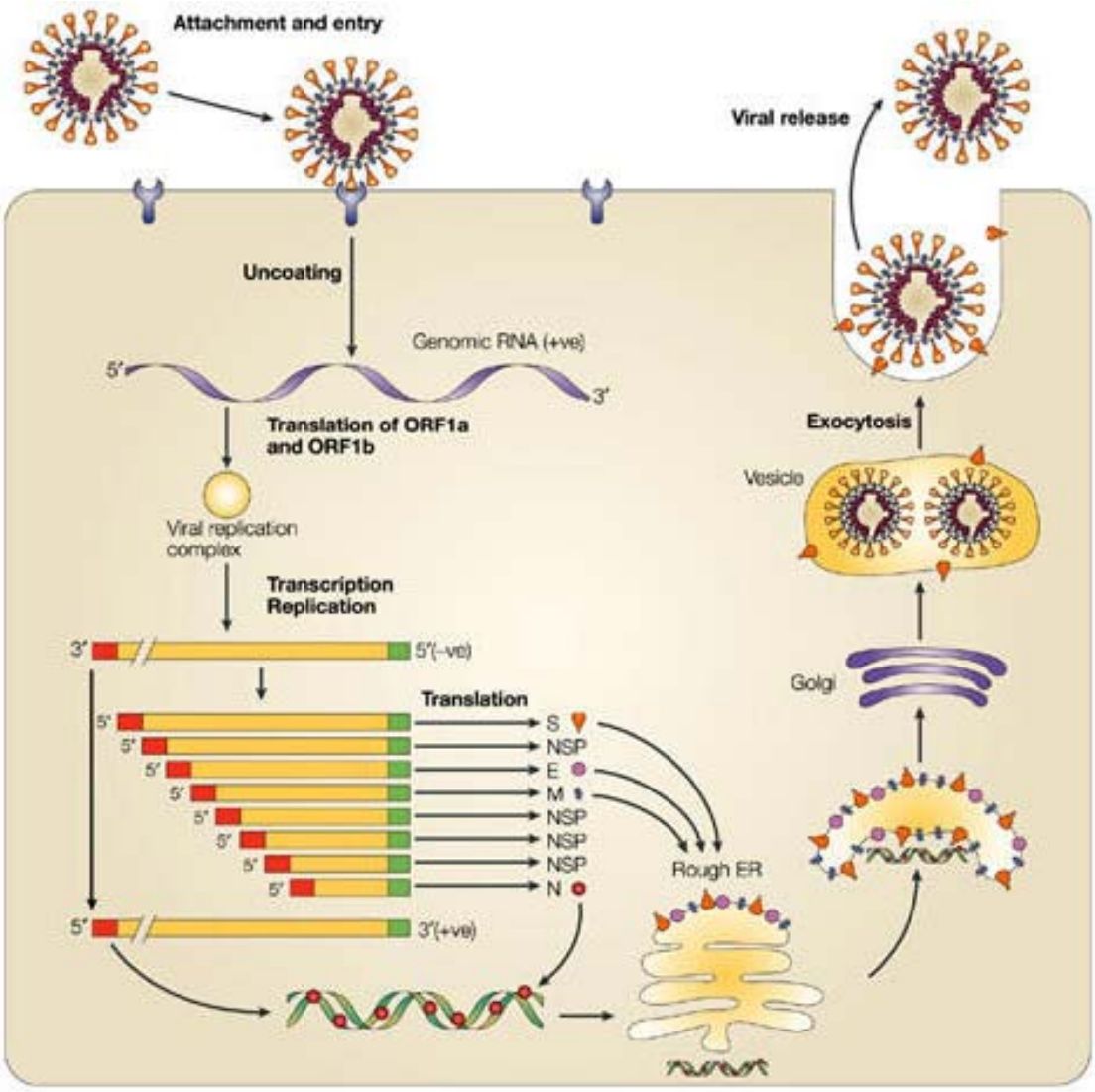

Figure 4. Life cycle of coronavirus (Taken from Stadler *et al*., 2003).

The viruses attach to and enter into the cell, then replicate using the host cell machinery. The figure shows the positive-sense mRNA products. The membrane (M), envelope (E) and spike (S) proteins, all of which will be lipid bilayer associated, are transported through the ER to the budding compartment, where the nucleocapsid probably interacts with the M protein to trigger assembly. During the transport of the virus through the Golgi apparatus, sugar moieties are modified. In some, but not all coronaviruses, the S protein is cleaved into S1 and S2 domains. Any S protein that is not incorporated into the virions is transported to the cell surface. Finally, the virus is released from the host cell by fusion of virion-containing vesicles with the plasma membrane. The common leader sequence on the 5′ end of each mRNA is shown in red.

iv. Spike protein of coronavirus

S protein is the outermost component of the virion, and is crucial for the entry of coronaviruses into host cells. Sequence variation is more frequent in S protein than in any other structural proteins in coronavirus species. Variations of S protein among strains of coronaviruses are responsible for host range and tissue tropism. Much of the species-specificity of infection depends on specific receptor interactions.

S protein is responsible for the attachment of virus to host cells (Gallagher and Buchmeier, 2001; Godet *et al.*, 1994; Kubo, Yamada, and Taguchi, 1994) and for instigating the fusion of the virus envelope with cell membrane. These properties result in S protein being a primary target for the host's immune responses. A number of studies have shown that S protein does contain epitopes to which neutralizing antibodies and cell-mediated cytotoxicity can be directed, and immunization of animals with S protein alone can induce protection from some coronaviruses (Ignjatovic and Galli, 1994; Torres *et al.*, 1995; Wang, Hong, and Seak, 2002).

S protein is relatively large, ranging from 1160 amino acids in infectious bronchitis virus (IBV) to 1452 amino acids in feline coronavirus (FCV). There are many potential N-linked glycosylation sites (21 to 35), most of which have glycans attached. S preprotein has an N-terminal signal peptide and a membrane-anchoring domain near the C-terminus. The S protein of SARS-CoV is ~200 kDa (Krokhin *et al.*, 2003) and contains 1255 amino acids with a 13-residue N-terminal signal peptide sequence. There are 23 potential N-glycosylation sites (Rota *et al.*, 2003), and most of them are located in the surface-exposed S1 subunit.

S protein may be cleaved into S1 and S2 subunits; the extent of its cleavage varies greatly among virus species. Whereas a high proportion of S protein is cleaved in IBV, MHV, BCV, TCV and PEDV (Jackwood *et al.*, 2001); none is cleaved in others: TGEV, FCV and CCV. It is uncertain whether the S protein of HCoV-229E and HCoV-OC43 are cleaved into S1 and S2 proteins. Recent studies show that most of the SARS-CoV S protein is not cleaved to any significant degree, although smaller proteins can be observed by SDS-PAGE analysis (Xiao *et al.*, 2003). Cleavage generates two glycoproteins, N-terminal S1 and C-terminal S2, the latter being acylated in cytoplasmic tail by palmitoylation (Thorp *et al.*, 2006). S1 is probably linked to S2 subunit by noncovalent linkage. Trypsin treatment of MHV virions causes cleavage of all S proteins without disrupting the spikes, however, S1 can be released from virions by either urea or mild alkali treatment (Weismiller *et al.*, 1990).

Among coronaviruses, S2 domain is much more conserved than S1. Regions of up to 30% amino acid identity (particularly in the transmembrane domain) exist between S2 proteins of different groups of coronaviruses, whereas there is almost no conservation in S1 sequence. The hypervariable regions of frequent deletions, mutations and recombinations suggest that this region is externally exposed and affects the virulence, pathogenesis and host cell tropism (Bos *et al.*, 1995; Haijema, Volders, and Rottier, 2003; Leparc-Goffart *et al.*, 1998; Sanchez *et al.*, 1999; Tsai *et al.*, 2003). S1 domain is predicted to form the globular portion of spikes, consistent with its highly variable nature. While the highly conserved sequence of S2 may play a role in forming the stalk with a more rigid structure (Figure 5).

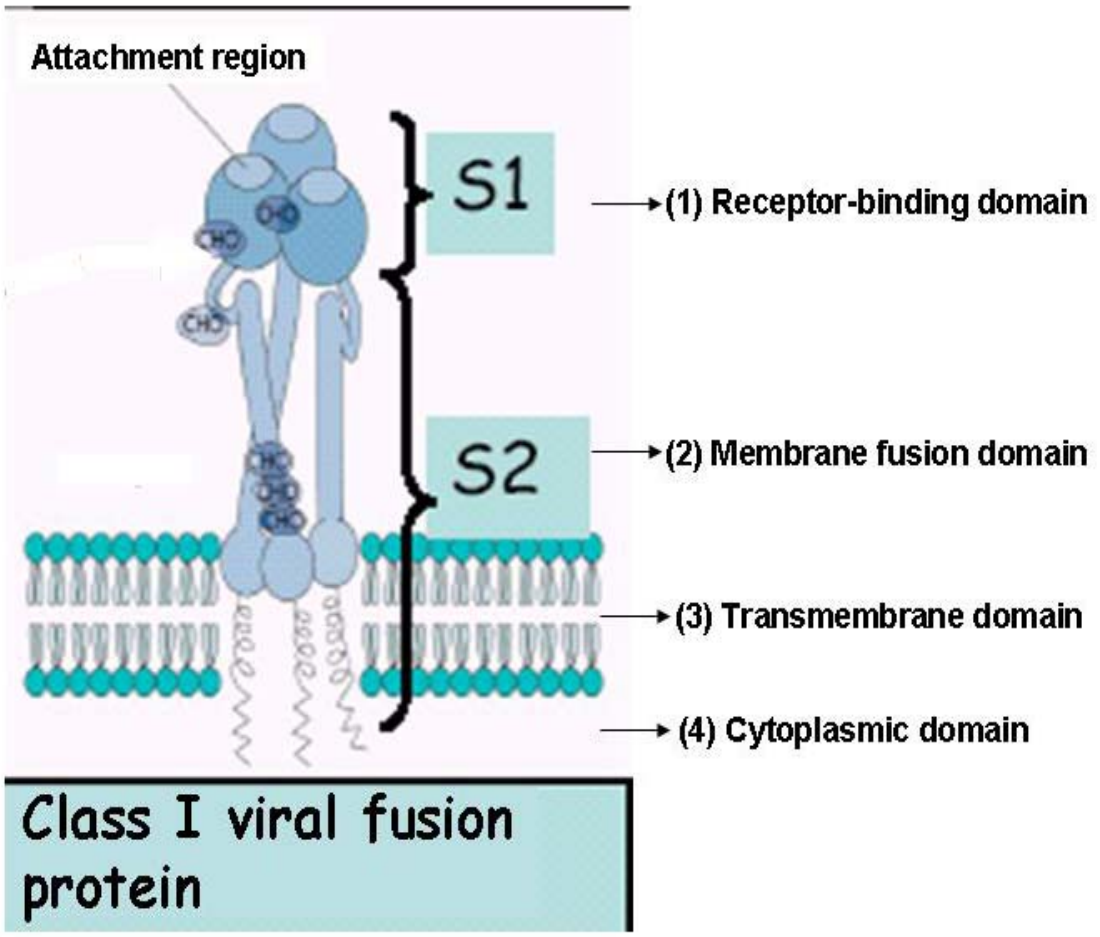

Figure 5. Schematic model of coronavirus S protein (Adapted from http:// www.sh.lsuhsc.edu/ intragrad/micro/Micro276.ppt.)

The coronavirus S protein is an about 180~220 kDa, heavily glycosylated protein. It is a class I viral fusion protein. S protein has four structural domains (from N to C-terminus): (1) S1 is receptor-binding domain which forms globular head. (2) S2 ectodomain is variable in length between isolates, containing membrane fusion region and forming external stalk. (3) Transmembrane domain. (4) Cytoplasmic domain. S protein forms trimer and exists on the viral surface. In some cases, such as IBV, TCV, BCV, MHV and PEDV, S protein is cleaved into S1 and S2, which remain associated. But in some coronaviruses, such as TGEV, FCV and CCV, S protein is not cleaved.

The coronavirus S2 protein has two Ser/Ile repeat regions with a seven-residue periodicity. These heptad repeats (HR) are predicted to form a coiled-coil structure known as a leucine zipper (Britton, 1991; Singh, Berger, and Kim, 1999). Indeed, current evidence suggests that mature S protein forms a trimer (Delmas and Laude, 1990). The trimeric S protein anchors in the viral membrane by an α-helical region near to the C-terminus of S2. Just beyond the outer membrane surface is the shorter repeat structure predicted to be an α-helix of 5-7 nm (HR2). The major repeat indicates a helix of 10-13 nm, which may form the narrow stalk of the spikes (HR).

In addition, all coronavirus S2 proteins have a highly conserved ten-residue sequence Y(V/I)KWPW(W/Y)VWL (Figure 6), which is rich in tryptophan (Trp) residues, the last five residues of which probably form the beginning of the membrane-spanning domain which is also called transmembrane (TM) domain, (Rota *et al.*, 2003; Sainz *et al.*, 2005). In the literature, this Trp-rich region has been referred to as aromatic domain, membrane-proximal external region (MPER), proximal-membrane region, or pretransmembrane region (preTM). This region will be referred to as the Trp-rich region in this study.

The functions of the Trp-rich region are intriguing and remain unknown. Similarly Trp-rich regions also exist in the TM of all lentiviruses (Figure 7), although they may differ somewhat with regard to the number of Trp residues contained, the length of the sequence in which these are interspersed, the properties of the other amino acids present, and the distances within the linear sequence of the Trp-residues between themselves and with the putative membrane-spanning domains.

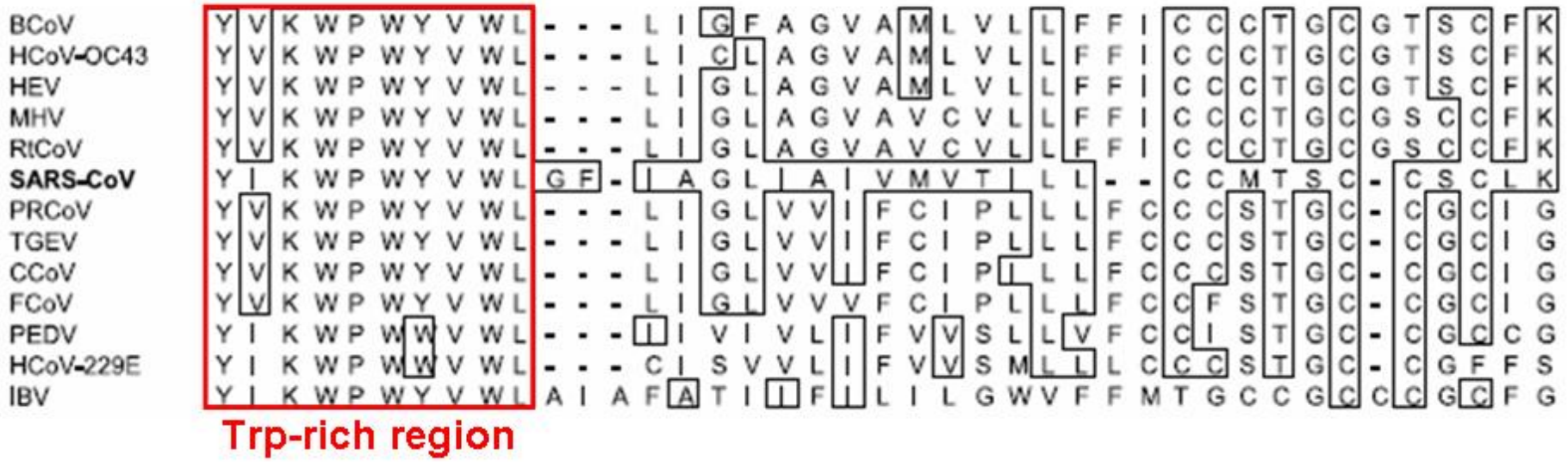

Figure 6. Conserved motifs in coronaviruses S protein (Adapted from Rota *et al.*, 2003).

Alignment of the C-terminal region of the SARS-CoV and reference coronavirus S proteins was generated with ClustalX 1.83. Residues that match the SARS-CoV sequence exactly are boxed. The red box indicates the amino acid sequence Y(V/I)KWPW(Y/W)VWL which is a conserved motif in all three coronavirus groups and SARS-CoV. This region is referred to as Trp-rich region in this study.

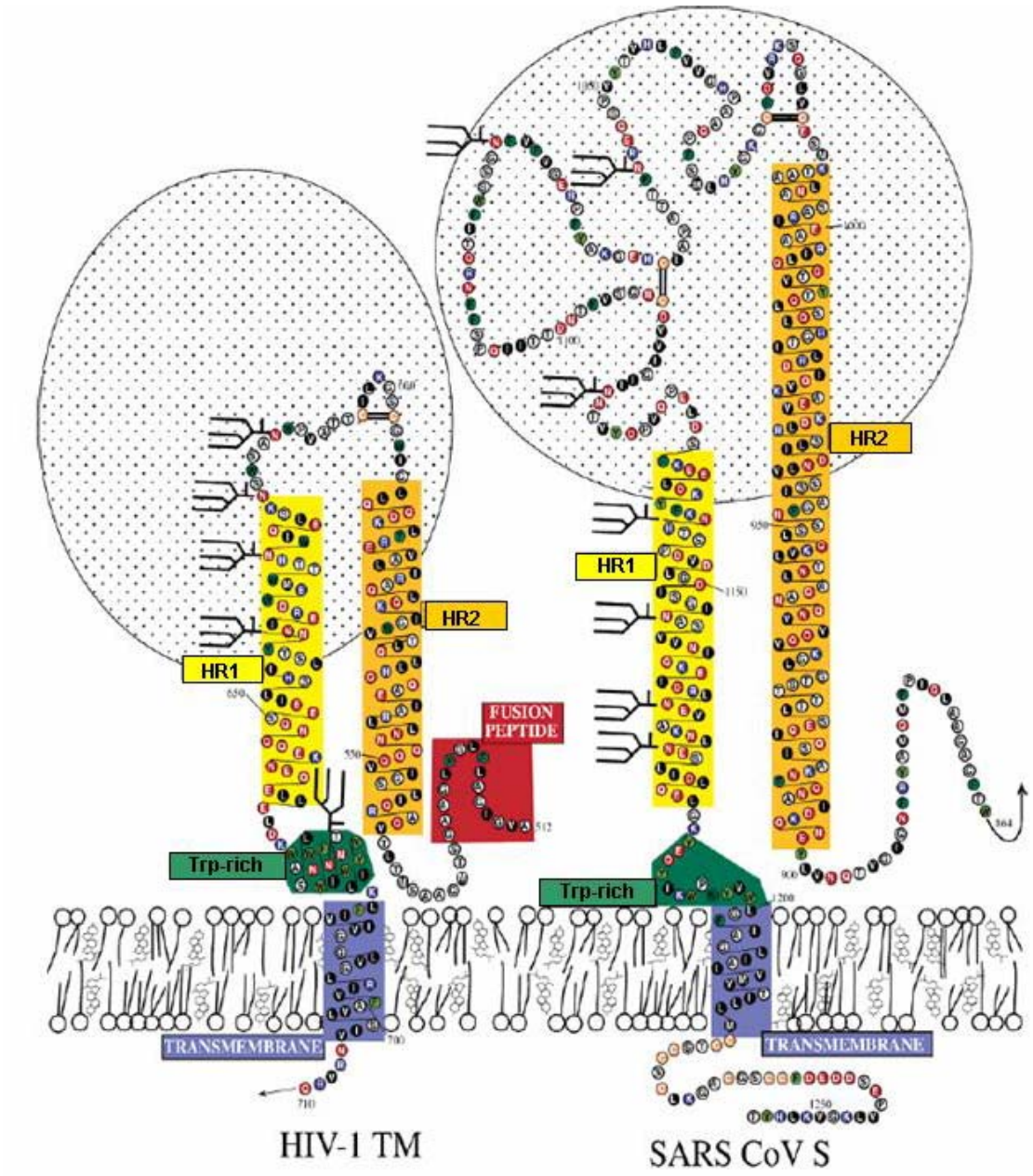

Figure 7. Schematic comparison of the HIV-1 TM with the SARS-CoV S protein (Adapted from Sainz *et al*., 2005).

Model of HIV-1 TM (Left). Hypothetical model of SARS-CoV S protein S2 subunit (Right) showing motifs shared with HIV-1 TM. The N-terminus of the S2 portion is not depicted. The two HR α-helical regions, HR1, (orange) and HR2, (yellow) are indicated. An interhelical domain (loop) is depicted between HR1 and HR2. This region is extremely similar to the interhelical region for retrovirus TM proteins and EboV GP2. Just prior to the TM anchor (indigo) of S2, there is a region enriched in aromatic amino acids. This region, termed Trp-rich region (green) is highly conserved throughout the coronaviruses and lies in an identical location to the Trp-rich regions of HIV and EboV.

According to current models for HIV entry (Chan and Kim, 1998; Doms and Moore, 2000; Gallo *et al.*, 2003; Lawless *et al.*, 1996; Munoz-Barroso *et al.*, 1998), and the high preference of Trp residues for residing at the external face of membranes (Schibli, Montelaro, and Vogel, 2001; Suarez *et al.*, 2000a; Suarez *et al.*, 2000b; Wimley and White, 1996; Yau *et al.*, 1998), the Trp-rich region appears to reside on the envelope surface at the membrane-water interface of the lipid bilayer. Following activation, the Trp-rich region is believed to undergo sequential conformational changes, from a reverse turn to an amphipathic helical structure capable of intimately interacting with the viral membrane through certain Trp residues and, concomitantly, with the cell membrane through others (Barbato *et al.*, 2003). With their relatively bulky indole side chains, the Trp residues, most likely synergizing with the fusion peptide (Suarez *et al.*, 2000a), would then destabilize both membranes and drive the energetically unfavorable lipid merging. This interaction permits the formation and expansion of the fusion pore in the late stages of the entry process (Melikyan *et al.*, 2000; Peisajovich and Shai, 2003). The putative functions of the Trp-rich region of SARS-CoV S protein are examined in Chapter 4.

Infection of host cells by coronavirus is initiated by the interaction of S1 protein with receptor(s) on the cell surface. Binding of S1 domain to cell surface receptor(s) induces a cascade of conformational changes, leading to membrane fusion and delivery of the genome into the cytoplasm (Dimitrov, 2000). The S protein therefore has two important functions: receptor-binding and inducing membrane fusion.

The S1 domains of all characterized coronaviruses mediate an initial high-affinity

association with their respective receptors (Bonavia *et al.*, 2003; Breslin *et al.*, 2003). Recently, the receptor-binding domain of SARS-CoV S protein has been located at the region between residues 303 and 537 (Xiao *et al.*, 2003). Wong *et al.* (Wong *et al.*, 2004) found that a 193-amino acid fragment (aa 318-510) of S1 subunit efficiently binds to the receptor. Li *et al.* (Li *et al.*, 2005) solved the crystal structure of SARS-CoV S protein receptor-binding domain complexed with its receptor angiotensin-converting enzyme 2 (ACE2), and found critical changes at positions 479 and 487 aa of S protein facilitate efficient cross-species infection and human-to-human transmission.

A major function of coronavirus S2 protein is membrane fusion. S2 contains functional elements involved in membrane fusion. Notably, glycoproteins (GPs) of highly divergent viruses, including HIV, SIV, influenza virus and EboV etc. exhibit a comparable architecture. These GPs, referred to as class I viral fusion proteins, use similar mechanisms to promote membrane fusion, which has important implications for therapeutic intervention.

The three dimensional structure of full-length S protein has not yet been defined. However, the crystal structure of the most important domains of S protein, such as receptor-binding domain and HR1-HR2 fusion core have already been solved (Li *et al.*, 2005; Xu *et al.*, 2004b), and the sequence similarities to other well-studied viruses have allowed modeling and predictive analysis. The schematic representative of SARS-CoV S protein is shown in Figure 8. Like other coronaviruses S protein, it also contains receptor-binding domain, fusion peptide region, HR1 and HR2 region,

TM domain and cytoplasmic domain (CP).

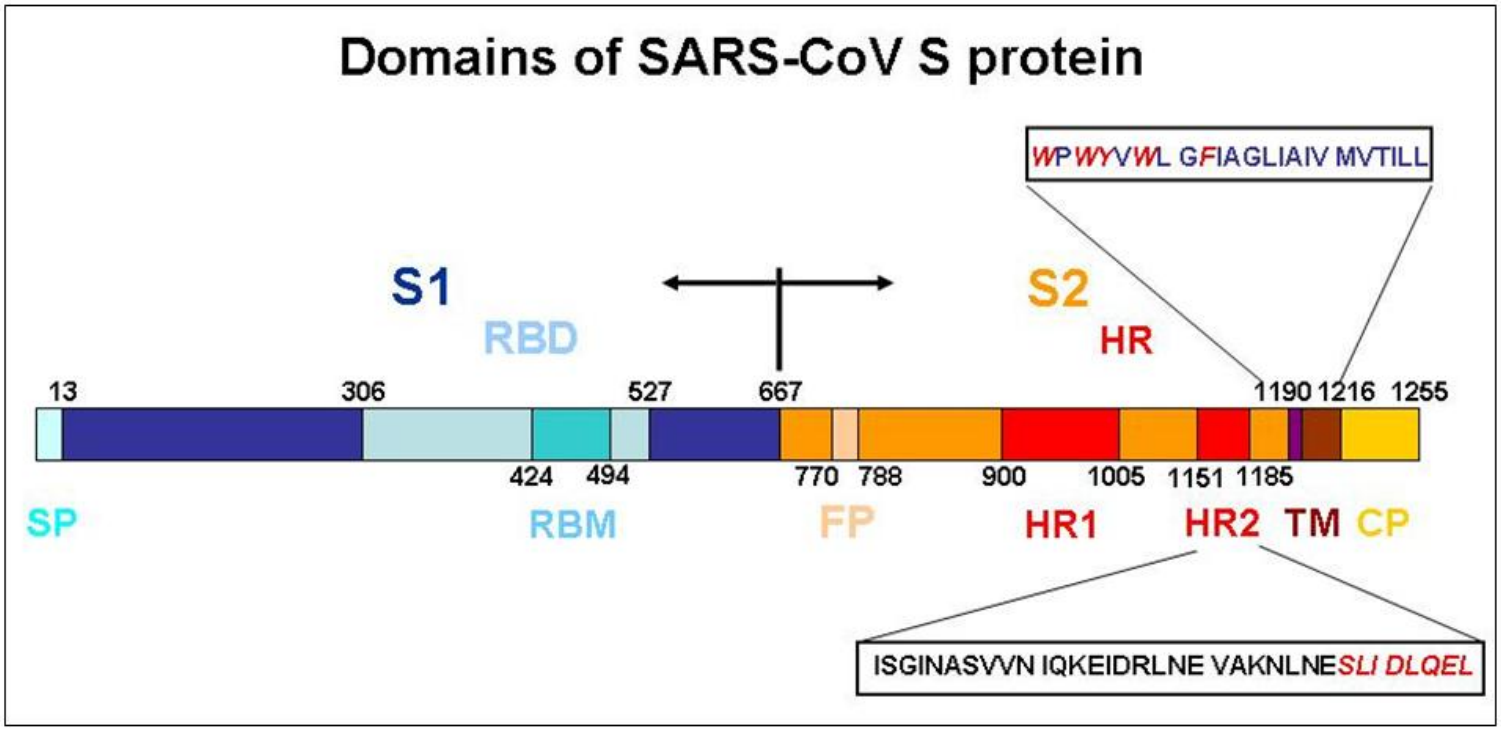

Figure 8. Schematic representative of SARS-CoV S protein (Adapted from Li *et al.*, 2005)

Modeling and predictive analysis of SARS-CoV S protein. SP: signal peptide, RBD: receptor-binding domain, RBM: receptor-binding motif, FP: fusion peptide, HR: heptad repeat, TM: transmembrane domain, CP: cytoplasmic domain. Residue numbers of each region correspond to their positions in S protein of SARS-CoV. The Trp-rich region is indicated in purple box. The italic red color labeled-residues in Trp-rich region and HR2 region are targeted for mutation and truncation in this study.

v. Receptors for coronaviruses

The distribution of coronavirus receptors is critical to the pathogenic outcome and coronavirus S proteins exhibit a range of receptor specificities. MHVs enter after binding carcinoembryonic antigen-cell adhesion molecules (CEACAMs) (Dveksler *et al.*, 1991); HCoV-229E, FCV, CCV and TGEV bind metalloproteases-aminopeptidase N (APN) (Benbacer *et al.*, 1997; Hegyi and Kolb, 1998; Tresnan, Levis, and Holmes, 1996). Li *et al*. (Li *et al.*, 2003) identified a metallopeptidase, ACE2 as a functional receptor for SARS-CoV. It has been suggested that SARS-CoV may use additional receptors because ACE2 is also found on number of cells that are not normally infected (Dimitrov, 2003). Jeffers *et al.* (Jeffers *et al.*, 2004) identified a different glycoprotein that can serve as an alternative receptor for SARS-CoV, CD209L, a C-type lectin (also called L-SIGN).

Zhu *et al*. (Zhu *et al*., 2006) examined the distribution and three-dimensional structure of AIDS virus envelope spikes using cryo-electron microscopy (cryoEM) tomography. They found that in HIV-1 and SIV, the Env spikes are not evenly distributed on the viral surface, but show some clustering patterns. They suggest the fusion efficiency might be dependent on local spike density (clustering). These findings shed light on our understanding of the class I virus fusion mechanism. Because of the similarity between SARS-CoV S protein and HIV Env glycoprotein (Kliger and Levanon, 2003), I hypothesize that SARS-CoV S protein may be clustered to gain efficient interaction with the receptors which may also concentrate in certain region of the host cell membrane. This hypothesis is examined in Chapter 3.

C. Current model of class I virus entry

The fusion glycoproteins from enveloped viruses belonging to the families of Orthomyxoviridae, Paramyxoviridae, Retroviridae, Filoviridae and Coronaviridae have been classified as class I fusion proteins based on their similarities in structure and function. In general, these viruses have a trimeric glycoprotein exposed on the surface. Their glycoproteins are either integral proteins (containing receptor-binding and fusion domains) or a separate trimeric fusion protein in complex with a separate receptor-binding protein on their surface. When these proteins are activated by their respective signals, such as a decrease of pH or binding to a specific cellular receptor, a conformational change is initiated. The result of this process is the fusion of viral with host-cell membranes (i.e. endosomal or plasma membranes) and infection ensues (Eckert and Kim, 2001; Skehel and Wiley, 2000; Wyatt *et al.*, 1998).

Influenza virus hemagglutin (HA) protein is a model for all class I virus fusion proteins. The most crucial conformational changes from the prefusion state to a fusion active state are the transition of a native HA2 loop region into a helical segment, and the reversal of the chain towards the end of the central triple-stranded coiled-coil structure. These two changes produce a typical six-helix bundle structure where the N-terminal end (fusion peptide) is located next to the C-terminal end (transmembrane anchor) (Bullough *et al.*, 1994; Chen, Skehel, and Wiley, 1999).

A number of fusion activated core structures from other enveloped viruses have been determined, such as gp41 from HIV and SIV (Caffrey *et al.*, 1998; Chan *et al.*, 1997; Malashkevich *et al.*, 1998; Tan *et al.*, 1997; Weissenhorn *et al.*, 1997), GP2

from EboV (Weissenhorn *et al.*, 1998), F1 from SV5 (Baker *et al.*, 1999) and F protein from respiratory syncytial virus (RSV) (Zhao *et al.*, 2000). Recently crystal structures of the fusion core of MHV and SARS-CoV have also been solved (Xu *et al.*, 2004a; Xu *et al.*, 2004b).

The common characteristic of all class I fusion protein structures is a central triple stranded coiled coil region with outer C-terminal anti-parallel layers, forming a trimer of helical hairpins (Dutch, Jardetzky, and Lamb, 2000; Eckert and Kim, 2001; Skehel and Wiley, 2000; Weissenhorn, 2002). This refolding process is thought to pull the two membranes into close proximity leading to membrane fusion by way of a number of lipidic intermediates including a fusion stalk, hemifusion diaphragm, fusion pore opening and expansion (Chernomordik and Kozlov, 2003; Chernomordik *et al.*, 1999). Numerous studies have shown that additional regions in the fusion proteins contribute to membrane fusion, including both membrane anchors, the glycine-rich fusion peptide itself, as well as the transmembrane and membrane-proximal regions (Salzwedel and Berger, 2000; Sanchez *et al.*, 1999; Schibli, Montelaro, and Vogel, 2001; Skehel and Wiley, 2000; Weber *et al.*, 1994; Weissenhorn, 2002; Wiley and Skehel, 1987).

Although debates remain for the proposed model (Figure 9), it is likely that SARS-CoV S protein mediates the membrane fusion in a similar way to cause infection. The proposed model provides a framework for my thesis work. Probing the putative protein-protein interactions of S protein functional domains as a result of

conformational change for developing potential reagents and inhibitors were studied by a new method in Chapter 5.

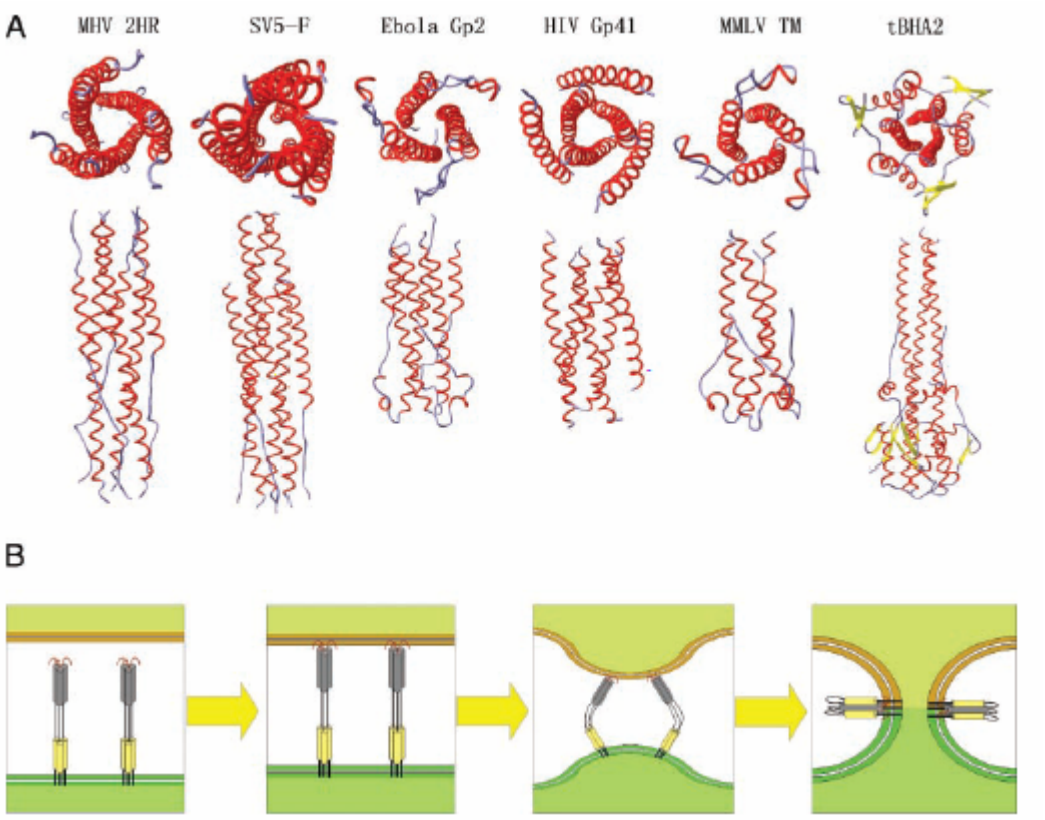

Figure 9. Viral fusion proteins and models for membrane fusion (Taken from Xu *et al*., 2004a).

A. Comparison of MHV fusion core with other viral fusion protein structures. The proteins under comparison include SV5F, EboV GP2, HIV gp41, MMLV Env-TM, and low pH-induced influenza virus HA, tBHA2. Top and side views are shown for the six fusion core structures. **B.** Model for coronavirus-mediated membrane fusion. The first state is the native conformation of coronavirus spike protein on the surface of viral membrane. It has been reported that the spike protein is trimeric in this conformation and about 200 Å in length (Davies and Macnaughton, 1979), but the exact structure of the full-length protein remains unknown. The second state is the prehairpin state of the S2 subunit. After several conformational changes, the fusion peptide inserts into the cellular membrane with the aid of other regions of S protein and possibly including the receptor. Although the internal fusion peptide is not exposed at the N-terminal of S2, it could insert into part of the target membrane by means of some hydrophobic residues. This insertion would be stable enough to drive the membrane motion with the conformational changes of HR1 region, which is adjacent to the fusion peptide. The third state is conformational change and juxtaposition of the target and viral membranes. With the help of other regions of S protein, the HR1 and HR2 regions move together and facilitate juxtaposition of the cellular and viral membrane. The last state is the postfusion conformation. The coiled coil will reorient with its long axis parallel to the membrane surface. The fused cellular and viral membranes make it possible for subsequent viral infections.

