Study of sars-cov entry mechanism and development of protein mimetic inhibitors

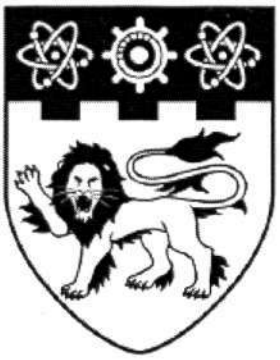
Ji, Zhe

2010

Ji, Z. (2010). Study of sars-cov entry mechanism and development of protein mimetic inhibitors. Doctoral thesis, Nanyang Technological University, Singapore.

https://hdl.handle.net/10356/47454

https://doi.org/10.32657/10356/47454



NANYANG
TECHNOLOGICAL
UNIVERSITY

STUDY OF SARS-CoV ENTRY MECHANISM AND DEVELOPMENT OF PROTEIN MIMETIC INHIBITORS

JI ZHE

SCHOOL OF BIOLOGICAL SCIENCES

2010

STUDY OF SARS-COV ENTRY MECHANISM AND DEVELOPMENT OF PROTEIN MIMETIC INHIBITORS

JI ZHE

School of Biological Sciences

A thesis submitted to the Nanyang Technological University

in fulfillment of the requirement for the degree of

Doctor of Philosophy

2010

This dissertation is dedicated to those people in my life who have made it possible, especially to my parents

Acknowledgements

When I sat down to decide who to acknowledge for their help with this work, I began to realize just how much help and encouragement I have received during the past five years. I am certainly indebted to a number of people in more ways than I can adequately recognize here. However, I would like to express my heartfelt thanks to those who have helped me to reach this point.

First of all, I wish to express my deepest appreciation to my supervisor, Professor James P. Tam, for his excellent scientific and technical guidance. His patience, enterprise, and remarkable insight have been invaluable to me over the years. This thesis would not have been possible without his great support, trenchant critiques, probing questions, cracking-of-the-whip and remarkable patience. He is my good mentor, not only in scientific research, but also in life. I benefited from his insight and wisdom these past years and will continue to do so in the future. In his daily life, he is also a charismatic leader, a humorous scholar, and a good cook. Professor Tam has talents in many aspects and it is hard to enumerate on them here.

I also wish to thank Dr. Liu Dingxiang from the Institute of Molecular Biology, Singapore, for his patient instruction and kindly help. When I first entered this field, he was the person who guided me by the hand through those first experiments. During these last five years, his suggestions have been invaluable and his support ever present. I also wish to thank Professor Michael Farzan, of Harvard University (U.S), and Professor Deng Hongkui, from Peking University (China). I will always be grateful to them for providing me with the necessary experiment materials.

I received countless help from my many professors at the university. I wish to give a heartfelt thanks to Dr. Li Jinming for his unending care and warmhearted help. To Dr. Kristein Sadler, she is like a big sister to me. Dr. Julien Lescar's guidance in structural biology was invaluable. I was taught tissue culture technology and the operation of FACS by Dr. Tan Suet Mien. Dr. Yoon Ho Sup, who we affectionately call Joe, helped me solve problems in protein expression and purification. When I started peptide synthesis, I received helpful instructions from Dr. Liu Chuanfa. Dr. Tang Kai provided me with the opportunity to operate the MALDI-TOF machine.

I also wish to express my appreciation to all the people in Professor Tam's group; Professor Liu Jianhua, from Prof. Tam's laboratory, for his patient help and guidance and Dr. Ma's, who was generous with himself and his time to help me with my experiments. Cai Lin, she is not only my colleague, but also my friend who offered unselfish help when I needed it, both with experiments, and life. Kelvin Poh offered me generous help and provided the necessary experiment materials; Dr. Wang Wei shared valuable information with me concerning experiments and I am deeply influenced by his hard work, dedication and spirit regarding scientific research. I also want to thank the following people from Prof. Tam's laboratory: Dr. Eom Khee Dong, Dr. Lu Yanning, Tin Twin, Goth Inn Fern, Leovy Francisca, Nguyen Kien Truc Giang, Tuan Lin, and Clarence.

I worked closely with Dr. Liu Dingxiang and would like to acknowledge his laboratory members here: Chen Bo, Frank Li Qisheng, Liao Ying, Wang Jibin, Wang Li and Yuan Quan. I deeply appreciated their constant help and friendship over these last five years. I also wish to thank Dr. Liu Bin and Dr. Fengling for their help. They are not only my classmates, but I would also describe them as my big brother and sister.

My family is the source of all my strength. Without them, I could not have achieved my goal. They have always encouraged me to do my best. I thank them for their emotional support and encouragement throughout this long process. My parents supported me more than anyone else, in the past twenty-eight years. I cannot imagine how I would have reached this point, without that support: They trusted me, supported me, and encouraged me countless times. It saddens me that they are so far away and could not be here with me to share in this happy and glorious moment. Still, they gladly take the disappointments and loneliness from me and keep them close to their hearts. I could not ask for better parents. There are no words to describe how thankful and grateful I am towards them, when I look back over these years. I also wish to thank my aunt, Ji Li, for her mother like care and consideration she gave to me. I wish to thank my girl friend, Chen Ying, for all the support, patience, and encouragement she gave to me during the past one-thousand days. I could not have achieved this, without her.

Finally, I wish a heartfelt thank you to all the members of my thesis committee, for their valuable comments and suggestions.

Abstract

The outbreak of severe acute respiratory syndrome (SARS) in 2002 and 2003 was caused by a novel and lethal virus, SARS-CoV. The existence of animal reservoirs of SARS-CoV and the threat of re-emergence of the disease intensified the urgency for clarifying its viral entry mechanism and searching for effective antiviral reagents. As one of the coronaviruses, its viral entry process is mediated by the S protein. Here, I present our study on SARS-CoV S protein and a potential approach for designing SARS-CoV viral entry inhibitors. Our studies, detailed in later chapters, furthered the understanding of the S protein mediated viral entry pathway in its early stage and shed light on possible countermeasures. Initially, we studied the S protein by identifying and characterizing its heptad repeat region and determining the potential role of pH during the viral fusion process. The two heptad repeat regions, corresponded to residues 900-1005 and 1151-1185, self-assembled into oligomers. Their oligomerization states were affected by pH. Peptide analogs derived from these regions were able to pull down the S protein, indicating the interaction between heptad repeat regions and the S protein. The interaction between the two heptad repeat regions was further confirmed in our study. To investigate the S protein mediated viral entry, the SARS-CoV receptor ACE2 expressing cell line was setup as a target for testing viral entry. To avoid handling the living virus, the S pseudotyped HIV virus was applied in this study. The SARS-CoV pseudovirus study showed that the viral entry is pH dependent. Through the S protein mediated membrane fusion assay, we found that the protease treatment enhanced the S protein mediated membrane fusion. It was found that the pH has little effect to this process. Based on these studies

and applying the design of fusion inhibitor on HIV-1, we proposed that protein mimetics can be produced to block the SARS-CoV viral entry base on the heptad repeat region sequences. Peptide analogs and protein mimetics of heptad repeat regions were developed and tested as inhibitors of the SARS-CoV infection. Peptides analogs showed the viral entry inhibitory effect in a dose-depended manner and blocked the viral entry at micromole grade. Summarizing our research findings and the most recent progresses, we proposed that the receptor binding and protease cleavage activation are the two essential factors for the early stage of the SARS-CoV viral entry; the low pH would facilitate the viral entry process, but was not necessary. The antiviral activity of the peptides and protein mimetics tested in our study provided an attractive starting point for the development of new fusion inhibitors for the SARS-CoV.