D. Lipid rafts and viral entry

Membrane-associated lipid rafts are currently an intensely investigated topic of cell biology. In addition to a demonstrated role in signal transduction of the host cell, lipid rafts serve as entry and exit sites for microbial pathogens, such as HIV, influenza virus, SV40 and measles virus.

i. Structure of lipid rafts

The fluid mosaic model of the plasma membrane has evolved considerably since its formulation 30 years ago. It is now accepted that membrane lipids do not form a homogeneous phase consisting of glycerophospholipids (GPLs) and cholesterol, but a mosaic of domains with unique biochemical compositions. Among these domains, those containing sphingolipids, glycosphingolipids (GSLs) and cholesterol are termed lipid rafts (Brown and London, 2000; Simons and Ikonen, 1997).

Lipid rafts possess unique physicochemical properties that direct their organization into liquid-order (lo) phase floating in a liquid-crystalline (lc) ocean of GPLs. These domains are resistant to detergent solubilization at 4°C but are destabilized by cholesterol- and sphingolipid-deleting agents (Ilangumaran and Hoessli, 1998; London and Brown, 2000). Lipid rafts have been morphologically characterized as small membrane patches that are tens of nanometers in diameter (Pralle *et al.*, 2000; Varma and Mayor, 1998; Zacharias *et al.*, 2002). Caveolae is a special form of lipid rafts domain. The schematic model of lipid raft and caveolae is shown in Figure 10.

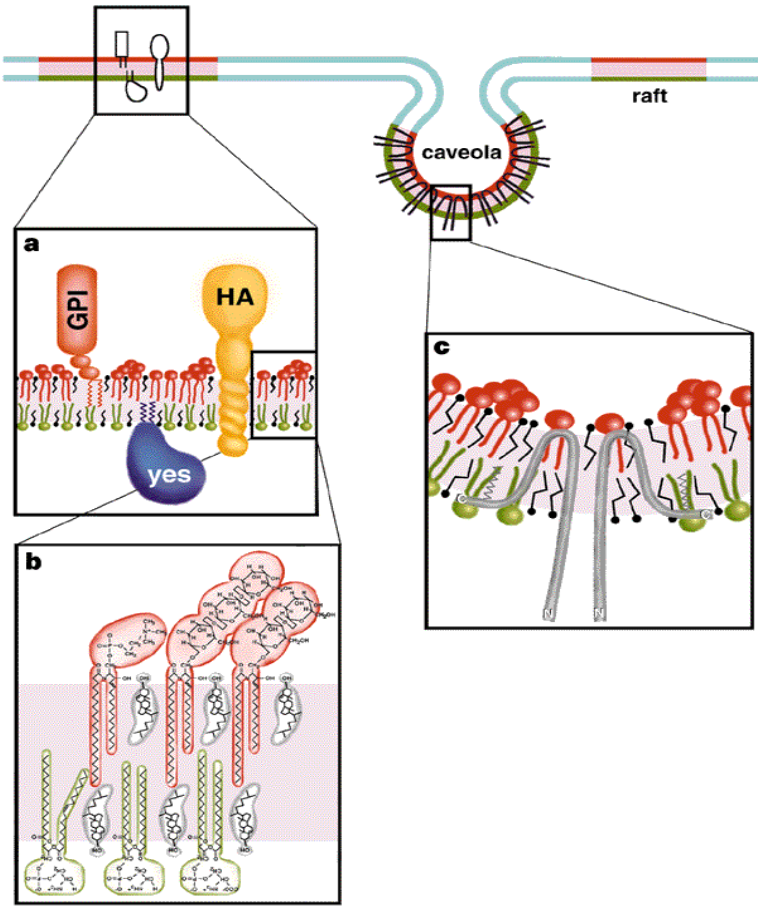

Figure 10. Schematic model of lipid rafts and caveolae (Taken from Simons and Ikonen, 1997).

The rafts (red) segregate from the other regions (blue) of the bilayer in which unsaturated phosphatidylcholine is predominantly in the exoplasmic leaflet, which has a different organization of intercalated cholesterol from that in the rafts.
a. Rafts contain proteins attached to the exoplasmic leaflet of the bilayer by their GPI anchors, proteins binding to the cytoplasmic leaflet by acyl tails (the Src-family kinase Yes is shown), or proteins associating through their transmembrane domains, like the influenza virus proteins neuraminidase and haemagglutinin (HA) (these proteins associate with DIGs in virus membranes). However, if cholesterol is removed by cyclodextrin extraction from the envelope, the virus glycoproteins become soluble in Triton X-100 at 4°C.
b. The lipid bilayer in rafts is asymmetric, with sphingolipids and glycosphingolipid (red) clustering in the exoplasmic leaflet and glycerolipids (e.g. phosphatidylserine and phosphatidylethanolamine; green) in the cytoplasmic leaflet. Cholesterol (grey) is present in both leaflets and intercalated between the predominantly saturated hydrocarbon chains of the sphingolipids.
c. Caveolae are formed by self-associating caveolin molecules making a hairpin loop in the membrane.

ii. Function of lipid rafts as a portal of entry for pathogens

As well as being utilized by cellular and/or exogenous proteins as transport shuttles on the cell surface, lipid rafts have been exploited as an entry point for pathogens. A broad range of pathogens, including viruses, parasites, bacteria and their toxins, use lipid rafts to enter host cells as an infection strategy, utilizing both cell-surface GPI-anchored proteins and raft lipids (GSLs, sphingolipid and cholesterol) as primary or accessory receptors (Brown, 2002; Shin and Abraham, 2001; van der Goot and Harder, 2001). The receptors required by several different viruses to achieve penetration of target cells are located in raft structures. For example, during attachment of HIV-1, lipid rafts are crucial elements in organizing the viral envelope and host cell receptors CD4 and the appropriate chemokine receptor into a membrane fusion complex leading to viral entry (Liao *et al.*, 2001; Manes *et al.*, 2000; Popik, Timothy M, and Au, 2002). The lipid rafts play an important role in this cluster formation by providing a microenvironment for virus-cell interactions to take place. Potential structural and functional similarities between SARS-CoV S protein and HIV-1 gp160 have led to the hypothesis that SARS-CoV enters the cell using a similar entry process as HIV-1, and lipid rafts may also play an important role in this process. In Chapter 3 of this study I investigated whether lipid rafts are involved in the SARS-CoV entry process.

E. Aims of this study

After the outbreak of SARS, an impetus was the characterization of, not only the SARS-CoV proteome but also other coronavirus proteomes, with the hope of

developing therapeutics and vaccines against highly infectious diseases. Because of its role in the attachment and fusion events, the SARS-CoV S protein has been initially identified as a major therapeutic and vaccine target. Although substantial efforts have been directed towards S protein, and the crystal structure of its receptor-binding domain and fusion core have been solved (Li *et al*., 2005; Xu *et al*., 2004b), much of its characteristics, functions and details concerning its interaction with the host cell receptor(s) and membrane during viral entry, assembly and release processes, remains to be determined.

The main goal of this study is to examine the early stage, prefusion state of S protein and its interaction with the host cell. It has been shown by cryoEM data that the spike proteins of HIV, SIV and Muloney murine leukemia virus (MMLV) are clustered on viral surfaces (Zhu *et al*., 2006), and based on the similarity of class I fusion proteins, it is suggested that the S protein of SARS-CoV is also clustered on the viral surfaces. One of my working hypotheses is that the entry process of SARS-CoV also requires clustering of the specific virus receptor, ACE2, anchored on the host cell membrane lipid rafts for cellular specificity and entry. This aspect will be examined in Chapter 3. In Chapter 4, I will investigate the putative role of the Trp-rich region of S protein to anchor the spike proteins on viral surfaces that may facilitate SARS-CoV entry into host cells. Finally, in Chapter 5, I will describe an approach to identify the exposed structural elements of S protein during the entry process. This approach is designed to use a peptide-affinity strategy to aid the development of antiviral therapeutics for SARS-CoV.

F. Experimental design and methods

For the purposes of my study, I expressed the ectodomain of S protein and Trp-rich region mutants of full-length S protein for functional characterization of SARS-CoV S protein. This work mainly focused on S protein interaction with host cell receptors and cellular membrane lipid rafts. My experimental strategy is delineated as follows.

1. Establishment of a system to mimic the attachment and entry process of SARS-CoV.

a) Expression of SARS-CoV S protein in mammalian cells. This expression system could produce S protein closely to native form which will be used for functional study and to produce pseudotyped SARS-CoV for infectivity assays.

b) Expression of S protein using insect cell expression system. This strategy allows expression of a large amount of gene products and provides an efficient method to purify S protein ectodomain for structure-function studies.

c) Construction of pseudovirus particles expressing SARS-CoV S protein based on lentivirus system in 293T cells containing luciferase reporter gene for single-cycle infectivity assay as a readout assay for viral entry and infection.

Vero E6 cell is used because of its provision of ACE2 and being permissive for the SARS-CoV. Purified S protein ectodomain could be added to the Vero E6 cells to induce the interaction of S protein with ACE2. The virus particles pseudotyped with SARS-CoV S protein will provide reliable models to mimic SARS-CoV entry process.

2. Isolation and detection of lipid rafts in host cell membrane.

The characteristic clustering of raft-associated lipids into ordered lipid phase renders them relatively insoluble in certain detergents such as Triton X-100 at 4°C. Accordingly, lipid rafts can be readily purified as detergent-resistant membranes (DRMs) by ultracentrifugation on sucrose density gradients. Under these conditions, the DRMs are recovered as molecular complexes from the buoyant fractions. The lipid raft marker caveolin-1 could be used to detect and identify lipid rafts after isolation by Western blot analysis according to the methods of Wakatsuki *et al.* (Wakatsuki, Kurisaki, and Sehara-Fujisawa, 2004).

3. Analysis of the association of ACE2 and S protein with lipid rafts.

a) Lipid rafts will be isolated from Vero E6 cells, and the localization of ACE2 in lipid rafts will be analyzed. At the same time, the purified S protein ectodomain will be used to bind the Vero E6 cell to mimic the attachment to activate S protein, and then lipid rafts of the cell membranes will be isolated and investigated to determine whether S protein can be clustered into lipid rafts. The purpose of this experiment is to analyze whether the receptor ACE2 is located in lipid rafts and whether S protein is located in lipid rafts after interacting with ACE2.

To confirm the co-localization of target proteins with lipid raft markers, confocal microscopy will be used. Florescence-conjugated CT-B and anti-ACE2, anti-HA antibody will be used to visualize the lipid rafts and target molecules. Such target molecules include ganglioside GM1 with ACE2; and GM1 with S protein ectodomain.

b) Effect of cholesterol deletion on the SARS-CoV infection. Another approach

to study the function of lipid rafts is to modulate their lipid composition. Molecules able to delete cholesterol from the plasma membrane, such as β-methyl-cyclodextrin (MβCD) or 2-hydroxypropyl-β-cyclodextrin (2OHpβCD) have been widely used as raft-disrupting agent (Thorp and Gallagher, 2004). The integrity of lipid rafts can also be affected by metabolic inhibitors of sphingolipid biosynthesis, such as L-cycloserine or fumonisin, (Mahfoud *et al.*, 2002). Thus, I will be able to delete cholesterol from the plasma membrane using MβCD to disrupt lipid rafts, and then investigate the importance of lipid rafts on S protein meditated SARS-CoV viral entry.

4. The functional importance of Trp-rich region of S protein on viral infectivity.

Trp-rich region appears to be a common feature of viral fusion proteins. The conservation of these regions across various virus families indicates that they may play a vital role in membrane fusion. There is a Trp-rich region in SARS-CoV S protein (Figure 6 and 7). Gene mutation method such as alanine scan mutagenesis will be used to mutate the Trp-rich region, substituting individually or in combinations with selected amino acid residues in this region (red color labeled- Trp, Phe and Tyr residues in Figure 8), pseudotyped SARS-CoV containing different mutated S protein will be produced and the infectivity of each pseudovirus will be analyzed using single-cycle infectivity assay.

5. Identification of the exposed and functional surfaces of S protein by recombinant HR2 peptides and peptide library screening using ectodomain of S protein.

A peptide library of SARS-CoV S protein will be set up on the cellulose

membrane using the SPOTs technology, a method similar to Western blot will be carried out to screen the membrane containing the peptide library using ectodomain of S protein as a probe to identify the exposed and functional motifs of S protein. Recombinant HR2 peptides with a maltose-binding protein will be used as controls to validate the concept of self-interaction of HR1-HR2 peptides in the ectodomain of S protein.

Chapter 2. Preparation of reagents

A. Introduction

The virulent nature of SARS-CoV precludes its handling under normal laboratory conditions. Therefore to establish a system for studying the entry process of SARS-CoV mediated by S protein, recombinant methods were used to express S protein. Several expression systems were tested to optimize the S protein expression. Attempts to express the full-length S protein in *E. coli.* were unsuccessful, but success was obtained using both mammalian and insect cell systems. Pseudotyped SARS-CoV was assembled to serve as a model to mimic live virus.

B. Materials and methods

i. Materials

Chemical Reagents. Tris (hydroxymethyl) aminomethane, glycine, sodium dodecyl sulfate (SDS), Triton X-100, glycerol, bromophenol blue, coomassie brilliant blue R-250 and Nonidet P 40 were purchased from Sigma-Aldrich. Calcium chloride ($CaCl_2$), sodium chloride (NaCl), potassium chloride (KCl), disodium hydrogen phosphate anhydrous (Na_2HPO_4), potassium phosphate monobasic (KH_2PO_4) and 4-(2-hydroxyethyl)-1-piperazineethane-sulphonic acid (HEPES) were all Ultra grade and purchased from Sigma-Aldrich. 30% acrylamide (37.5:1), ammonium persulfate (APS), TEMED, dithiothreitol (DTT), β-mercaptoethanol Tween 20 and precision plus all blue protein marker were from Bio-Rad. Prestained protein marker (broad range) and 1kb DNA ladder were purchased from New England Biolabs (NEB).

Methanol and acetic acid glacial (>99.8%) were purchased from Merck Biosciences.

Antibodies. Rabbit anti-HA (Y-11) polyclonal antibody was purchased from Santa Cruz Biotechnology. Mouse anti-penta-His monoclonal antibody was purchased from QIAGEN. HRP-conjugated polyclonal rabbit anti-mouse antibody and HRP-conjugated polyclonal swine anti-rabbit antibody were purchased from DakoCytomation. Rabbit anti-S antibody was from Prof. Ding Xiang Liu (IMCB).

Cell lines. 293T cells (CRL-11268), HeLa cells (CCL-2) and Sf9 cells (CRL-1711) were purchased from ATCC.

Plasmids. pCDNA3.1(+) was purchased from Invitrogen. pJX40-S which encodes SARS-CoV spike gene of Sin 2774 (AY283798) was a kind gift from IMCB in Singapore. pNL4-3Luc^{+}Env^{-}Vpr^{-} and pcDNA3.1-OPT9-S were kindly provided by Prof. Zhang Linqi (Aaron Diamond AIDS Research Center, Rockefeller University, New York 10016). pVPack-VSV-G which carries vesicular stomatitis virus envelope glycoprotein (VSV-G) was purchased from Stratagene. pVL1392 and linearized BaculoGold^{TM} baculovirus viral DNA were purchased from BD Biosciences.

Kits and Enzymes. Anti-HA affinity matrix, HA peptide, DOTAP liposomal transfection reagent, complete protease inhibitor cocktail tablets and complete EDTA-free protease inhibitor cocktail tablets were purchased from Roche Diagnostics GmbH, Roche Applied Sciences. EnzChek® reverse transcriptase assay kit was purchased from Molecular Probes. ECL Plus Western blotting detection reagent was purchased from Amersham Biosciences. BaculoGold^{TM} starter package and transfection kit were purchased from BD Biosciences. Plasmid DNA Mini, Midi and

Maxi preparation kits, PCR purification kit, QIAEX II gel extraction kit, Ni-NTA agarose, restriction enzymes were purchased from QIAGEN. Native membrane protein extraction kit was purchased from Merck Biosciences. Pfu DNA polymerase was purchased from Stratagene.

Solutions.

Solutions for calcium phosphate transfection. 0.5 M $CaCl_2$ contains 0.5 M $CaCl_2$ in double-distilled H_2O (ddH_2O). Hepes buffered saline 2x (HeBS 2x) contains 0.28 M NaCl, 50 mM HEPES and 1.5 mM Na_2HPO_4 in ddH_2O, pH7.0. Plasmid mixing solution contains 2.5 mM HEPES in ddH_2O, pH7.3.

Solutions for cell culture. DMEM, Sf-900 II SFM culture media, 5mg/ml penicillin-streptomycin and 1× Trypsin-EDTA were purchased from Gibco. New born calf serum was purchased from Gemini Bio-Products (Calabasas, Calif.). Phosphate buffered saline (PBS) buffer contains 0.137 M NaCl, 2.7 mM KCl, 2 mM KH_2PO_4 and 10 mM Na_2HPO4 in ddH_2O, pH7.3.

Solutions for protein purification. (1) Solutions for anti-HA affinity column purification: Lysis buffer contains 50 mM Tris, 150 mM NaCl and 0.1% Nonidet P 40, pH7.5. Washing buffer contains 20 mM Tris, 0.1 M NaCl, 0.1 mM EDTA and 0.05% Tween-20, pH7.5. Equilibration buffer contains 20 mM Tris, 0.1 M NaCl and 0.1 mM EDTA, pH7.5. Elution buffer contains 1 mg/ml HA peptide in equilibration buffer. (2) Solutions for Ni-NTA column purification: Lysis buffer contains 50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole and 1% Nonidet P 40, pH 8.0. Resuspension buffer contains 50 mM NaH_2PO_4, 300 mM NaCl, and 10 mM imidazole,

pH 8.0. Washing buffer contains 50 mM NaH_2PO_4, 300 mM NaCl and 20 mM imidazole, pH 8.0. Elution buffer contains 50 mM NaH_2PO_4, 300 mM NaCl and 250 mM imidazole, pH 8.0.

Solutions for SDS-PAGE. 10% APS contains 1g ammonium persulfate in 10 ml ddH_2O, and was stored at 4°C, protected from light. 10% SDS contains 1g SDS in 10 ml ddH_2O, and was stored at room temperature. Staining solution contains 3 mM coomassie brilliant blue R-250 dissolved in destaining solution and filtered through 0.22μ filter and stored at room temperature. Destaining solution contains 10% (v/v) glacial acetic acid, 30% (v/v) methanol in ddH_2O. Running buffer (1×) contains 25 mM Tris base, 0.25 M glycine and 0.1% SDS (w/v) dissolved in ddH_2O. Sample loading buffer (4×) contains 0.2 M Tris-HCl (pH 6.8), 8% SDS (w/v) and 40% glycerol (v/v), 0.4 M DTT or 5% β-mercaptoethanol (v/v) and 0.008% bromophenol blue (w/v). It was aliquoted in 0.5 ml and stored at -20°C.

Solutions for Western blot. Transfer buffer (1×) contains 25.6 mM Tris base and 186 mM glycine. Washing buffer (PBST) is PBS containing 0.1% Tween 20 (v/v). Blocking buffer contains 5% nonfat milk (w/v) in washing buffer.

Primers. Primers designed and used in PCR are shown in Table 2.

Instruments. The following instruments were used in this study: Mini-PROTEAN® 3 electrophoresis cell and semi-dry transfer instrument (BioRad), Optima™ L-XP series preparative ultracentrifuge and Avanti™ J25 centrifuge (Beckman Coulter); Sigma 3-18K centrifuge (Sigma); PTC-100® Peltier Thermal cycler (MJ Research) and Kodak X-OMAT processor (Kodak).

Table 2. Primers used in PCR

Name	Sequence 5'-3'*	Product size (bp)
pcDNA3.1-5	CGGGAGATCTCCCGATCCCCTATGGTGCAC	-
pcDNA3.1HA-3	GCCCTCTAGATT***AAGCGTAATCTGGAACATC GTATGGGTA***CATCTCGAGCGGCCGCCACT	1025
Mammalian cell expression		
F33EcoRI	AGTCGAATTCCGAACATGTTTATTTTCTTA	-
S540XhoI	ACGTCTCGAGAGGAGTTAACACACCAGTAC	1648
S663XhoI	ACGT CTCGAGAACTGTATGGTAACTAGCAC	2014
S795XhoI	ACGTCTCGAGAGTTGGCTTTAGAGGGTCAG	2410
S1255XhoI	ACGTCTCGAGTGTGTAATGTAATTTG ACAC	3790
Insect cell expression		
F33EcoRI	AGTCGAATTCCG AACATGTTTATTTTCTTA	-
S1188HABamHI	GCCCGGATCCTTA***AGCGTAATCTGGAACATCG TATGGGTA***CATCTCGAGATATTTTCCCAATTCTT	3628
S1190HisBamHI	GCCCGGATCCTTA***GTGATGATGATGATGATG***T TGCTCATATTTTCCCA	3616
S1255HABamHI	GCCCGGATCCTTAAGCGTAATCTGGAACAT	3829
Bacteria expression		
pMALC2EcoRI	TTCAGAATTCATTTCAGGCATTAAC	-
HR2-35PstI	TTGCCTGCAGTTACAATTCTTGAAG	128
pMALC2BglII	TAACAAAGATCTGCTGCCGAACCCG	-
HR2-34PstI	ACGTCTGCAGTTATTCTTGAAGGTCAATGA	933
HR2-33PstI	ACGTCTGCAGTTATTGAAGGTCAATGAGTG	930
HR2-32PstI	ACGTCTGCAGTTAAAGGTCAATGAGTGATT	927
HR2-31PstI	ACGTCTGCAGTTAGTCAATGAGTGATTCAT	924
HR2-30PstI	ACGTCTGCAGTTAAATGAGTGATTCATTTA	921
HR2-29PstI	ACGTCTGCAGTTAGAGTGATTCATTTAAAT	918
HR2-28PstI	ACGTCTGCAGTTATGATTCATTTAAATTTT	915
HR2-27PstI	ACGTCTGCAGTTATTCATTTAAATTTTTAG	912

***** Underlined bases are restriction enzyme sites, the boldfaced and italic bases encode HA tag and His tag.

ii. Methods

Cell culture. 293T cells and HeLa cells were maintained in DMEM medium supplemented with 10% new born calf serum and 50 μg/ml penicillin-streptomycin at 37°C, 5% CO_2. Sf9 cells were maintained in Sf-900 II SFM at 27°C, and protected from light. Cells were passaged twice a week.

Plasmids construction.

(1) Construction of Plasmid pcDNA3.1(+)HA. The fragment containing the HA tag (influenza hemagglutin epitope, YPYDVPDYA) was amplified from pcDNA3.1(+) plasmid using Pfu DNA polymerase with primers: pcDNA3.1-5 and pcDNA3.1HA-3 which comprise a Bgl II site upstream with HA tag between Xho I and Xba I sites downstream. PCR reactions were performed by heating the reaction mixture at 95ºC for 45 s, followed by 30 cycles at 95ºC for 45 s, subsequently, an annealing and elongation step were carried out at 65ºC for 45 s and at 72ºC 2 min respectively. After 30 cycles, a final elongation step at 72ºC was performed for 10 min. The PCR products were cloned into pcDNA3.1(+) vector between Bgl II and Xba I sites to produce the plasmid pcDNA3.1(+)HA.

(2) Construction of clones containing the full-length and truncated N-terminal fragments of SARS-CoV S protein. Different fragments of S gene were amplified directly using Pfu DNA polymerase from plasmid pJX40-S. The 5' primer F33EcoRI comprises an EcoR I site upstream of the ATG codon followed by the S gene sequences. The 3' primers S540XhoI, S663XhoI, S795XhoI and S1255XhoI, all contain the Xho I site downstream. PCR reactions were performed by heating the

reaction mixture at 95ºC for 45 s, followed by 30 cycles at 95ºC for 45 s, subsequently, annealing and elongation steps were carried out at 50ºC for 45 s and at 72ºC for 3 min (S540), 4 min (S663), 5 min (S795) and 8 min (S1255) respectively. After 30 cycles, an elongation step at 72ºC was performed for 10 min. The PCR products were cloned into the pcDNA3.1(+)HA vector between EcoR I and Xho I sites. The resulting plasmids were named as pcDNA3.1(+)HA-S540, pcDNA3.1(+)HA-S663, pcDNA3.1(+)HA-S795 and pcDNA3.1(+)HA-S1255.

The nucleotide fragment corresponding to amino acid residues 1-1255 of S protein with HA tag at the C-terminus (SHA) was amplified by PCR from the plasmid pcDNA3.1(+)HA-S1255 using primers F33EcoRI and S1255HABamHI. The nucleotide fragment corresponding to amino acid residues 1-1188 of S protein was amplified by PCR from the plasmid pJX40-S and HA tag was added at the C-terminus of S protein (referred to as S1188HA) using primers F33EcoRI and S1188HABamHI. The nucleotide fragment corresponding to amino acid residues 1-1190 of S protein was amplified by PCR from plasmid pJX40-S and a His tag was added at the C-terminus of S protein (referred to as S1190His) using primers F33EcoRI and S1190HisBamHI. The amplification program for all these fragments was 95ºC, 45 s, then 30 cycles of 95ºC, 45 s, 50 ºC, 45 s and 72 ºC, 8 min, followed by the last elongation step at 72ºC, 10 min. An upstream EcoR I site and a downstream BamH I site as well as a stop codon preceding the BamH I site were designed to insert these fragments into transfer vector pVL1392, under the control of a strong viral promoter, polyhedrin promoter. The resulting plasmids were named as pVL1392SHA,

pVL1392S1188HA and pVL1392S1190His.

Protein Expression.

(1) Transient expression of SARS-CoV S protein in HeLa cell using the recombinant vaccinia virus/T7 system. Expression of S protein in HeLa cells using the recombinant vaccinia virus/T7 system is regulated by T7 RNA polymerase promoter. HeLa cells were grown to 80% confluence in 75 cm^2 flasks (Nunc). After trypsinization using Trypsin-EDTA for 3 min, cells were resuspended in DMEM containing 10% new born calf serum, plated into 6-well plates (5 x 10^5 cells/well) and cultured at 37°C, 5% CO_2, overnight. Following addition of recombinant vaccinia virus-T7 (VTF7.3) and 2 h incubation, the HeLa cells were transfected with plasmids containing different fragments of S gene using DOTAP transfection reagent. The plasmids pcDNA3.1(+)HA-S1255, pcDNA3.1(+)HA-S795, pcDNA3.1(+)HA-S663 and pcDNA3.1(+)HA-S540 directed the synthesis of different fragments of S protein with a HA tag at their C-termini to facilitate identification. After 48 h of transfection, cells were collected, cell lysates were resolved by 8% SDS-PAGE and the expression of S protein was confirmed by Western blot using rabbit anti-HA antibody as primary probe, and HRP-conjugated swine anti-rabbit antibody as secondary probe.

(2) Transient expression of S protein in 293T cell using codon-optimized plasmid. 293T cells were grown to 80% confluence in 75 cm^2 flasks (Nunc), and after trypsinization with Trypsin-EDTA 3 min; cells were resuspended, plated into 6-well plates (5 x 10^5 cells/well) and cultured at 37°C, 5% CO_2, overnight. The 293T cells were transfected with plasmid pcDNA3.1-OPT9-S using DOTAP

transfection reagent. After transfection 48 h, cells were collected, cell lysate was resolved by 8% SDS-PAGE, and the expression of S protein was detected by Western blot using rabbit anti-S antibody as primary antibody.

(3) Expression and purification of SARS-CoV S protein in Sf9 cells. To generate recombinant baculoviruses, the Autographa Californica Multiple Nucleopolyhedrovirus (AcMNPV) viral genome was used. The individual plasmids pVL1392SHA, pVL1392S1188HA, pVL1392S1190His which contain SHA, S1188HA and S1190His fragments were co-transfected with linearized baculovirus viral DNA using transfer buffer A and B into Sf9 cells separately following the manufacturer's instructions. Successful recombinants were identified by detecting the protein expression using Western blot. Recombinant viruses were subjected to three or four rounds of amplification to prepare the viral stock. Sf9 cells were infected with the recombinant baculovirus at a multiplicity of infection (moi) of 3-10. At 4 days post-infection (dpi), cells were collected by a cell scraper (Costar) and then re-suspended in lysis buffer which contains complete protease inhibitor cocktail tablets, lysed on ice for 30 min, vortex several times during incubation, then centrifuged (4°C, 30,000 g, 30 min, Avanti^{TM} J25 centrifuge) to remove the cell debris. The supernatant was stored at -20°C for analysis.

To extract full-length S protein using native membrane protein extraction kit, Sf9 cells were cultured in 75 cm^2 flask (Nunc) to 80% confluence, and then infected with recombinant baculovirus containing the full-length S gene. The cells were collected 4 dpi by a cell scraper, and were pelleted by centrifugation at 100-300 g, 10 min at

4°C. The cell pellet was washed with 8 ml ice cold wash buffer two times, and then were resuspended into 8 ml extraction buffer I (provided in the native membrane protein extraction kit) containing protease inhibitor cocktail 40 µl. After incubation of 10 min at 4°C under gentle agitation, the cells were centrifuged at 16,000 g, 4°C for 15 min. The supernatant (enriched with "soluble" proteins) was removed and kept for analysis. The pellet was resuspended in 4 ml extraction buffer II (provided in the native membrane protein extraction kit) containing protease inhibitor cocktail 20 µl. After incubation for 30 min at 4°C under gentle agitation, cells were centrifuged at 16,000 x g, 4°C for 15 min. The supernatant (enriched with integral membrane and membrane-associated proteins) was collected for analysis.

To purify S1188HA, the protein extract was loaded on the column containing anti-HA affinity matrix. After washing the column with a minimum of 20 bed-volumes of wash buffer at 15 to 25°C to remove nonspecifically bound protein, the protein was eluted using elution buffer. Eluate was collect and analyzed by SDS-PAGE, aliquoted and stored at -20°C for later use.

To purify S1190His, 5 ml protein extract was loaded on the affinity column which contained 1 ml Ni-NTA, incubated at 4°C for 2-4 hours. After removal of the bottom cap from the column to let the lysate flow through, the column was washed twice with 4 times bed volume of the washing buffer. The protein was eluted 4 times with 0.5 ml elution buffer, the eluate was collect and analyzed by SDS-PAGE, aliquoted and stored at -20°C for later use.

Preparation of pseudotyped SARS-CoV. Pseudovirus was generated by

co-transfection of 293T cells using calcium phosphate transfection method with pNL4-3Luc^{+}Env^{-}Vpr^{-} and pcDNA3.1-OPT9-S or pVPack-VSV-G. The culture supernatant containing virus was collected on day 2 and 3 after transfection and clarified by filtration through a 0.45 μm filter. The virus was concentrated by ultracentrifugation at 26,000 rpm for 3 h (Zhang *et al*., 2004). The pelleted virus was resuspended in DMEM with 10% new born calf serum, aliquoted, and stored at -80°C. The virus concentration was determined by the reverse transcriptase (RT) activity assay using EnzChek® reverse transcriptase assay kit following the manufacturer's instructions.

Western blot analysis. After separating expressed protein on 8% SDS-PAGE, proteins were electrically transferred to a PVDF membrane (20V, 25 min) using a semi-dry transfer instrument. The membrane was incubated in blocking buffer on a rocking platform overnight at 4ºC, and then was probed with 1:1000 dilution of corresponding primary antibody on the rocker plate at room temperature for 1 h. After washing 3 times (15 min each) with washing buffer, the membrane was subsequently incubated with corresponding HRP-conjugated secondary antibody (1:1000 in blocking buffer) for 1 h at room temperature, and then washed 3 times (15 min each). After addition of the ECL Plus solution for 5 min, the membrane was exposed to a film (Kodak) and developed using the Kodak X-OMAT processor.

C. Results and discussion

i. Construction of plasmids.

Construction of plasmid pcDNA3.1(+)HA. The PCR fragment contains HA tag

between Xho I and Xba I sites is shown in Figure 11, lane 4. The vector pcDNA3.1(+) was digested with Bgl II and Xba I, after digestion, two fragments were obtained, the longer fragment (>4 kb) was extracted and purified from the 1% agarose gel (Figure 11, lane 3), and ligated with the Bgl II / Xba I digested PCR fragment to obtain the clone. The sequencing result showed that HA tag was inserted into plasmid pcDNA3.1(+) successfully and the restriction map of the original plasmid was not changed. This plasmid is named as pcDNA3.1(+)HA and used as the vector for later cloning.

Construction of clones containing full-length and different N-terminal fragments of SARS-CoV S protein. The S protein of SARS-CoV contains two subunits: S1 and S2. S1 subunit is responsible for receptor-binding, and a major function of S2 subunit is membrane fusion. Recently, the receptor-binding domain (RBD) of SARS-CoV S protein has been located between residues 303 and 537 (Xiao *et al*., 2003), but the boundary between S1 and S2 of S protein remains to be determined. Sequence alignment with other coronaviruses together with preliminary experimental evidence suggest that the putative boundary lies around 600-700 aa (Li *et al*., 2005). In this study, I chose three different N-terminal fragments of the putative S1 protein containing the RBD for mammalian cell expression system. Since the insect cell expression system could express S protein more easily than mammalian cell, I chose N-terminal fragments of S protein which deleted TM and cytoplamic domain for insect cell expression system. These truncations could be expressed and produce S protein as long as possible which remains most of its

domains, and forms soluble form at the same time. This is very important for the subsequent purification procedure and functional study of S protein.

The PCR fragments containing N-terminal segments of S gene: S540, S663, S795 and S1255 which encode the N-terminal 540, 663 and 795 amino acids and full-length of S protein are shown in Figure 12. These fragments were inserted into pcDNA3.1(+)HA plasmid between EcoR I and Xho I sites. The sequencing result showed that all the clones contained the correct fragments of S gene with a HA tag at their C-termini. All these clones were used for expressing S protein in mammalian cells.

The PCR fragments containing N-terminal segments and full-length of S gene: S1188HA, S1190His and S1255HA (Figure 13) were inserted into pVL1392 plasmid between EcoR I and BamH I sites. The correct clones were selected by double restriction enzyme digestion and confirmed by sequencing. All these clones were used for expressing S protein in insect cells.

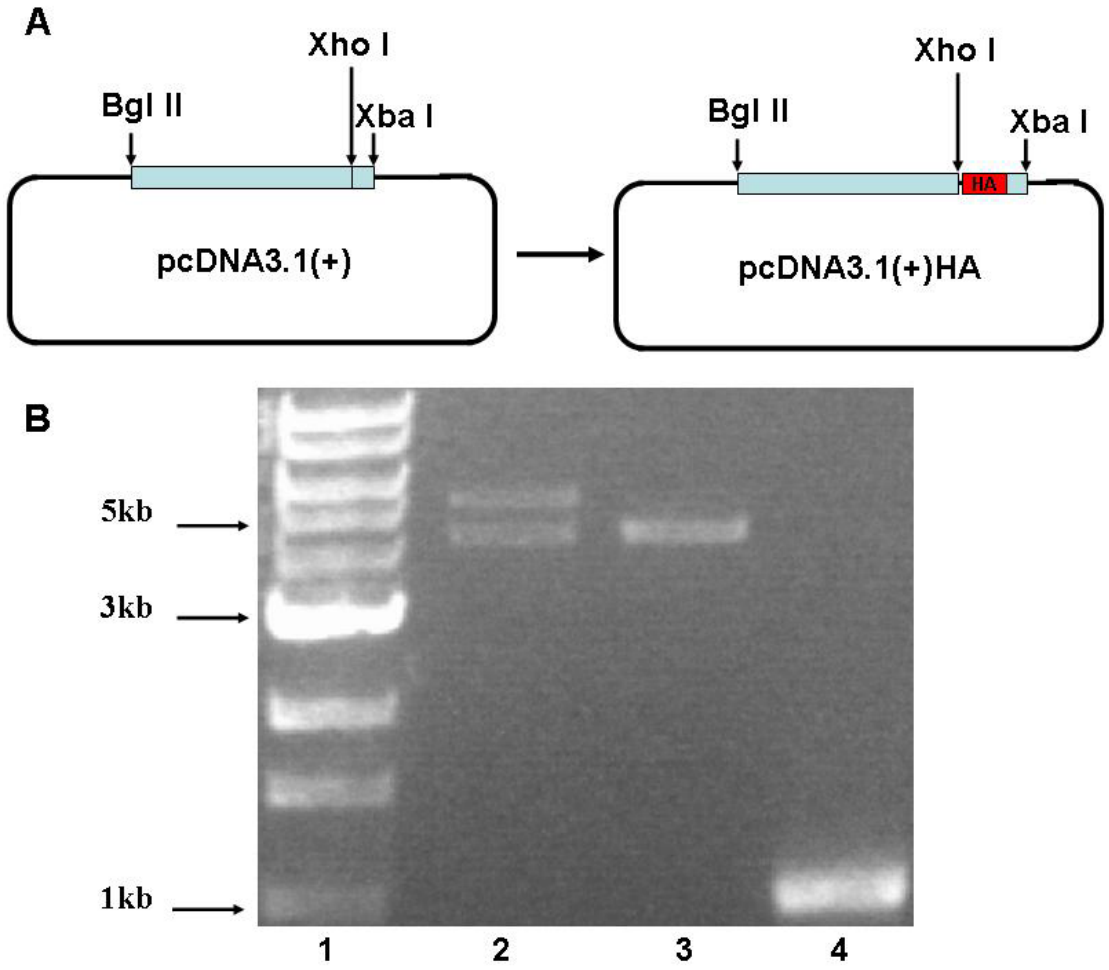

Figure 11. Construction of plasmid pcDNA3.1(+)HA.

A. Strategy of construction of plasmid pcDNA3.1(+)HA.
B. The fragment containing HA tag was obtained from PCR reaction, and digested with Bgl II and Xba I. The vector pcDNA3.1(+) was digested by Bgl II and Xba I, and the larger fragment was extracted from agarose gel to ligate with digested PCR fragment.

Lane 1: 1kb DNA standard ladder (NEB),
Lane 2: purified fragment of pcDNA3.1(+) vector which was digested first with Bgl II and then with Xba I, the digestion was not completed,
Lane 3: purified fragment of pcDNA3.1(+) vector which was digested first with Xba I and then with Bgl II, the digestion was completed, and this fragment was used as vector to ligate with PCR fragment,
Lane 4: PCR product containing HA tag.

All these DNA fragments were run in 1% agarose gel in TAE buffer at a constant voltage of 120 V for 30 min. The arrows indicate DNA marker.

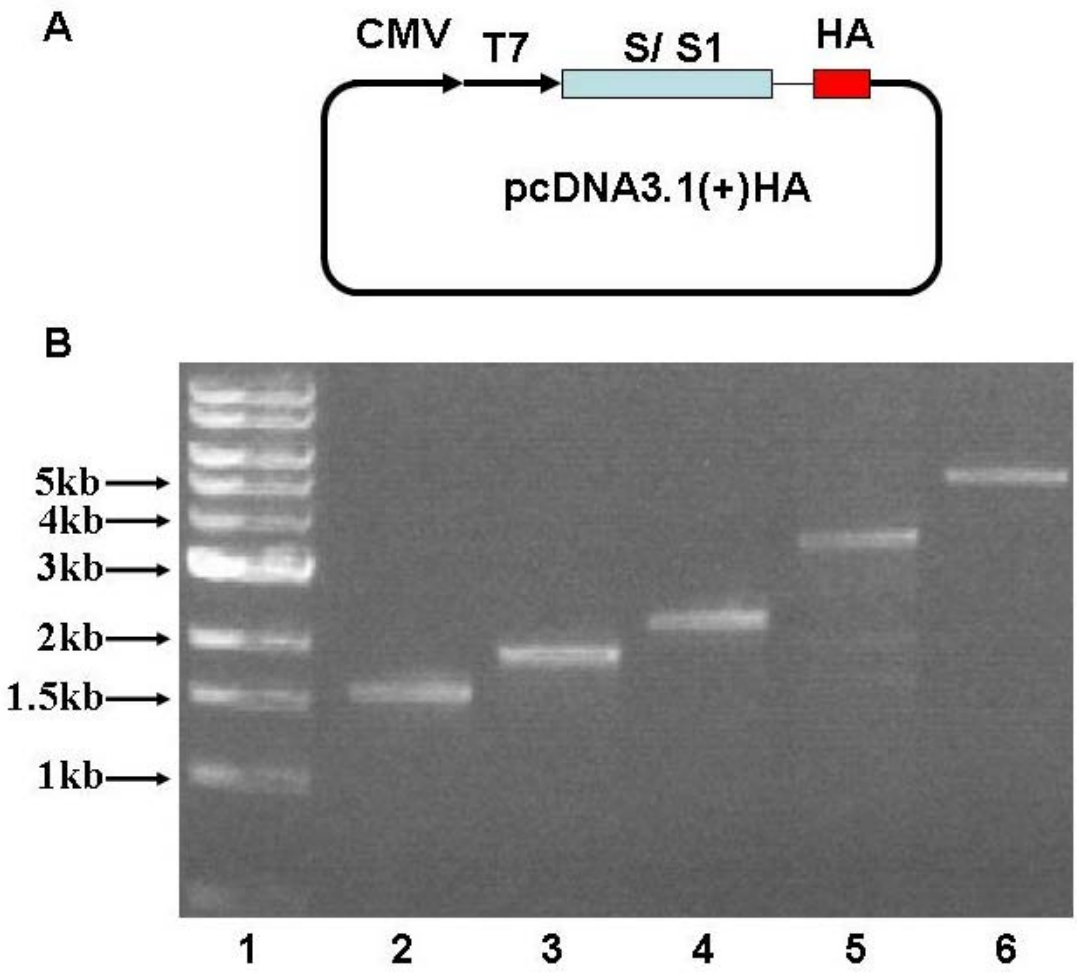

Figure 12. Construction of plasmids containing different fragments of SARS-CoV S protein for mammalian cell expression.

A. Strategy of construction of plasmids containing different fragments of S protein using pcDNA3.1(+)HA vector.
B. The fragments of different length of S gene were obtained from PCR reaction, and digested with EcoR I and Xho I. The vector pcDNA3.1(+)HA was digested by EcoR I and Xho I, and ligated with digested PCR fragments.

Lane 1: 1kb DNA standard ladder (NEB),
Lane 2: PCR product of N-terminal 540 aa of S protein, S540,
Lane 3: PCR product of N-terminal 663 aa of S protein, S663,
Lane 4: PCR product of N-terminal 795 aa of S protein, S795,
Lane 5: PCR product of full-length S protein,
Lane 6: purified fragment of pcDNA3.1(+)HA vector digested with EcoR I and Xho I.

All these DNA fragments were run in 1% agarose gel in TAE buffer at a constant voltage of 120 V for 25 min. The arrows indicate DNA marker.

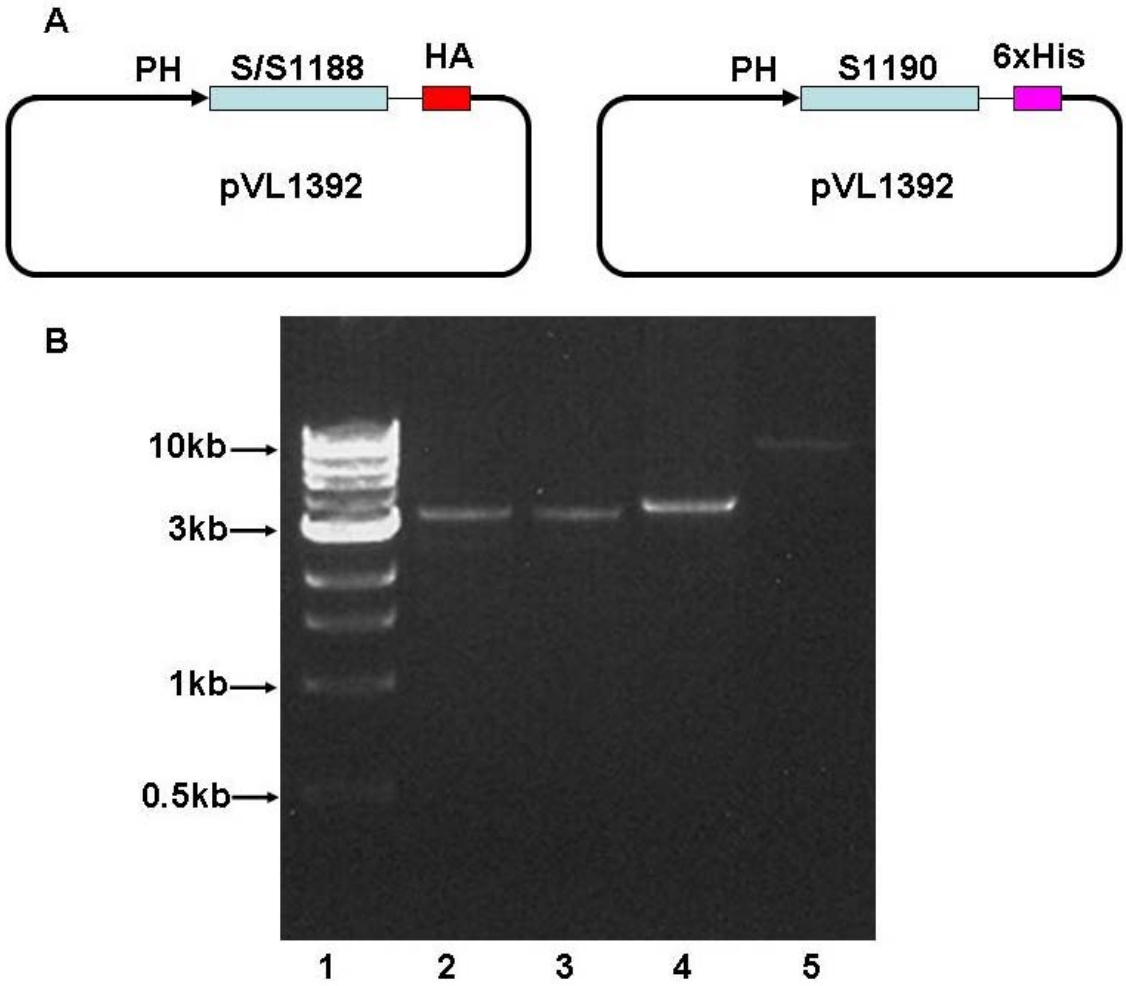

Figure 13. Construction of plasmids containing SARS-CoV full-length and ectodomain of S protein for insect cell expression

A. Strategy of construction of plasmids containing S protein with different tag for insect cell expression (PH means polyhedrin promoter).
B. Different fragments of S gene were obtained from PCR reaction, and digested with EcoR I and BamH I. The vector pVL1392 was digested by EcoR I and BamH I, and ligated with digested PCR fragments.

Lane 1: 1kb DNA standard ladder (NEB),
Lane 2: PCR product of N-terminal 1188 aa of S protein with a HA tag at its C-terminus, S1188HA,
Lane 3: PCR product of N-terminal 1190 aa of S protein with a His tag at its C-terminus, S1190His,
Lane 4: PCR product of full-length S protein with a HA tag at its C-terminus, SHA,
Lane 5: Purified fragment of pVL1392 vector digested with EcoR I and BamH I.

All these DNA fragments were run in 1% agarose gel in TAE buffer at a constant voltage of 120 V for 25 min. The arrows indicate DNA marker.

ii. Expression of protein.

(1) Transient expression of S protein in HeLa cells using the recombinant vaccinia virus/T7 system. The proteins encoded in plasmids pcDNA3.1(+)HA-S1255, pcDNA3.1(+)HA-S795, pcDNA3.1(+)HA-S663 and pcDNA3.1(+)HA-S540 were expressed as HA-tagged proteins in HeLa cells, and identified by Western blot analysis (Figure 14). These results indicate that all clones could be expressed in VTF7.3 vaccinia virus infected HeLa cells. The shorter fragments exhibited better expression than longer fragments. Two bands were observed in each clone expression product. The high molecular-weight band corresponded to the glycosylated form of these proteins and the other corresponded to unglycosylated form with the estimated molecular weights of S1255: 139 kDa, S795: 87 kDa, S663: 73 kDa and S540:59 kDa.

To study the characteristics and functions of S protein, significant quantities of full-length S protein and S protein fragments are required. Firstly, I attempted to express the S protein in *E. coli.* by varying several factors, including the concentration of IPTG, the induction time and temperature. However, no expression was observed in the bacterial host. Possible reasons contributing to my failure include the large size of S protein and its toxicity to bacteria. Then I utilized a mammalian expression system using the CMV promoter of pcDNA3.1(+) in an attempt to produce S protein. The plasmid containing the S gene was directly transfected into HeLa cells and cultured for 24, 36 or 48 h, but again S protein expression was unsuccessful. It is possible that the CMV promoter is not powerful enough or suitable for initiating the

expression of the SARS-CoV S gene. Finally, I examined a second mammalian expression system - the recombinant vaccinia virus/T7 system. After HeLa cells were infected with vaccinia virus/T7 for 2 h, the cells were then transfected with plasmid DNA containing the target gene. After 36 h of transfection, protein expression could be detected by Western blot analysis (Figure 14). In each case, expression from single clones resulted in two different forms of S protein, glycosylated and unglycosylated. Cells transfected with plasmid containing full-length S gene produced a protein estimated to be ~200 kDa, consistent with the size of S protein isolated from SARS-CoV virions (Rota *et al*., 2003).

From these results, I conclude that the recombinant vaccinia virus/T7 system is suitable for expressing our target proteins in HeLa cells, which suggests that T7 promoter is stronger than CMV promoter to initiate the expression of S gene. Specifically, this system can generate proteins with post-translational modification to produce samples that closely mimic native proteins for our further studies. One disadvantage of this system is that the amount of expressed protein is not enough for purification, receptor-binding and other characterization studies. A second disadvantage of this system is that it introduces another virus into the host cell, which deems the system unsuitable for pseudovirus production. The vaccinia virus infection was observed to cause many changes in cell morphology and cell surface characteristics, which make this system unsuitable for microscopy imaging studies.

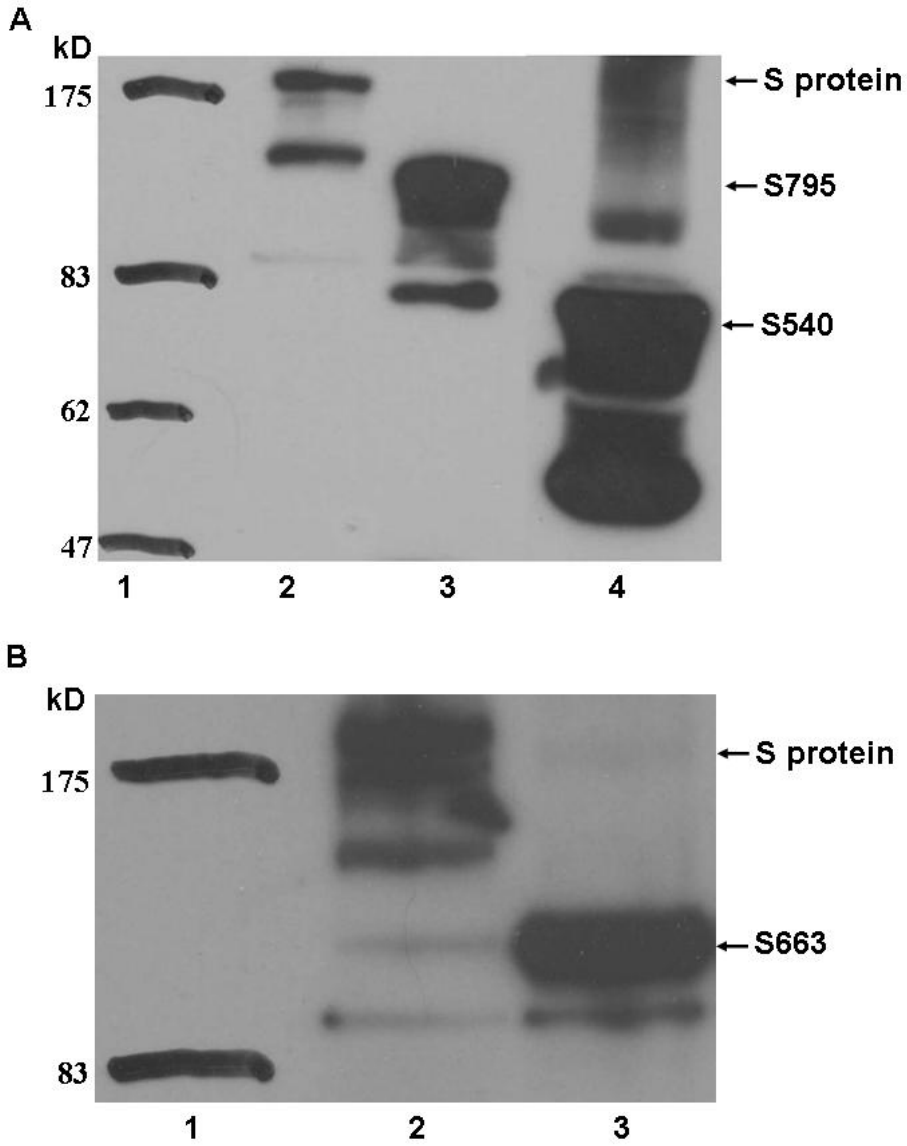

Figure 14. Expression of different fragments of SARS-CoV S protein in HeLa cells.

Full-length S protein, N-terminal fragments of S protein, S795, S540 (A) and S663 (B) were expressed as HA-tagged products. The cell lysates were resolved on 8% SDS-PAGE, transferred onto PVDF membrane, and subjected to Western blot analysis using rabbit anti-HA antibody as the primary probe.
Lane A1: Protein molecular weight marker (NEB),
Lane A2: Full-length S protein,
Lane A3: N-terminal 795 aa fragment of S protein, S795,
Lane A4: N-terminal 540 aa fragment of S protein, S540,
Lane B1: Protein molecular weight marker (NEB),
Lane B2: Full-length S protein,
Lane B3: N-terminal 663 aa fragment of S protein, S663.
The upper band of each lane is the glycosylated form of S protein and the lower band is the unglycosylated form of S protein.

(2) Transient expression of S protein in 293T cells using codon-optimized plasmids. The full-length S protein was expressed in 293T cells under the CMV promoter after transfection with pcDNA-OPT9-S which encodes a codon-optimized S gene. The expression product was detected by Western blot using rabbit anti-S antibody. The results showed that using codon-optimized plasmid, the expression of full-length S protein under the CMV promoter could be detected, but the plasmid containing the original S gene could not initiate detectable expression (Figure 15).

"Codon optimization" is a general approach that seeks to improve heterologous protein expression when a gene is moved into a foreign genome that exhibits different codon usage from its native genome. It has been widely applied to improve heterologous protein expression in the past two decades (Gustafsson, Govindarajan, and Minshull, 2004).

Using the original gene codon of SARS-CoV, it was very difficult to express the full-length S protein in *E. coli* and mammalian cells. However it could be expressed using codon-optimized plasmids (Figure 15). The failure of expression in *E. coli* and mammalian cells using the original codon could be due to the codon bias. Codon optimization is necessary for the expression of S protein. The advantage of this system is that protein expression could be induced without introducing another virus into the host cell, which makes the system most suitable for producing pseudotyped SARS-CoV containing S protein to investigate the viral infection. The main disadvantage of this system is that the amount of expressed protein is still not enough for purification, receptor-binding and other characterization studies.

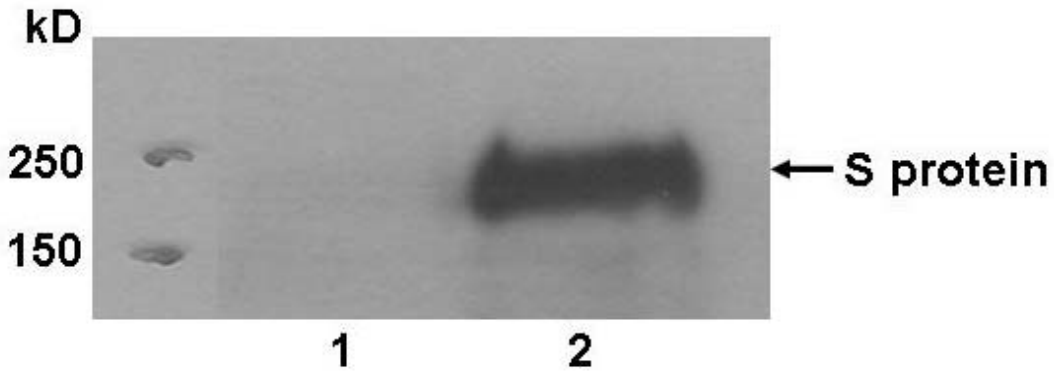

Figure 15. Expression of SARS-CoV S protein in 293T cells using plasmids containing original or codon-optimized S gene.

The full-length S protein was expressed in 293T cells using pcDNA3.1(+)HA-S1255 and pcDNA-OPT9-S respectively. After transfection 48 h, cells were collected and cell lysates were resolved on 8% SDS-PAGE. Proteins were transferred onto PVDF membrane (20V, 25 min), and subjected to Western blot using rabbit anti-S antibody as the primary probe.

Lane 1: pcDNA3.1(+)HA-S1255 (containing original S gene) expression product,
Lane 2: pcDNA-OPT9-S (containing codon-optimized S gene) expression product.

Protein molecular weight marker is from Bio-Rad. Protein molecular weight is indicated at the left side.

(3) Expression and purification of S protein in Sf9 cells. Different fragments of S gene with different of tags at their C-termini were recombined with linearized baculovirus DNA to produce recombinant baculoviruses containing different fragments of S gene. After infection with recombinant baculoviruses, higher amount of S protein expression was obtained in Sf9 cells than in mammalian cells. The ectodomain of S protein, N-terminal 1188 aa with a HA tag at its C-terminus (referred to as S1188HA, Figure 16A) could secrete into the culture medium (Figure 17), and this fragment could be purified using anti-HA affinity column to near homogeneity (Figure 16B, lane 3). The purified ectodomain of S protein was used in later experiments for characterization and receptor-binding assay.

Even though S1188HA was secreted into the culture media, the full-length S protein was not, and it was difficult to be extracted from the cell membrane when using the same lysis buffer as to extract S1188HA. I tried to use native membrane protein extraction kit to extract the full-length S protein from Sf9 cells. However, the full-length S protein could only be obtained in the membrane fraction solution, but not in the soluble protein fraction (Figure 17). This further confirmed that S protein was a membrane protein. S1190His, N-terminal 1190 aa of S protein with a His tag at its C-terminus (referred to as S1190His, Figure 18A) could be expressed in Sf9 cells (Figure 18B), but it is difficult to be purified using Ni-NTA column.

Using mammalian cell expression system, either by recombinant vaccinia virus/T7 system or by codon-optimized plasmids, S protein most closely to the native viral protein could be obtained, but the amount of the protein was not large enough for

further study. The baculovirus expression vector system (BEVS) is one of the most powerful and versatile eukaryotic expression systems. The BEVS typically produces overexpressed recombinant proteins containing proper folding, disulfide bond formation and oligomerization (Kidd and Emery, 1993). Additionally, this system is capable of performing several post-translational modifications, including N- and O-linked glycosylation, phosphorylation, acylation, amidation, signal peptide cleavage etc. This leads to a protein that is similar to its native counterpart, both structurally and functionally. Due to the difference between the N-glycosylation pathways and their enzyme systems of mammalian and insect cell, the major processed N-glycan typically produced by insect cells is the paucimannose structure, $Man_3GlcNAc_2$(± Fuc)-N-Asn. Most mammalian recombinant glycoproteins produced using baculovirus-insect cell expression system usually have $Man_3GlcNAc_2$(±Fuc)-N-Asn in place of complex, terminally sialylated glycans found on the native product (Kulakosky, Hughes, and Wood, 1998; Luckow, 1995; März, 1995). Although differences exist in N-linked glycosylation between insect and mammalian cells, the full extent and consequence of these differences have not yet been fully determined. However, most of the proteins produced thus far retain full biological activity (Luckow, 1990; Wickham et al., 1992).

Using this system higher expression of S protein was obtained. Purified S1188HA was gained to near homogeneity using anti-HA affinity column, and suitable for use in receptor-binding assay. S1190His was also successfully expressed in Sf9 cells, but could not be successfully purified. Different conditions were tested,

such as different concentration of imidazole in resuspension buffer and washing buffer, different incubation time with Ni-NTA resin, different ratio between cell lysate and Ni-NTA resin, even incubating the protein with copper-NTA, zinc-NTA or cobalt-NTA resin for purification, but high purity protein still could not be obtained. In addition, the column flow through still contained S1190His, indicating that S1190His could not bind to the Ni-NTA resin very well. This may be due to the His tag being buried inside or masked by the large protein after expression and folding. Therefore the His tag could not be exposed very well to interact with Ni-NTA. Another possible reason is that comparing with *E. coli*, insect cells contains much more different kinds of proteins and most of the proteins have been glycosylated, other proteins in Sf9 cells or the glycans attached to the proteins could bind to the Ni-NTA resin nonspecifically and occupy the binding sites, and decrease the efficiency of S1190His binding to Ni-NTA resin.

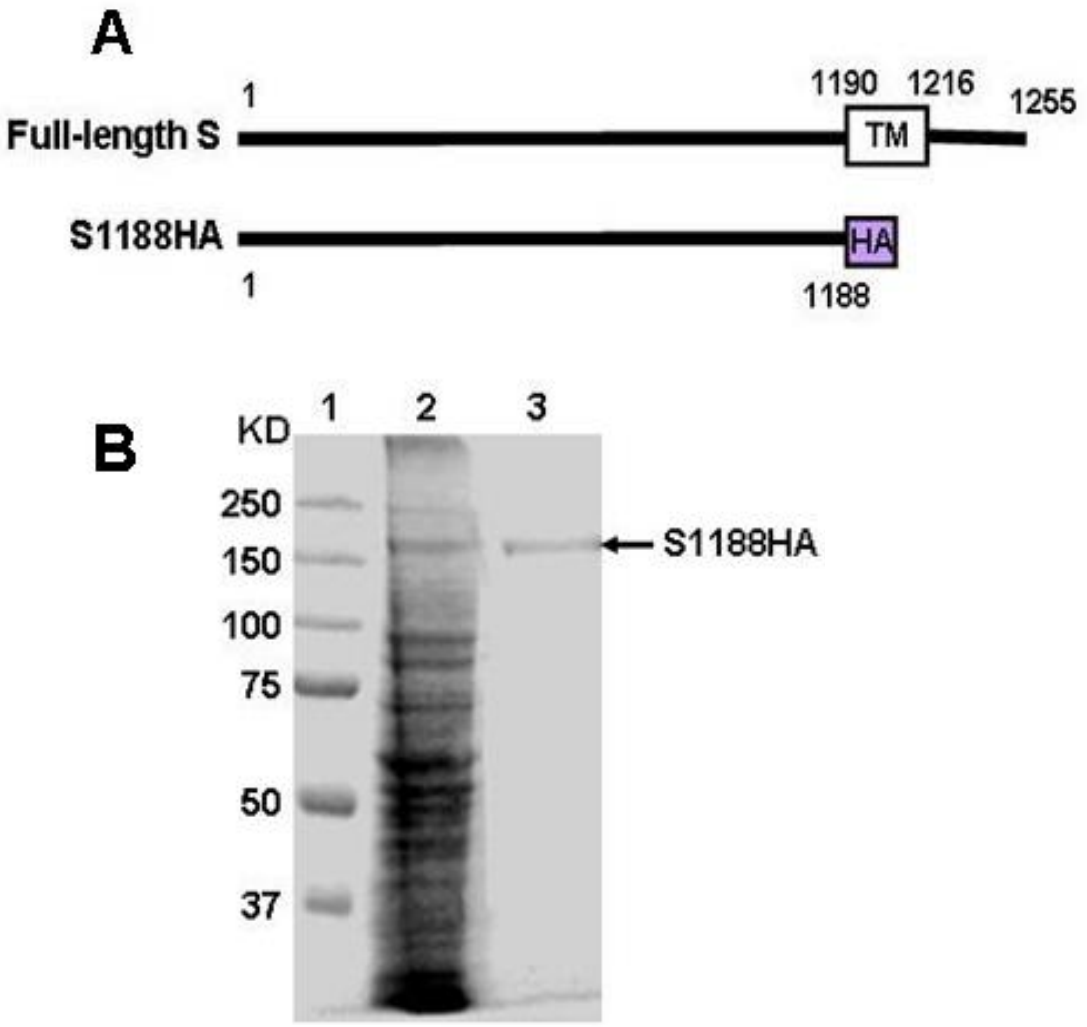

Figure 16. Expression and purification of S1188HA protein in Sf9 cells.

The ectodomain of SARS-CoV S protein with a HA tag at its C-terminus, S1188HA (A) was expressed in Sf9 cells using baculovirus expression system and purified by anti-HA affinity column. The cell lysate and purified protein were resolved by 8% SDS-PAGE, and stained with Coomassie blue R250 (B).

Lane 1: Protein marker (Bio-Rad),
Lane 2: Cell lysate containing S1188HA,
Lane 3: Purified S1188HA.

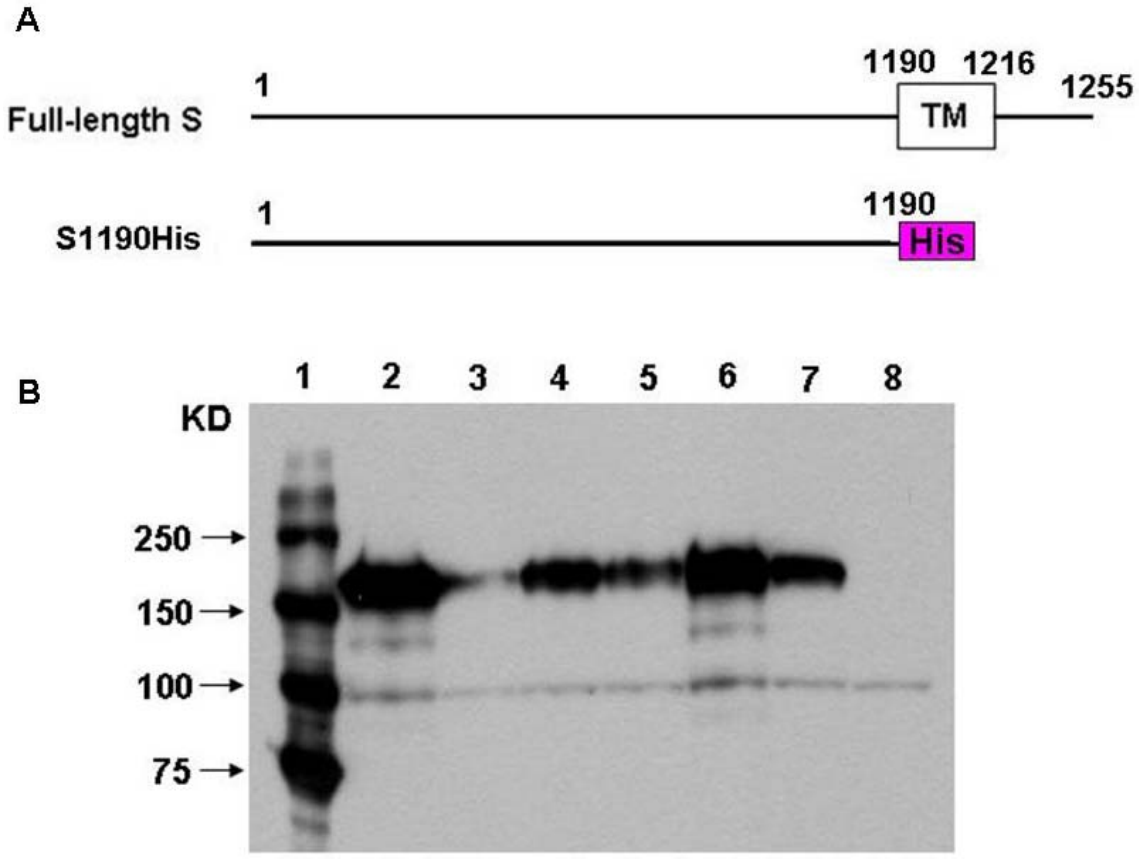

Figure 18. Expression of S1190His protein in Sf9 cells.

The ectodomain of SARS-CoV S protein, S1190 with a 6x His tag at its C-terminus (A) was expressed in Sf9 cells using baculovirus expression system. Sf9 cells were infected with different amount recombinant baculoviruses containing S1190His gene, the cell lysate were resolved on 8% SDS-PAGE. Proteins were transferred onto the PVDF membrane, and subjected to Western blot using mouse anti-penta-His antibody as primary probe (B).

Lane 1: Protein molecular weight marker (Bio-Rad),
Lane 2-7: S1190His expressed in Sf9 cells (infected with different amount of recombinant virus),
Lane 8: Negative control, Sf9 cells without infection.

(4) Summary of S protein expression in different systems.

Summing up all the systems I used to express SARS-CoV S protein (Table 3), the insect cell system could provide the highest expression level and enough amount of protein for purification and function study. Codon-optimization of S gene is necessary for the mammalian cell system to express this protein under the CMV promoter, while the original gene codon could only be expressed in mammalian cells under the T7 promoter. The failure of expression of S protein in *E. coli* may be due to the codon bias. In subsequent studies, I used insect cells to express soluble form of S protein to obtain enough purified S protein ectodomain to do the receptor-binding study, and used codon-optimized construct and expressed S protein in mammalian cells to assemble SARS-CoV pseudovirus.

Table 3. Summary of S protein expression in different systems

Host cell	Mammalian cell					Insect cell		
Codon	Original codon			Codon-optimized		Original codon		
Constructs	pcDNA3.1(+)HA-S1255	pcDNA3.1(+)HA-S795	pcDNA3.1(+)HA-S663	pcDNA3.1(+)HA-S540	pcDNA-OPT9-S	pVL1392 SHA	pVL1392 S1188HA	pVL1392 S1190His
Promoter	T7	T7	T7	T7	CMV	polyhedrin	polyhedrin	polyhedrin
Encoding	Full-length S protein	N-terminal 795 aa of S protein	N-terminal 663 aa of S protein	N-terminal 540 aa of S protein	Full-length S protein	Full-length S protein	N-terminal 1188 aa of S	N-terminal 1190 aa of S
Expression level	+	++	++	+++	++	++++	++++	++++

iii. Assembly of pseudotyped SARS-CoV.

Due to the highly infectious and virulent nature of SARS-CoV, a safe model that could mimic the live virus became very critical to my study. A pseudovirus carrying the SARS-CoV S protein is a suitable alternative.

The lentiviral expression vector contains the genetic elements required for packaging, transduction, stable integration of the viral expression construct into genomic DNA, and expression of the target protein, siRNA, cDNA or reporter. The lentiviral packaging plasmids provide all of the proteins essential for transcription and packaging of an RNA copy of the expression construct into recombinant pseudoviral particles. For production of high-titer of pseudoviral particles, producer cells (*e.g.*, 293T cells) require to be transiently co-transfected with the expression and packaging vectors. Expression constructs packaged in pseudoviral particles are secreted by producer cells in culture media and can be used directly to transduce expression constructs into target cells. Following transduction into the target cells, this expression construct is reverse transcribed and integrated into the genome of the target cell, providing a high level of expression of effector or reporter molecules.

The most popular lentiviral expression system is HIV based (Federico, 2003; Heiser, 2004; Machida, 2003). Many groups have used this system to assemble pseudotyped SARS-CoV. It has become a safe and mature method to study SARS-CoV (He *et al*., 2004; Ni *et al*., 2005; Yi *et al*., 2005; Zhang *et al*., 2004).

In this study, an HIV-based lentiviral system was used to assemble viruses pseudotyped with SARS-CoV S protein. pNL4-3Luc^{+}Env^{-}Vpr^{-} is the packaging

plasmid which contains the packaging signal and firefly luciferase reporter gene. This reporter gene could provide a sensitive indicator for infectivity assay. pcDNA3.1-OPT9-S is the plasmid containing codon-optimized SARS-CoV S gene which can produce S protein. After co-transfection of 293T cells with these two plasmids, the expressed S protein can be incorporated into the viral particle to produce pseudovirus carrying the SARS-CoV S protein. Two frameshifts (5′ Env and Vpr aa 26) of HIV gene in pNL4-3 $Luc^{+}Env^{-}Vpr^{-}$ render the pseudovirus to complete only a single round of replication. All these characters make this pseudotyped SARS-CoV become a suitable model for the entry study.

D. Conclusions

In conclusion, a range of expression systems were examined to obtain adequate amounts of S protein and establish a suitable model to mimic viral entry. After comparing all these systems and weighing up their advantages and disadvantages, I found that the baculovirus expression system could produce a high amount of S protein with post-translational modification (Table 3). The ectodomain of S protein with an HA tag (S1188HA) expressed in this system could be purified using anti-HA affinity column. Mammalian cell system could produce the native form of S protein, and the codon-optimized S gene could be expressed in mammalian cells without infection of T7 vaccinia virus, this advantage makes this system become most suitable for producing pseudotyped SARS-CoV, which could be used as a good model to study the viral entry and infection.

Chapter 3. Clustering of host receptors by lipid rafts in SARS-CoV entry

A. Introduction

Lipid rafts are functional lipid microdomains in the cell membrane, consisting of cholesterol, sphingolipid, and their associated proteins (Simons and Ikonen, 1997). In polarized cells, lipid rafts are concentrated at the apical surface, whereas in nonpolarized cells they are dispersed over the cell surface as small (about 30 to 50 nm in diameter), highly dynamic domains able to include about 10 to 30 protein molecules (Pralle *et al.*, 2000). The raft size can be modulated by oligomerization of raft components. However, the precise mechanism and signals that may lead to raft aggregation are not clear. Lipid rafts are involved in membrane trafficking, cell morphogenesis, and signal transduction (Simons and Ikonen, 1997; Simons and Toomre, 2000). Numerous signaling molecules are concentrated in raft domains, including Src family kinases and heterotrimeric G proteins, as well as molecules involved in Ca^{2+} influx (Simons and Ikonen, 1997). It has been shown that recruitment of the T-cell receptor (TCR) to lipid rafts upon receptor stimulation coincided with the aggregation of lipid rafts and triggering of signaling cascades (Janes, Ley, and Magee, 1999).

In addition to the demonstrated role in signal transduction, lipid rafts serve as a docking site for entry of microbial pathogens such as viruses, bacteria and toxins as well as for virus assembly and budding (Kovbasnjuk, Edidin, and Donowitz, 2001; Rawat *et al.*, 2003; Suomalainen, 2002). It has been shown that the assembly and

budding of different viruses, including measles virus (Manie *et al.*, 2000), Semliki Forest virus (Lu and Kielian, 2000), influenza virus (Scheiffele *et al.*, 1999), Sindbis virus (Lu, Cassese, and Kielian, 1999), murine leukemia virus (Lu and Silver, 2000), and HIV-1 (Hermida-Matsumoto and Resh, 2000; Nguyen and Hildreth, 2000; Ono and Freed, 2001), take place in lipid rafts.