Table of Contents

List of Figures

Abbreviation List

6HB	six-helix bundle
aa	amino acid
ACE2	angiotensin-converting enzyme 2
ACN	acetonitrile
AMLV	amphotropic murine leukemia virus
APN	aminopeptidase N
BCoV	bovine coronavirus
BSA	bovine serum albumin
C-terminal	Carboxyl terminal
C/Cys	Cysteine
CD	circular dichroism
CHR2/Cys-HR2/CL36	peptide on SARS-CoV S protein amino acid 1151-1185 with N-terminal cysteine
CL26	peptide on SARS-CoV S protein amino acid 1161-1185 with N-terminal cysteine
CoV	coronavirus
Da	Dalton
DCM	Dichloromethane
DEAE	Diethylaminoethyl
DIEA	Diisopropylethylamine
DMEM	Dulbecco's Modified Eagle's Medium
DMF	Dimethylformamide
EDT	1,2-ethanedithiol
EDTA	ethylenediaminetetraacetic acid
EGS	ethylene glycol bis (succinimidyl succinate)
EM	electron microscopy
ESI-MS	electrospray ionisation mass spectrometry
FDA	Food and Drug Administration

FIPV	feline infectious peritonitis virus
FITC	fluorescein isothiocyanate
Fmoc	9-fluorenylmethoxy carbonyl
GP	Glycoprotein
HCoV	human coronavirus
HE	hemaglutinin esterase
HIV	human immunodeficiency virus
HPLC	high performance liquid chromatography
HR	heptad repeat
HR1/2	HR1 or HR2
HR1+2	HR1 and HR2 mixture
HR1	peptide on SARS-CoV S protein amino acid 900-1005
HR1C	peptide on SARS-CoV S protein amino acid 900-949
HR2 / IL35	peptide on SARS-CoV S protein amino acid 1151-1185
HRP	horseradish peroxidase
IBV	infectious bronchitis virus
IL25	peptide on SARS-CoV S protein amino acid 1161-1185
IL29	peptide on SARS-CoV S protein amino acid 1151-1179
IPTG	isopropyl-β-D-thiogalactopyranoside
GST	glutathione S-transferase
kb	kilo base pairs
kDa	kilo Dalton
LB	Luria Broth
M.W.	molecular weight
MALDI	matrix-assisted laser desorption ionization
MBP	maltose binding protein
MHV	mouse hepatitis virus
MS	mass spectrometry
Mtt	methyltrityl
N-terminal	aminoterminal, NH_2-terminal

NA	Neuraminidase
NCBI	National Center for Biotechnology Information
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PBS-T	phosphate buffered saline with 0.1% Tween 20
PCR	polymerase chain reaction
PDB	protein data bank
PL	pallet from the cell lysate
PVDF	polyvinylidene fluoride
RBC	red blood cell
RP	reverse phase
RPM	rotate per minute
SARS	severe acute respiratory syndrome
SARS-CoV	severe acute respiratory syndrome coronavirus
SDS	sodium dodecyl sulfate
SEC	size-exclusion chromatography
SP	supernatant of the total cell lysate
tBu	tert-butyl
TCP	total cell protein
TEMED	Tetramethylethylenediamine
TFA	trifluoroacetic acid
TGEV	transmissible gastroenteritis coronavirus
TIS	Triisopropylsilane
TPCK	L-1-tosylamido-2-phenylethyl chloromethyl ketone
TRS	transcription regulatory sequence
v:v	volume : volume
VSV	vesicular stomatitis virus

Chapter 1

General Introduction

General Introduction

The Coronavirus

The coronaviruses (CoV), a genus of *Coronaviridae* family, are large, enveloped, positive-strand RNA viruses that cause highly prevalent diseases in humans and domestic animals. Their viral RNA genome is between 27 to 32 kb, which is the largest genomes of all RNA viruses. It replicates by a unique mechanism, which results in a high frequency of recombination. The virions mature by budding at intracellular membranes, and infection of some coronaviruses inducing cell fusion.

The coronaviruses were first recognized as a distinct virus group by their characteristic virion morphology in negatively stained preparations (D.A.J.Tyrrell. 1965; F.A. Murphy 1995). The viral envelopes are studded with long, petal-shaped spike (S) glycoproteins (Cavanagh 1995; Lai 2001), giving coronaviruses the unique appearance of a crown (Latin, *corona*), from which their name is originated.

The *Coronaviridae* family members have a broad host range, including numerous mammalian and avian species. But, most coronaviruses naturally infect only one animal species or a limited number of closely related species. The narrow host range is thought due to the viruses' ability to interact with species-specific receptors (Michael M. C. Lai and 1996).

The *Coronaviridae* family is divided into three distinct groups (Table 1) according to their nucleotide sequences, host type, and serologic relationships. Group 1 and 2 include both human and animal coronavirus. Group 3 includes the avian coronavirus (Spaan, Cavanagh et al. 1988; Siddell 1995). Before the outbreak of SARS, the human coronavirus, HCoV-229E (group 1) and human HCoV-OC43 (group 2) were the only two known coronaviruses which can cause upper respiratory

tract infections in humans (Navas-Martin and Weiss 2004). About thirty percent of the common colds reported are caused by these human viruses, but no serious illness has been reported until now. In murine, porcine, feline, bovine and avian species, coronavirus can cause respiratory, enteric or neurological infections, which can result in a significant economic loss in the livestock and poultry industry.

Coronavirus Genome Organization, Replication and Transcription

Using viral RNA genome as a template, the coronavirus viral mRNA is subsequently transcribed. The RNA is capped at the 5' end and polyadenylated at its 3' end. The untranslated regions (UTR) at the 5' end and 3' end harbor cis-acting and structural elements essential for viral replication and transcription. The coronavirus contains a genome with remarkably conserved organizations (Figure 1-1).

Group	Virus	Host	Disease
1	Human coronavirus (HCoV)-229E	Human	Respiratory infection
	Transmissible gastroenteritis virus (TGEV)	Pig	Respiratory / enteric infection
	Porcine respiratory coronavirus (PRCoV)	Pig	Respiratory / enteric infection
	Canine coronavirus (CCoV)	Dog	Enteric infection
	Feline coronavirus (FCoV)	Cat	Enteric infection
	Rabbit coronavirus (RbCoV)	Rabbit	Enteric infection
2	Human coronavirus (HCoV)-OC43	Human	Respiratory infection
	Murine hepatitis virus (MHV)	Mouse	Respiratory / enteric / neurological infection, hepatitis
	Porcine hemagglutinating encephalomyelitis virus (HEV)	Pig	Respiratory / enteric infection
	Bovine coronavirus (BCoV)	Cow	Respiratory / enteric infection
3	Turkey coronavirus (TCoV)	Turkey	Respiratory / enteric infection
	Avian infectious bronchitis virus (IBV)	Chicken	Respiratory infection, hepatitis
?	Severe acute respiratory-associated coronavirus (SARS-CoV)	human	Respiratory infection (SARS)

Table 1. Coronavirus Groups

Coronaviruses can be divided into three groups based on their serological and genetic properties. Within each group, the viruses are classified according to their natural hosts, nucleotide sequences, and serologic relationships. The main representatives, hosts and associated diseases from each group, are listed in the table. With accumulated experiment evidence, the newly identified SARS-CoV is prone to be classified in the group 2 coronavirus.

Figure 1-1

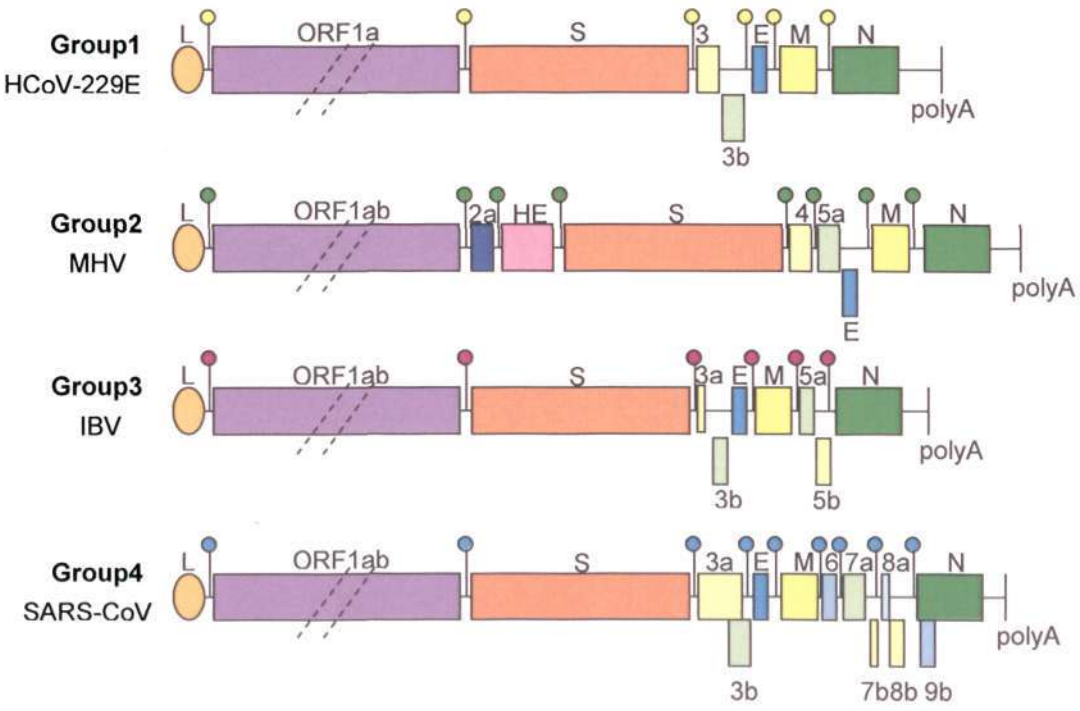


Figure 1-1. Genomic Organization of Coronaviruses

The genomic structure of coronavirus representatives from group 1 to 3 and the SARS-associated coronavirus (SARS-CoV) are shown. Group 1: human coronavirus 229E, (HCoV-229E); group 2: mouse hepatitis virus, (MHV); group 3: avian infectious bronchitis virus, (IBV). Except for gene 1, the sizes of the genes are drawn about to scale. The 5′ end of the RNA genome consists of a cap and a 65- to 98-base long leader sequence (L). The coronavirus genomes are polyadenylated (polyA) at the 3′ end. The circles of different colors represent the group-specific transcription regulatory sequences (TRS).

From the 5' end, the pol1ab gene takes about two-thirds of the viral genome. It encodes two precursors (pp1a and pp1ab). The po1ab is transcribed directly from the genome, while the pol1b gene is only translated through a -1 ribosome frameshift regulated by a pseudoknot structural element. The remaining part of the genome contains the S, E, M and N genes, respectively. Through a discontinuous transcription process, they are generated as the nested set of subgenomic RNAs, with a leader sequence at the 5' end. Following the leader sequence, there exists a transcription regulatory sequence (TRS). In most cases of the discontinuous negative strand RNA synthesis, the TRS serves as signals for the generation of subgenomic RNAs. The existence of subgenomic-size negative RNA strands containing anti-leader sequences within the cell has been confirmed. The negative stand subgenomic RNAs supposedly serve as templates for the positive strands subgenomic RNAs synthesis.

Except for the essential structural genes, one-third of the genome at 3' end contains several non-essential genes. They are called group specific genes because their identity and position is similar within the coronavirus groups, but changes between groups. The functions of these genes are currently unknown, but the studies on those genes show that they are dispensable for viral infectivity (Fischer, Peng et al. 1997; de Haan, Masters et al. 2002; Ortego, Sola et al. 2003; Haijema, Volders et al. 2004). The group 2 coronaviruses express the structural HE glycoprotein and the membrane-associated I protein. The latter one is translated from an internal open reading frame within the N gene (Senanayake, Hofmann et al. 1992; Fischer, Peng et al. 1997).

Coronavirus Structural Proteins

Following the mRNA transcription, the viral structural proteins are produced by translation. It involves cap-dependent protein synthesis mechanisms in the host cell cytoplasm. RNA synthesis also occurs in cytoplasm using negative-strand RNA intermediate. After the RNA and protein synthesis, the nucleocapsid protein associated with viral RNA genome to form the viral nucleocapsid. Further assembled newly produced virions are budded to the intracellular membranes to obtain viral glycoproteins (Figure 1-2). Then, they are released to the cell through secretary pathway (Siddell 1995; Michael M. C. Lai and 1996; Lai 2001).

As the replication order, from 5' to 3', the coronavirus genes are arranged as hemagglutinin esterase (HE) glycoprotein, (Group 2 only), spike (S) protein, envelope (E) glycoprotein, membrane (M) glycoprotein and nucleocapsid (N) phosphoprotein (Spaan, Cavanagh et al. 1988). There also exist interspersed non-conserved open reading frames (ORF). Their locations and numbers vary between different *Coronaviridae* family members (Holmes and Enjuanes 2003). The function for those non-conserved ORF still needs to be determined.

Figure 1-2

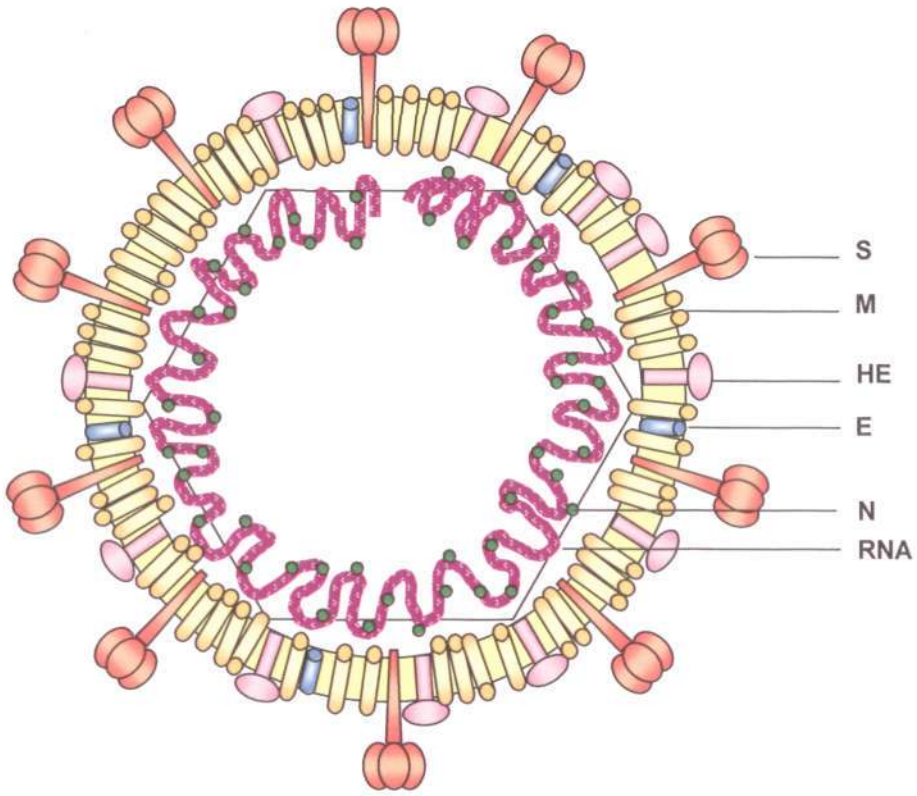


Figure 1-2. Schematic Diagram of the Coronavirus Virion Structure

Virions are spherical enveloped particles about 100 to 120 nm in diameter. Inside the virion is a single-stranded, positive-sense genomic RNA, 27 to 32 kb in size. The nucleocapsid protein (N) is associated with the RNA genome to form a long, flexible, and helical nucleocapsid. The nucleocapsid is surrounded by a lipid bilayer containing the spike protein (S), the membrane glycoprotein (M) and the envelope protein (E). Some group 2 coronaviruses contain a fourth membrane protein, the hemaglutinin esterase (HE) in their lipid envelope.

N protein

The N protein is the most abundantly expressed viral protein in infected cells. It ranges from 377-455 residues in length and has a molecular weight varying between 45 and 60 kDa. Consistent with its role as an RNA binding protein, it is abundant of arginine and lysine residues therefore it is highly basic (pI 10.3-10.7). The nucleocapsid protein is the only coronavirus structural protein known to be phosphorylated, but the exact function of this modification remains unknown. The overrepresented serine residues (7-11 %) are the potential targets for phosphorylation. Three structural domains in the N protein have been identified. The functions of the N-terminal conserved domains are not yet clear. The middle domain is an RNA-binding domain. The C-terminal domain may be involved in self-association to form the oligomers. The primary function of the N protein is the formation of the viral ribonucleoprotein complex, but the protein also plays a facilitating role in coronavirus replication (Laude 1995).