Recent studies have implicated lipid rafts involved in the cell uptake of *Mycobacterium bovis* (Gatfield and Pieters, 2000), the malaria parasite *Plasmodium falciparum* (Samuel *et al.*, 2001) and some viruses, including HIV-1 (Popik, Timothy M, and Au, 2002), simian virus 40 (Parton and Lindsay, 1999; Pelkmans, Kartenbeck, and Helenius, 2001), murine leukemia virus (Lu and Silver, 2000), and MHV (Choi, Aizaki, and Lai, 2005).

Raft-colocalized receptors may be crucial for entry of different pathogens. The entry of enveloped viruses involves the attachment of virus to the receptor, followed by fusion between virus and cell membrane, which can be either plasma or endosomal membrane. Lipid rafts in association with cellular receptors are known to be involved in the viral entry process through several different ways. These include the association of viral glycoproteins with lipid rafts of either the viral envelope or the target membrane. Hemagglutinin of influenza virus (Scheiffele, Roth, and Simons, 1997), gp120-gp41 of HIV-1 (Pickl, Pimentel-Muinos, and Seed, 2001) and glycoprotein of Ebola virus (Bavari *et al.*, 2002) are associated with lipid rafts in the virion. The E1 protein of Semliki Forest virus is inserted selectively into the cholesterol-rich microdomains of the target cell membrane (Ahn, Gibbons, and

Kielian, 2002). CD4 and CCR5, the receptor and coreceptor, respectively, of HIV-1 are associated with lipid rafts (Del Real *et al.*, 2002; Kozak, Heard, and Kabat, 2002; Popik, Timothy M, and Au, 2002).

Although there was no direct evidence that lipid rafts are involved in coronavirus replication, previous studies have implied that cholesterol and cholesterol-related environments may regulate coronavirus replication; supplementation of cholesterol in the culture medium resulted in marked enhancement of MHV-induced cell fusion (Daya, Cervin, and Anderson, 1988), and a hypercholesterolemic diet increased the susceptibility of mice to MHV-3 infection (Braunwald *et al.*, 1991). In the case of HCoV-229E, virus entry was inhibited by depletion of cholesterol, resulting in the disruption of viral association with the cellular receptor, CD13 (Nomura *et al.*, 2004). Moreover, knockdown of caveolin-1 affected the entry of HCoV-229E but not its binding (Nomura *et al.*, 2004), although the significance of caveolin-1 in virus entry has yet to be demonstrated.

Despite this evidence, the role of lipid rafts in mediating coronavirus entry into host cells remains controversial. Thorp and Gallagher (Thorp and Gallagher, 2004) showed that cholesterol-rich microdomains were crucial for the entry and fusion of MHV, but the MHV receptor (MHVR) did not associate with lipid rafts; anchoring of the MHV receptor to lipid rafts did not enhance MHV infection. These results indicate that cholesterol-rich microdomains are implicated in the viral entry in yet another uncharacterized mechanism.

Culmulative experimental evidence indicates that lipid rafts provide a

microenvironment for the interaction of viral proteins with their cellular receptor(s) to take place, either as signaling platforms or as concentration devices. Because the role of lipid rafts in SARS-CoV entry is mostly unknown, my aim was to determine whether lipid rafts are involved in the SARS-CoV entry process. In this study, I used the African green monkey kidney cell line Vero E6 which permits replication of SARS-CoV (Ksiazek *et al.*, 2003) as host cell. I found that productive entry of SARS-CoV pseudotyped virus into Vero E6 cells requires the presence of intact lipid rafts, potentially due to the significant concentration and localization of SARS-CoV receptor ACE2 in these microdomains.

B. Materials and methods

i. Materials

Chemical Reagents. In addition to the materials listed in Chapter 2, the following reagents were utilized in this study. Methyl-β-cyclodextrin (MβCD), paraformaldehyde and enzyme-linked immunosorbent assay (ELISA)-grade bovine serum albumin (BSA) were all purchased from Sigma-Aldrich.

Antibodies. In addition to the antibodies listed in Chapter 2, the following antibodies were utilized in this study. Alexa Fluor 594-conjugated CT-B, rabbit anti-CT-B antibody, Alexa Fluor 488-conjugated donkey anti-goat antibody, Alexa Fluor 488-conjugated donkey anti-rabbit antibody and Alexa Fluor 488-conjugated donkey anti-mouse antibody were all purchased from Molecular Probes. Mouse anti-HA (F-7) monoclonal antibody and rabbit anti-caveolin-1 (N-20) polyclonal antibody were purchased from Santa Cruz Biotechnology. Goat anti-ACE2 (AF933)

polyclonal antibody was purchased from R&D systems. Mouse anti-CD71 monoclonal antibody was purchased from Zymed Laboratories Inc. HRP-conjugated polyclonal rabbit anti-goat antibody was purchased from DakoCytomation.

Cell lines. In addition to the cells listed in Chapter 2, Vero E6 cell was utilized in this study. Vero E6 cells (Vero C1008 [CRL-1586]) were purchased from ATCC.

Kits and Enzymes. In addition to the kits and enzymes listed in Chapter 2, the following kits were utilized in this study. The luciferase assay system was purchased from Promega. The ProLong antifade kit was purchased from Molecular Probes.

Solutions. In addition to the solutions listed in Chapter 2, the following solutions were utilized in this study. TNE buffer contains 25 mM Tris, 150 mM NaCl and 5 mM EDTA, pH 7.5. TNE lysis buffer contains 1% Triton X-100 (v/v) in TNE buffer with complete protease cocktail. 80% Sucrose contains 80% sucrose (w/v) in TNE buffer. 30% Sucrose contains 30% sucrose (w/v) in TNE buffer. 5% Sucrose contains 5% sucrose (w/v) in TNE buffer.

Instruments. The following instruments were utilized in this study: TD-20/20 Luminometer (Tuner Designs); LSM 510 META laser scanning confocal microscope (Zeiss) and Becton Dickinson FACSCalibur system (BD Biosciences).

ii. Methods

Cell culture. Vero E6 cells and 293T cells were maintained in DMEM medium supplemented with 10% new born calf serum and 50 μg/ml penicillin-streptomycin.

Expression and purification of S protein ectodomain. Described in Chapter 2.

Preparation of pseudotyped SARS-CoV. The pseudovirus was generated by

co-transfecting 293T cells using calcium phosphate transfection method with pNL4-3Luc$^+$Env$^-$Vpr$^-$ and pcDNA3.1-OPT9-S or pVPack-VSV-G. The detail information could be found in Chapter 2.

Single-cycle infectivity assay. Vero E6 cells (3 x 10^4 cells/well in 300 µl) were plated in 48-well plates (Nunc) and cultured at 37°C, 5% CO_2, overnight. On the following day, Vero E6 cells were incubated for 1 h with standardized amounts of pseudotyped viruses in DMEM (based on RT Assay, 0.5 U of RT/well). After washing, the cultures were propagated for 48 h. After 48 h of incubation, the cells were washed, and lysed with 100 µl of lysis buffer (provided in luciferase assay system kit). Luciferase activity was determined after addition of 100 µl of luciferase assay reagent (provided in luciferase assay system kit) into 20 µl aliquot of cell lysate and counting the resultant scintillation for 15 s using a TD-20/20 Luminometer.

Cholesterol depletion from Vero E6 cell membrane. Before extraction of cholesterol, Vero E6 cells were washed twice with DMEM. The cells were incubated with 10 mM MβCD in DMEM supplemented with 0.1% BSA for 30 min at 37°C. Under these conditions, cell viability was not significantly affected, as determined by trypan blue exclusion. As controls, cells were exposed to medium alone at 37°C. After incubation the cells were washed twice with ice-cold DMEM-0.1% BSA before use.

Biochemical isolation of lipid rafts. Cells (2-5 ×10^7) 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 supplemented with complete protease inhibitor cocktail. The cell lysates were

homogenized with 80 strokes of a Dounce homogenizer and centrifuged (5 min, 720 g, 4°C, Sigma 3-18K centrifuge) to remove insoluble material and nuclei. The supernatant was mixed with 1 ml of 80% sucrose in TNE buffer, placed at the bottom of an ultracentrifuge tube, and overlaid with 6 ml of 30% and 3 ml of 5% sucrose in TNE buffer. The lysates were ultracentrifuged at 4°C in a SW41 rotor (Beckman) for 18 h at 38,000 rpm. After centrifugation, the Triton X-100-insoluble, low-density material containing rafts was visible as a band migrating on the boundary between 5 and 30% sucrose. Eleven 1 ml fractions were collected from the top to the bottom of the tube and analyzed immediately by Western blot using rabbit anti-caveolin-1, mouse anti-CD71 and goat anti-ACE2 antibodies or stored at -80°C.

To investigate the presence of SARS-CoV S protein ectodomain in lipid rafts after binding to ACE2 receptor, Vero E6 cell were incubated with S1188HA (10μg/ml) for 30 min at 4°C, after extensive washing with PBS, cells were collected, lysed and lipid rafts were isolated using the same procedure, S1188HA was detected using rabbit anti-HA and HRP-conjugated swine anti-rabbit antibody by Western blot.

Western blot analysis. Aliquots of 20 μl of individual sucrose gradient fractions were resolved by 8% or 15% SDS-PAGE. The proteins were transferred onto PVDF membranes and probed with primary antibody: rabbit anti-caveolin-1, mouse anti-CD71 and goat anti-ACE2 or rabbit anti-HA antibodies, and following with second antibody: HRP-conjugated swine anti-rabbit, rabbit anti-mouse and rabbit anti-goat antibodies, as described previously. Bound antibodies were detected with the ECL Plus Western blotting detection reagents.

Rafts aggregation and confocal microscopy imaging. To detect ACE2 in lipid rafts, Vero E6 cells were cultured on 4-well chamber slide (Iwaki) to a 50% confluence. The cells were first incubated for 30 min at 4°C with both Alexa Fluor 594-conjugated CT-B (10μg/ml) and goat anti-ACE2 antibodies (1:50 in PBS supplemented with 0.1% BSA) at the same time, followed by extensive washing with ice-cold PBS.

To detect S protein ectodomain bound to Vero E6 cells, the cells were incubated with S1188HA (10μg/ml) for 30 min at 4°C. After extensive washing with PBS, the cells were incubated with Alexa Fluor 594-conjugated CT-B (10μg/ml) and mouse anti-HA antibodies (1:50 in PBS supplemented with 0.1% BSA, 30 min, 4°C), followed by extensive washing with ice-cold PBS.

Lipid rafts aggregation was initiated by incubation with rabbit anti-CT-B antibody in the presence of Alexa Fluor 488-conjugated donkey anti-goat antibody or Alexa Fluor 488-conjugated donkey anti-mouse antibody (1:50 v/v, 30 min, 4°C). After being washed twice with ice-cold PBS, the cells were fixed (15 min, 4°C) with fresh 4% paraformaldehyde. In addition, lipid raft ganglioside GM1 was identified by incubation with Alexa Fluor 594-conjugated CT-B (10μg/ml) alone. The cells were washed, fixed, and mounted with ProLong antifade kit.

Fluorescently labeled cells were analyzed with a Zeiss LSM 510 META laser scanning confocal microscope with 100x objective lenses. The images were processed using LSM 510 META software.

Flow cytometry. Expression of ACE2 on the cell surface of Vero E6 and the

binding of S1188HA to Vero E6 and 293T cells were evaluated by flow cytometry. The cells were treated or untreated with MβCD (30 min, 37°C) in DMEM, subsequently were trypsinized and collected by centrifugation (800 rpm, 5 min, 4°C, Sigma 3-18K centrifuge). The cells were incubated for 30 min at 4°C with or without S1188HA (10μg/ml). After three washes with PBS, the cells were incubated with goat anti-ACE2 antibody or rabbit anti-HA antibody (1:50, v/v). After three times washes with PBS, the cells were incubated with Alexa Fluor 488-conjugated donkey anti-goat or donkey anti-rabbit antibodies (1:50, v/v) for 30 min at room temperature, followed by three times washes with PBS. Samples were analyzed on a Becton Dickinson FACS Calibur within 2 h of immunofluorescent staining.

C. Results and discussion

i. ACE2, the receptor for SARS-CoV, is localized in lipid rafts in Vero E6 cells.

ACE2 has been identified as a functional receptor for SARS-CoV (Li *et al.*, 2003). To determine the relationship between ACE2 and lipid rafts, raft isolation was performed using cold nonionic detergents by ultracentrifugation. Proteins detected in low-density fractions during sucrose gradient centrifugation are considered to be associated with rafts (Simons and Toomre, 2000). Using this standard biochemical isolation method, lipid rafts were found in gradient fractions 3 and 4 after ultracentrifugation. To confirm these results, the distribution of ACE2 with raft and nonraft markers in live cells was analyzed. In normal Vero E6 cells, the raft marker caveolin-1 was predominantly localized in fractions 3 and 4 (Figure 19A). In contrast, a nonraft marker, transferrin receptor, CD71 (Janes, Ley, and Magee, 1999)

was confined to detergent-soluble fractions 10 and 11 (Figure 19A). Western blot analysis of the same fractions with goat anti-ACE2 antibody showed that approximately 70% of ACE2 was co-fractionated with the raft marker caveolin-1 in fraction 3 and 4 (Figure 19A), revealing that ACE2 is largely associated with lipid rafts in Vero E6 cells.

To confirm the biochemical results obtained using flotation gradients, the association of ACE2 with lipid rafts was investigated in intact cells by confocal microscopy using fluorescently labeled CT-B, which specifically binds to a raft-resident ganglioside GM1 (Janes, Ley, and Magee, 1999). To examine the presence of ACE2 in rafts, I analyzed whether ligation of GM1 with CT-B subunit and the consequent aggregation of rafts affects the localization of ACE2. Treatment of Vero E6 cells with Alexa Fluor 594-conjugated CT-B alone at 4°C resulted in a uniform staining pattern of the plasma membrane (Figure 19B, GM1, unpatched). However, a 30 min-incubation with rabbit anti-CT-B at 4°C led to the formation of large GM1 patches (Figure 19B, GM1, patched). Incubation of cells simultaneously with Alexa Fluor 594-conjugated CT-B/rabbit anti-CT-B and goat anti-ACE2/Alexa Fluor 488-conjugated donkey anti-goat antibodies, resulted in substantial colocalization of ACE2 with GM1 patches (Figure 19B overlay). The observed sequestration of ACE2 by lateral cross-linking of GM1 demonstrates that virus receptor and GM1 reside in the same lipid environment in the plasma membrane and further supports that SARS-CoV receptor, ACE2, is associated with lipid rafts.

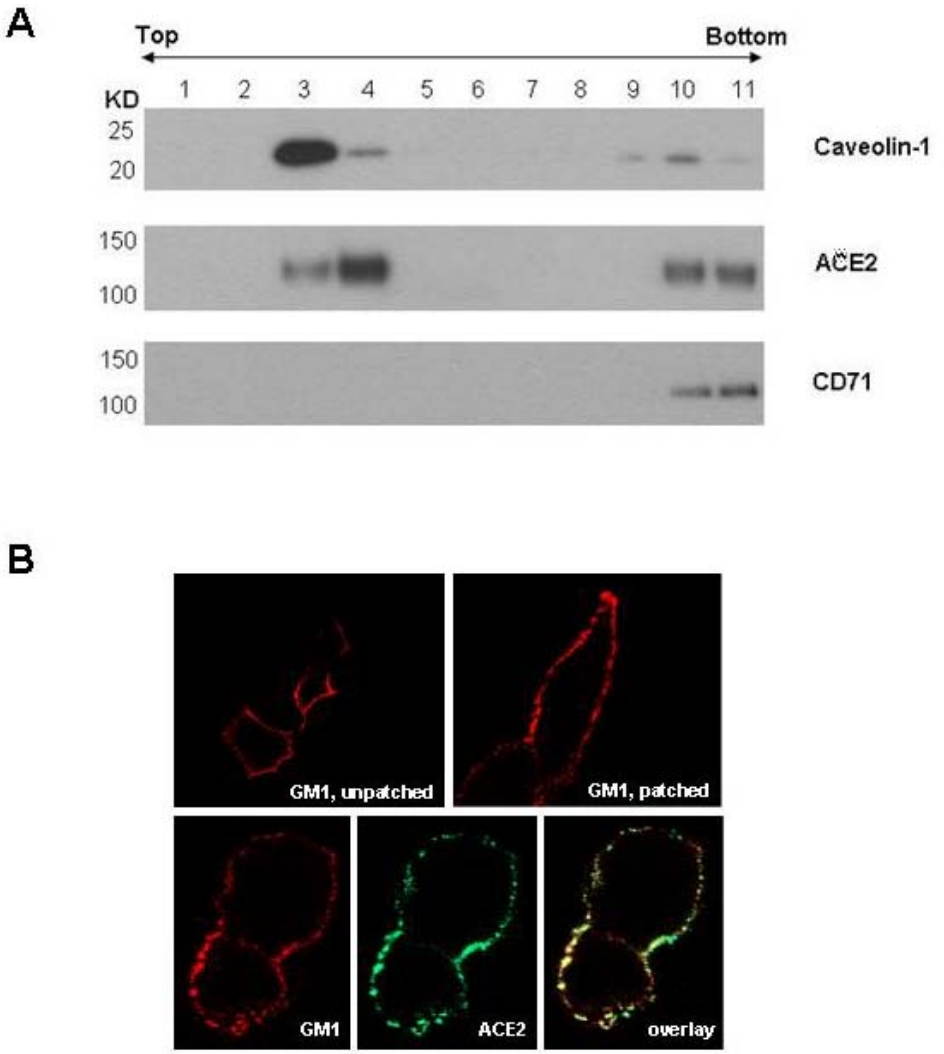

Figure 19. ACE2 is localized in lipid rafts in Vero E6 cells.

Vero E6 cells were lysed with Triton X-100 lysis buffer, and cell postnuclear extracts were fractionated on a sucrose density gradient as described in Materials and Methods. Equal volumes (20 µl) of individual fractions were resolved by SDS-PAGE, caveolin-1, ACE2 and CD71 were detected separately by western blot using respective antibodies (A). For the confocal microscopy imaging, Vero E6 cells were incubated at 4°C with Alexa Fluor 594-conjugated CT-B alone, which specifically binds to the raft-resident ganglioside GM1, or together with goat anti-ACE2 antibody. Raft aggregation was initiated by incubation of the labeled cells at 4°C with rabbit anti-CT-B antibody for 30 min with or without Alexa Fluor 488-conjugated donkey anti-goat antibody as described in Materials and Methods. Cells were fixed and visualized by confocal microscopy. Single confocal sections are shown (B). The colocalization of SARS-CoV receptor ACE2 with the raft marker GM1 is presented (B, overlay). Control cells incubated with Alexa Fluor 594-conjugated CT-B alone at 4°C show uniform distribution of GM1 before raft aggregation (B, GM1, unpatched). While incubation with anti-CT-B antibody led to formation of large GM1 patches (B, GM1, patched).

Although ACE2 was identified as a SARS-CoV receptor just a few months after the virus itself was discovered, it is not clear whether coreceptor is required for the entry of this virus; and whether lipid rafts are involved to cluster these receptors for specific docking of host cells as found in HIV-1 and some other viruses. The results presented here show that using the biochemical isolation method, about more than 70% of ACE2 is present in gradient fractions 3 and 4 along with the raft marker caveolin-1, while the nonraft marker CD71 is found in fractions 10 and 11. These results indicate that most of the ACE2 in Vero E6 cells is associated with lipid rafts. Confocal microscopy imaging results confirm the presence of ACE2 in lipid rafts. When using Alexa Fluor 594-conjugated CT-B alone to label the cell, a uniform staining pattern of plasma membrane was observed. However a 30 min-incubation with anti-CT-B antibody led to formation of large GM1 patches which indicates the aggregation of lipid rafts. When the cell was simultaneously labeled with anti-ACE2 antibody, the colocalization of ACE2 with GM1 patches was observed. These data support the conclusion that ACE2 localizes in lipid rafts in Vero E6 cells.

ii. Depletion of cholesterol disintegrates lipid rafts in Vero E6 cells, but does not affect the expression of ACE2 on cell surface.

The association of ACE2 with lipid rafts was further investigated by depletion of cholesterol from cell membrane using MβCD which leads to disintegration of lipid rafts (Kabouridis *et al.*, 2000; Xavier *et al.*, 1998). Previous studies have confirmed that membrane cholesterol is required to maintain lipid rafts integrity and extraction of cholesterol from plasma membrane with MβCD can destroy the lipid rafts and

relocalize raft markers into detergent-soluble fractions (Popik, Timothy M, and Au, 2002). Here the effectiveness of this treatment was examined through analysis of the distribution of raft-specific components in cell membrane microdomains. Extraction of cholesterol was found to have different effects on these raft and nonraft markers. Expression and localization of CD71 in Vero E6 cell membranes remained unchanged after cholesterol extraction (Figure 20A). In contrast, cholesterol was required to maintain the association of caveolin-1 with lipid rafts. Extraction of cholesterol from cell membranes by MβCD efficiently destabilized lipid raft integrity, resulting in the re-localization of both ACE2 and caveolin-1 into fractions 10 and 11 (Figure 20A) after ultracentrifugation.

As cell surface expression of ACE2 is essential for its function as an efficient receptor for SARS-CoV, experiments were designed to determine if the treatment of Vero E6 cell with MβCD would affect cell surface expression of ACE2. Vero E6 cells either treated or untreated with MβCD were labeled with goat anti-ACE2 antibody followed with Alexa Fluor 488-conjugated rabbit anti-goat antibody, and analyzed by flow cytometry. No significant change in the cell surface expression of ACE2 was observed after depletion of cholesterol (Figure 20B), indicating that disruption of lipid rafts does not block the translocation of ACE2 to the cell surface.

Taken together, these results suggest that lipid rafts are important in SARS-CoV entry and cholesterol preserves raft integrity to permit the accumulation of ACE2 receptor in lipid rafts, while depletion of cholesterol does not down-regulate ACE2 surface expression.

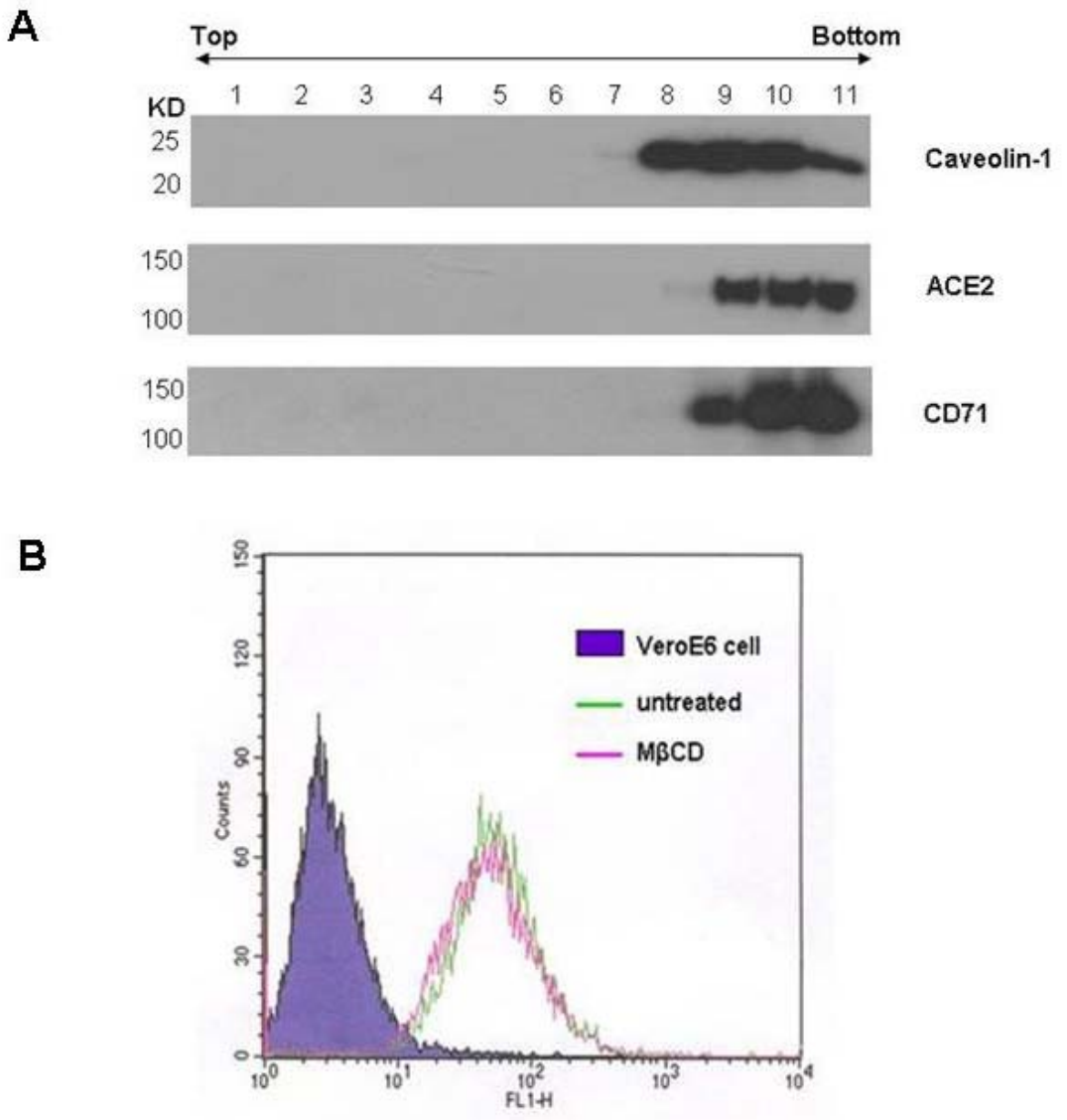

Figure 20. Cholesterol depletion disintegrates lipid rafts in Vero E6 cells, but does not significantly change surface expression of ACE2.

Vero E6 cells were treated with 10 mM MβCD to extract cholesterol. Triton X-100 cell lysates were fractionated on a sucrose density gradient as described in Materials and Methods. Equal volumes (20 µl) of individual fractions were resolved by SDS-PAGE, caveolin-1, ACE2 and CD 71 were detected by Western blot using respective antibodies (A). For flow cytometry, Vero E6 cells untreated or treated with 10 mM MβCD were collected and labeled with goat anti-ACE2 antibody, following with Alexa Fluor 488-conjugated-donkey anti-goat antibody, and analyzed by flow cytometry as described in Materials and Methods (B). The VeroE6 cell peak is detected from the Vero E6 cells without antibody labeling which serve as the blank of this experiment.

iii. Ectodomain of S protein (S1188HA) is associated with lipid rafts in Vero E6 cells upon binding to its receptor.

Since ACE2 was shown to be present in lipid rafts, I next investigated whether the S protein was able to interact with raft microdomains. Vero E6 cells were incubated at 4°C for 30 min with S protein ectodomain S1188HA (aa 1-1188 of S protein with a , and lipid rafts were prepared. Analysis of fractions by Western blot using rabbit anti-HA antibody showed that a majority of S1188HA was co-fractionated with caveolin-1 in detergent-insoluble rafts (Figure 21A, fractions 3 and 4). A minor proportion of the protein was also detected in soluble fractions (Figure 21A, fraction 11). These results demonstrate that S1188HA is colocalized with raft marker caveolin-1 in lipid rafts.

The association of S1188HA with lipid rafts was further investigated in intact cells by confocal microscopy. Vero E6 cells were incubated with S1188HA at 4°C for 30 min. The cells were extensively washed and exposed to Alexa Fluor 594-conjugated CT-B and mouse anti-HA antibody at 4°C for 30 min. After washing, raft aggregation was initiated by reaction with rabbit anti-CT-B antibody in the presence of Alexa Fluor 488-conjugated donkey anti-mouse antibodies. The sequestration of S1188HA after lateral cross-linking of GM1 was observed (Figure 21B, overlay). These results suggest that the ectodomain of S protein and GM1 reside in the same lipid rafts environment and further support that S1188HA is functional and could bind to its receptor which associates with lipid rafts. Taken together, these data indicate that the docking sites of SARS-CoV on Vero E6 cells are mostly located in lipid rafts.

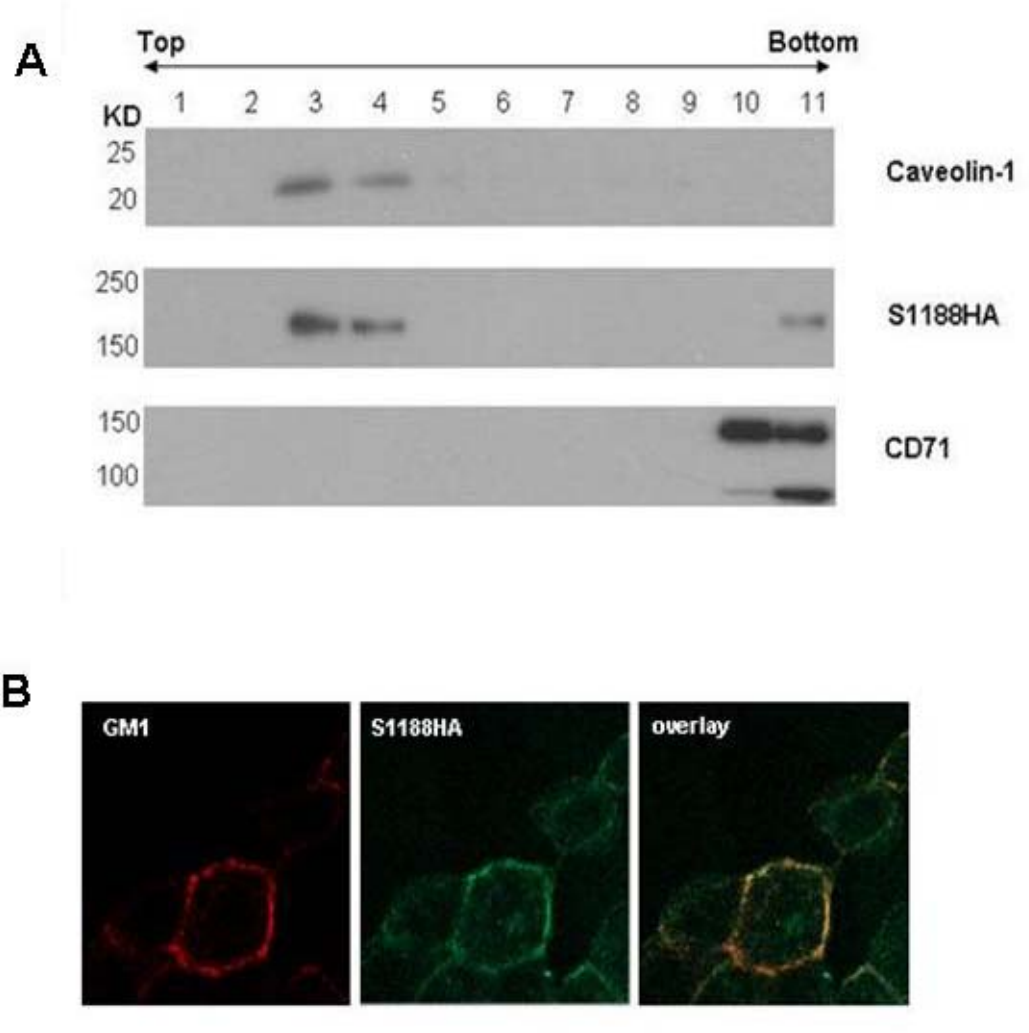

Figure 21. S1188HA associates with lipid rafts in Vero E6 cells.

Vero E6 cells were incubated for 30 min at 4°C with S1188HA (10μg/ml). The cells were washed to remove unbound S1188HA. For biochemical isolation, cells were lysed in Triton X-100 lysis buffer, and the postnuclear extracts were fractionated in a sucrose density gradient. An aliquot of each fraction (20 μl) was resolved on SDS-PAGE, caveolin-1, S1188HA and CD71 were detected by Western blot using respective antibodies as described in Materials and Methods. The results showed that membrane-bound S1188HA colocalizes preferentially with raft marker caveolin-1, but not the nonraft marker CD71 (A). For confocal microscopy imaging, cells were exposed for 30 min at 4°C to mouse anti-HA antibody and Alexa Fluor 594-conjugated CT-B, following with rabbit anti-CT-B and Alexa Fluor 488-conjugated donkey anti-mouse antibody to induce cross-linking. Distribution and colocalization of the raft marker GM1 with S1188HA were analyzed by confocal microscope as described in Materials and Methods. Single confocal sections are shown (B). The colocalization of SARS-CoV S protein ectodomain S1188HA with the raft marker GM1 is presented (B, overlay).

iv. Depletion of cholesterol does not affect the binding of ectodomain of S protein to Vero E6 cells.

To examine whether the association of S1188HA with lipid rafts is due to S1188HA directly interacting with lipid rafts or interacting with its receptor first, and then associated with lipid rafts, the effect of depletion of cholesterol on the binding of S1188HA to Vero E6 cells was investigated by biochemical fractionation of MβCD-treated Vero E6 cells. In contrast to the results obtained from untreated cells, depletion of cholesterol resulted in the detection of both S1188HA and caveolin-1 in the soluble fractions together with CD71 (Figure 22A). This indicates that S1188HA could still bind to the cells but shifted to the nonraft environment along with raft marker caveolin-1 after depletion of cholesterol and S1188HA associates with lipid rafts upon its binding to its receptor, but does not directly bind to lipid rafts.

The binding of S1188HA to MβCD-treated Vero E6 cells was further analyzed by flow cytometry. In agreement with the biochemical data, incubation of S1188HA with MβCD-treated or untreated Vero E6 cells showed that the protein bound equally well to both treated and untreated cells, no obvious change of protein binding was observed (Figure 22B). The negative control, 293T cells showed no protein binding after incubation with S1188HA (Figure 22C), confirming that the increase in fluorescence observed in Vero E6 cells was due to the specific binding of S1188HA to its receptor on the cell surface. These results suggest that binding of S1188HA to its receptor determines its association with lipid rafts, and the ectodomain of S protein could efficiently bind to its receptor regardless of the integrity of lipid rafts.

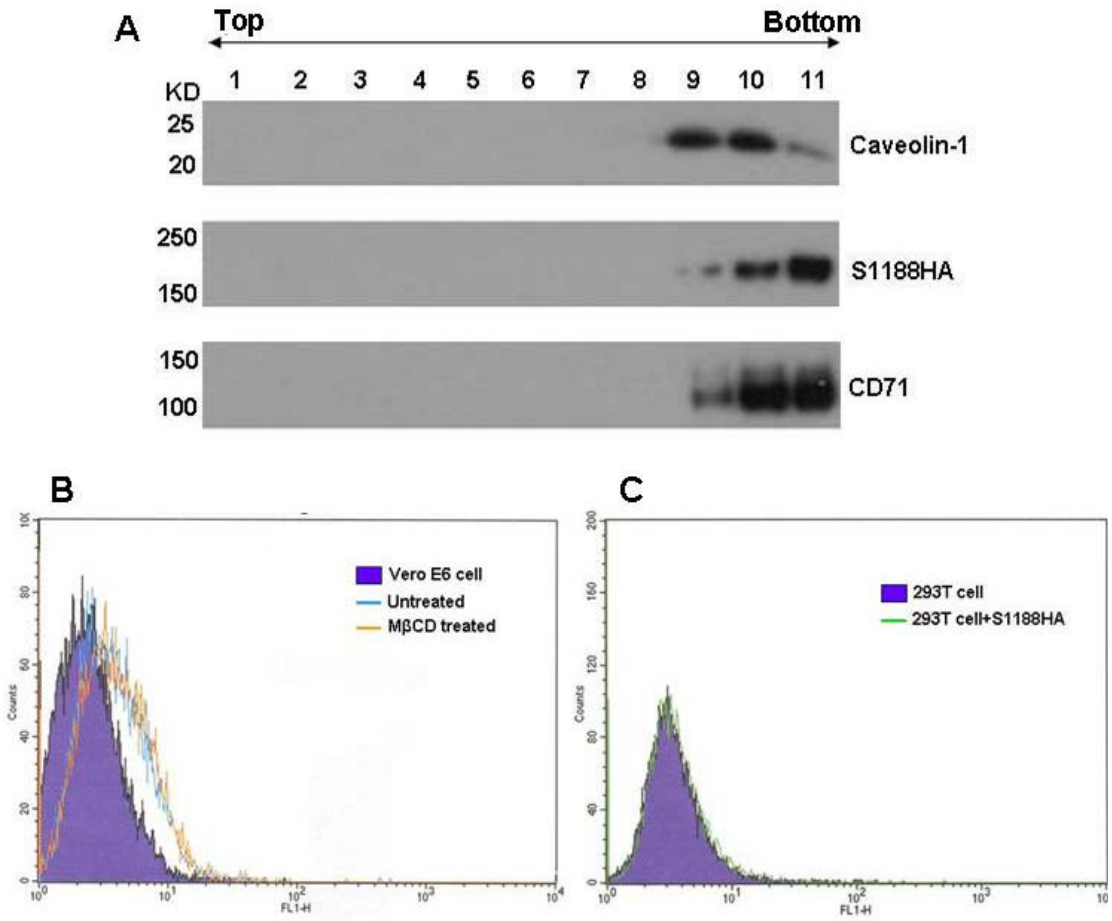

Figure 22. Depletion of cholesterol shifts S1188HA to nonraft environment, but does not significantly change its binding to Vero E6 cells.