M protein

The M protein is the majority protein of the coronavirus membrane. It is a class III integral membrane protein around 25 kDa. It is a highly hydrophobic glycoprotein, with only a short aminoterminal domain of M exposed on the exterior of the viral envelope. Glycosylation of the aminoterminal domain is either O-linked (group 2) or N-linked (group 1 and 3) (Holmes, Doller et al. 1981; Niemann, Geyer et al. 1984; Deregt, Sabara et al. 1987; Laude, Rasschaert et al. 1987). Following the aminoterminal domain, a triple-membrane–spanning domain makes it across the membrane three times. After that, there is an α-helical domain, and a large

carboxylterminal domain inside the viral envelope (Armstrong, Niemann et al. 1984; Routledge, Stauber et al. 1991; Locker, Rose et al. 1992; Machamer, Grim et al. 1993). The M protein interacts with each other, through multiple contact sites, especially at the transmembrane level (de Haan, Vennema et al. 2000). Interactions between the M protein molecules are thought to be the driving force in coronavirus envelope assembly (de Haan, Kuo et al. 1998). In addition, the protein's carboxyl terminus is associated with the virion core and is believed to function in the maintenance of the core structure (Escors, Camafeita et al. 2001; Escors, Ortego et al. 2001).

E Protein

The E protein is a small glycoprotein about 76 to 109 amino acids (Siddell 1995). It is low in abundance within the virion envelope. However, it plays an essential role in the assembly of the coronavirus. Like the M protein, it is highly hydrophobic. It contains a large hydrophobic domain in the N-terminal, followed by a cysteine-rich domain and a short hydrophilic tail. Topologically, the E protein is buried in the membrane and spans the membrane one or two times. Its C-terminal is exposed outside of the lipid bilayers (Raamsman, Locker et al. 2000). Electron microscopy (EM) shows that the E protein is supposed to induce the membrane curvature during budding or formation of the rounded, stable and infectious particles (Fosmire, Hwang et al. 1992; Vennema, Godeke et al. 1996; Raamsman, Locker et al. 2000).

HE Protein

The HE protein is found as a glycoprotein in viral membrane of the group 2 coronavirus, including HCoV-229E, murine hepatitis virus (MHV), bovine

coronavirus (BCV) and hemagglutinating encephalomyelitis virus (HEV) (Siddell 1995). It's about 424 to 429 amino acids in length, and exists as a dimer in the viral membrane (Lai 2001). The HE protein binds to 9-O-acetyl groups from the sialic acid-containing cellular receptor glycoproteins and cleaves the acetyl group from the neuraminic acid (Liao, Zhang et al. 1995; Regl, Kaser et al. 1999). The role of HE is currently under study, but it is not required for viral infectivity. It may play a role in viral pathogenicity (Yokomori, Baker et al. 1992).

S Protein

The coronavirus S protein is a large, heavily glycosylated, trimerized membrane protein. It is responsible for virus-cell attachment, virus-cell and cell-cell fusion, and is the major determinant of host-range, pathogenesis and virulence. Under electron microscopes, it can be easily observed as large, spike-shaped structures on the surface of the virion, which are approximately 20 nm long (Xu, Liu et al. 2004) and 10 nm wide (Cavanagh 1995). Because of this, in some cases, it is named spike protein. It is a type I membrane protein of 1160 to 1452 amino acids in length and 170 to 220 kDa in size. The S protein is synthesized as a precursor with an N-terminal signal sequence that is cleaved upon translocation and a transmembrane domain followed by a short hydrophilic cytoplasmic tail at its C-terminal. The cytoplasmic tail contains a cysteine-rich region.

Usually, the S protein contains 21 (MHV) to 35 (FIPV) potential N-glycosylation sites. The protein is only N-glycosylated, as the oligosaccharides can be removed biochemically using specific glycosidase (Cavanagh 1995). Glycosylation is believed to facilitate proper folding of the spike protein during biosynthesis. The spike protein can oligomerize to trimers (Eckert and Kim 2001), which is thought to

be needed for transport out of the ER. The S protein can induce cell-cell fusion, once it is transported to the plasma membrane (Vennema, Rottier et al. 1990; Opstelten, Raamsman et al. 1995). In group 2 and 3 coronavirus, the spike protein is cleaved, probably by cellular proteases, into a soluble S1 subunit and a membrane-anchored S2 subunit. The S1 subunit represents the globular head, while the S2 subunit forms the stalk-like region of the spike (de Groot, Luytjes et al. 1987). S1 and S2 are associated together by non-covalent interactions (Cavanagh 1995). The S1 subunit mediates the receptor binding (Kubo, Yamada et al. 1994), while the S2 subunit is responsible for viral and cellular membrane fusion (Yoo, Parker et al. 1991; Taguchi 1995). The S2 subunit contains two so-called heptad repeat (HR) regions (de Groot, Luytjes et al. 1987), which have a strong propensity to form coiled-coils and is attributed to the stalk-like structure of the S2 subunit. The HR regions are implicated in oligomerization and could well play a role in membrane fusion, as it is observed in other viral membrane fusion proteins (Chan and Kim 1998; Weissenhorn, Dessen et al. 1999).

Severe acute respiratory syndrome (SARS)

In November of 2002, an atypical pneumonia was first reported in Guangdong Province, China. Outbreak of this respiratory disease, in that area, rapidly leads to the worldwide spread and eight thousand people were infected in twenty-nine countries and the death toll from the virus reached 774 (Goldsmith, Tatti et al. 2004). The new emerging infectious disease, known as SARS, is identified to be caused by a newly emerged coronavirus (Drosten, Gunther et al. 2003; Ksiazek, Erdman et al. 2003; Peiris, Lai et al. 2003). It is named SARS-associated coronavirus (SARS-CoV) eventually.

The SARS-CoV infection results in an atypical pneumonia, causing pulmonary destruction in some people. Although the exact mechanism of death caused by SARS-CoV infection remains unclear, a disease model is proposed that consists of three phases: (1) viral replication, (2) immune hyperactivity and (3) pulmonary destruction (Tsui, Kwok et al. 2003). The hyperactive immune response to the SARS-CoV infection is believed to contribute to the advanced stages of SARS.

Severe Acute Respiratory Syndrome-associated Coronavirus (SARS-CoV)

The severe acute respiratory syndrome coronavirus (SARS-CoV) is the causative agent of SARS (Ksiazek, Erdman et al. 2003; Rota, Oberste et al. 2003). This virus is an enveloped, positive-strand RNA coronavirus. Its complete RNA genome contains 29,727 nucleotides and it only shares less than seventy-percent sequence similarity with the other coronaviruses. Different strands are isolated and show little or no variation in their genome sequence, which suggests that the viral genome is rather stable during human passage. Compared to the other family members of *Coronaviridae*, it shows a high level of apparent difference. Its polymerase sequence is similar to the MHV and the BCV (group 2) and the avian infectious bronchitis virus (IBV) (group 3). Like group 3 coronavirus, the SARS-CoV encodes a single papain-like protease. Unlike the group 2 coronavirus, it has no HE gene. Those features cause it to be listed in a new group of coronaviruses, distinguished from the existing three known groups. Recent studies tend to classify it into a subgroup within the group 2 coronavirus (Gorbalenya, Snijder et al. 2004; Magiorkinis, Magiorkinis et al. 2004). Until now, the data shows it is not originated from the mutated human coronavirus or recombination between other coronavirus. The SARS-CoV-like viruses are isolated from palm civets (*Payuma iaravta*), raccoon

dogs (*Nyctereutes procynonoides*) and Chinese ferret badgers (*Melogale moschata*) (Guan, Zheng et al. 2003). Furthermore, it is reported that ferrets and domestic cats are susceptible to SARS-CoV infection and capable of transmitting the virus to uninfected animals (Martina, Haagmans et al. 2003). Therefore, the SARS-CoV most likely comes from animals.

As the other coronaviruses, the spike protein of SARS-CoV mediates viral and target cell membrane fusion. Understanding this process can benefit the development of therapeutics to prevent SARS-CoV entry and replication. Previous studies suggested that CoV S proteins may represent a unique type of viral fusion proteins. More recent studies indicated that the SARS-CoV S proteins shares structural features with class I viral fusion proteins (Chambers, Pringle et al. 1990; Garry 2003). The position of the two heptad repeat regions were mapped to be the residue of 900 to 1005 (HR1) and 1151 to 1185 (HR2), respectively. In addition, it showed that peptides analogs to these regions are capable of self-assembling into anti-parallel oligomeric complexes. Formation of the complexes resembles the classic six-helix bundle (6HB) structure representative of the class I viral fusion proteins (Zhu, Xiao et al. 2004).