Vero E6 cells were treated with 10 mM MβCD, and then incubated for 30 min at 4°C with S1188HA (10μg/ml). After washing to remove unbound S1188HA, cells were lysed in Triton X-100 lysis buffer, and the postnuclear extracts were fractionated in a sucrose density gradient. An aliquot of each fraction (20 μl) was resolved on SDS-PAGE, caveolin-1, S1188HA and CD71 were detected by Western blot as described in Materials and Methods. The results showed that membrane-bound S1188HA shifted to soluble fractions with the raft marker caveolin-1, and was detected at same fractions as the nonraft marker CD71 (A). For flow cytometry, Vero E6 cells untreated or treated with 10 mM MβCD and 293T cells were collected and incubated for 30 min at 4°C with S1188HA (10μg / μl), after three times washing with PBS, cells were labeled with rabbit anti-HA antibody and Alexa Fluor 488-conjugated donkey anti-rabbit antibody, and subsequently analyzed by flow cytometry as described in Materials and Methods. There is a peak shift in the Vero E6 cells incubated with S1188HA comparing with the negative control (solid peak, Vero E6 cell labeled with first and secondary antibody, but no incubation with S1188HA), but there is no peak shift comparing the treated with untreated Vero E6 cells (B). While in 293T cells, there is no peak shift in S1188HA incubated and no S1188HA incubated cells (C).

v. Depletion of cholesterol inhibits productive infection of SARS-CoV pseudovirus in Vero E6 cells.

The effect of depletion of cholesterol on the entry of pseudotyped SARS-CoV into Vero E6 cells was investigated using a single-cycle infectivity assay. In this assay, virions pseudotyped with SARS-CoV S protein or envelope glycoprotein G from VSV (VSV-G) were used to infect Vero E6 cells. Luciferase activity would serve as an indirect estimation of viral entry, integration, and transcriptional activity in Vero E6 cells. Cells were left untreated or treated with 10 mM MβCD to extract cholesterol. The background signal of this assay was monitored and found to be at a negligible level. Luciferase activity could be detected in lysates of cells infected with virions generated by reporter vector alone (Figure 23A, NL4-3). As expected, productive infection by pseudotyped SARS-CoV was inhibited by 90% after extraction of cholesterol (Figure 23. S), but did not affect expression of luciferase significantly in cells infected with virus pseudotyped with VSV-G envelope (Figure 23. VSV-G). These results demonstrate that extraction of cholesterol and subsequent disruption of lipid rafts dramatically inhibits productive infection of viruses pseudotyped with SARS-CoV S protein but does not affect the viruses pseudotyped with VSV-G which targets virions for entry via a raft-independent endocytosis pathway (Aiken, 1997).

SARS-CoV may enter the cells by both receptor-mediated membrane fusion and endocytosis (Ng et al., 2003; Qinfen et al., 2004; Yang et al., 2004). Lipid rafts could be involved in both of these pathways. Semliki Forest virus and Sindbris virus are envelope viruses that infect their host cells by receptor-mediated endocytosis, and

subsequently fusion with the acidic endosome membrane. Wilscut's group (Waarts, Bittman, and Wilschut, 2002) found that fusion of viral envelope requires the presence of both sphingolipid and cholesterol in the target membranes for these viruses, but lacks correlation with lipid rafts formation in target liposomes using virus-liposome model system. In other words, the Semliki Forest virus fusion does not require target membrane sphingolipids to reside in raft-like microdomains with cholesterol. For SARS-CoV, the situation is different; the receptor for SARS-CoV reside in lipid rafts, and the disruption of lipid rafts does not affect the surface expression of ACE2 but inhibited the viral infection, that means the accumulation of ACE2 into lipid rafts is important for SARS-CoV entry into the host cell. My results strongly support that cholesterol is required for efficient entry of SARS-CoV pseudovirus, possibly by maintaining lipid raft integrity to provide a platform for ACE2 accumulating to high concentration at these microdomains and support the virus-receptor interaction, although the detailed mechanism need to be studied further.

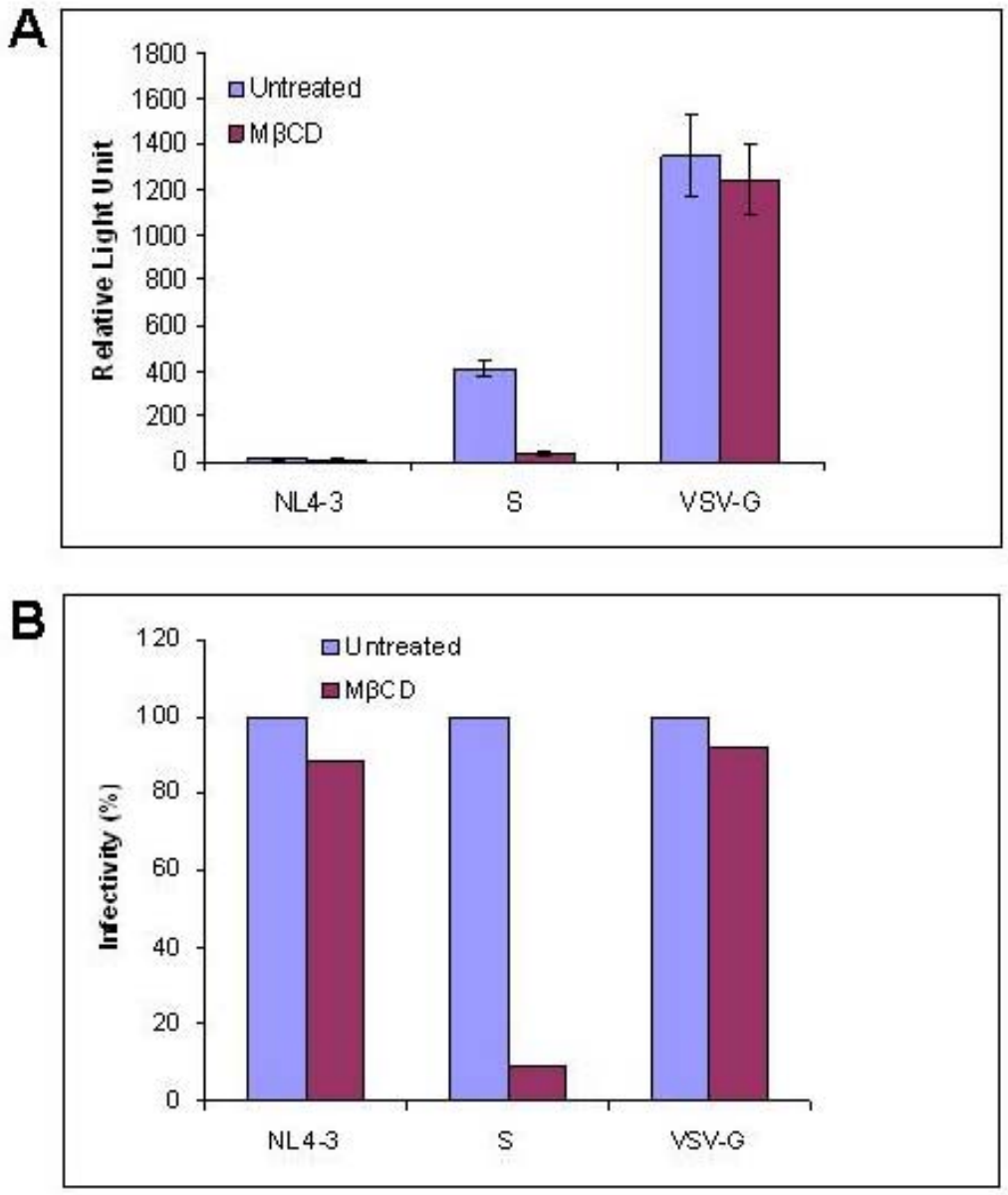

Figure 23. Depletion of cholesterol inhibits productive infection of SARS-CoV pseudovirus with Vero E6 cells.

Vero E6 cells, untreated (control) or treated for 30 min with 10 mM MβCD, were infected for 1 h with envelope-defective pNL4-3Luc^{+}Env^{-}Vpr^{-} virions alone, (NL4-3) or with virions pseudotyped with SARS-CoV S protein or VSV-G protein (0.5 u of RT/well). After extensive washing, the cells were cultivated and collected 48 h later. Luciferase activity in 20 µl of cell lysates (expressed as relative light units) was determined using luciferase assay reagent as described in Materials and Methods. The results are representative of three independent experiments (A). The infectivity of untreated cell of each group served as 100% infectivity, the infectivity (%) of MβCD-treated cell of each group was calculated based on their own negative control (untreated cell) (B).

D. Conclusions

Entry of SARS-CoV into permissive cells, like other human and animal coronaviruses, is mediated through the binding of its S protein to cellular receptor (Hofmann *et al.*, 2004; Simmons *et al.*, 2004). Thus, the differential expression of ACE2, a functional cellular receptor for SARS-CoV (Li *et al.*, 2003; Wang *et al.*, 2004), could contribute to the entry process of SARS-CoV. Indeed, it has been shown that ACE2 is abundantly expressed on the surface of lung and intestinal epithelial cells, consistent with their susceptibility to SARS-CoV (Hamming *et al.*, 2004). Additionally, epithelial cells are highly polarized cells. ACE2 is expressed predominately on the apical surface of these cells, which may play a crucial role in dictating viral infection (Tseng *et al.*, 2005; Warner *et al.*, 2005). The entry and release of SARS-CoV in highly polarized Calu-3 cells was shown to occur at the apical surface (Tseng *et al.*, 2005). In polarized cells, lipid rafts are concentrated at the apical surface, whereas in nonpolarized cells they are dispersed over the cell surface (Pralle *et al.*, 2000). Available evidence shows that lipid rafts play an important role in some coronaviruses entry either by clustering of cell surface receptors into these microdomains, such as HCoV-229E (Nomura *et al.*, 2004) or other uncharacterized mechanism such as MHV (Choi, Aizaki, and Lai, 2005). In this study, lipid rafts were shown to be engaged in the entry process of pseudotyped SARS-CoV into Vero E6 cells. The supporting evidence and possible mechanism are discussed below.

Biochemical isolation and confocal microscopy results showed that SARS-CoV

receptor ACE2 clusters in lipid rafts predominantly, and after binding to the host cell, S protein ectodomain also clusters in lipid rafts. Although I could not confirm whether the S protein binding can increase the ACE2 association with the lipid raft, the specific concentration of ACE2 in lipid rafts may explain why SARS-CoV selects these microdomains for entry. The preferential association of the ectodomain of SARS-CoV S protein with lipid rafts further supports our hypothesis that SARS-CoV receptor localizes in these microdomains represent the population of membrane receptor engaged in virus binding and entry.

How does localization of the receptor in lipid rafts affect the efficiency of SARS-CoV entry and productive infection? Coronavirus enters the cell by membrane fusion on cell surface or receptor mediated endocytosis. For SARS-CoV, detailed information of its entry pathway is still unclear. Lipid rafts could be important in both of these pathways, and membrane fusion process also takes place in both of these pathways, one kind of membrane fusion takes place on the cell surface, another takes place inside of the endosome. Membrane fusion is a cooperative process (Doms, 2000), and recent evidence suggests that lipid rafts may concentrate components of the membrane docking and fusion machinery for exocytosis and endocytosis (Chamberlain, Burgoyne, and Gould, 2001; Lang *et al*., 2001). Targeting of signaling proteins to rafts may affect their functions in different ways. First, increased concentration of these proteins in rafts may facilitate their intermolecular interactions by proximity. Second, the ordered lipid environment may affect functional properties of the receptors (Brown and London, 2000; Gimpl

and Fahrenholz, 2000). There is also evidence showing that the distribution of the spikes on the viral surface is not even, but displays some clustering pattern (Zhu *et al*., 2006), although I could not obtain satisfactory results to show that S protein of SARS-CoV is clustering in lipid rafts. These clustered spikes might need clustered receptors or trigger the clustering of receptors to achieve best interaction during viral attachment and docking. Lipid rafts provide a convenient platform. I speculate that lipid rafts may be efficiently aggregated by the stimulation of the signaling pathways emanating from these microdomains and induced by virus binding. Importantly, entry through rafts may direct SARS-CoV preintegration complexes into the correct cytoplasmic compartment. Accordingly, rafts have been shown to be sites of the active signal induced actin polymerization (Caroni, 2001; Rozelle *et al*., 2000) which is important for the membrane fusion and endocytosis. Holmes' group claimed that CD209L (L-SIGN), serve as an alternative receptor for SARS-CoV (Jeffers *et al*., 2004). CD209L is a C-type lectin, which is associated with lipid rafts (Cambi, Koopman, and Figdor, 2005). This does also support my findings.

In conclusion, I demonstrate that SARS-CoV receptor ACE2 is localized in lipid rafts in Vero E6 cells. Upon binding to Vero E6 cells, ectodomain of S protein also localizes in lipid rafts. After depletion of cholesterol, both ACE2 and S protein shift to nonraft fraction following with caveolin-1, but the surface expression of ACE2 and the binding of S protein are not changed significantly. However, the infectivity of pseudotyped SARS-CoV is dramatically decreased. These data demonstrate the essential role of lipid rafts and SARS-CoV receptor colocalization with lipid rafts to

support productive virus infection. My observations could lead to the development of effective anti-SARS-CoV drug therapies that interfere with ACE2 receptor localization in rafts and/or raft aggregation.

Chapter 4. Role of Trp-rich region of S protein in SARS-CoV entry

A. Introduction

Nearly all class I viral fusion proteins possess a membrane-proximal external region (MPER) that is Trp-rich (called Trp-rich region in this thesis). This region exists in all coronaviruses S protein C-terminal domain, the putative S2 protein. They share a highly conserved ten-residue sequence Y(V/I)KWPW(W/Y)VWL that is rich in aromatic amino acids with 3 to 4 Trp residues (Figure 6). Surprisingly, this Trp-rich region is exposed on the protein surface and lies at the water-lipid interface of the viral membrane.

Many laboratories have shown that the Trp-rich region plays an important role in HIV infection (Peisajovich *et al*., 2003; Schibli, Montelaro, and Vogel, 2001; Suarez *et al*., 2000a). More importantly, this exposed region contains epitopes to which broadly neutralizing antibodies could be raised (Schibli, Montelaro, and Vogel, 2001; Suarez *et al.*, 2000a; Suarez *et al.*, 2000b). Indeed, Sainz, B. *et al.* (Sainz *et al.*, 2005) found that the Trp-rich region of coronavirus class I viral fusion protein induces membrane permeabilization to facilitate the viral entry. Epand *et al.* (Epand, Sayer, and Epand, 2005) claimed that the Trp-rich region of HIV gp41 could promote the formation of cholesterol-rich domains which could facilitate membrane fusion.

Results presented in chapter 3 show that cholesterol-enriched microdomains, known as lipid rafts, are required for SARS-CoV entry. The cryoEM structures of HIV and SIV reveal that the Trp-rich region forms the “feet” of a tripod-like structure

of the spike proteins, anchoring the legs (HR1 and HR2 regions) firmly on the lipid rafts of viral surface. Because of their similarity in fusion mechanism of SARS-CoV and HIV, I hypothesized that the highly conserved Trp-rich region in SARS-CoV S protein plays a similar role in viral entry as in HIV by anchoring the tripod-like structure of S protein on a lipid rafts-like environment on the viral surface. Such a mechanism of anchoring S protein promotes clustering of S proteins, an effect which was observed in cryoEM structures of HIV, and which in turn facilitates docking of host receptors, and subsequently membrane fusion by merging the cholesterol-enriched microdomains of both virus and host membranes.

In this study, I used site-directed mutagenesis and in collaboration with Ms Neo Tuan Ling, to produce a series of S protein mutants by Ala- and Phe-scans to assemble pseudotyped SARS-CoV using an HIV-1 expressing the luciferase reporter gene in a single-cycle infectivity assay for readouts to determine the S protein-mediated infectivity. Our results show that replacement of Trp, Phe or Tyr with Ala dramatically decreases the pseudotyped SARS-CoV viral infectivity. In contrast, Phe-substituted mutants could restore 10-25% infectivity, suggesting that the Trp-rich region and the maintenance of its aromaticity is essential for SARS-CoV infectivity.

B. Materials and methods

i. Materials

Plasmids. pNL4-3Luc^{+}Env^{-}Vpr^{-} and pcDNA3.1-OPT9-S were kindly provided by Prof. Zhang Linqi (Aaron Diamond AIDS Research Center, Rockefeller University,

New York 10016).

Primers. Primers designed to encode S protein mutants are listed in Table 4.

Kits and Enzymes. The QuickChange® II XL site-directed mutagenesis kit was purchased from Stratagene.

ii. Methods

Cell culture. 293T cells were maintained in DMEM medium supplemented with 10% new born calf serum and 50 µg/ml penicillin-streptomycin at 37°C, 5% CO_2.

Construction of clones containing S protein mutants. Based on plasmid pcDNA3.1-OPT9-S which contains codon-optimized S gene, a series of plasmids containing mutated S gene (Table 5) were produced using QuickChange® II XL site-directed mutagenesis kit, following the manufacturer's instructions. The PCR reaction contains 5 µl of 10x reaction buffer, 3 µl (10 ng) of dsDNA template, 3 µl (125 ng) of primers (forward and reverse each), 1 µl of dNTP mix and 3 µl of QuickSolution. ddH_2O was added to a final volume of 50 µl, then with 1 µl of PfuUltra HF DNA polymerase (2.5 U/µl) added. The amplification program was performed at 95°C, 1 min for 18 cycles of 95°C, 50 S, 60°C, 50 S, 68°C, 12 min, followed by a final elongation step at 68°C for 7 min. The PCR amplification products were digested using Dpn I (1 µl of 10 U/µl Dpn I in 50 µl amplification products) at 37°C for 1 h. The Dpn I-treated DNA was transformed into XL10-Gold® Ultracompetent cells, and then the cells were grown on LB-ampicillin agar plates at 37°C for >16 h to produce clones.

Table 4. Primers used to produce S protein mutants

Name	Sequence 5'-3'[1]
optS1194WA5	GTACGAGCAGTACATCAAG**GCC**CCCTGGTACGTGTGGCTGGGC
optS1194WA3	GCCCAGCCACACGTACCAGGG**GGC**CTTGATGTACTGCTCGTAC
optS1196WA5	TACATCAAGTGGCCC**GCC**TACGTGTGGCTGGGC
optS1196WA3	GCCCAGCCACACGTA**GGC**GGGCCACTTGATGTA
optS1197YA5	ATCAAGTGGCCCTGG**GCC**GTGTGGCTGGGCTTC
optS1197YA3	GAAGCCCAGCCACA**GGC**CCCAGGGCCACTTGAT
optS1199WA5	CAAGTGGCCCTGGTACGTG**GCC**CTGGGCTTCATCGCCGGCCTG
optS1199WA3	CAGGCCGGCGATGAAGCCCAG**GGC**CACGTACCAGGGCCACTTG
optS1202FA5	GCCCTGGTACGTGTGGCTGGGC**GCC**ATCGCCGGCCTGATCGCC
optS1202FA3	GGCGATCAGGCCGGCGAT**GGC**GCCCAGCCACACGTACCAGGGC
optS11946WA5	CGAGCAGTACATCAAG**GCC**CCC**GCC**TACGTGTGGCTGGGCTTCATCG
optS11946WA3	CGATGAAGCCCAGCCACACGTA**GGC**GGG**GGC**CTTGATGTACTGCTCG
optS11949WA5	AAG**GCC**CCCTGGTACGTG**GCC**CTGGGCTTCATC
optS11949WA3	GATGAAGCCCAG**GGC**CACGTACCAGGG**GGC**CTT
optS11969WA5	AAGTGGCCC**GCC**TACGTG**GCC**CTGGGCTTCATC
optS11969WA3	GATGAAGCCCAG**GGC**ACAGTA**GGC**GGGCCACTT
optS119469WA5	AAG**GCC**CCC**GCC**TACGTG**GCC**CTGGGCTTCATC
optS119469WA3	GATGAAGCCCAG**GGC**CACGTA**GGC**GGG**GGC**CTT
optS119469WA 1197YA5	GTACATCAAG**GCC**CCC**GCCGCC**GTG**GCC**CTGGGCTTCATCGCCGG
optS119469WA 1197YA3	CCGGCGATGAAGCCCAG**GGC**CAC**GGCGGC**GGG**GGC**CTTGATGTAC
optS119469WA 1197YA1202FA5	CC**GCCGCC**GTG**GCC**CTGGGC**GCC**ATCGCCGGCCTGATCGCCATCG
optS119469WA 1197YA1202FA3	CGATGGCGATCAGGCCGGCGAT**GGC**GCCCAG**GGC**CAC**GGCGGC**GG

[1] Mutated base pairs are shown in red.

Table 5. Mutations introduced in Trp-rich region of S protein

Name/Abbreviations	Description	Sequence (aa 1190-1204)
Wild type	Control	**Q Y I K W P W Y V W L G F I I**
1. Global mutations		
WYF5A	W1194A+W1196A+ Y1197A+W1199A +F1202A	**Q Y I K A P A A V A L G A I I**
WY4A	W1194A+W1196A+ Y1197A+W1199A	**Q Y I K A P A A V A L G F I I**
W3A	W1194A+W1196A+ W1199A	**Q Y I K A P A Y V A L G F I I**
W2A-4/6	W1194A+W1196A	**Q Y I K A P A Y V W L G F I I**
W2A-4/9	W1194A+ W1199A	**Q Y I K A P W Y V A L G F I I**
W2A-6/9	W1196A+ W1199A	**Q Y I K W P A Y V A L G F I I**
2. Single mutations		
W1194A	W1194A	**Q Y I K A P W Y V W L G F I I**
W1196A	W1196A	**Q Y I K W P A Y V W L G F I I**
Y1197A	Y1197A	**Q Y I K W P W A V W L G F I I**
W1199A	W1199A	**Q Y I K W P W Y V A L G F I I**
F1202A	F1202A	**Q Y I K W P W Y V W L G A I I**
W1194F	W1194F	**Q Y I K F P W Y V W L G F I I**
W1196F	W1196F	**Q Y I K W P F Y V W L G F I I**
Y1197F	Y1197F	**Q Y I K W P W F V W L G F I I**
W1199F	W1199F	**Q Y I K W P W Y V F L G F I I**

Note: Wild type aromatic residues are shown in red. Mutated aromatic residues are shown in blue.

Transient expression of different mutants of S protein in 293T cells using codon-optimized plasmids. 293T cells were grown to 80% confluence in 75 cm^2 flasks (Nunc), and after trypsinization with Trypsin-EDTA for 3 min, cells were resuspended, plated into 6-well plates (5 x 10^5 cells/well) and cultured at 37°C, 5% CO_2, overnight. The 293T cells were transfected with plasmids pcDNA3.1-OPT9-S and S mutants using DOTAP transfection reagent. After transfection 48 h, cells were collected, cell lysates were resolved by 8% SDS-PAGE, and the expression of S protein and its mutants were investigated by Western blot using rabbit anti-S antibody as primary probe, and HRP-conjugated swine anti-rabbit as secondary probe.

Preparation of pseudotyped SARS-CoV containing different S protein mutants. Pseudoviruses were generated by co-transfection of 293T cells using calcium phosphate transfection method with pNL4-3Luc$^+$Env$^-$Vpr$^-$ and pcDNA3.1-OPT9-S mutants. The culture supernatant containing virus was collected on day 2 and 3 after transfection and clarified by filtration through a 0.45 μm filter. The virus was concentrated by ultracentrifugation at 26,000 rpm for 3 h (Zhang *et al.*, 2004). The pelleted virus was resuspended in DMEM with 10% new born calf serum, aliquoted, and stored at -80°C. The concentration of the viruses was determined by the reverse transcriptase (RT) activity assay using EnzChek® reverse transcriptase assay kit following the manufacturer's instructions, and the viruses were standardized according the RT assay.

Single cycle infectivity assay. Vero E6 cells (3 x 10^4 cells/well in 300 μl) were plated in 48-well plates (Nunc) and cultured at 37°C, 5% CO_2, overnight. On the

following day, Vero E6 cells were incubated for 1 h with standardized amounts of pseudotyped viruses containing different S protein mutants in DMEM (based on RT Assay, 0.5 U of RT/well). After washing, the cultures were propagated for 48 h. After 48 h of incubation, the cells were washed, and lysed with 100 µl of lysis buffer (provided in luciferase assay system kit). Luciferase activity was determined after addition of 100 µl of luciferase assay reagent (provided in luciferase assay system kit) into 20 µl aliquot of cell lysate and counting the resultant scintillation for 15 s using a TD-20/20 Luminometer.

C. Results

To investigate the importance of Trp-rich region in SARS-CoV viral entry, S proteins with mutations in Trp-rich region were prepared. All clones containing different mutated S genes were selected and confirmed by sequencing.

The expression of S protein mutants in 293T cells was detected by rabbit anti-S protein antibody. The results showed that mutants W1194A, W1199A, W2A-4/6, W2A-4/9, WY4A and WYF5A (see Table 5 for abbreviations) were expressed at high level, whereas others, such as Y1197A, W2A-6/9 and W3A were expressed at low level. No protein expression was detected under my experimental condition for W1196A and F1202A containing only single mutation (Figure 24).

An HIV-1 expressing luciferase reporter gene and pseudotyped with SARS-CoV S protein and its mutants was used to determine S protein-mediated infectivity. Pseudotyped retroviruses containing S protein mutants were first generated using a global site-directed mutagenesis and then Ala scan of aromatic amino acids in the

Trp-rich region to determine the positional importance of each Trp, Phe or Tyr. To restore the aromaticity of Ala-mutants, we performed rescue experiments using a Phe-scan (This part of work was performed in collaboration with Ms Neo Tuan Ling). The results show that global substituted mutants, tri-, tetra- and penta-substitution with Ala (WYF5A, WY4A, W3A) completely abrogate infectivity, while single- and double-substitution with Ala (W2A-4/6, W2A-4/9, W2A-6/9, W1194A, W1196A, Y1197A, W1199A and F1202A) substantially decrease infectivity by >90% (Table 6). In contrast, Phe-substituted mutants are able to restore 10-25% infectivity comparing to the wild type (Table 7). These results suggest that maintenance of the aromaticity of the Trp-rich region of S protein is essential for SARS-CoV infectivity.

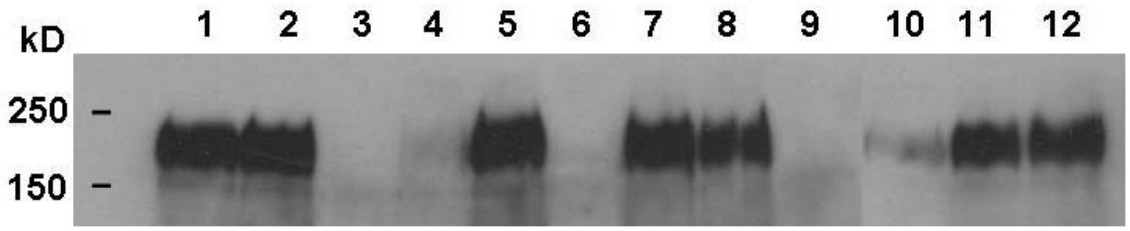

Figure 24. Expression of wild type and mutants of S protein in 293T cells.

Wild type S protein and its mutants were expressed in 293T cell. The cells were transfected with plasmids pcDNA3.1-OPT9-S and S mutants using DOTAP transfection reagent. After 48 h of transfection, the cells were collected, and cell lysates were resolved on 8% SDS-PAGE, proteins were transferred onto PVDF membrane, and subjected to Western blot using rabbit anti-S antibody as primary probe.

Lane 1: Wild type S protein
Lane 2: W1194A
Lane 3: W1196A
Lane 4: Y1197A
Lane 5: W1199A
Lane 6: F1202A
Lane 7: W2A-4/6
Lane 8: W2A-4/9
Lane 9: W2A-6/9
Lane 10: W3A
Lane 11: WY4A
Lane 12: WYF5A

The red color indicates the higher expression level, and the blue color indicates the lower expression level or no detectable expression.

Table 6. Infectivity of different SARS-CoV pseudotyped with S protein mutants

Mutation on S protein	Name	Infectivity (%)
Wild type		100*
Penta-mutation	WYF5A	0
Tetra-mutation	WY4A	0
Tri-mutation	W3A	0
Double-mutation	W2A-4/6	0.32
	W2A-4/9	0.08
	W2A-6/9	0.12
Single-mutation	W1194A	0.55
	W1196A	0.61
	Y1197A	0.59
	W1199A	0.59
	F1202A	0.02

* The infectivity of SARS-CoV pseudotyped with wild type S protein was set as 100%; the infectivity of each SARS-CoV pseudotyped with S protein mutants was compared with wild type pseudovirus.

Table 7. Comparison of the infectivity of SARS-CoV pseudotyped with S protein Ala- and Phe-mutants

Position	Infectivity %			Infectivity (%) restored by Phe-mutant
	W/Y	A	F	
1194	100	0.55	0.24	0
1196	100	0.61	23.97	23.36
1197	100	0.59	9.13	8.54
1199	100	0.59	11.22	10.73

Note: Phe-mutants data were from Ms Neo Tuan Ling.

D. Discussion

Coronavirus entry is mediated by class I viral envelope S protein. S protein of SARS-CoV is responsible for receptor-binding and membrane fusion. S protein contains several functional domains to support its correct folding and conformation which are crucial for its function. The Trp-rich region is absolutely conserved in members of coronaviruses and highly conserved in other RNA viruses such as HIV, FIV and EboV. The importance of the Trp-rich region in vaccine and therapeutics development has been demonstrated in HIV research as its epitopes are recognized by broadly neutralizing antibodies from human sera. It may play important roles during the whole fusion process. In cryoEM (Zhu *et al*., 2006), in the prefusion state, the Trp-rich region appears as "feet" to anchor the tripod-like structure of class I viral fusion protein and resides on the envelope surface at the membrane-water interface of the lipid bilayer (Figure 25).

In the post-fusion state, the two heptad-repeat regions (HR1 and HR2) form a 6-helix bundle, positioning the fusion peptide closely to the Trp-rich region to drive apposition and subsequent fusion of viral and cell membranes. Following TM activation, the Trp-rich region is believed to undergo sequential conformational transitions, from a closed turn to an extended arrangement and finally to an amphipathic helical structure capable of intimately interacting with the viral membrane through certain Trp residues and, concomitantly, with the cell membrane through others residues (Barbato *et al.*, 2003).

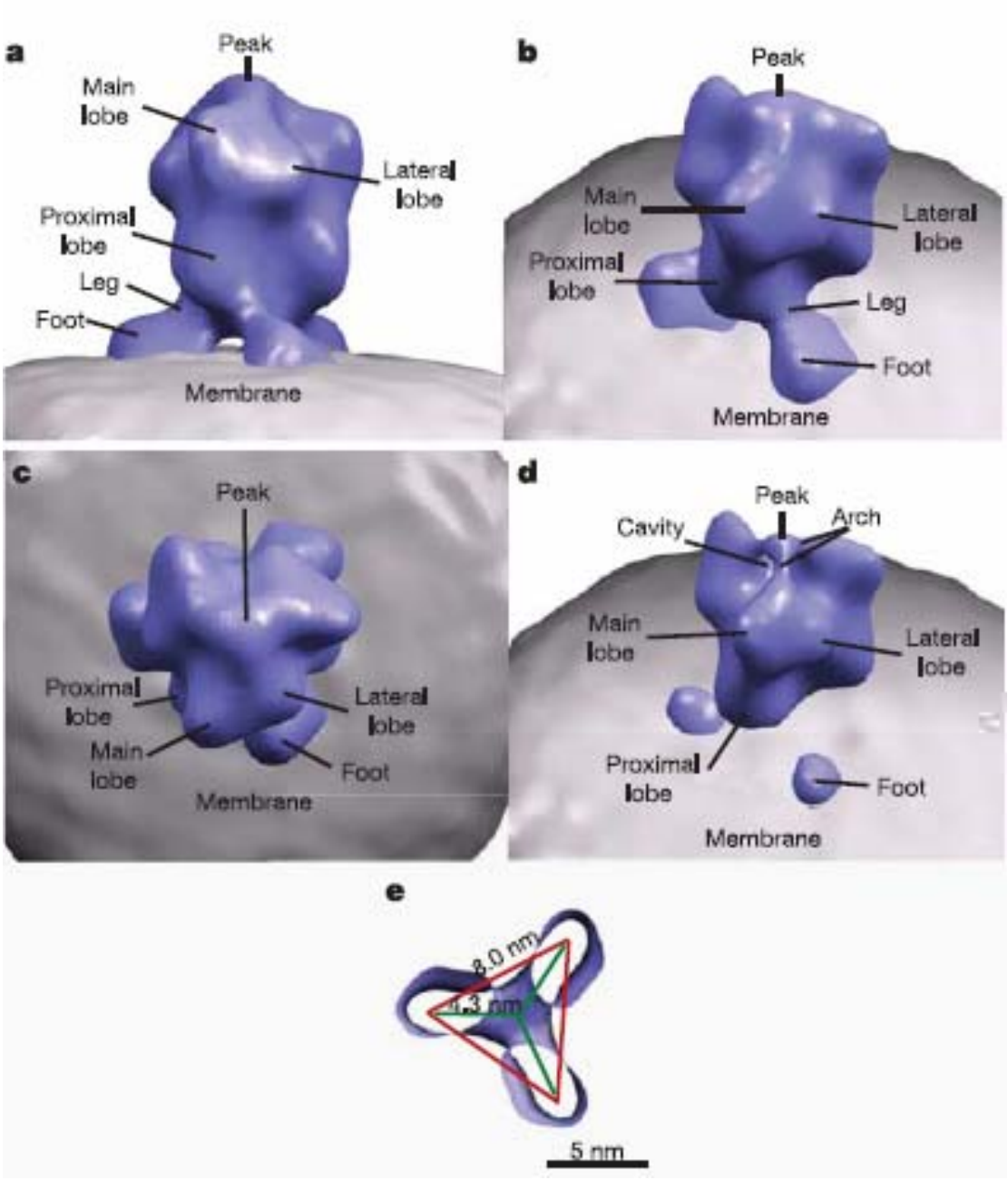

Figure 25. Surface-rendered model of averaged SIV Env spike tomograms (Taken from Zhu *et al*., 2006).

Env spike (blue) and membrane (grey) in side (a), intermediate (b), and top (c) views and an intermediate view at a higher density threshold (d). Panel e represents a membrane-parallel slice encompassing the legs and feet (MPER) as viewed from below. The centre-to-centre distance between possible membrane-spanning domain (MSD) attachment sites (red lines) and the medial axis to MSD sites (green lines) are indicated. The spike and membrane surfaces (in a–d) were rendered at different thresholds, the latter adjusted to conform to the known thickness of the lipid bilayer (~ 4 nm).

To investigate the importance of Trp-rich region in viral entry of SARS-CoV, pseudotyped SARS-CoV containing S protein mutants were used to determine S protein-mediated infectivity by Ala- and Phe-scans of the Trp-rich region. The results show that global-substituted Ala-mutants completely abrogate infectivity, while double- and single-substituted Ala-mutants substantially decrease infectivity by >90%. In contrast, Phe-substituted mutants aiming to restore the aromaticity of the Trp-rich region are able to restore 10-25% infectivity comparing to wild type.

Why is this Trp-rich region so important for the infection of SARS-CoV and other class I viruses? Trp, Tyr and Phe are aromatic amino acids; their phenyl rings could interact stably under lipidic environment and could form channels. These characters are not found in hydrophobic aliphatic amino acids such as Ile. For these reasons, I chose Phe to replace the Trp, but not other more hydrophobic amino acids, such as Ile to do the rescue experiments. At a later date, aromatic to aliphatic amino acid mutations will be further investigated and completed in the future studies. The importance of maintaining aromaticity of Trp, Phe or Tyr clusters at the interface between the hydrophilic and hydrophobic surfaces is a typical biophysical phenomenon in Class I envelope proteins. Indeed, in membrane proteins Trp residues are often located at the lipid bilayer interface but rarely within the hydrophobic core of the membrane (Ridder *et al.*, 2000; Yau *et al.*, 1998). For example, in the membrane-proximal Trp-rich region of HIV glycoprotein gp41 four of five Trp residues and one Tyr residue form a collar of aromatic residues along the axial length of the amphipathic helix, resulting in a Velcro-like interaction on the

outer viral membrane (Schibli, Montelaro, and Vogel, 2001). The role of the Trp-rich region in participating in the clustering of gp41 monomer within the HIV-1 envelope has been shown by Saez-Cirion *et al.* (Saez-Cirion *et al*., 2003b). They reported that the Trp-rich region of EboV glycoprotein predicted to bind the membrane interface, adopted an α-helical structure. This pre-transmembrane sequence might target membranes inherently prone to destabilization of the membrane. This group also reported that sphingomyelin and cholesterol promote HIV-1 gp41 Trp-rich region surface aggregation and membrane restructuring (Saez-Cirion *et al.*, 2002). Furthermore, Guillen, J. *et al.* (Guillen *et al.*, 2005) identified three membrane-active regions of SARS-CoV S protein using a 16/18-mer peptide scan, one of them is located in the Trp-rich region. Peptides corresponding to this region exert a dramatic effect on leakage for different model membranes, suggesting that this region might be involved in the promotion of the membrane destabilization required for fusion, as well as in fusion pore formation and enlargement.