Viral Fusion Protein

The fusion glycoprotein of enveloped viruses from different families shares several similar features (Figure 1-3). Such as orthomyxoviruses, paramyxoviruses, retroviruses, arenaviruses and filoviruses, which are all encoded as class I integral membrane proteins, containing a cleavable signal sequence at the N-terminal, a large extracellular domain, a transmembrane domain followed by a C-terminal cytoplasmic tail. Usually, these fusion proteins are generated as a precursor, oligomerize into

trimers, and are then cleaved into two closely associated subunits. These fusion proteins contain: (1) a short hydrophobic region within their sequence, called the fusion peptide. The fusion peptide interacts with the host cell membrane at an early stage of the membrane-fusion process; (2) two α-helices, specifically leonine zipper-like motifs, heptad repeat (HR) regions, one (HR1) located downstream of the fusion peptide and the other (HR2) usually located upstream of the transmembrane region; (3) a cluster of aromatic amino acids proximal to the transmembrane domain (Wilson, Skehel et al. 1981; Chambers, Pringle et al. 1990; Carr and Kim 1993; Suarez, Gallaher et al. 2000). Those fusion proteins with these conserved structure features, are classified as Class I viral fusion proteins, which are distinguished from class II viral fusion protein. Class II viral fusion proteins contain an internal fusion peptide and are composed of three anti-parallel β-sheet domains, instead of helical secondary structures (Heinz and Allison 2001; Garry and Dash 2003; Heinz and Allison 2003). Besides, many of these proteins are generated as a complex with a second membrane glycoprotein. The second associated glycoprotein is necessary for the fusogenic potential of the fusion protein.

Figure 1-3

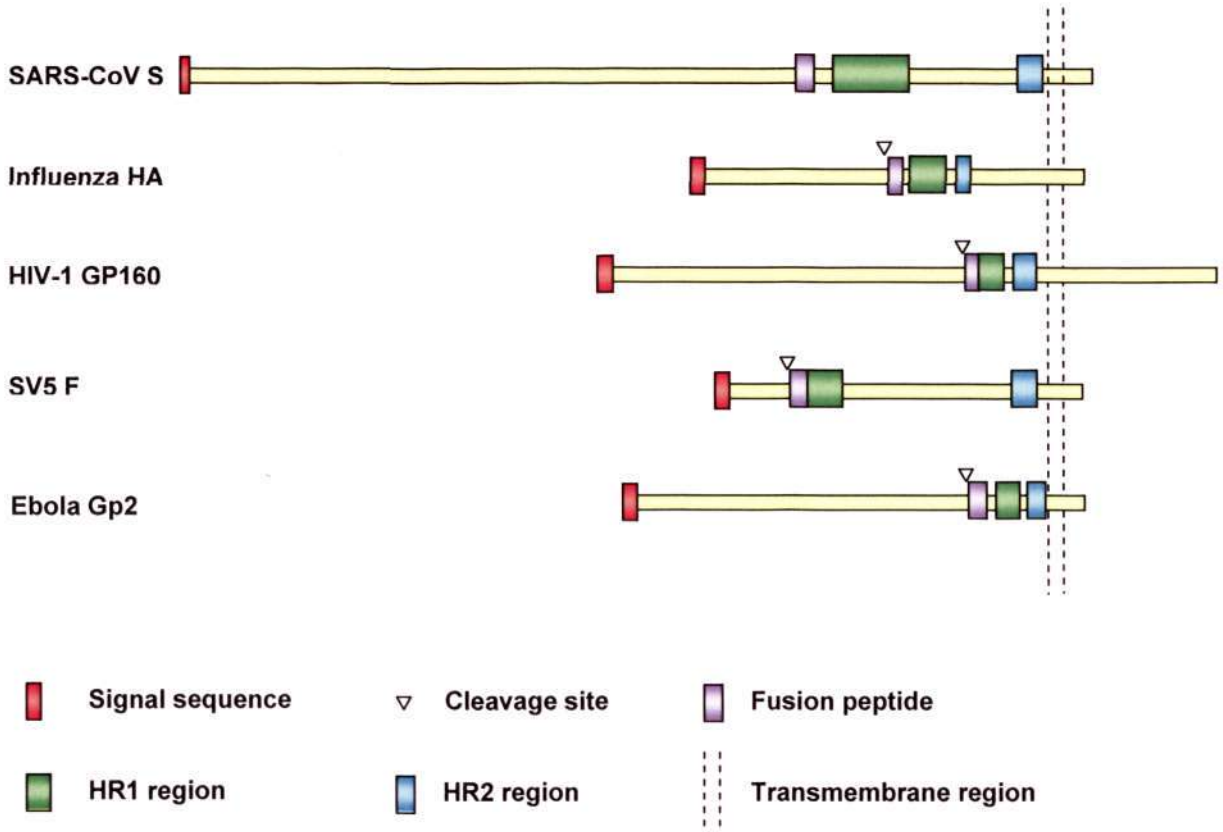


Figure 1-3. Class I Viral Fusion Proteins

Viral fusion proteins from five different virus families are shown in the diagram: SARS-CoV S (*Coronaviridae*), influenza HA (*Orthomyxoviridae*), HIV-1 gp160 (*Retroviridae*), SV5 F (*Paramyxoviridae*) and Ebola Gp2 (*Filoviridae*). Class I viral envelope fusion glycoproteins are synthesized as single-chain precursors, which then are assembled into trimers. The single-chain precursors contain a signal sequence, a proteolytic cleavage site, a hydrophobic sequence that is known as the fusion peptide, two heptad repeat regions, and a C-terminal transmembrane domain followed by a cytoplasmic tail. (The diagram is drawn according to the information provided by (Bosch, van der Zee et al. 2003))

The similarities in class I viral fusion proteins of these enveloped viruses lead to a common membrane fusion model (Figure. 1-4). During the membrane fusion process, the fusion proteins experience a series of conformational changes. Cleavage of the precursor leaves the fusion protein in a metastable state. Receptor binding or protonation after endocytosis triggers conformational changes of the fusion protein, causing the expose of the hidden fusion peptide towards the target host cell membrane. Insertion of the fusion peptide into the target cellular membrane, while the transmembrane domain is still anchored in the viral membrane, forms a pre-hairpin intermediate. This process is followed by a structural reorganization of the fusion protein, in which the interaction between two heptad repeat regions plays a critical role. Ultimately, the separated HR1 and HR2 fold together and assemble into a highly stable six-helix bundle structure, with the HR1 and HR2 helices oriented in an anti-parallel manner. In this conformation, the fusion peptide inserted in the target membrane, is brought into close proximity with the transmembrane region of the protein. The free energy change associated with this structural transition of the fusion protein is thought to be sufficient enough to cause lipid mixing and viral and cellular membrane fusion. The S protein of the coronavirus shares similar structural features with the fusion proteins of orthomyxoviruses, paramyxoviruses, retroviruses, arenaviruses and filoviruses. This indicates that a similar entry mechanism is applied to the coronavirus during the viral entry process (Bosch, van der Zee et al. 2003; Zhu, Xiao et al. 2004).

Figure 1-4

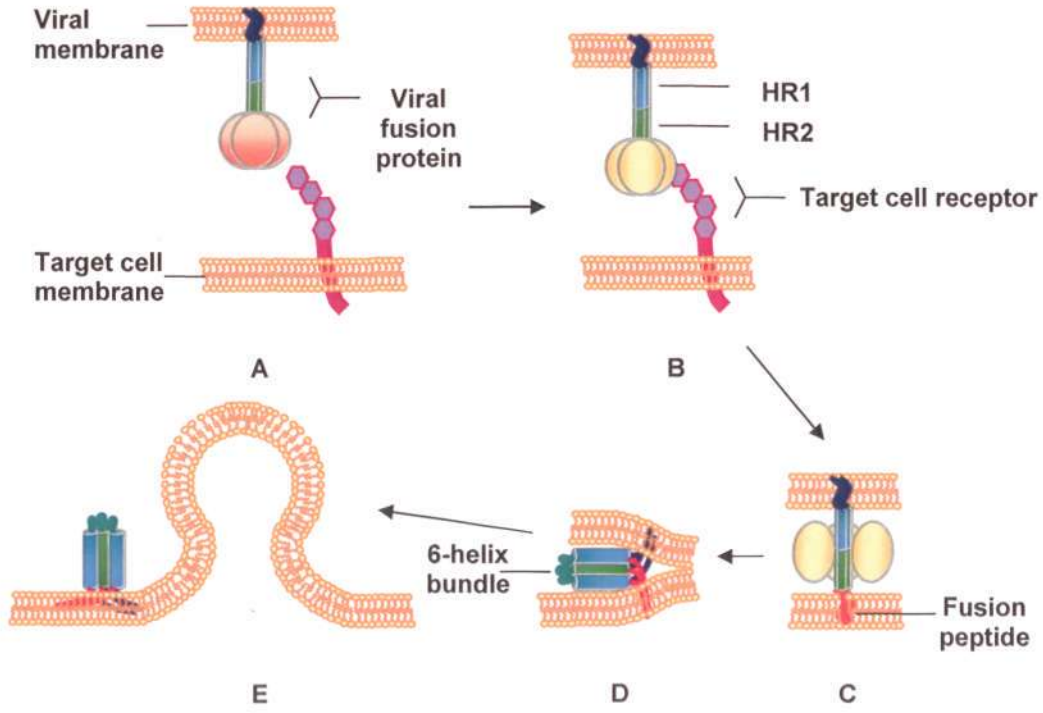


Figure 1-4. A Model Pathway of Class I Viral Fusion Protein Mediated Viral and Cellular Membrane Fusion

A. Class I viral fusion proteins are synthesized as single-chain precursors, but exist as trimers on the virion with their C-terminals anchoring in the viral membrane. B. The viral fusion process starts with the viral envelope protein attached to the receptor on the target cell surface. C. The receptor binding or protonation after endocytosis, triggers conformational changes of the fusion proteins, causing the exposure of the hidden fusion peptides towards the target host cell membrane. Insertion of the fusion peptides into the target cellular membrane while the transmembrane domains are still anchored in the viral membrane form the pre-hairpin intermediates. D. The two separated HR1 and HR2 on each trimer fold together and assemble into a highly stable six-helix bundle structure, with the HR1 and HR2 helices oriented in an anti-parallel manner. In this conformation, the fusion peptides inserted in the target membrane is brought into close proximity with the transmembrane region of the protein. E. The free energy change associated with this structural transition of the fusion protein is thought to be sufficient to cause lipid mixing and viral and cellular membrane fusion. The pathway is drawn according to our understanding to viral entry process based on model virus (Colman and Lawrence 2003; Harrison 2005).

The virus infection starts with the attachment of the virus to the cellular receptor. Following receptor binding, the viral fusion protein mediates the viral membrane fusion, with either the plasma membrane or the endosomal membrane. For fusion at the plasma membrane, the attachment of the virus with its receptor generates sufficient energy to trigger membrane fusion, while the viruses fusing with the endosomal membrane are usually triggered by the protonation of the fusion protein during acidification of the endosome. Using retroviruses (HIV-1) as an example, its membrane fusion occurs at the plasma membrane. Plasma membrane fusion is a pH independent process and it is triggered by receptor binding. HIV-1 receptor binding and membrane fusion functions are combined in the envelope fusion glycoprotein. The receptor binding domain (gp120) can bind to its primary receptor CD4 and then bind to its co-receptor. Primary receptor binding arouse the conformational change of gp120. Its hidden co-receptor binding site is exposed as a consequence. The CD4 binding also induces conformational changes in the envelop protein which leads to detachment of the gp120 subunit from the membrane fusion subunit gp41. This process activates the membrane fusion subunit gp41. Using the influenza virus as another example, it's HA protein mediates the attachment of the influenza virus to its sialic acid-containing receptors on the cell surface through weak interactions. The virus then enters the cell through endocytosis process. Acidification inside the endosomes results in a conformational change of the HA protein because of specific protonation, which triggers membrane fusion (Lamb 2001).

The coronavirus binds to specific cellular receptors for viral entry. Some of the receptors have been found. In group 1, the TGEV, FIPV and HCoV-229E use the cell membrane bound metalloprotease and aminopeptidase N (APN) of their host species. In group 1, MHV use murine biliary glycoprotein, which belongs to the

carcinoembryonic antigen family, as its receptor. It is reported that the HCoV-OC43, BCoV, TGEV and IBV (Schultze, Gross et al. 1991; Schultze, Cavanagh et al. 1992; Kunkel and Herrler 1993; Schultze, Krempl et al. 1996) bind 9-O-acetylated sialic acid-containing glycans and possess a hemagglutinating activity. They might use sialic acids as an additional receptor (Schwegmann-Wessels and Herrler 2006; Winter, Schwegmann-Wessels et al. 2006). Group 2 coronaviruses have an additional membrane glycoprotein, the hemaglutinin esterase, HE (Vennema, Godeke et al. 1996; de Haan, Vennema et al. 2000). Comparing to the NA protein of influenza A and B viruses, which have neuraminidase activity, the HE glycoprotein has an esterase activity that cleaves 9-O-acetylated sialic acid. So, it may function as a detachment protein.

The cellular receptor for SARS-CoV was originally found to be the metallopeptidase angiotensin-converting enzyme 2 (ACE2) (Li, Moore et al. 2003; Wang, Chen et al. 2004). Later, an additional receptor, CD209L (L-SIGN), is shown to function as a receptor for the SARS-CoV (Jeffers, Tusell et al. 2004).

Whether the coronavirus spike mediated membrane fusion is low pH dependent or independent is still under investigation. In a number of coronaviruses (For example, MHV-A59 and FIPV 79-1146), cell-cell fusion can occur at basic pH. The study on MHV-A95 virion show that the conformational changes of spike protein could occur at alkaline pH (Sturman, Ricard et al. 1990). Its virions maintain stable at acidic pH but rapidly lose their infectivity at alkaline (pH 8. 0). Basic pH 8.0 caused the release of S1 and aggregation of S2. This suggests that fusion of the viral and cellular membrane could take place at plasma membrane under normal physiological pH of the extracellular environment. Although, there also exist some viruses that use the endocytosis pathway for infection (Gallagher, Escarmis et al. 1991; Hansen,

Delmas et al. 1998), coronaviruses generally seem to enter cells at the plasma membrane.

Membrane fusion is an irreversible process during the whole process of viral entry. It is also composed of a serial of regulated timing steps, including cleavage activation by cellular proteases, receptor binding or low pH protonation. Proteolytic cleavage is essential for fusion activation of influenza viruses, retroviruses, filoviruses and paramyxoviruses. It converts the fusion protein into a metastable state. In contrast, the proteolytic cleavage in coronaviruses seems not as critical as it does in those viruses mentioned above. Cleavage does not occur in group 1 coronavirus S proteins, such as FIPV (Vennema, Heijnen et al. 1990). Group 2 and 3 coronavirus spikes are cleaved to different extents depending on the viral strain and the host cell type. The data indicates that cleavage of the MHV spike is important for cell-cell fusion, but is not required for virus-cell fusion (Cavanagh 1995).

Viral fusion proteins contain hydrophobic fusion peptides. These fusion peptides are inserted into the target cellular membrane during the initiation of membrane fusion. The location of the fusion peptide is ahead of the HR1 region and adjacent to the cleavage site. Cleavage of the fusion protein places the fusion peptide at the one end of the membrane anchored subunit. In coronaviruses, some of the spike proteins, which are cleaved, do not contain a hydrophobic sequence near the cleavage site, and the group 1 coronaviruses is not cleaved at all. It appears that the coronaviruses probably have an internal fusion peptide.

Heptad repeat regions in viral fusion proteins were show to play a critical role in the viral fusion process of orthomyxoviruses, retroviruses, filoviruses and paramyxoviruses (Skehel and Wiley 1998). For the coronaviruses, the defective mutant viruses in spike protein oligomerization and fusion (Luo, Matthews et al. 1999)

were found to have mutations in the HR2 region. Mutations located within the HR1 region are correlated with a low pH requirement for fusion (Gallagher, Escarmis et al. 1991). These suggest that the HR regions also play a critical role in the viral fusion process of the coronavirus. Although, there are many viral or cellular proteins or molecules involved in the entry process, many of these are potential targets for the viral fusion inhibitor. Applied the experience of viral inhibition study in the HIV-1 and the influenza virus, the HR regions of coronavirus are believed to be good targets for designing inhibitors to block the viral and membrane fusion and thus the viral entry process.