CryoEM (Zhu *et al*., 2006) data show that Trp-rich region form the "feet" to interact with the lipid-water interface. New predictive approaches including computation of interfacial affinity and the corresponding hydrophobic moments also suggest that this region is functionally segmented into two consecutive subdomains: one amphipathic at the N-terminal side and one fully interfacial at the C-terminus (Saez-Cirion *et al*., 2003a). The N-terminal subdomain would extend α-helices from the preceding carboxyl-terminal heptad repeat (HR2) and provide, at the same time, a hydrophobic-at-interface surface. Saez-Cirion *et al.* (Saez-Cirion *et al.*, 2003b)

reported that gp41 Trp-rich peptides have the ability to oligomerize and insert into the viral membrane interface. In our study, mutation of three or more aromatic residues in this region that functions as "feet" in pre-fusion state would irreversibly change its hydrophobic moments and its ability to anchor stably on the membrane to support the fusion process. Mutation of one or two aromatic amino acid residues in this region partially affects the "feet"-like conformation and virus infectivity.

In the post-fusion state, removing the bulky indole side chains and its weakly charged amine of the Trp residues in the Trp-rich region could disrupt its role in synergizing with the fusion peptide (Suarez *et al.*, 2000a) to destabilize both membranes and drive the energetically unfavorable lipid-lipid merging that permits the formation and expansion of the fusion pore in the late stages of the entry process (Melikyan *et al.*, 2000; Peisajovich and Shai, 2003).

The fusion process is likely aided by the presence of cholesterol-binding motifs found in the S protein sequences. Vincent *et al.* and Epand *et al.* (Epand, Sayer, and Epand, 2005; Epand *et al.*, 2006; Epand, 2004; Epand, Sayer, and Epand, 2003; Vincent, Genin, and Malvoisin, 2002) reported that a consensus sequence found in a group of proteins that sequester to cholesterol-rich regions of membranes has the pattern –L/V-(X)(1-5)-Y-(X)(1-5)-R/K-, in which (X)(1-5) represents 1-5 residues of any amino acid. This sequence can induce formation of cholesterol-rich domains. The HIV-1 fusion protein gp41 has a LWYIK motif, consistent with this consensus sequence for this protein sequestering into cholesterol-rich domains and promoting membrane fusion. The SARS-CoV S protein Trp-rich region **Q Y I K W P W Y V**

W L G F I I also has similar motif. The data presented here show that lipid rafts are involved in the virus entry, and it is likely that the Trp-rich region interacts with lipid rafts in triggering the aggregation of lipid rafts and destabilizing the host cell membrane to promote membrane fusion.

The literature precedents suggest that the Trp-rich region plays an important role in maintaining the trimeric conformation of S protein and in interacting with the lipid rafts to promote the membrane fusion and viral entry. My data support the literature precedents and show that the crucial roles played by aromatic amino acid residues in the Trp-rich region in S protein for viral infectivity and the maintenance of its aromaticity is essential but not sufficient for fully restoring the SARS-CoV infectivity.

Chapter 5. Peptide-affinity strategy for probing exposed functional domains of S protein in prefusion state

A. Introduction

The previous two chapters focus on the effects of clustering ACE2 receptors on host cells by lipid rafts and the corresponding clustering of S proteins on viral surface to facilitate infectivity. This chapter focuses on probing the exposed surfaces of class I fusion protein represented by SARS-CoV S protein using peptides based on the protein-protein complementary interactions which are the characteristics of S protein.

In the prefusion state, S protein forms parallel homo-trimer through self-interaction. Interactions of cellular receptor(s) with the trimeric quaternary structure of S protein trigger extensive conformational changes within different domains. The best known example of self-interaction is the fusion-active state in which the HR1 and HR2 domains of the S protein interact with each other to form a stable 6-helix bundle structure. This interaction is based on the complementarity of the HR1 and HR2 regions. Another example of complementary interactions is its exposed surface(s) for receptor-binding. In S protein, the receptor-binding domain (RBD) provides protein-protein interaction with ACE2 anchored on the host cell membrane. The exposed surfaces and the helical domains of S protein provide an opportunity to utilize peptides as affinity probes to dissect potential protein-protein interactions of different domains during the viral entry process.

The peptide-affinity probes were designed to contain a putative affinity-active

peptide segment of S protein appended with a capturing moiety. They were generated using two approaches. In the first approach, recombinant methods were used to prepare affinity probes of MBP-HR2 constructs. The HR2 segments were 27-35 aa in length and were appended at their N-termini to a maltose-binding protein (MBP) as capturing moiety. The use of HR2 peptides is appropriate to serve as a proof-of-concept because HR2 is known to bind to the corresponding HR1 region during the fusion-active state.

In the second approach, a global strategy was used to screen for affinity-active peptides. SPOT synthesis was used to prepare a library of overlapping peptides of S protein consisting of 15 amino acid residues anchored onto cellulose membrane. S protein ectodomain S1188HA was used to screen for "hot spots" to provide clues of self-interacting sites leading to trimerization and conformational transitions of different fusion states as well as putative exposed regions that could interact with host receptors. These "hot spots" could also be used as inhibitors or "synthetic antibody" for pull-down experiments as the recombinant MBP-HR2 which will be described later in this chapter.

B. Materials and methods

i. Materials

Chemical reagents. Ampicillin sodium salt was purchased from Sigma-Aldrich. Isopropyl-β-D-thiogalactopyranoside (IPTG) was purchased from Fermentas. Yeast extract, tryptone and agarose (molecular biology grade) were purchased from USB. Bacto agar was purchased from Difco. Acetic acid glacial, (>99.8%) and ethanol

(99.5-99.8%), absolute (GR grade for analysis) were purchased from Merck Biosciences. Urea (molecular biology grade) and β-mercaptoethanol were purchased from Vivantis.

Plasmids. pMAL-C2X was purchased from New England Biolabs.

Primers. Primers designed and used in PCR to produce HR2 and its C-terminal truncated gene are shown in Table 2.

Kits and Enzymes. Amylose resin was purchased from New England Biolabs. PNGase F was purchased from QIAGEN. Protein assay kit was purchased from Bio-Rad.

Solutions. Solutions used in this study are listed below. (1) For bacteria culture: LB medium contains 10 g tryptone, 5 g yeast extract and 10 g NaCl in 1 liter ddH_2O. LB-ampicillin medium contains 100μg/ml filter-sterilized ampicillin. LB-ampicillin agar plate was prepared by mixing 15g agar with 1 liter LB, autoclaved, and after cooling to 55°C, filter-sterilized ampicillin was added to gain 100μg/ml final concentration. The solution was poured into petri dishes to make agar plates. (2) For regenerating cellulose membrane: regeneration buffer A contains 8M urea, 1% SDS (w/v), 0.1% β-mercaptoethanol (v/v), and regeneration buffer B contains 50% ethanol (v/v) and 10% acetic acid (v/v) in ddH_2O. TAE buffer contains 40 mM Tris-acetate 1 mM EDTA in ddH_2O.

Instruments. MultiPep (SPOT) synthesizer (Intavis AG, MultiPep, Germany).

ii. Methods

Construction of clones containing HR2 and C-terminal truncated HR2 region

of S protein. The nucleotide fragment corresponding to amino acid residues 1151-1185 of S protein (HR2) was amplified by PCR from the plasmid pJX40-S. Primers used were pMALC2EcoRI and HR2-35PstI. The amplification program was 95ºC, 45 s, then 30 cycles of 95ºC, 45 s, 50 ºC, 45 s, 72 ºC, 30 s, followed by the last elongation step at 72ºC, 10 min. An upstream EcoR I site and a downstream Pst I site as well as a stop codon preceding the Pst I site were designed to insert the fragment into vector pMAL-C2X plasmid. The resulting plasmid was named as pMAL-C2X-HR2.

The C-terminal truncated HR2 genes were amplified by PCR from the plasmid pMAL-C2X-HR2. The forward primer was pMALC2BglII and the reverse primers were HR2-34PstI, HR2-33PstI, HR2-32PstI, HR2-31PstI, HR2-30PstI, HR2-29PstI, HR2-28PstI, and HR2-27PstI. The amplification program was 95ºC, 45 s, then 30 cycles of 95ºC, 45 s, 50 ºC, 45 s, 72 ºC, 2 min, followed by the last elongation step at 72ºC, 10 min. An upstream Bgl II site and a downstream Pst I site as well as a stop codon preceding the Pst I site were designed to insert the fragments into the vector pMAL-C2X plasmid which was digested by Bgl II and Pst I, and removed the smaller fragment. The resulting plasmids containing the HR2 region were named as pMAL-C2X-HR2-34, pMAL-C2X-HR2-33, pMAL-C2X-HR2-32, pMAL-C2X-HR2-31, pMAL-C2X-HR2-30, pMAL-C2X-HR2-29, pMAL-C2X-HR2-28, and pMAL-C2X-HR2-27, each was truncated at the C-terminus by one amino acid residue.

Probing S1188HA using MBP-HR2 and its truncated forms. MBP-HR2 fusion protein expression plasmids containing full-length or truncated HR2 were individually

transformed into *E. coil.* BL21 (DE3) competent cells and cultured on the LB-ampicillin agar plate for 14 h. Single clone was picked for culturing in LB-ampicillin media at 37˚C overnight. The overnight-culture was transferred to LB media for large-scale protein expression at 37˚C. IPTG was added into the culture media to a final concentration of 0.15 mM to induce protein expression when culture density (OD_{600}) reached 0.6. Cultured bacteria were harvested after 5 h of induction.

The harvested bacteria were separated by centrifugation and the pellet was resuspended in PBS and homogenized by sonication. The lysate was loaded into amylose resin. After binding for 30 min at room temperature, the resin was washed three times with PBS. The insect-cell lysate containing S1188HA was applied to the MBP-HR2 and truncated HR2 loaded amylose resin. After incubation for 30 min, the resin was washed three times with PBS to remove the unbounded protein. The cell lysate flow through, wash buffer and the amylose resin are collected respectively for Western blot assay.

Deglycosylation of S1188HA protein. 9 µl of purified S1188HA protein (10µg/ml, measured by protein assay kit) was mixed with 1 µl glycoprotein denaturing buffer. The mixture was boiled for 10 min at 100°C. 2 µl PNGase F, 1.5 µl 10x reaction buffer (G7 buffer) and 1.5 µl 10% NP40 was added to the mixture and incubated for 1h at 37°C. The sample was stored in -20°C for analysis.

Probing Peptide library using S1188HA protein.

(1) Synthesis of membrane-bound peptide library. The cellulose membrane-bound peptide library synthesized on a commercial membrane

(Amino-PEG_{500}-UC540 Sheet, Intavis AG, Germany) were semi-automatically prepared by Mr. Kelvin Poh in Prof. Tam's laboratory at NTU using MultiPep (SPOT) synthesizer (Intavis AG, MultiPep, Germany) using Fmoc-chemistry. The peptide library representing SARS-CoV S protein (Sin 2774, Accession AY283798) consisted of overlapping peptides, each 15aa in length with a 2-3 amino acids shift to the C-terminal direction.

(2) Screening of peptide library with S1188HA. The membrane-bound peptide library were screened for binding of SARS-CoV S protein ectodomain S1188HA by Western blot adapted from Mine & Rupa (Mine and Rupa, 2003). As negative controls, to visualize non-specific binding of antibodies to peptides, two experiments were performed. Firstly, the membrane was blocked by 5% non-fat milk in PBST overnight at 4°C), and then subjected to HRP-conjugated swine anti-rabbit antibody (1:1000 in blocking buffer) for 1 h at room temperature. After washing with PBST for 3 times, 15 min each, positive spots were detected by ECL Plus Western blotting detection reagent. Secondly, the membrane was regenerated to remove all protein reagents (detail procedure is described below). The same blotting procedure was repeated, first blotting with rabbit anti-HA antibody (1:1000 v/v in blocking buffer), and then followed with HRP-conjugated swine anti-rabbit antibody (1:1000 v/v in blocking buffer). To detect specific positive spots, the regenerated membrane was blotted with different forms of S1188HA (native form, denatured form and deglycosylated form), followed by rabbit anti-HA and HRP-conjugated swine anti-rabbit antibody. For the native form of S1188HA, two different binding

conditions were tested in PBST with and without 2M urea.

(3) Membrane regeneration. Membrane-bound peptides were regenerated using a method adapted from Mine & Rupa (Mine and Rupa, 2003). Membranes were washed 3 times, 5 min each with distilled water, DMF and then distilled water. Membranes were treated with regeneration buffer A 3 times, 10 min each, followed by regeneration buffer B 3 times, 10 min each. Finally membranes were washed with ethanol 3 times, 10 min each, air-dried and stored at -20^{o}C.

C. Results

i. Experimental design. To test the concept of peptide-affinity probes, the necessary reagents were prepared using both recombinant and chemical methods. Using recombinant methods, I prepared nine MBP-HR2 probes which contain HR2 segments ranging from 27 to 35 amino acids, the S protein ectodomain (S1188HA, aa 1-1188 of S protein with a HA tag) as well as the denatured and deglycosylated S1188HA as substrates. Using chemical methods, I collaborated with my coworkers to prepare peptide library.

ii. Construction of clones for expression of SARS-CoV S protein HR2 and C-terminal truncated HR2 region.

To prepare various forms of MBP-HR2 segments, PCR fragments containing full-length of HR2 gene (35aa) were inserted into pMAL-C2X plasmid between EcoR I and Pst I sites (Figure 26). The sequencing result showed that the clone pMAL-C2X-HR2 contained the correct fragment of HR2 gene fused with MBP at its N-terminus. pMAL-C2X was digested with Bgl II and Pst I, and then the larger

band was extracted from the gel to serve as the vector for cloning of the truncated HR2 into pMAL-C2X. The C-terminal truncated fragments of HR2 were amplified from pMAL-C2X-HR2 (Figure 26), and were inserted into this vector between Bgl II and Pst I sites. The sequencing results showed that all the clones contain the correct fragments of C-terminal truncated HR2 gene fused with MBP at their N-termini.

iii. Probing S1188HA protein using MBP-HR2 and its truncated forms.

To exploit the known interaction of HR1 and HR2 regions during the fusion-active state to form a 6-helix bundle that enables membrane fusion, all MBP-HR2 constructs were used for pull-down experiments with S1188HA to validate the design concept and to confirm the robustness of MBP-HR2 constructs as affinity-probes.

My experimental design was conceptually similar to the tandem purification method. First, MBP-HR2 constructs expressed in bacteria were absorbed onto amylose resin. After washing with PBS to remove unbound bacterial products, the insect-cell lysate containing S1188HA was applied to the amylose column. After incubation for 30 min, the resin was washed to remove unbound protein. The binding of S1188HA to MBP-HR2 was detected by Western blot assay.

The results show that all MBP-HR2 constructs are capable of binding S1188HA (Figure 27, 28). Constructs of MBP-containing C-terminal truncated HR2 were less efficient in these pull-down experiments and their potency appeared to decrease with increasing truncation (Figure 28). These results suggest that S1188HA can interact with HR2 through its HR1 region. The interaction intensity decreases when the HR2 fragment becomes shorter.

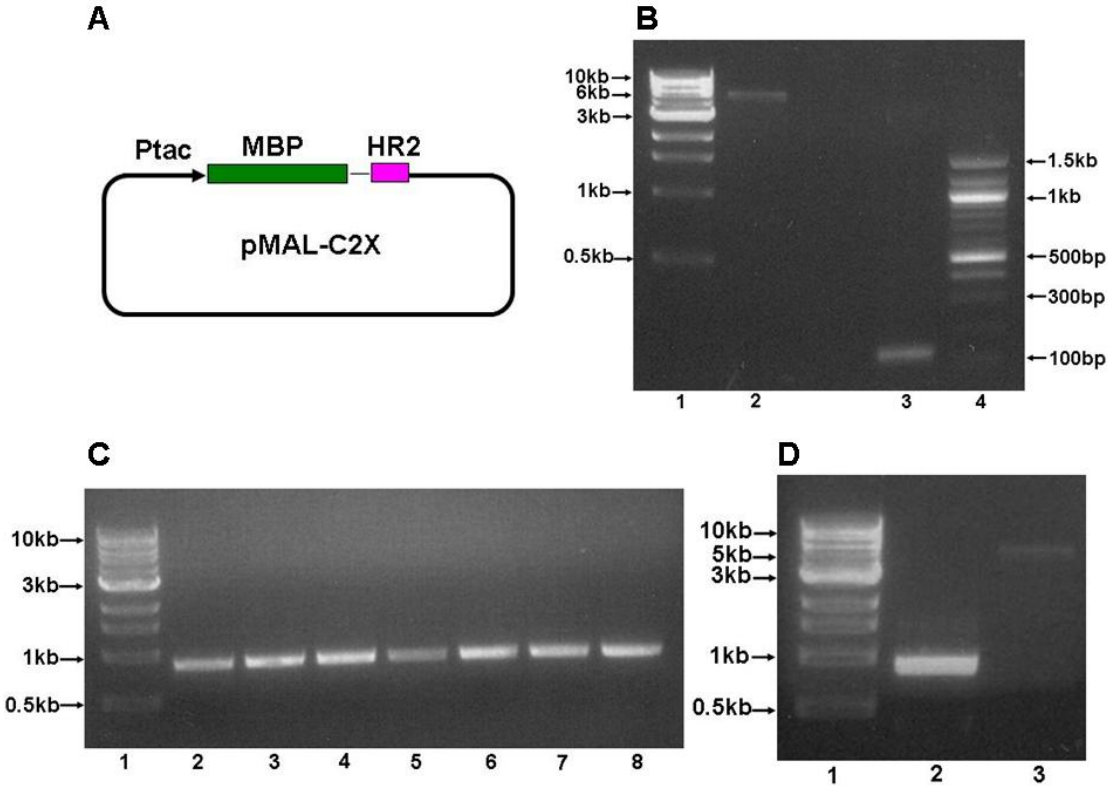

Figure 26. Construction of plasmids containing different fragments of HR2 region for bacteria expression.

A. Strategy of construction of plasmids containing MBP-HR2 and MBP-HR2 truncated fusion protein.
B. PCR reactions were performed to gain HR2 gene fragment which was to be inserted into pMAL-C2X plasmid to obtain clones of pMAL-C2XHR2.
Lane 1: 1kb DNA ladder (NEB),
Lane 2: Purified fragment of pMAL-C2X vector digested with EcoR I and Pst I,
Lane 3: PCR product of HR2 gene of S protein,
Lane 4: 100bp DNA ladder (NEB).
C and D. PCR reactions were performed to gain fragments containing C-terminal truncated HR2 gene which were to be inserted into pMAL-C2X plasmid to obtain clones containing truncated HR2 gene.
Lane 1 (C and D): 1kb DNA ladder (NEB),
Lane 2-8 (C): PCR products of fragments containing HR2 N-terminal 27-33 aa gene,
Lane 2 (D): PCR product of fragment containing HR2 N-terminal 34 aa gene,
Lane 3 (D): The larger fragment of pMAL-C2X digested with Bgl II and Pst I
All these DNA fragments were run in 1.5% agarose gel (B) or 1% agarose gel (C and D) in TAE buffer at a constant voltage of 120 V for 20-30 min. The arrows indicate DNA marker.

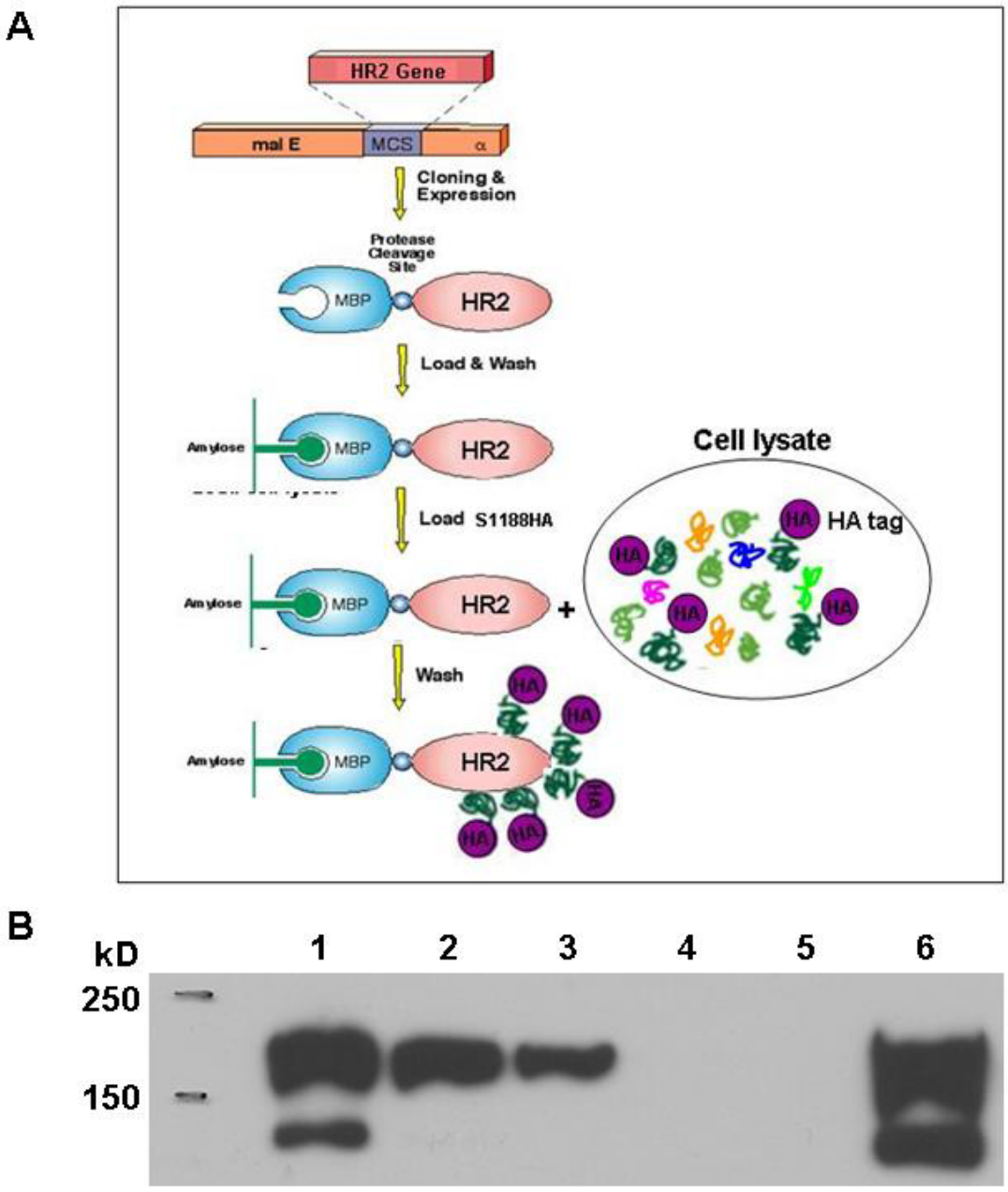

Figure 27. Interaction of MBP-HR2 with S1188HA protein.

A. The principle of the pull-down assay used to examine protein-protein interactions. B. The MBP-HR2 was expressed in *E. coli.*, the cell lysate containing MBP-HR2 was loaded on the amylose resin column and incubated for 30 min at room temperature, the supernatant was removed and the resin was washed with PBS for three times, then subjected to the cell lysate containing S1188HA, after binding for 30 min at room temperature, the supernatant was removed, the resin was washed with PBS for three times. All the samples from each step were collected and analyzed by Western blot.
Lane 1: Sf9 cell lysate containing S1188HA protein,
Lane 2: Sf9 cell lysate flow through from the column containing MBP-HR2,
Lane 3-5: First to third washes after binding,
Lane 6: Resin which contains MBP-HR2-S1188HA complex after wash.

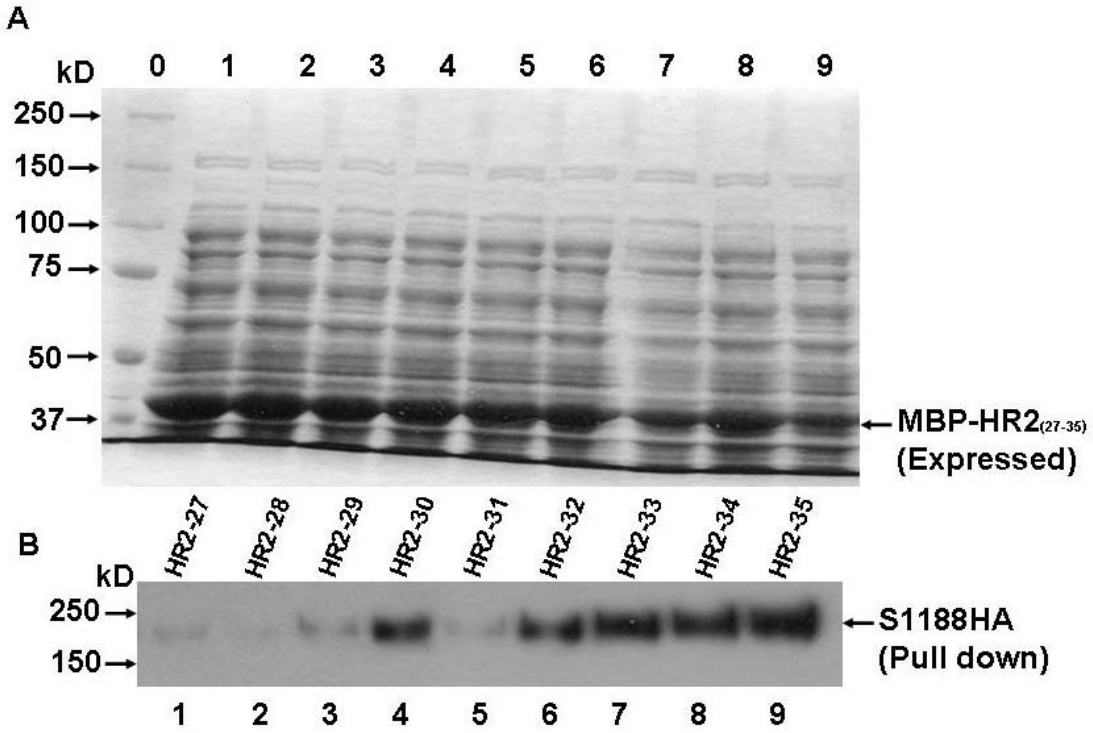

Figure 28. Interaction of MBP-HR2 and its truncated form with S1188HA protein.

A. The C-terminal truncated HR2 region fused with MBP was expressed in *E. coli.*, the expression of each fusion protein is shown.

B. The cell lysate containing MBP-HR2 or its C-terminal truncated forms was loaded onto the amylose resin column and incubated for 30 min at room temperature, the supernatant was removed and the resin was washed three times with PBS, then subjected to S1188HA, after binding for 30 min at room temperature, the supernatant was removed and the resin was washed three times with PBS. S1188HA pulled down by MBP-HR2 and its truncated forms is shown. Each lane was the resin which contains MBP-HR2 (27-35 aa) -S1188HA complex after extensively washing.

The aa sequences of HR2 and its truncated forms are listed below.

MBP-HR2-35	ISGINASVVN IQKEIDRLNE VAKNLNESLI DLQEL
MBP-HR2-34	ISGINASVVN IQKEIDRLNE VAKNLNESLI DLQE
MBP-HR2-33	ISGINASVVN IQKEIDRLNE VAKNLNESLI DLQ
MBP-HR2-32	ISGINASVVN IQKEIDRLNE VAKNLNESLI DL
MBP-HR2-31	ISGINASVVN IQKEIDRLNE VAKNLNESLI D
MBP-HR2-30	ISGINASVVN IQKEIDRLNE VAKNLNESLI
MBP-HR2-29	ISGINASVVN IQKEIDRLNE VAKNLNESL
MBP-HR2-28	ISGINASVVN IQKEIDRLNE VAKNLNES
MBP-HR2-27	ISGINASVVN IQKEIDRLNE VAKNLNE

iv. Deglycosylation of S1188HA protein.

PNGase F, an amidase which cleaves between the innermost GlcNAc and asparagine residues of high mannose hybrid and complex oligosaccharides from N-linked glycoproteins (Maley *et al.*, 1989) was used to remove glycan from the S1188HA protein. After treatment of PNGase, the larger molecular-weight band disappeared and a smaller molecular-weight band appeared and its estimated molecular weight agreed with the molecular weight of the unglycosylated S1188HA protein (Figure 29), which suggested that the lower molecular weight band of S1188HA protein shown in some overexpression samples is the unglycosylated form.

Post-translational glycosylation can affect the biological activity of proteins. Carbohydrate units on glycoproteins can regulate the folding, conformation and intracellular traffic and localization of the protein; such as their transport towards the cell surface and the stabilization of their functional conformation (Chu, Trimble, and Maley, 1978; Dube, Fisher, and Powell, 1988; Gibson, Schlesinger, and Kornfeld, 1979; Matzuk, Keene, and Boime, 1989; Semenkovich et al., 1990). N-linked glycosylation of some proteins is required for proper folding *in vivo*, maintaining its conformation (Lis and Sharon, 1993; Rademacher, Parekh, and Dwek, 1988). In addition, Glycosylation may act as recognition determinants or contribute to binding sites in protein-protein, protein-cell and cell-cell interactions.

My result showed that in mammalian and insect cell expression systems, the glycosylated form of S protein could be obtained, and could be used for binding assay, producing pseudovirus and screening the peptide library.

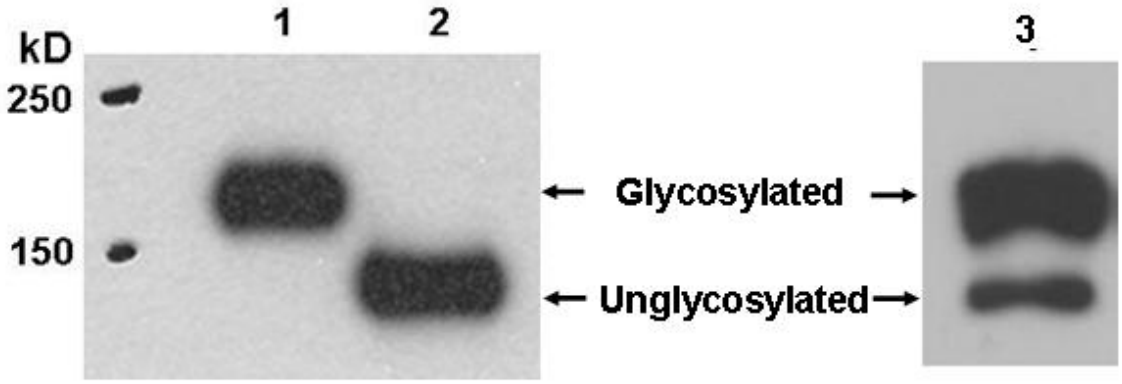

Figure 29. Deglycosylation of S1188HA protein using PNGase F.

The purified S1188HA protein (10µg/ml) was treated with PNGase F for 1h at 37°C. The sample was analyzed by Western blot using rabbit anti-HA antibody.
Lane 1: S1188HA protein before PNGase F treatment,
Lane 2: S1188HA protein after PNGase F treatment,
Lane 3: S1188HA overexpressed in Sf9 cells which contained both glycosylated and unglycosylated form.

v. Screening the peptide library using S1188HA protein.

To investigate the feasibility of peptides binding to S protein, synthetic peptide library derived from S protein were prepared. The library was synthesized by SPOT method as 452 overlapping peptides covering the C-terminal 116-1255 amino acids of S protein, each 15aa in length, 2-3 amino acids shift. Screenings were performed on the membrane-bound peptide library using a modified Western blot protocol. Positive results were visualized as spots by the enhanced chemiluminescence test (Figure 30-33).

To test for non-specific binding of antibodies to peptides, when blotted with secondary antibody, HRP-conjugated swine anti-rabbit antibody, one positive spot R25 (T-Y-V-T-Q-Q-L-I-R-A-A-E-I-R-A, 988-1002 aa) was observed. Similarly, when blotted with primary antibody rabbit anti-HA followed by secondary antibody HRP-conjugated swine anti-rabbit antibody, another positive spot P15 (G-A-E-H-V-D-T-S-Y-E-C- D-I-P-I, 638-652 aa) was observed. These spots were subtracted from the subsequent pepscan assays (Figure 32).

Comparing with the controls, 14 positive spots were identified and the results are summarized in Table 8. The results show that most of the peptides that could bind to S1188HA are located in the RBD and HR1 regions. Deglycosylation and denaturation of S protein removed most of the positive spots that could bind to S1188HA protein, but also new positive spots were observed. Addition of 2M urea to the binding buffer eliminated three sequential spots which in HR1 region, but not others (Figure 33).

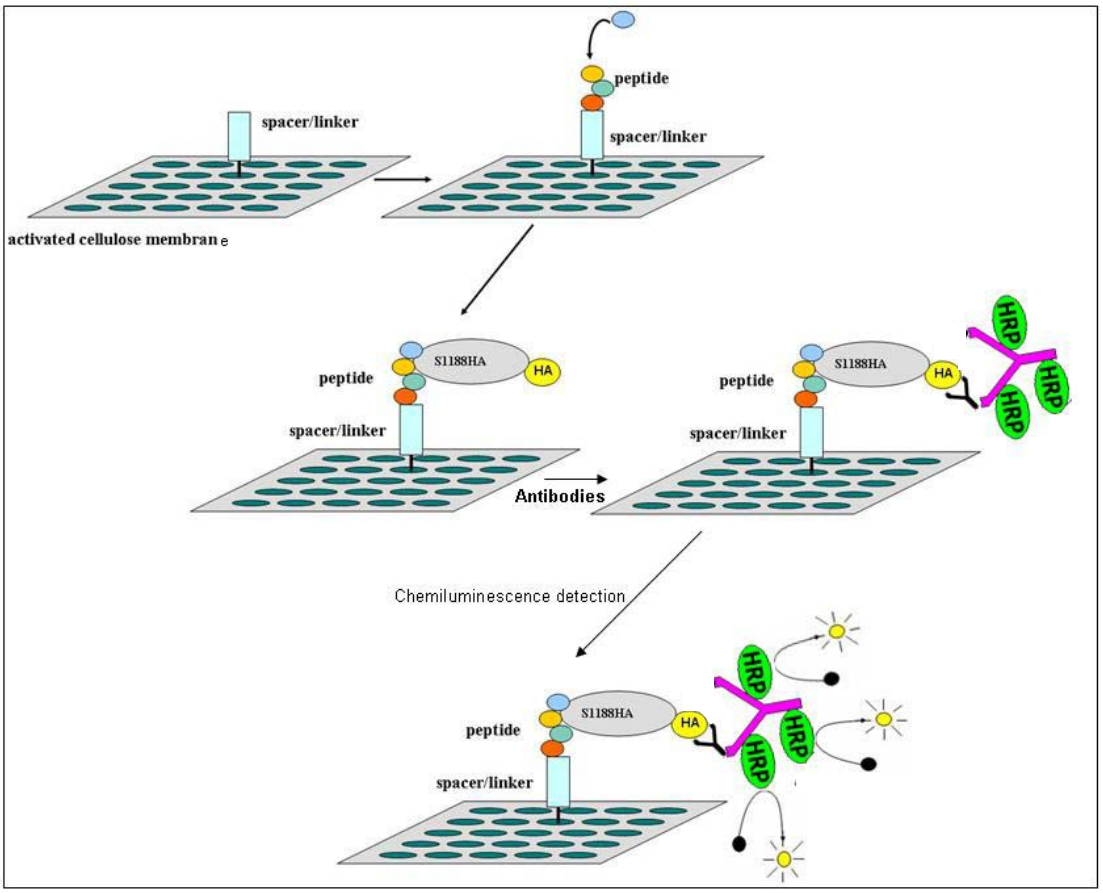

Figure 30. Schematic diagram of the SPOT pepscan.

The principle of the peptide library screening. Peptides are first synthesized on the cellulose membrane using a SPOT synthesis method to make a peptide library. The library is then probed with S1188HA protein as substrate, followed by rabbit anti-HA antibody and HRP-conjugated swine anti-rabbit antibody. The positive spots are visualized by chemiluminescence test.