Fusion Inhibitors

Early studies show that the heptad repeat region derived peptides have anti-fusion activity (Wild, Oas et al. 1992; Wild, Shugars et al. 1994). These observations later extended to the homologous peptides of the F protein in the paramyxoviral virus (Rapaport, Ovadia et al. 1995; Lambert, Barney et al. 1996; Yao and Compans 1996). A double-layered rod structure was observed in both cases. The corresponding HR1 region of those viruses formed coiled-coils in the solution, but they are less potent in the viral inhibition comparing the HR2 region, which show loose structure in solution (Wild, Shugars et al. 1994; Lu, Blacklow et al. 1995; Lambert, Barney et al. 1996). Later studies demonstrate the inhibitory effect of the HR2 region originates from its ability to bind to the coiled-coil trimer HR1 region on the viral fusion protein, thus interfering with the folding of pre-hairpin intermediate during the viral fusion process, and blocks the formation of the six-helix bundle subsequently. Amino acid substitution study on the HR2 region, shows the correlation between the inhibitory potency and the stability of the corresponding six-helix bundle (Chan, Chutkowski et al. 1998) they formed.

The HR region-derived peptides, especially the HR2 region, serve as a pre-entry inhibitor which targets the pre-hairpin intermediate, showing a reasonable inhibition effect in the HIV-1. The 36 amino acids peptide (T20) derived from the HR2 of the HIV-1 gp41 are potent, with viral inhibition effect at nanomolar, which shows a reasonable success in human clinical trials and has been approved as the first HIV-1 entry drug (Kilby, Hopkins et al. 1998). However, T20 and the other entry inhibitory peptides are amphipathic and exhibit certain drawbacks that include proteolytic instability and the propensity for aggregation in aqueous solutions. To improve on these deficiencies, several laboratories have designed peptide mimetics of T20 to increase proteolytic stability (Eckert and Kim 2001; Louis, Bewley et al. 2001; Root, Kay et al. 2001; Sia and Kim 2003). Dimerized or trimerized peptides decrease the aggregation of its monomer form and significantly increase the inhibitory activity.

The coronavirus spike protein also belongs to the class I viral fusion protein, and some of them share a similar entry mechanism of the HIV-1. The device used for producing the HIV-1 fusion inhibitor can also be used for a reference to generate the SARS-CoV fusion inhibitor.

Specific Aims

The computational analysis of the SARS-CoV spike proteins demonstrate similarities to the HIV-1 gp41 (Garry 2003), a representative class I viral fusion protein. Later crystallization studies on the SARS-CoV spike protein fusion cores clearly show the formation of a six-helix coiled-coil bundle, similar to that of the other class I viral fusion proteins. The structure of the fusion core is consistent with earlier computer predictions and suggests the S protein belongs to the class I viral fusion protein.

Based on the similarities between the CoV S protein and that of class I viral fusion proteins (Figure 1-3), we hypothesize that the SARS-CoV S protein is a class I viral fusion protein. Since the HR regions play a critical role in the formation of the six-helix bundle, as well as to the viral and cellular membrane fusion process, further identification and characterization of these domains will increase our understanding of the S-mediated viral entry process. Thus, the overall objectives of this dissertation are: (1) to provide further information for understanding the SARS-CoV viral entry process; (2) to design peptide inhibitors of the SARS-CoV infection based on the approach used to produce fusion inhibitors in the other virus. These objectives are summarized to the following three specific aims:

Specific Aim 1: To identify and characterize the heptad repeat region of the SARS-CoV S protein.

Specific Aim 2: To characterize the SARS-CoV S protein mediated viral entry process.

Specific Aim 3: Using the HIV-1 gp41 as a model, design and develop protein mimetics based on the HR region as fusion inhibitors to block the SARS-CoV viral entry.

Chapter 2

Identification and Characterization of the Heptad Repeat Region of the SARS-CoV S Protein

Identification and Characterization of the Heptad Repeat Region of the SARS-CoV S Protein

Introduction

The accumulated evidence indicated that the SARS-CoV uses similar membrane fusion mechanisms for viral entry, as do the other enveloped virus encoding class I viral fusion proteins (Xu, Lou et al. 2004; Zhu, Xiao et al. 2004).

Class I viral fusion proteins, like the HIV-1 env, the influenza virus hemaglutinin, the paramyxovirus F protein, in coronavirus, murine hepatitis virus S protein, have common structural features. They are type I membrane glycoproteins and exist in trimer form in the viral membrane. They contain a protease cleavage site, a fusion peptide and at least two heptad repeat regions. Protease cleavage processes the viral fusion protein into two closely associated subunits, one is the receptor binding unit and one is the membrane anchored unit. The fusion peptide located at the N-terminal of the membrane anchored subunit. The two heptad repeat regions also exist in the membrane anchored subunit, which are designated HR1 and HR2 from the N-terminal to the C-terminal here. Usually the HR1 is adjacent to the fusion peptide and the HR2 is located in the contiguity of the transmembrane domain (Lescar, Roussel et al. 2001). The cellular protease cleavage events promote the viral fusion proteins into a metastable state. Receptor binding, or the protonation during endocytosis, triggers a series of conformational changes to mediate viral and cellular membrane fusion. Initially, the fusion peptide became exposed and then inserted into the cellular membrane. The C-terminal of the membrane anchored unit is still in the viral membrane while the fusion peptide at its N-terminal is inserted into the cellular

membrane, a pre-hairpin intermediate is formed. The transient is followed by the folding the HR2 region into the HR1 region to form a stable structure. This structure, a homo-trimeric HR1 coiled-coil surrounded by three HR2 helices in an anti-parallel way, is a six-helix bundle. The formation of the six-helix bundle leads to a close approximation of the fusion peptide with the transmembrane domain. It helps to pull the viral and cellular membrane together, facilitating membrane fusion (Eckert and Kim 2001).

In the coronavirus, the spike protein plays critical roles in the viral entry, tropism and virulence of the infection (Li, Li et al. 2005). It has several similarities with the class I viral fusion proteins. It is a trimerized type I membrane protein (Xu, Lou et al. 2004). The N-terminal half of the protein (S1) provides host specificity and functions for the attachment and receptor binding (Taguchi 1995; Bonavia, Zelus et al. 2003). The C-terminal half (S2) is membrane anchored and mainly responsible for virus-cell membrane fusion (Yoo, Parker et al. 1991). Similar to the other class I viral fusion proteins, the S protein also contain two Ile / Leu heptad repeat regions. Both regions are shown to possess the helical structure (Xu, Lou et al. 2004) in monomeric and trimeric form in protein prediction models or in synthesized peptides form. The HR region of the SARS-CoV spike protein is believed to play a critical role during the SARS-CoV viral entry process.

Figure 2-1

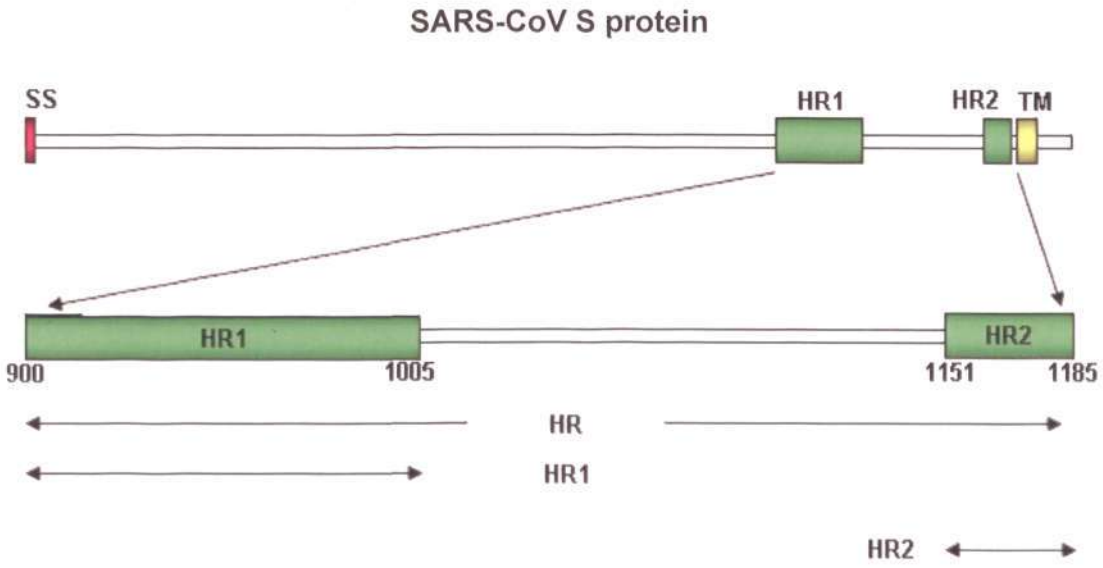


Figure 2-1. A schematic view of the S protein of SARS-CoV.