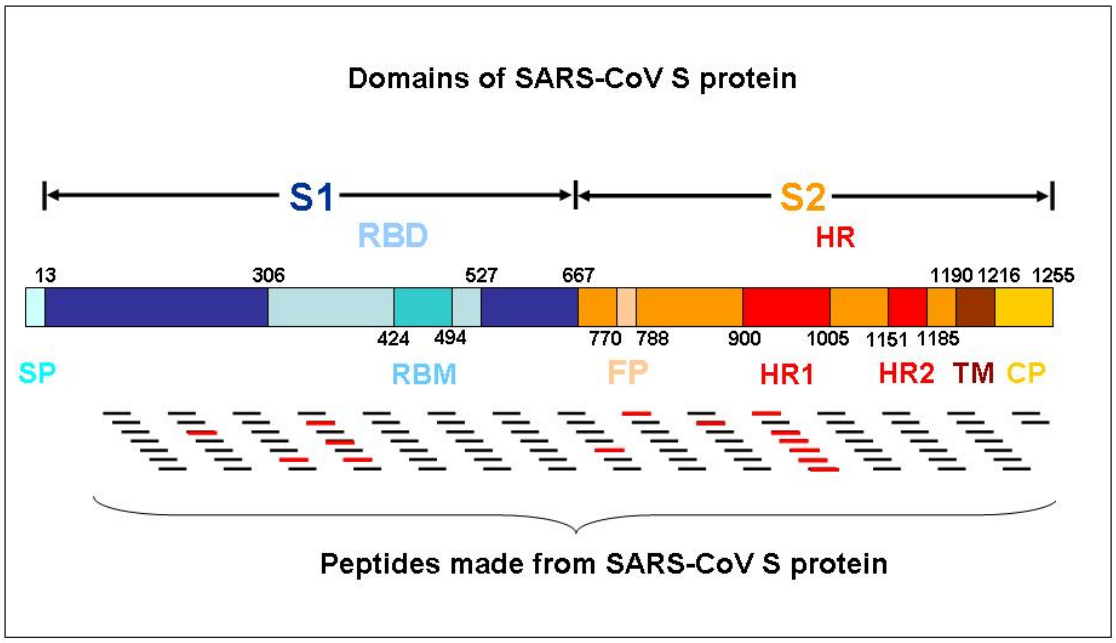

Figure 31. Schematic diagram of preparation of peptide library from S protein.

The entire peptide library was synthesized by SPOT method as 452 overlapping peptides of the 116-1255 amino acids of S protein, each 15aa in length on the cellulose membrane. The red color-labeled peptides indicate the position of peptides identified to bind S1188HA during the screening. Most of these peptides are located in the RBD and HR1 regions.

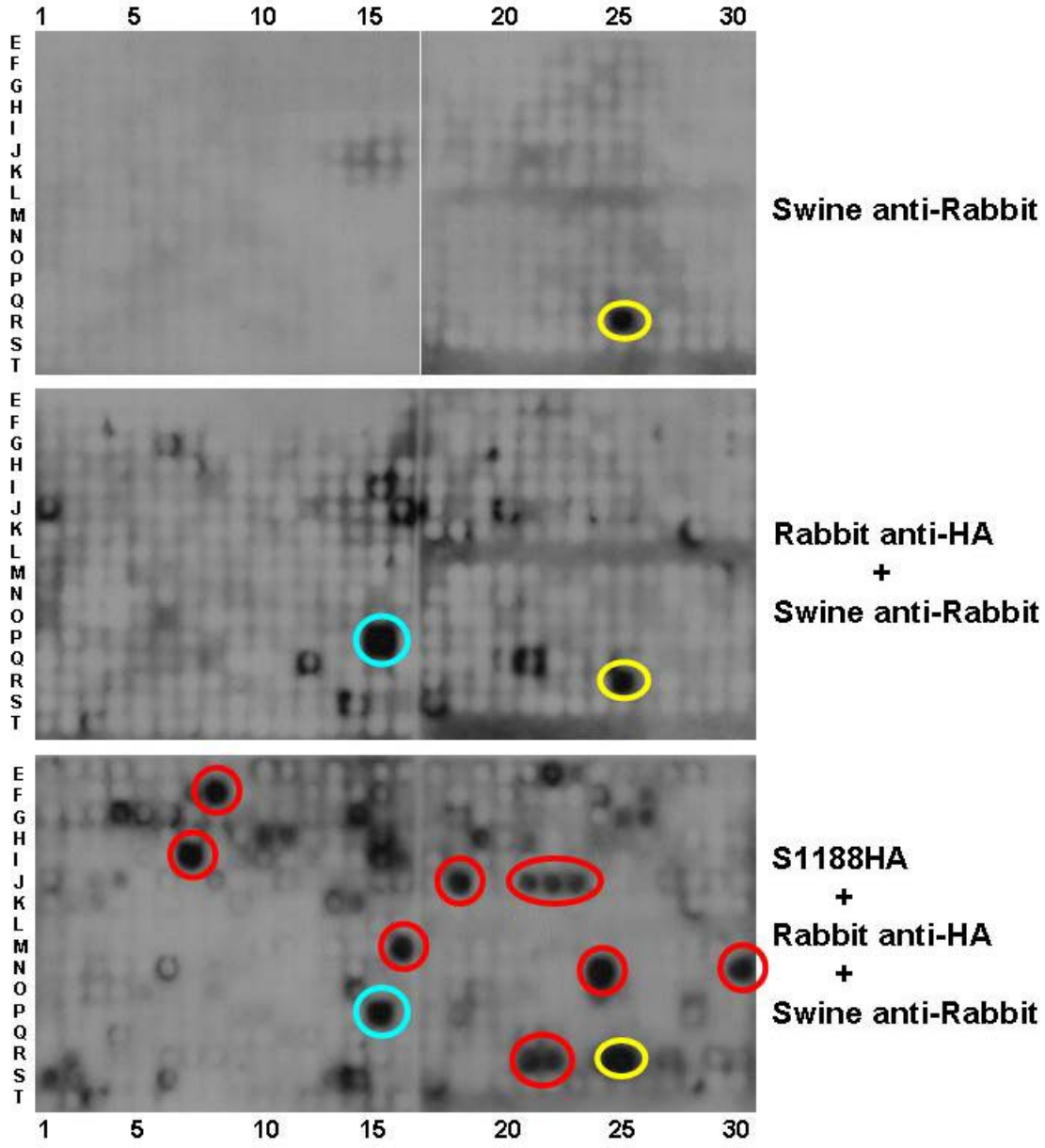

Figure 32. Peptide library screening with native S1188HA protein.

Results of screening with or without S1188HA protein under normal condition are shown. The yellow color-labeled spot is false positive spot when using the secondary antibody to probe the membrane. The blue color-labeled spot is false positive spot when using the primary antibody, following the secondary antibody to probe the membrane. The red color-labeled spots are positive spots when using native form of S1188HA protein under normal binding condition.

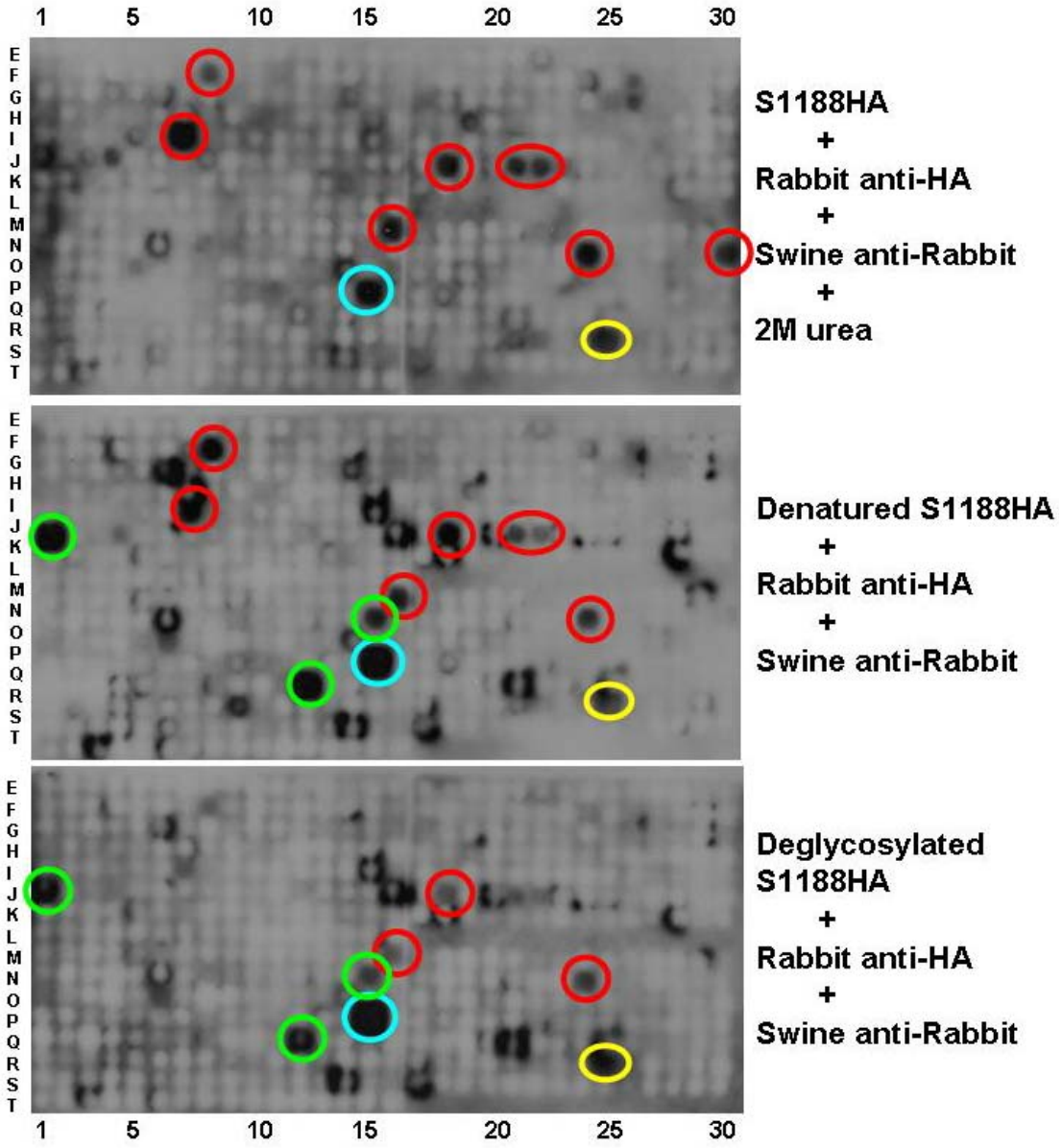

Figure 33. Peptide library screening with different forms of S1188HA protein.

The peptide library was screened with native S1188HA in 2M urea, or denatured S1188HA and deglycosylated S1188HA protein in PBST. The positive spots under each condition were shown. The yellow and blue color-labeled spots are false positive spots as described in Figure 32. The red and green color-labeled spots indicate peptides that could bind to S1188HA protein. The green color-labeled spots indicate the peptides that could bind to S1188HA protein only when blotting with denatured and deglycosylated form of S1188HA protein.

Table 8. Peptides identified to bind to S1188HA protein in screening the peptide library

Spot	Amino acid sequence	Position	Region*	S1188HA protein form			
				Native		Denatured	Deglycosylated
				PBST	2M Urea		
M16	F-L-Y-V-Y-K-G-Y-Q-P-I-D-[illegible]-V-R	193-207	S1	+	+	+	+
F 8	Q-T-S-N-F-R-V-V-P-S-G-D-V-[illegible]-D-Y	301-315	RBD	+	+	+	-
N15	Y-A-W-E-R-K-K-I-S-N-C-V-A-[illegible]-Q-I	338-353	RBD	-	-	+	+
N24	Y-A-D-S-F-V-I-K-G-D-D-V-R-[illegible]-T-R-N	383-397	RBD	+	+	+	+
N30	P-D-D-F-M-G-C-V-L-A-W-N-[illegible]-N-T	413-427	RBD	+	+	-	-
I7	N-R-A-L-S-G-I-A-A-E-Q-D-R-[illegible]-Q-I	746-760	S2	+	+	+	-
Q12	T-P-T-L-K-Y-F-G-G-F-N-F-S-[illegible]-I-P-F	773-787	FP	-	-	+	+
J1	A-G-W-T-F-G-A-G-A-A-L-Q-[illegible]-R	866-880	S2	-	-	+	+
J18	N-F-G-A-I-S-S-V-L-N-D-I-L-[illegible]-I-T	951-965	HR1	+	+	+	+
J21	L-D-K-V-E-A-E-V-Q-I-D-R-L-[illegible]-R	966-980	HR1	+	+	+	-
R21	K-V-E-A-E-V-Q-I-D-R-L-I-T-[illegible]-Q-S	968-982	HR1	+	-	-	-
J22	A-E-V-Q-I-D-R-L-I-T-G-R-L-[illegible]-Q	971-985	HR1	+	+	+	-
R22	V-Q-I-D-R-L-I-T-G-R-L-Q-S-[illegible]-Y-V	973-987	HR1	+	-	-	-
J23	D-R-L-I-T-G-R-L-Q-S-L-Q-T-[illegible]	976-990	HR1	+	-	-	-

*RBD: receptor-binding domain; FP: fusion peptide; HR1: heptad repeat 1.

D. Discussion

Cumulative evidence suggests that SARS-CoV S protein shares features similar to other class I fusion proteins to mediate viral entry (Bosch *et al.*, 2003). These include HA protein of influenza virus, gp160 of HIV and glycoprotein of Ebola virus. The helical domains of the prefusion state of S protein form the stalk and contain two heptad repeat regions, HR1 and HR2. These two regions rearrange to a hairpin-like fusion core in the post-fusion state. This hairpin structure likely juxtaposes the viral and cellular membranes, thus facilitating membrane fusion and subsequent viral entry. The fusion core of S protein has been crystallized, and its structure was determined at 2.8 Å of resolution by Rao's group (Xu *et al.*, 2004b). It is a 6-helix bundle with three HR2 helices packed against the hydrophobic grooves on the surface of central coiled coil formed by three parallel HR1 helices in an oblique antiparallel manner. This structure shares significant similarity with the fusion core structure of MHV S protein and other class I viral fusion proteins, suggesting a conserved mechanism of membrane fusion (Tripet *et al.*, 2004). Strategies aimed at inhibiting viral entry by blocking hairpin formation have been successfully developed for entry inhibitors in a Class I Env virus (Kliger and Levanon, 2003). For example, a 36-residue HR2 peptide T20 has been approved as the first pre-entry drug against HIV. It shows efficacy comparable to other post-entry HIV-1 drugs in reducing viral loads in AIDS patients (Kilby *et al.*, 1998).

Thus far, pre-entry synthetic peptide inhibitors are generally prepared chemically. To provide a recombinant method to develop pre-entry peptide inhibitors and to

design new affinity peptide-probes to study HR1-HR2 interactions, various HR2 segments appended with a MBP at their N-termini were prepared and used to dissect their interactions with S protein ectodomain S1188HA.

My results show that all these nine MBP-HR2 fusion proteins were capable of binding to S1188HA, the S protein ectodomain containing both HR1 and HR2 regions. Sequential C-terminal truncations of HR2 from 35 to 27 residues in the MBP-HR2 constructs diminished the binding of S1188HA, suggesting that the peptide length may play a role in conformational stability which may affect its binding to the putative HR1 region. Thus, the MBP-HR2 constructs can be used as probes to pull down S proteins for identification and purification similar to use of the conventional antibodies. Such peptides may also have potentials to be used as tools to screen for anti-SARS entry inhibitors.

SPOT synthesis is an easy-to-handle and flexible technique for parallel chemical synthesis on membrane supports (Frank, 2002). SPOTs technology was originally developed as a system for determining peptide antigens. Peptide sequences corresponding to the specificity of an antibody are identified by conventional enzyme-linked or autoradiographic detection methods. This technology has since progressed to include a variety of new applications for research in many areas of biology and biochemistry.

SPOTs technology enables researchers to generate the enormous sequence diversity of custom peptide library for mapping applications. A major advantage of this technology is that it facilitates identification of the functional domains of proteins

by conducting functional assays to be performed on peptide libraries attached to cellulose membranes.

Exploiting the known characteristics of S protein, I used this technology to investigate the potential S peptide fragments that bind to recombinant S protein itself. The positive spots provide clues about the self-interaction sites of S protein which could be further developed as inhibitors, synthetic antibodies or as probes for diagnostic applications or to provide some motif for investigating co-receptors.

The results show that the majority of the peptides that bound native S1188HA protein are located in RBD and HR1 regions. Deglycosylation or denaturation of S1188HA removes the binding to most of these peptides, but also new interactions are observed. These results suggest that specific interactions between S1188HA protein and positive peptides largely depend on glycosylation and the conformational state of the S1188HA protein. The affinity between native S1188HA protein and these positive peptides appear to be high because screening under more stringent conditions using 2M urea results in removing only three overlapping peptide spots located in HR1 region, but not other positive spots.

These positive spots also indicate that these peptide sequences most possibly are exposed surfaces of S protein which could support self-interaction. Prabakaran *et al*. (Prabakaran *et al*., 2006) defined 7 exposed β-sheets in RBD which supports my findings that positive spots N15 (338-353 aa) located in β1-sheet and N24 (383-397 aa) located in β3-sheet are exposed surfaces of S protein. RBD contains more than 200 aa, the crystal structure shows that the receptor-binding motif contains about 70

aa, the other amino acids form tertiary and quaternary structures to maintain the trimeric states of S protein. Thus, RBD peptides associate with RBD domain could be attributed to these interactions, e.g. complementarity of anti-parallel strands of RBD with our peptide probes. More importantly, the RBD is exposed for receptor binding to facilitate these interactions. It should be noted that aggregation and recognition of like sequences in peptides and proteins are well known and common occurrences.

There is an apparent contradiction of my results using MBP-HR2 constructs and the peptide library. Whereas the MBP-HR2 constructs are able to pull down S1188HA protein, no positive spot is observed in the HR2 region when screening the peptide library. A possible reason could be the lengths of HR2 peptides which are longer (27-35 residues) than those synthesized in the peptide library (15 residues). My results in the pull-down experiments show that sequential truncation of HR2 peptides diminish their ability to pull down S1188HA protein. For HR2 peptides in gp41 of HIV, HR2 >30 amino acids have nmolar affinity to HR1 peptides and HR2<30 amino acids have mmolar affinity. These results support that the length of HR2 peptides may affect their conformational states as affinity probes.

Another possible reason could be the polarity imposing steric factor of the peptide probes to limit its interaction with the HR1 region. In the class I fusion model, the HR1 and HR2 interaction are antiparallel, with HR2 reversing back to bind to the HR1 region. This interaction may require a flexible C-terminal end of HR2. In MBP-HR2 constructs, the capture moiety MBP is appended at the N-terminal of the

HR2 peptides which is free at their carboxyl ends. In contrast, the peptide probes in the peptide library are appended to the C-termini with their N-termini free. Unpublished results in Prof. Tam's laboratory provide support for this conclusion as HR2 peptides constrained at their C-terminal ends are 5-30 fold less potent than HR2 peptides as pre-entry inhibitors and in their binding affinity to the HR1 peptides.

The effect of polarity on peptide-protein interactions is well known. For example, PDZ domains bind to peptides only if the ligands have a free C-terminus. Unfortunately, the constraint of the present protocol in SPOT synthesis permits only growing peptides at their N-termini with their C-terminal ends attached to the cellulose support.

Although my results are preliminary, the proof-of-concept is promising. This technology is currently being pursued by my colleagues in Prof. Tam's laboratory. Thus far, follow-up studies pertinent to this work have confirmed my preliminary results. For example, the SPOT-library have been repeated with new syntheses and longer overlapping peptides (20 aa). Screenings with S1188HA protein on these libraries generally agree with my results. Furthermore, the positive peptides appended with a biotin probe and prepared as free peptides are able to pull down S protein by streptavidin beads.

Although there are unanswered questions and rules to be learned regarding this platform technology, such as whether there is any conserved motif in these peptides, whether these peptides bind to themselves or any other sites of S protein, whether these binding sites are responsible for the trimerization, this strategy may provide a

new platform for our study, and "hot spots" to provide information for the subsequent study to develop antiviral inhibitors, diagnostics peptide chips, synthetic antibodies, and screening for co-receptors.

In summary, my results give support for the following conclusions in developing peptides as affinity probes to S protein.

1. The complementarity of HR1 and HR2 regions to form a 6-helix bundle in the post-fusion state provide useful structural clues for developing peptides as affinity probes to study and characterize S protein and in general, class I viral fusion proteins.
2. The length and polarity of the peptide constructs appear to play important roles in designing affinity probes. MBP-HR2 probes, with MBP fused at the N-termini of HR2 segments, proved to be effective to bind to HR1 region of S protein ectodomain in pull-down experiments using amylose resin. HR2-peptides with its C-terminus anchored on membrane were ineffective in binding to HR1 region.
3. The positive peptides found to bind S1188HA in the screening are largely limited to the exposed surfaces with α-helices and β-sheets structures that are located in RBD and HR1 regions. This provides information to further understand the structural and functional characteristics of S protein.

Chapter 6 Conclusions and perspectives

SARS was first reported in Asia in February 2003. In the following few months, the illness rapidly spread to more than two dozen countries in North America, South America, Europe and Asia before SARS was first recognized as a global threat. With concerted efforts of laboratories world-wide, the pathogen causing SARS was identified as SARS-CoV, and the genome of SARS-CoV was sequenced (Marra *et al.*, 2003; Rota *et al.*, 2003). In late 2003, its receptor was identified as ACE2 (Li *et al.*, 2003), and the mapping of the receptor-binding domain (Wong *et al.*, 2004; Xiao *et al.*, 2003), the crystal structures of fusion core and the receptor-binding domain complex with ACE2 soon followed (Li *et al.*, 2005; Xu *et al.*, 2004b).

The chains of events on SARS outbreak altered the pace of research on coronavirus and my thesis. Conceived in the fall of 2002 before the SARS outbreak, my thesis initially focused on coronavirus affecting chickens was refocused onto humans after the outbreak. This change brought mixed blessings. The global efforts of SARS-CoV produced publications at such a rapid pace that necessitated continuing adjustments of my research plan during the past few years. More seriously, it was not possible to develop infectivity assays using the highly infectious SARS-CoV under normal laboratory conditions. To adapt to these changes, I learned different experimental approaches and laboratory skills. A case in point is that I have broadened my skills in biochemical techniques, peptide and protein chemistry instead of relying solely on molecular biology in the experimental design and methods of this thesis.

With this background, I would like to summarize my conclusions as follows.

1. For structure-function studies of S protein and many of its site-directed mutants, the mammalian expression system and codon-optimized gene was found to be suitable for obtaining functional S protein with post-translational modification close to its native form for produce pseudovirus to mimic viral entry in the single-cycle infectivity assay. The insect cell expression system was suitable to obtain relatively large amounts of glycosylated S protein and its ectodomains for various functional and binding assays.

2. Receptor-associated lipid rafts which are exploited by viruses to gain entry into host cells are also involved in SARS-CoV entry process. Clustering ACE2 receptors on host epithelial cell membranes facilitate the entry process. This finding adds detail of the requirement and specificity of SARS-CoV in the early stage of viral infection.

3. To facilitate entry, the clustering of ACE2 receptors on the host cell membrane may be complemented by clustering of spike proteins (e.g. S protein) on viral surface as observed by cryoEM in HIV and SIV. The highly conserved aromatic amino acids and Trp-rich region of the S protein proximal to the external membrane may play an important role in this process. Altering its aromatic character decreases infectivity and replacing three or more aromatic amino acid residues with Ala abrogates infectivity. This region likely serves as a “foot” in combination with the TM region to anchor the tripod-like spike protein in clusters. It may also serve additional roles to trigger the aggregation of lipid rafts after binding

to its receptor and interacting with lipid to form and enlarge the fusion pore to facilitate viral entry. Understanding this region may provide useful clues to design fusion inhibitors and to develop synthetic vaccines to elicit broadly neutralizing antibodies.

4. The concept of using peptide library to probe self-interaction was successfully validated. This strategy departs from the conventional mapping strategy of epitope mapping using antibodies. The predicted interaction of the HR1 and HR2 peptides based on the postfusion state to form a 6-helix bundle was confirmed using peptide-affinity probes of HR2 produced by recombinant method. The use of more global screening by SPOTs pepscan with overlapping peptides covering almost entire length of S protein revealed that self-interacting peptides largely limited to the exposed surfaces with ordered structures of the receptor-binding domain and the HR1 regions. These “hot spots” could provide clues for understanding oligomeric states of S protein and for designing inhibitors, synthetic antibodies, and diagnostics peptide chips for diagnostics of various SARS-CoV.

Clearly, many questions in the viral entry process remain unanswered. For example, the working mechanisms of the Trp-rich region with lipid rafts on host cells during fusion are intriguing and unknown. Furthermore, the multiple roles and conformations of S protein during viral entry and budding, such as its interaction with the lipid through palmitoylation will need further studies.

References

Adungo, F. O., Adungo, N. I., Bedno, S., and Yingst, S. L. (2005). Influenza: the next pandemic?: A review. *East Afr Med J* **82**(9)**,** 477-81.

Ahn, A., Gibbons, D. L., and Kielian, M. (2002). The fusion peptide of semliki forest virus associates with sterol-rich membrane domains. *Journal of virology* **76**(7)**,** 3267-3275.

Aiken, C. (1997). Pseudotyping human immunodeficiency virus type 1 (HIV-1) by the glycoprotein of vesicular stomatitis virus targets HIV-1 entry to an endocytic pathway and suppresses both the requirement for Nef and the sensitivity to cyclosporin A. *J Virol* **71**(8)**,** 5871-7.

Antoine A. F. de Vries, M. C. H., and Peter J. M. Rottier, a. R. J. d. G. (1997). The Genome Organization of the Nidovirales: Similarities and Differences between Arteri-, Toro-, and Coronaviruses. *Seminars in VIROLOGY* **8,** 33-47.

Baker, K. A., Dutch, R. E., Lamb, R. A., and Jardetzky, T. S. (1999). Structural basis for paramyxovirus-mediated membrane fusion. *Mol Cell* **3**(3)**,** 309-19.

Baltimore, D. (1971). Expression of animal virus genomes. *Bacteriol Rev* **35**(3)**,** 235-41.

Bamford, D. H., Grimes, J. M., and Stuart, D. I. (2005). What does structure tell us about virus evolution? *Curr Opin Struct Biol* **15**(6)**,** 655-63.

Barbato, G., Bianchi, E., Ingallinella, P., Hurni, W. H., Miller, M. D., Ciliberto, G., Cortese, R., Bazzo, R., Shiver, J. W., and Pessi, A. (2003). Structural analysis of the epitope of the anti-HIV antibody 2F5 sheds light into its mechanism of neutralization and HIV fusion. *J Mol Biol* **330**(5)**,** 1101-15.

Bavari, S., Bosio, C. M., Wiegand, E., Ruthel, G., Will, A. B., Geisbert, T. W., Hevey, M., Schmaljohn, C., Schmaljohn, A., and Aman, M. J. (2002). Lipid raft microdomains: a gateway for compartmentalized trafficking of Ebola and Marburg viruses. *J Exp Med* **195**(5)**,** 593-602.

Benbacer, L., Kut, E., Besnardeau, L., Laude, H., and Delmas, B. (1997). Interspecies aminopeptidase-N chimeras reveal species-specific receptor recognition by canine coronavirus, feline infectious peritonitis virus, and transmissible gastroenteritis virus. *J Virol* **71**(1)**,** 734-7.

Bonavia, A., Zelus, B. D., Wentworth, D. E., Talbot, P. J., and Holmes, K. V. (2003). Identification of a receptor-binding domain of the spike glycoprotein of human coronavirus HCoV-229E. *J Virol* **77**(4)**,** 2530-8.

Bos, E. C., Heijnen, L., Luytjes, W., and Spaan, W. J. (1995). Mutational analysis of the murine coronavirus spike protein: effect on cell-to-cell fusion. *Virology* **214**(2)**,** 453-63.

Bosch, B. J., van der Zee, R., de Haan, C. A., and Rottier, P. J. (2003). The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex. *J Virol* **77**(16)**,** 8801-11.

Braunwald, J., Nonnenmacher, H., Pereira, C. A., and Kirn, A. (1991). Increased susceptibility to mouse hepatitis virus type 3 (MHV3) infection induced by a

hypercholesterolaemic diet with increased adsorption of MHV3 to primary hepatocyte cultures. *Res Virol* **142**(1)**,** 5-15.

Breslin, J. J., Mork, I., Smith, M. K., Vogel, L. K., Hemmila, E. M., Bonavia, A., Talbot, P. J., Sjostrom, H., Noren, O., and Holmes, K. V. (2003). Human coronavirus 229E: receptor binding domain and neutralization by soluble receptor at 37 degrees C. *J Virol* **77**(7)**,** 4435-8.

Brian, D. A., B. G. Hogue, and T. E. Kienzle. (1995). The coronavirus hemagglutinin esterase glycoprotein. *In* "The Coronaviridae" (S. G. Siddell, Ed.), pp. 165-179. Plenum Press,, New York, N. Y.

Britton, P. (1991). Coronavirus motif. *Nature* **353**(6343)**,** 394.

Brown, D. (2002). Structure and function of membrane rafts. *Int J Med Microbiol* **291**(6-7)**,** 433-7.

Brown, D. A., and London, E. (2000). Structure and function of sphingolipid- and cholesterol-rich membrane rafts. *J Biol Chem* **275**(23)**,** 17221-4.

Bullough, P. A., Hughson, F. M., Skehel, J. J., and Wiley, D. C. (1994). Structure of influenza haemagglutinin at the pH of membrane fusion. *Nature* **371**(6492)**,** 37-43.

Caffrey, M., Cai, M., Kaufman, J., Stahl, S. J., Wingfield, P. T., Covell, D. G., Gronenborn, A. M., and Clore, G. M. (1998). Three-dimensional solution structure of the 44 kDa ectodomain of SIV gp41. *Embo J* **17**(16)**,** 4572-84.

Cambi, A., Koopman, M., and Figdor, C. G. (2005). How C-type lectins detect pathogens. *Cell Microbiol* **7**(4)**,** 481-8.

Caroni, P. (2001). New EMBO members' review: actin cytoskeleton regulation through modulation of PI(4,5)P(2) rafts. *Embo J* **20**(16)**,** 4332-6.

Chamberlain, L. H., Burgoyne, R. D., and Gould, G. W. (2001). SNARE proteins are highly enriched in lipid rafts in PC12 cells: implications for the spatial control of exocytosis. *Proc Natl Acad Sci U S A* **98**(10)**,** 5619-24.

Chan, D. C., Fass, D., Berger, J. M., and Kim, P. S. (1997). Core structure of gp41 from the HIV envelope glycoprotein. *Cell* **89**(2)**,** 263-73.

Chan, D. C., and Kim, P. S. (1998). HIV entry and its inhibition. *Cell* **93**(5)**,** 681-4.

Chen, J., Skehel, J. J., and Wiley, D. C. (1999). N- and C-terminal residues combine in the fusion-pH influenza hemagglutinin HA(2) subunit to form an N cap that terminates the triple-stranded coiled coil. *Proc Natl Acad Sci U S A* **96**(16)**,** 8967-72.

Chernomordik, L. V., and Kozlov, M. M. (2003). Protein-lipid interplay in fusion and fission of biological membranes. *Annu Rev Biochem* **72,** 175-207.

Chernomordik, L. V., Leikina, E., Kozlov, M. M., Frolov, V. A., and Zimmerberg, J. (1999). Structural intermediates in influenza haemagglutinin-mediated fusion. *Mol Membr Biol* **16**(1)**,** 33-42.

Choi, K. S., Aizaki, H., and Lai, M. M. (2005). Murine coronavirus requires lipid rafts for virus entry and cell-cell fusion but not for virus release. *J Virol* **79**(15)**,** 9862-71.

Chu, F. K., Trimble, R. B., and Maley, F. (1978). The effect of carbohydrate depletion on the properties of yeast external invertase. *J Biol Chem* **253**(24)**,** 8691-3.

Davies, H. A., and Macnaughton, M. R. (1979). Comparison of the morphology of three coronaviruses. *Arch Virol* **59**(1-2)**,** 25-33.

Daya, M., Cervin, M., and Anderson, R. (1988). Cholesterol enhances mouse hepatitis virus-mediated cell fusion. *Virology* **163**(2)**,** 276-83.

Del Real, G., Jimenez-Baranda, S., Lacalle, R. A., Mira, E., Lucas, P., Gomez-Mouton, C., Carrera, A. C., Martinez, A. C., and Manes, S. (2002). Blocking of HIV-1 infection by targeting CD4 to nonraft membrane domains. *J Exp Med* **196**(3)**,** 293-301.

Delmas, B., and Laude, H. (1990). Assembly of coronavirus spike protein into trimers and its role in epitope expression. *J Virol* **64**(11)**,** 5367-75.

Dimitrov, D. S. (2000). Cell biology of virus entry. *Cell* **101**(7)**,** 697-702.

Dimitrov, D. S. (2003). The secret life of ACE2 as a receptor for the SARS virus. *Cell* **115**(6)**,** 652-3.

Doms, R. W. (2000). Beyond receptor expression: the influence of receptor conformation, density, and affinity in HIV-1 infection. *Virology* **276**(2)**,** 229-37.

Doms, R. W., and Moore, J. P. (2000). HIV-1 membrane fusion: targets of opportunity. *J Cell Biol* **151**(2)**,** F9-14.

Drosten, C., Gunther, S., Preiser, W., van der Werf, S., Brodt, H. R., Becker, S., Rabenau, H., Panning, M., Kolesnikova, L., Fouchier, R. A., Berger, A., Burguiere, A. M., Cinatl, J., Eickmann, M., Escriou, N., Grywna, K., Kramme, S., Manuguerra, J. C., Muller, S., Rickerts, V., Sturmer, M., Vieth, S., Klenk, H. D., Osterhaus, A. D., Schmitz, H., and Doerr, H. W. (2003). Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *N Engl J Med* **348**(20)**,** 1967-76.

Dube, S., Fisher, J. W., and Powell, J. S. (1988). Glycosylation at specific sites of erythropoietin is essential for biosynthesis, secretion, and biological function. *J Biol Chem* **263**(33)**,** 17516-21.

Dutch, R. E., Jardetzky, T. S., and Lamb, R. A. (2000). Virus membrane fusion proteins: biological machines that undergo a metamorphosis. *Biosci Rep* **20**(6)**,** 597-612.

Dveksler, G. S., Pensiero, M. N., Cardellichio, C. B., Williams, R. K., Jiang, G. S., Holmes, K. V., and Dieffenbach, C. W. (1991). Cloning of the mouse hepatitis virus (MHV) receptor: expression in human and hamster cell lines confers susceptibility to MHV. *J Virol* **65**(12)**,** 6881-91.

Eckert, D. M., and Kim, P. S. (2001). Mechanisms of viral membrane fusion and its inhibition. *Annu Rev Biochem* **70,** 777-810.

Epand, R. F., Sayer, B. G., and Epand, R. M. (2005). The tryptophan-rich region of HIV gp41 and the promotion of cholesterol-rich domains. *Biochemistry* **44**(14)**,** 5525-31.

Epand, R. F., Thomas, A., Brasseur, R., Vishwanathan, S. A., Hunter, E., and Epand, R. M. (2006). Juxtamembrane protein segments that contribute to recruitment of cholesterol into domains. *Biochemistry* **45**(19)**,** 6105-14.

Epand, R. M. (2004). Do proteins facilitate the formation of cholesterol-rich domains?

Biochim Biophys Acta **1666**(1-2)**,** 227-38.

Epand, R. M., Sayer, B. G., and Epand, R. F. (2003). Peptide-induced formation of cholesterol-rich domains. *Biochemistry* **42**(49)**,** 14677-89.

Federico, M. (2003). Lentivirus gene engineering protocols. *In* "Methods in Molecular Biology." Vol. 229. Humana Press.

Frank, R. (2002). The SPOT-synthesis technique. Synthetic peptide arrays on membrane supports--principles and applications. *J Immunol Methods* **267**(1)**,** 13-26.

Gallagher, T. M., and Buchmeier, M. J. (2001). Coronavirus spike proteins in viral entry and pathogenesis. *Virology* **279**(2)**,** 371-4.

Gallo, S. A., Finnegan, C. M., Viard, M., Raviv, Y., Dimitrov, A., Rawat, S. S., Puri, A., Durell, S., and Blumenthal, R. (2003). The HIV Env-mediated fusion reaction. *Biochim Biophys Acta* **1614**(1)**,** 36-50.

Gatfield, J., and Pieters, J. (2000). Essential role for cholesterol in entry of mycobacteria into macrophages. *Science* **288**(5471)**,** 1647-50.

Gibson, R., Schlesinger, S., and Kornfeld, S. (1979). The nonglycosylated glycoprotein of vesicular stomatitis virus is temperature-sensitive and undergoes intracellular aggregation at elevated temperatures. *J Biol Chem* **254**(9)**,** 3600-7.

Gimpl, G., and Fahrenholz, F. (2000). Human oxytocin receptors in cholesterol-rich vs. cholesterol-poor microdomains of the plasma membrane. *Eur J Biochem* **267**(9)**,** 2483-97.

Godet, M., Grosclaude, J., Delmas, B., and 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**(12)**,** 8008-16.

Gorbalenya, A. E., Enjuanes, L., Ziebuhr, J., and Snijder, E. J. (2006). Nidovirales: evolving the largest RNA virus genome. *Virus Res* **117**(1)**,** 17-37.

Griffiths, G., and Rottier, P. (1992). Cell biology of viruses that assemble along the biosynthetic pathway. *Semin Cell Biol* **3**(5)**,** 367-81.

Guillen, J., Perez-Berna, A. J., Moreno, M. R., and Villalain, J. (2005). Identification of the membrane-active regions of the severe acute respiratory syndrome coronavirus spike membrane glycoprotein using a 16/18-mer peptide scan: implications for the viral fusion mechanism. *J Virol* **79**(3)**,** 1743-52.

Gustafsson, C., Govindarajan, S., and Minshull, J. (2004). Codon bias and heterologous protein expression. *Trends Biotechnol* **22**(7)**,** 346-53.

Haijema, B. J., Volders, H., and Rottier, P. J. (2003). Switching species tropism: an effective way to manipulate the feline coronavirus genome. *J Virol* **77**(8)**,** 4528-38.

Hambleton, S. (2005). Chickenpox. *Curr Opin Infect Dis* **18**(3)**,** 235-40.

Hamming, I., Timens, W., Bulthuis, M. L., Lely, A. T., Navis, G. J., and van Goor, H. (2004). Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first step in understanding SARS pathogenesis. *J Pathol* **203**(2)**,** 631-7.

He, Y., Zhou, Y., Siddiqui, P., and Jiang, S. (2004). Inactivated SARS-CoV vaccine elicits high titers of spike protein-specific antibodies that block receptor binding and virus entry. *Biochem Biophys Res Commun* **325**(2)**,** 445-52.

Hegyi, A., and Kolb, A. F. (1998). Characterization of determinants involved in the feline infectious peritonitis virus receptor function of feline aminopeptidase N. *J Gen Virol* **79 (Pt 6),** 1387-91.

Heiser, W. C. (2004). Gene delivery to mammalian cells. *In* "Methods in Molecular Biology." (W. C. Heiser, Ed.), Vol. 246. Humana Press.

Hermida-Matsumoto, L., and Resh, M. D. (2000). Localization of human immunodeficiency virus type 1 Gag and Env at the plasma membrane by confocal imaging. *J Virol* **74**(18)**,** 8670-9.

Hofmann, H., Hattermann, K., Marzi, A., Gramberg, T., Geier, M., Krumbiegel, M., Kuate, S., Uberla, K., Niedrig, M., and Pohlmann, S. (2004). S protein of severe acute respiratory syndrome-associated coronavirus mediates entry into hepatoma cell lines and is targeted by neutralizing antibodies in infected patients. *J Virol* **78**(12)**,** 6134-42.

Hofmann, H., and Pohlmann, S. (2004). Cellular entry of the SARS coronavirus. *Trends Microbiol* **12**(10)**,** 466-72.

Hsieh, P. K., Chang, S. C., Huang, C. C., Lee, T. T., Hsiao, C. W., Kou, Y. H., Chen, I. Y., Chang, C. K., Huang, T. H., and Chang, M. F. (2005). Assembly of severe acute respiratory syndrome coronavirus RNA packaging signal into virus-like particles is nucleocapsid dependent. *J Virol* **79**(22)**,** 13848-55.

Ignjatovic, J., and Galli, L. (1994). The S1 glycoprotein but not the N or M proteins of avian infectious bronchitis virus induces protection in vaccinated chickens. *Arch Virol* **138**(1-2)**,** 117-34.

Ilangumaran, S., and Hoessli, D. C. (1998). Effects of cholesterol depletion by cyclodextrin on the sphingolipid microdomains of the plasma membrane. *Biochem J* **335 (Pt 2),** 433-40.

Jackwood, M. W., Hilt, D. A., Callison, S. A., Lee, C. W., Plaza, H., and Wade, E. (2001). Spike glycoprotein cleavage recognition site analysis of infectious bronchitis virus. *Avian Dis* **45**(2)**,** 366-72.

Janes, P. W., Ley, S. C., and Magee, A. I. (1999). Aggregation of lipid rafts accompanies signaling via the T cell antigen receptor. *J Cell Biol* **147**(2)**,** 447-61.

Jeffers, S. A., Tusell, S. M., Gillim-Ross, L., Hemmila, E. M., Achenbach, J. E., Babcock, G. J., Thomas, W. D., Jr., Thackray, L. B., Young, M. D., Mason, R. J., Ambrosino, D. M., Wentworth, D. E., Demartini, J. C., and Holmes, K. V. (2004). CD209L (L-SIGN) is a receptor for severe acute respiratory syndrome coronavirus. *Proc Natl Acad Sci U S A* **101**(44)**,** 15748-53.

Kabouridis, P. S., Janzen, J., Magee, A. L., and Ley, S. C. (2000). Cholesterol depletion disrupts lipid rafts and modulates the activity of multiple signaling pathways in T lymphocytes. *Eur J Immunol* **30**(3)**,** 954-63.

Kidd, I. M., and Emery, V. C. (1993). The use of baculoviruses as expression vectors. *Appl Biochem Biotechnol* **42**(2-3)**,** 137-59.

Kilby, J. M., Hopkins, S., Venetta, T. M., DiMassimo, B., Cloud, G. A., Lee, J. Y., Alldredge, L., Hunter, E., Lambert, D., Bolognesi, D., Matthews, T., Johnson, M. R., Nowak, M. A., Shaw, G. M., and Saag, M. S. (1998). Potent suppression of HIV-1 replication in humans by T-20, a peptide inhibitor of gp41-mediated virus entry. *Nat Med* **4**(11)**,** 1302-7.

Klausegger, A., Strobl, B., Regl, G., Kaser, A., Luytjes, W., and Vlasak, R. (1999). Identification of a coronavirus hemagglutinin-esterase with a substrate specificity different from those of influenza C virus and bovine coronavirus. *J Virol* **73**(5)**,** 3737-43.

Kliger, Y., and Levanon, E. Y. (2003). Cloaked similarity between HIV-1 and SARS-CoV suggests an anti-SARS strategy. *BMC Microbiol* **3**(1)**,** 20.

Kovbasnjuk, O., Edidin, M., and Donowitz, M. (2001). Role of lipid rafts in Shiga toxin 1 interaction with the apical surface of Caco-2 cells. *J Cell Sci* **114**(Pt 22)**,** 4025-31.

Kozak, S. L., Heard, J. M., and Kabat, D. (2002). Segregation of CD4 and CXCR4 into distinct lipid microdomains in T lymphocytes suggests a mechanism for membrane destabilization by human immunodeficiency virus. *J Virol* **76**(4)**,** 1802-15.

Krokhin, O., Li, Y., Andonov, A., Feldmann, H., Flick, R., Jones, S., Stroeher, U., Bastien, N., Dasuri, K. V., Cheng, K., Simonsen, J. N., Perreault, H., Wilkins, J., Ens, W., Plummer, F., and Standing, K. G. (2003). Mass spectrometric characterization of proteins from the SARS virus: a preliminary report. *Mol Cell Proteomics* **2**(5)**,** 346-56.

Ksiazek, T. G., Erdman, D., Goldsmith, C. S., Zaki, S. R., Peret, T., Emery, S., Tong, S., Urbani, C., Comer, J. A., Lim, W., Rollin, P. E., Dowell, S. F., Ling, A. E., Humphrey, C. D., Shieh, W. J., Guarner, J., Paddock, C. D., Rota, P., Fields, B., DeRisi, J., Yang, J. Y., Cox, N., Hughes, J. M., LeDuc, J. W., Bellini, W. J., and Anderson, L. J. (2003). A novel coronavirus associated with severe acute respiratory syndrome. *N Engl J Med* **348**(20)**,** 1953-66.

Kubo, H., Yamada, Y. K., and 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**(9)**,** 5403-10.

Kulakosky, P. C., Hughes, P. R., and Wood, H. A. (1998). N-Linked glycosylation of a baculovirus-expressed recombinant glycoprotein in insect larvae and tissue culture cells. *Glycobiology* **8**(7)**,** 741-5.

Lai, M. M. (1992). RNA recombination in animal and plant viruses. *Microbiol Rev* **56**(1)**,** 61-79.

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

Lang, T., Bruns, D., Wenzel, D., Riedel, D., Holroyd, P., Thiele, C., and Jahn, R. (2001). SNAREs are concentrated in cholesterol-dependent clusters that define docking and fusion sites for exocytosis. *Embo J* **20**(9)**,** 2202-13.

Lawless, M. K., Barney, S., Guthrie, K. I., Bucy, T. B., Petteway, S. R., Jr., and Merutka, G. (1996). HIV-1 membrane fusion mechanism: structural studies of

the interactions between biologically-active peptides from gp41. *Biochemistry* **35**(42)**,** 13697-708.

Lee, H. J., Shieh, C. K., Gorbalenya, A. E., Koonin, E. V., La Monica, N., Tuler, J., Bagdzhadzhyan, A., and Lai, M. M. (1991). The complete sequence (22 kilobases) of murine coronavirus gene 1 encoding the putative proteases and RNA polymerase. *Virology* **180**(2)**,** 567-82.

Leparc-Goffart, I., Hingley, S. T., Chua, M. M., Phillips, J., Lavi, E., and Weiss, S. R. (1998). Targeted recombination within the spike gene of murine coronavirus mouse hepatitis virus-A59: Q159 is a determinant of hepatotropism. *J Virol* **72**(12)**,** 9628-36.

Li, F., Li, W., Farzan, M., and Harrison, S. C. (2005). Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. *Science* **309**(5742)**,** 1864-8.

Li, W., Moore, M. J., Vasilieva, N., Sui, J., Wong, S. K., Berne, M. A., Somasundaran, M., Sullivan, J. L., Luzuriaga, K., Greenough, T. C., Choe, H., and Farzan, M. (2003). Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* **426**(6965)**,** 450-4.

Liao, Z., Cimakasky, L. M., Hampton, R., Nguyen, D. H., and Hildreth, J. E. (2001). Lipid rafts and HIV pathogenesis: host membrane cholesterol is required for infection by HIV type 1. *AIDS Res Hum Retroviruses* **17**(11)**,** 1009-19.

Lis, H., and Sharon, N. (1993). Protein glycosylation. Structural and functional aspects. *Eur J Biochem* **218**(1)**,** 1-27.

Loeffler F., F. P. (1897). Summarischer bericht uber die ergebnisse der untersuchungen der kommoission zur erforchung der maul-und-klamenseuche. *Zentralbl. Bakteriol. Parasitenk. Infektionskr* **22,** 257-259.

London, E., and Brown, D. A. (2000). Insolubility of lipids in triton X-100: physical origin and relationship to sphingolipid/cholesterol membrane domains (rafts). *Biochim Biophys Acta* **1508**(1-2)**,** 182-95.

Lu, X., and Silver, J. (2000). Ecotropic murine leukemia virus receptor is physically associated with caveolin and membrane rafts. *Virology* **276**(2)**,** 251-8.

Lu, Y. E., Cassese, T., and Kielian, M. (1999). The cholesterol requirement for sindbis virus entry and exit and characterization of a spike protein region involved in cholesterol dependence. *J Virol* **73**(5)**,** 4272-8.

Lu, Y. E., and Kielian, M. (2000). Semliki forest virus budding: assay, mechanisms, and cholesterol requirement. *J Virol* **74**(17)**,** 7708-19.

Luckow, V. A. (1990). Cloning and expression of heterologous genes in insect cells with baculovirus vectors. *In* "Recombinant DNA Technology and Applications" (C. Ho, Prokop, A., Baipai, R., Ed.), pp. 1-24, McGraw-Hill, New York.

Luckow, V. A. (1995). Protein production and processing from baculovirus expression vectors. *In* "Baculovirus Expression Systems and Biopesticides." (M. L. Shuler, Wood,H.A., Granados,R.R. and, and D. A. Hammer, Eds.), pp. 51-90. John Wiley & Sons, New York,.

Machida, C. A. (2003). Viral vectors for gene therapy. *In* "Methods and Protocols." (C.

A. Machida, Ed.). Humana Press.

Mackie, P. L. (2003). The classification of viruses infecting the respiratory tract. *Paediatr Respir Rev* **4**(2)**,** 84-90.

Mahfoud, R., Garmy, N., Maresca, M., Yahi, N., Puigserver, A., and Fantini, J. (2002). Identification of a common sphingolipid-binding domain in Alzheimer, prion, and HIV-1 proteins. *J Biol Chem* **277**(13)**,** 11292-6.

Malashkevich, V. N., Chan, D. C., Chutkowski, C. T., and Kim, P. S. (1998). Crystal structure of the simian immunodeficiency virus (SIV) gp41 core: conserved helical interactions underlie the broad inhibitory activity of gp41 peptides. *Proc Natl Acad Sci U S A* **95**(16)**,** 9134-9.

Maley, F., Trimble, R. B., Tarentino, A. L., and Plummer, T. H., Jr. (1989). Characterization of glycoproteins and their associated oligosaccharides through the use of endoglycosidases. *Anal Biochem* **180**(2)**,** 195-204.

Manes, S., del Real, G., Lacalle, R. A., Lucas, P., Gomez-Mouton, C., Sanchez-Palomino, S., Delgado, R., Alcami, J., Mira, E., and Martinez, A. C. (2000). Membrane raft microdomains mediate lateral assemblies required for HIV-1 infection. *EMBO Rep* **1**(2)**,** 190-6.

Manie, S. N., Debreyne, S., Vincent, S., and Gerlier, D. (2000). Measles virus structural components are enriched into lipid raft microdomains: a potential cellular location for virus assembly. *J Virol* **74**(1)**,** 305-11.

Markowska, J., Fischer, N., Markowski, M., and Nalewaj, J. (2005). The role of Chlamydia trachomatis infection in the development of cervical neoplasia and carcinoma. *Med Wieku Rozwoj* **9**(1)**,** 83-6.

Marra, M. A., Jones, S. J., Astell, C. R., Holt, R. A., Brooks-Wilson, A., Butterfield, Y. S., Khattra, J., Asano, J. K., Barber, S. A., Chan, S. Y., Cloutier, A., Coughlin, S. M., Freeman, D., Girn, N., Griffith, O. L., Leach, S. R., Mayo, M., McDonald, H., Montgomery, S. B., Pandoh, P. K., Petrescu, A. S., Robertson, A. G., Schein, J. E., Siddiqui, A., Smailus, D. E., Stott, J. M., Yang, G. S., Plummer, F., Andonov, A., Artsob, H., Bastien, N., Bernard, K., Booth, T. F., Bowness, D., Czub, M., Drebot, M., Fernando, L., Flick, R., Garbutt, M., Gray, M., Grolla, A., Jones, S., Feldmann, H., Meyers, A., Kabani, A., Li, Y., Normand, S., Stroher, U., Tipples, G. A., Tyler, S., Vogrig, R., Ward, D., Watson, B., Brunham, R. C., Krajden, M., Petric, M., Skowronski, D. M., Upton, C., and Roper, R. L. (2003). The Genome sequence of the SARS-associated coronavirus. *Science* **300**(5624)**,** 1399-404.

März, L., Altmann,F., Staudacher,E., Kubelka,V. (1995). Protein glycosylation in insects. *In* "Glycoproteins" (J. Montreuil, Vliegenthart,J.F.G. and Schachter,H., Ed.), pp. 543-563. Elsevier,, Amsterdam,.

Matzuk, M. M., Keene, J. L., and Boime, I. (1989). Site specificity of the chorionic gonadotropin N-linked oligosaccharides in signal transduction. *J Biol Chem* **264**(5)**,** 2409-14.

Melikyan, G. B., Markosyan, R. M., Hemmati, H., Delmedico, M. K., Lambert, D. M., and Cohen, F. S. (2000). Evidence that the transition of HIV-1 gp41 into a six-helix bundle, not the bundle configuration, induces membrane fusion. *J*

Cell Biol **151**(2)**,** 413-23.

Mine, Y., and Rupa, P. (2003). Fine mapping and structural analysis of immunodominant IgE allergenic epitopes in chicken egg ovalbumin. *Protein Eng* **16**(10)**,** 747-52.

Minor, P. D. (2004). Polio eradication, cessation of vaccination and re-emergence of disease. *Nat Rev Microbiol* **2**(6)**,** 473-82.

Mullins, J. I., and Jensen, M. A. (2006). Evolutionary dynamics of HIV-1 and the control of AIDS. *Curr Top Microbiol Immunol* **299,** 171-92.

Munoz-Barroso, I., Durell, S., Sakaguchi, K., Appella, E., and Blumenthal, R. (1998). Dilation of the human immunodeficiency virus-1 envelope glycoprotein fusion pore revealed by the inhibitory action of a synthetic peptide from gp41. *J Cell Biol* **140**(2)**,** 315-23.

Nal, B., Chan, C., Kien, F., Siu, L., Tse, J., Chu, K., Kam, J., Staropoli, I., Crescenzo-Chaigne, B., Escriou, N., van der Werf, S., Yuen, K. Y., and Altmeyer, R. (2005). Differential maturation and subcellular localization of severe acute respiratory syndrome coronavirus surface proteins S, M and E. *J Gen Virol* **86**(Pt 5)**,** 1423-34.

Narayanan, K., Chen, C. J., Maeda, J., and Makino, S. (2003). Nucleocapsid-independent specific viral RNA packaging via viral envelope protein and viral RNA signal. *J Virol* **77**(5)**,** 2922-7.

Narayanan, K., Maeda, A., Maeda, J., and Makino, S. (2000). Characterization of the coronavirus M protein and nucleocapsid interaction in infected cells. *J Virol* **74**(17)**,** 8127-34.

Narayanan, K., and Makino, S. (2001). Cooperation of an RNA packaging signal and a viral envelope protein in coronavirus RNA packaging. *J Virol* **75**(19)**,** 9059-67.

Ng, M. L., Tan, S. H., See, E. E., Ooi, E. E., and Ling, A. E. (2003). Early events of SARS coronavirus infection in vero cells. *J Med Virol* **71**(3)**,** 323-31.

Nguyen, D. H., and Hildreth, J. E. (2000). Evidence for budding of human immunodeficiency virus type 1 selectively from glycolipid-enriched membrane lipid rafts. *J Virol* **74**(7)**,** 3264-72.

Ni, L., Zhu, J., Zhang, J., Yan, M., Gao, G. F., and Tien, P. (2005). Design of recombinant protein-based SARS-CoV entry inhibitors targeting the heptad-repeat regions of the spike protein S2 domain. *Biochem Biophys Res Commun* **330**(1)**,** 39-45.

Nomura, R., Kiyota, A., Suzaki, E., Kataoka, K., Ohe, Y., Miyamoto, K., Senda, T., and Fujimoto, T. (2004). Human coronavirus 229E binds to CD13 in rafts and enters the cell through caveolae. *J Virol* **78**(16)**,** 8701-8.

Ono, A., and Freed, E. O. (2001). Plasma membrane rafts play a critical role in HIV-1 assembly and release. *Proc Natl Acad Sci U S A* **98**(24)**,** 13925-30.

Parton, R. G., and Lindsay, M. (1999). Exploitation of major histocompatibility complex class I molecules and caveolae by simian virus 40. *Immunol Rev* **168,** 23-31.

Peiris, J. S., Guan, Y., and Yuen, K. Y. (2004). Severe acute respiratory syndrome. *Nat*

Med **10**(12 Suppl)**,** S88-97.

Peisajovich, S. G., Blank, L., Epand, R. F., Epand, R. M., and Shai, Y. (2003). On the interaction between gp41 and membranes: the immunodominant loop stabilizes gp41 helical hairpin conformation. *J Mol Biol* **326**(5)**,** 1489-501.

Peisajovich, S. G., and Shai, Y. (2003). Viral fusion proteins: multiple regions contribute to membrane fusion. *Biochim Biophys Acta* **1614**(1)**,** 122-9.

Pelkmans, L., Kartenbeck, J., and Helenius, A. (2001). Caveolar endocytosis of simian virus 40 reveals a new two-step vesicular-transport pathway to the ER. *Nat Cell Biol* **3**(5)**,** 473-83.

Pickl, W. F., Pimentel-Muinos, F. X., and Seed, B. (2001). Lipid rafts and pseudotyping. *J Virol* **75**(15)**,** 7175-83.

Popik, W., Timothy M, A., and Au, W.-C. (2002). Human Immunodeficiency virus type 1 uses lipid rafts-colocalized CD4 and chemokine receptors for productive entry into CD4+ T cells. *Journal of virology* **76**(10)**,** 4709-4722.

Pourrut, X., Kumulungui, B., Wittmann, T., Moussavou, G., Delicat, A., Yaba, P., Nkoghe, D., Gonzalez, J. P., and Leroy, E. M. (2005). The natural history of Ebola virus in Africa. *Microbes Infect* **7**(7-8)**,** 1005-14.

Prabakaran, P., Gan, J., Feng, Y., Zhu, Z., Choudhry, V., Xiao, X., Ji, X., and Dimitrov, D. S. (2006). Structure of severe acute respiratory syndrome coronavirus receptor-binding domain complexed with neutralizing antibody. *J Biol Chem* **281**(23)**,** 15829-36.

Pralle, A., Keller, P., Florin, E. L., Simons, K., and Horber, J. K. (2000). Sphingolipid-cholesterol rafts diffuse as small entities in the plasma membrane of mammalian cells. *J Cell Biol* **148**(5)**,** 997-1008.

Puscas, L. (2005). The role of human papilloma virus infection in the etiology of oropharyngeal carcinoma. *Curr Opin Otolaryngol Head Neck Surg* **13**(4)**,** 212-6.

Qinfen, Z., Jinming, C., Xiaojun, H., Huanying, Z., Jicheng, H., Ling, F., Kunpeng, L., and Jingqiang, Z. (2004). The life cycle of SARS coronavirus in Vero E6 cells. *J Med Virol* **73**(3)**,** 332-7.

Rademacher, T. W., Parekh, R. B., and Dwek, R. A. (1988). Glycobiology. *Annu Rev Biochem* **57,** 785-838.

Rawat, S. S., Viard, M., Gallo, S. A., Rein, A., Blumenthal, R., and Puri, A. (2003). Modulation of entry of enveloped viruses by cholesterol and sphingolipids (Review). *Mol Membr Biol* **20**(3)**,** 243-54.

Ridder, A. N., Morein, S., Stam, J. G., Kuhn, A., de Kruijff, B., and Killian, J. A. (2000). Analysis of the role of interfacial tryptophan residues in controlling the topology of membrane proteins. *Biochemistry* **39**(21)**,** 6521-8.

Rota, P. A., Oberste, M. S., Monroe, S. S., Nix, W. A., Campagnoli, R., Icenogle, J. P., Penaranda, S., Bankamp, B., Maher, K., Chen, M. H., Tong, S., Tamin, A., Lowe, L., Frace, M., DeRisi, J. L., Chen, Q., Wang, D., Erdman, D. D., Peret, T. C., Burns, C., Ksiazek, T. G., Rollin, P. E., Sanchez, A., Liffick, S., Holloway, B., Limor, J., McCaustland, K., Olsen-Rasmussen, M., Fouchier, R., Gunther, S., Osterhaus, A. D., Drosten, C., Pallansch, M. A., Anderson, L. J.,

and Bellini, W. J. (2003). Characterization of a novel coronavirus associated with severe acute respiratory syndrome. *Science* **300**(5624)**,** 1394-9.

Rozelle, A. L., Machesky, L. M., Yamamoto, M., Driessens, M. H., Insall, R. H., Roth, M. G., Luby-Phelps, K., Marriott, G., Hall, A., and Yin, H. L. (2000). Phosphatidylinositol 4,5-bisphosphate induces actin-based movement of raft-enriched vesicles through WASP-Arp2/3. *Curr Biol* **10**(6)**,** 311-20.

Saez-Cirion, A., Arrondo, J. L., Gomara, M. J., Lorizate, M., Iloro, I., Melikyan, G., and Nieva, J. L. (2003a). Structural and functional roles of HIV-1 gp41 pretransmembrane sequence segmentation. *Biophys J* **85**(6)**,** 3769-80.

Saez-Cirion, A., Gomara, M. J., Agirre, A., and Nieva, J. L. (2003b). Pre-transmembrane sequence of Ebola glycoprotein. Interfacial hydrophobicity distribution and interaction with membranes. *FEBS Lett* **533**(1-3)**,** 47-53.

Saez-Cirion, A., Nir, S., Lorizate, M., Agirre, A., Cruz, A., Perez-Gil, J., and Nieva, J. L. (2002). Sphingomyelin and cholesterol promote HIV-1 gp41 pretransmembrane sequence surface aggregation and membrane restructuring. *J Biol Chem* **277**(24)**,** 21776-85.

Sainz, B., Jr., Rausch, J. M., Gallaher, W. R., Garry, R. F., and Wimley, W. C. (2005). The aromatic domain of the coronavirus class I viral fusion protein induces membrane permeabilization: putative role during viral entry. *Biochemistry* **44**(3)**,** 947-58.

Saitou, N., and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**(4)**,** 406-25.

Salzwedel, K., and Berger, E. A. (2000). Cooperative subunit interactions within the oligomeric envelope glycoprotein of HIV-1: functional complementation of specific defects in gp120 and gp41. *Proc Natl Acad Sci U S A* **97**(23)**,** 12794-9.

Samuel, B. U., Mohandas, N., Harrison, T., McManus, H., Rosse, W., Reid, M., and Haldar, K. (2001). The role of cholesterol and glycosylphosphatidylinositol-anchored proteins of erythrocyte rafts in regulating raft protein content and malarial infection. *J Biol Chem* **276**(31)**,** 29319-29.

Sanchez, C. M., Izeta, A., Sanchez-Morgado, J. M., Alonso, S., Sola, I., Balasch, M., Plana-Duran, J., and Enjuanes, L. (1999). Targeted recombination demonstrates that the spike gene of transmissible gastroenteritis coronavirus is a determinant of its enteric tropism and virulence. *J Virol* **73**(9)**,** 7607-18.

Scheiffele, P., Rietveld, A., Wilk, T., and Simons, K. (1999). Influenza viruses select ordered lipid domains during budding from the plasma membrane. *J Biol Chem* **274**(4)**,** 2038-44.

Scheiffele, P., Roth, M. G., and Simons, K. (1997). Interaction of influenza virus haemagglutinin with sphingolipid-cholesterol membrane domains via its transmembrane domain. *Embo J* **16**(18)**,** 5501-8.

Schibli, D. J., Montelaro, R. C., and Vogel, H. J. (2001). The membrane-proximal tryptophan-rich region of the HIV glycoprotein, gp41, forms a well-defined

helix in dodecylphosphocholine micelles. *Biochemistry* **40**(32)**,** 9570-8.

Semenkovich, C. F., Luo, C. C., Nakanishi, M. K., Chen, S. H., Smith, L. C., and Chan, L. (1990). In vitro expression and site-specific mutagenesis of the cloned human lipoprotein lipase gene. Potential N-linked glycosylation site asparagine 43 is important for both enzyme activity and secretion. *J Biol Chem* **265**(10)**,** 5429-33.

Serraino, D., Piselli, P., Angeletti, C., Scuderi, M., Ippolito, G., and Capobianchi, M. R. (2005). Infection with Epstein-Barr virus and cancer: an epidemiological review. *J Biol Regul Homeost Agents* **19**(1-2)**,** 63-70.

Shin, J. S., and Abraham, S. N. (2001). Caveolae as portals of entry for microbes. *Microbes Infect* **3**(9)**,** 755-61.

Simmons, G., Reeves, J. D., Rennekamp, A. J., Amberg, S. M., Piefer, A. J., and Bates, P. (2004). Characterization of severe acute respiratory syndrome-associated coronavirus (SARS-CoV) spike glycoprotein-mediated viral entry. *Proc Natl Acad Sci U S A* **101**(12)**,** 4240-5.

Simons, K., and Ikonen, E. (1997). Functional rafts in cell membranes. *Nature* **387**(6633)**,** 569-72.

Simons, K., and Toomre, D. (2000). Lipid rafts and signal transduction. *Nat Rev Mol Cell Biol* **1**(1)**,** 31-9.

Singh, M., Berger, B., and Kim, P. S. (1999). LearnCoil-VMF: computational evidence for coiled-coil-like motifs in many viral membrane-fusion proteins. *J Mol Biol* **290**(5)**,** 1031-41.

Skehel, J. J., and Wiley, D. C. (2000). Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. *Annu Rev Biochem* **69,** 531-69.

Snijder, E. J., Bredenbeek, P. J., Dobbe, J. C., Thiel, V., Ziebuhr, J., Poon, L. L., Guan, Y., Rozanov, M., Spaan, W. J., and Gorbalenya, A. E. (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**(5)**,** 991-1004.

Sobrino, F., Saiz, M., Jimenez-Clavero, M. A., Nunez, J. I., Rosas, M. F., Baranowski, E., and Ley, V. (2001). Foot-and-mouth disease virus: a long known virus, but a current threat. *Vet Res* **32**(1)**,** 1-30.

Stadler, K., Masignani, V., Eickmann, M., Becker, S., Abrignani, S., Klenk, H. D., and Rappuoli, R. (2003). SARS--beginning to understand a new virus. *Nat Rev Microbiol* **1**(3)**,** 209-18.

Suarez, T., Gallaher, W. R., Agirre, A., Goni, F. M., and Nieva, J. L. (2000a). Membrane interface-interacting sequences within the ectodomain of the human immunodeficiency virus type 1 envelope glycoprotein: putative role during viral fusion. *J Virol* **74**(17)**,** 8038-47.

Suarez, T., Nir, S., Goni, F. M., Saez-Cirion, A., and Nieva, J. L. (2000b). The pre-transmembrane region of the human immunodeficiency virus type-1 glycoprotein: a novel fusogenic sequence. *FEBS Lett* **477**(1-2)**,** 145-9.

Suomalainen, M. (2002). Lipid rafts and assembly of enveloped viruses. *Traffic* **3**(10)**,** 705-9.

Tan, K., Liu, J., Wang, J., Shen, S., and Lu, M. (1997). Atomic structure of a

thermostable subdomain of HIV-1 gp41. *Proc Natl Acad Sci U S A* **94**(23)**,** 12303-8.

Thorp, E. B., Boscarino, J. A., Logan, H. L., Goletz, J. T., and Gallagher, T. M. (2006). Palmitoylations on murine coronavirus spike proteins are essential for virion assembly and infectivity. *J Virol* **80**(3)**,** 1280-9.

Thorp, E. B., and Gallagher, T. M. (2004). Requirements for CEACAMs and cholesterol during murine coronavirus cell entry. *J Virol* **78**(6)**,** 2682-92.

Torres, J. M., Sanchez, C., Sune, C., Smerdou, C., Prevec, L., Graham, F., and Enjuanes, L. (1995). Induction of antibodies protecting against transmissible gastroenteritis coronavirus (TGEV) by recombinant adenovirus expressing TGEV spike protein. *Virology* **213**(2)**,** 503-16.

Tresnan, D. B., Levis, R., and Holmes, K. V. (1996). Feline aminopeptidase N serves as a receptor for feline, canine, porcine, and human coronaviruses in serogroup I. *J Virol* **70**(12)**,** 8669-74.

Tripet, B., Howard, M. W., Jobling, M., Holmes, R. K., Holmes, K. V., and Hodges, R. S. (2004). Structural characterization of the SARS-coronavirus spike S fusion protein core. *J Biol Chem* **279**(20)**,** 20836-49.

Tsai, J. C., Zelus, B. D., Holmes, K. V., and Weiss, S. R. (2003). The N-terminal domain of the murine coronavirus spike glycoprotein determines the CEACAM1 receptor specificity of the virus strain. *J Virol* **77**(2)**,** 841-50.

Tseng, C. T., Tseng, J., Perrone, L., Worthy, M., Popov, V., and Peters, C. J. (2005). Apical entry and release of severe acute respiratory syndrome-associated coronavirus in polarized Calu-3 lung epithelial cells. *J Virol* **79**(15)**,** 9470-9.

Tyrrell, D. A., and Bynoe, M. L. (1965). Cultivation of a Novel Type of Common-Cold Virus in Organ Cultures. *Br Med J* **5448,** 1467-70.

van der Goot, F. G., and Harder, T. (2001). Raft membrane domains: from a liquid-ordered membrane phase to a site of pathogen attack. *Semin Immunol* **13**(2)**,** 89-97.

Van Regenmortel, M. H. (1999). How to write the names of virus species. *Arch Virol* **144**(5)**,** 1041-2.

Varma, R., and Mayor, S. (1998). GPI-anchored proteins are organized in submicron domains at the cell surface. *Nature* **394**(6695)**,** 798-801.

Vincent, N., Genin, C., and Malvoisin, E. (2002). Identification of a conserved domain of the HIV-1 transmembrane protein gp41 which interacts with cholesteryl groups. *Biochim Biophys Acta* **1567**(1-2)**,** 157-64.

Waarts, B. L., Bittman, R., and Wilschut, J. (2002). Sphingolipid and cholesterol dependence of alphavirus membrane fusion. Lack of correlation with lipid raft formation in target liposomes. *J Biol Chem* **277**(41)**,** 38141-7.

Wakatsuki, S., Kurisaki, T., and Sehara-Fujisawa, A. (2004). Lipid rafts identified as locations of ectodomain shedding mediated by Meltrin beta/ADAM19. *J Neurochem* **89**(1)**,** 119-23.

Wang, C. H., Hong, C. C., and Seak, J. C. (2002). An ELISA for antibodies against infectious bronchitis virus using an S1 spike polypeptide. *Vet Microbiol* **85**(4)**,** 333-42.

Wang, P., Chen, J., Zheng, A., Nie, Y., Shi, X., Wang, W., Wang, G., Luo, M., Liu, H., Tan, L., Song, X., Wang, Z., Yin, X., Qu, X., Wang, X., Qing, T., Ding, M., and Deng, H. (2004). Expression cloning of functional receptor used by SARS coronavirus. *Biochem Biophys Res Commun* **315**(2)**,** 439-44.

Warner, F. J., Lew, R. A., Smith, A. I., Lambert, D. W., Hooper, N. M., and Turner, A. J. (2005). Angiotensin-converting enzyme 2 (ACE2), but not ACE, is preferentially localized to the apical surface of polarized kidney cells. *J Biol Chem* **280**(47)**,** 39353-62.

Wat, D. (2004). The common cold: a review of the literature. *Eur J Intern Med* **15**(2)**,** 79-88.

Weber, T., Paesold, G., Galli, C., Mischler

Wyatt, R., Kwong, P. D., Desjardins, E., Sweet, R. W., Robinson, J., Hendrickson, W. A., and Sodroski, J. G. (1998). The antigenic structure of the HIV gp120 envelope glycoprotein. *Nature* **393**(6686)**,** 705-11.

Xavier, R., Brennan, T., Li, Q., McCormack, C., and Seed, B. (1998). Membrane compartmentation is required for efficient T cell activation. *Immunity* **8**(6)**,** 723-32.

Xiao, X., Chakraborti, S., Dimitrov, A. S., Gramatikoff, K., and Dimitrov, D. S. (2003). The SARS-CoV S glycoprotein: expression and functional characterization. *Biochem Biophys Res Commun* **312**(4)**,** 1159-64.

Xu, Y., Liu, Y., Lou, Z., Qin, L., Li, X., Bai, Z., Pang, H., Tien, P., Gao, G. F., and Rao, Z. (2004a). Structural basis for coronavirus-mediated membrane fusion. Crystal structure of mouse hepatitis virus spike protein fusion core. *J Biol Chem* **279**(29)**,** 30514-22.

Xu, Y., Lou, Z., Liu, Y., Pang, H., Tien, P., Gao, G. F., and Rao, Z. (2004b). Crystal structure of severe acute respiratory syndrome coronavirus spike protein fusion core. *J Biol Chem* **279**(47)**,** 49414-9.

Yang, Z. Y., Huang, Y., Ganesh, L., Leung, K., Kong, W. P., Schwartz, O., Subbarao, K., and Nabel, G. J. (2004). pH-dependent entry of severe acute respiratory syndrome coronavirus is mediated by the spike glycoprotein and enhanced by dendritic cell transfer through DC-SIGN. *J Virol* **78**(11)**,** 5642-50.

Yau, W. M., Wimley, W. C., Gawrisch, K., and White, S. H. (1998). The preference of tryptophan for membrane interfaces. *Biochemistry* **37**(42)**,** 14713-8.

Yi, C. E., Ba, L., Zhang, L., Ho, D. D., and Chen, Z. (2005). Single amino acid substitutions in the severe acute respiratory syndrome coronavirus spike glycoprotein determine viral entry and immunogenicity of a major neutralizing domain. *J Virol* **79**(18)**,** 11638-46.

Zacharias, D. A., Violin, J. D., Newton, A. C., and Tsien, R. Y. (2002). Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells. *Science* **296**(5569)**,** 913-6.

Zhang, H., Wang, G., Li, J., Nie, Y., Shi, X., Lian, G., Wang, W., Yin, X., Zhao, Y., Qu, X., Ding, M., and Deng, H. (2004). Identification of an antigenic determinant on the S2 domain of the severe acute respiratory syndrome coronavirus spike glycoprotein capable of inducing neutralizing antibodies. *J Virol* **78**(13)**,** 6938-45.

Zhao, X., Singh, M., Malashkevich, V. N., and Kim, P. S. (2000). Structural characterization of the human respiratory syncytial virus fusion protein core. *Proc Natl Acad Sci U S A* **97**(26)**,** 14172-7.

Zhu, P., Liu, J., Bess, J., Jr., Chertova, E., Lifson, J. D., Grise, H., Ofek, G. A., Taylor, K. A., and Roux, K. H. (2006). Distribution and three-dimensional structure of AIDS virus envelope spikes. *Nature* **441**(7095)**,** 847-52.

Zuniga, S., Sola, I., Alonso, S., and Enjuanes, L. (2004). Sequence motifs involved in the regulation of discontinuous coronavirus subgenomic RNA synthesis. *J Virol* **78**(2)**,** 980-94.