The locations of the heptad repeat regions (HR1 and HR2), transmembrane domain (TM) in SARS-CoV S protein are shown in the diagram. There is a signal sequence (SS) followed by a cleavage site at the start of the sequence. HR1 starts from amino acid 900 to 1005 and the HR2 from 1151 to 1185, which were selected according to the prediction results by LearnCoil-VMF program: http://groups.csail.mit.edu/cb/learncoil-vmf/cgi-bin/vmf.cgi (Singh, Berger et al. 1999). The diagram is draw according to the S protein sequence from SARS-CoV Singapore strand 2774.

Biochemical and conformational studies of the HR1 and HR2 of the SARS-CoV S protein may add to our understanding on the possible intermediate states of the fusion process and its fusion mechanism. Here we report that the pH is critical in the oligomerization of the HR regions, which indicate its impact in the conformational changes of the viral fusion core complex of the S protein. We expressed HR1 (aa 900 to 1005), HR2 (aa 1151 to 1185), Cys-HR2 (HR2 containing an additional N-terminal cysteine) and full-length HR (aa 900 to 1185) in *Escherichia coli*. Although, both HR1 and HR2 were able to form homo-oligomers, their oligomer states were disrupted by the acidic pH. The homo-oligomers were stable at the neutral or basic pH. HR1 and HR2 can also form the heteromeric oligomers in a 1:1 ratio. The hetero-oligomer states of HR1 and HR2 complexes were not affected by the pH in the conditions we tested. Through the CD study, we found that the acidic pH decreases the helical state of the HR1. Our findings suggested that the pH may be a crucial factor during the S protein fusion core formation. This result supported the experiment results that the SARS-CoV takes the pH dependent endocytosis pathway for viral entry. The maltose binding protein (MBP) fused with HR regions were able to pull down the S protein, which indicates that there exists interactions between the HR regions derived peptides and the S protein.

Materials and Methods

Plasmid Construction

The nucleotide fragments corresponding to amino acid residues 900 to 1005 (HR1), 1151 to 1185 (HR2) and 900 to 1185 (HR) of the SARS-CoV S protein were amplified by the PCR from the plasmid pJX40-S, which contains the SARS-CoV spike gene (gift from Institute of Molecular Cell Biology, Singapore). Primers used

are A (5'-CCG GAA TTC GAG AAC CAA AAA CAA A-3') and B (5'-AAA ACT GCA GTT AAT TAG CAG AAG CCC T-3') for HR1 (Amino acid sequence: ISEFENQKQIANQFNKAISQIQESLTTTSTALGKLQDVVNQNAQALNTLVKQL SSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRAS AN), C (5'-CCG GAA TTC ATT TCA GGC ATT AAC G-3') and D (5'-AAA ACT GCA GTT ACA ATT CTT GAA GGT C-3') for the HR2 (Amino acid sequence: ISEFISGINASVVNIQKEIDRLNEVAKNLNESLIDLQEL), and A and D for the HR. An upstream EcoRI site and a downstream PstI site, as well as a stop codon preceding the PstI site, were designed to insert the fragments into the pMAL-C2 vector (pMAL™ Protein Fusion and Purification System, New England BioLabs, UK). HR1, HR2 and HR were expressed as maltose binding protein (MBP) fusion proteins. After Factor Xa cleavage, it leaves four additional amino acids Ile-Ser-Glu-Phe at the N-terminal of those peptides.

Point mutation was performed by using a QuickChange® site-directed mutagenesis kit (Stratagene, La Jolla, CA) to introduce a cysteine residue at the N-terminal of HR2, producing the C-HR2 (Amino acid sequence: CISGINASVVNIQ KEIDRLNEVAKNLNESLIDLQEL). The mutagenic oligonucleotide primers were 5'-GGG ATC GAG GGA AGG TGC ATT TCA GGC ATT AAC G-3' and 5'-C GTT AAT GCC TGA AAT GCA CCT TCC CTC GAT CCC-3'. C-HR2 was expressed and purified similar to HR2, using the MBP-fusion protein strategy.

Expression and Purification of the Peptides and Proteins

The freshly transformed *E. coli* BL21 cells (Novagen, CA, U.S.) were grown in Luria Broth (LB) medium to log phase and subsequently induced by adding IPTG (isopropyl-β-D-thiogalactopyranoside, GIBCO BRL, U.S.) to a final concentration of

0.15 mM. Six hours later, the cells were pelleted, and re-suspended in 1/10 volume of the column buffer (20 mM Tris-HCl pH 7.4, 200 mM NaCl, 1mM EDTA), and sonicated on ice (20 minutes). The cell debris was removed by centrifugation and the supernatant was loaded onto an amylose resin column (New England BioLabs, U.K.). The protein-loaded column was then washed with 12 × column volume of the column buffer and eluted with 10 mM maltose. Fractions containing the MBP-fusion protein were identified using the SDS-PAGE. These fractions were pooled and proteolyzed with Factor Xa (Qiagen, Japan) at 37°C.

For the first purification step of the HR1, HR2 and Cys-HR2, methanol was added to the Factor Xa digested protein samples at a ratio of 1:1 (v:v). After precipitation of the MBP and other byproducts, the samples are centrifuged (4,000 × g; 20 minutes; 4°C) and the supernatant dried by speed vacuum (SPD2010 SpeedVac® system, Thermo Savant). Samples were reconstituted with double-distilled H_2O and target proteins were further purified using reversed-phase high pressure liquid chromatography (RP-HPLC).

The HR samples were purified following the Factor Xa digestion. The HR region was first cleaved from the MBP-fusion protein, and the MBP was removed using an amylose resin column. The flow through contains little uncleaved MBP-fusion protein and cleaved MBP and HR region. The HR region is separated from the mixture by size exclusion chromatography (SEC).

Expression and purification of SARS-CoV S protein ectodomain.

The nucleotide fragment corresponding to 1-1188 amino acid residues of SARS-CoV S protein were amplified with a HA tag at the C-terminus (S1188HA).

Primers used were A (5'-AGTCGAATTCCGAAC ATGTTTATTTTCTTA-3') and B (5'-GCCCTCTAGATTAAGCGTAATCTGGAACATCGTATGGGTACATCTCGA GATATTTTCCCAATTCTT-3'). The fragment was inserted into the transfer vector pVL1392 (BD Bioscience Pharmingen). The resulted plasmid was co-transfected with linearized baculovirus DNA (BD Biosciences) into Sf9 cells. Successful recombinant baculoviruses were selected and amplified. Sf9 cells were infected with the recombinant baculovirus at a multiplicity of infection (moi) of 3–10. At 4 days post-infection, cells were collected and lyzed in lysis buffer (50 mM Tris, pH 7.5; 150 mM NaCl; 0.1% Nonidet P40; complete protease inhibitor cocktail). The protein was purified using the anti-HA affinity column (Roche applied Science).

S protein pull down and Western Blot detection

Amylose column was bound with protein MBP-HR1 or MBP-HR2, which was the purified fusion protein from pMAL-C2 expression system (New England Biolabs). The MBP tag is in the N-terminal. HR1 or HR2 is fused at C-terminal. The purified MBP from empty vector pMAl-C2's expression was also loaded to the amylose column as a negative control. The existence of MBP tag ensured the binding of MBP-HR1 or MBP-HR2 protein to the amylose column. S protein was expressed in insect cell SF9 and the lysate were then loaded to the MBP-HR1 / 2 protein pre-bound amylose column and washed, respectively. Flow through from the MBP-HR1 / 2 pre-bound amylose column and washing fractions were collected. The MBP-HR1 / 2 bound amylose resin, after insect cell lysate loading, were taken out and boiled with the SDS-PAGE loading buffer. The supernatant of cell lysate, flow through, and washing fractions and the boiled amylose resin were subjected to the SDS–PAGE for later western blot assay to detect the existence of S protein in each sample. The same

operation was done on the MBP pre-bound column and the amylase resin was also subjected to SDS-PAGE for western blot assay.

The samples on the SDS-PAGE gel were transferred to polyvinylidene fluoride (PVDF) membrane (Millipore, Singapore), using the semi-dry transfer method. The membrane was blocked in 5% low fat milk containing PBS-T for one hour. The membrane was then incubated with diluted anti-serum or antibody in 5% low-fat milk containing PBS-T at room temperature for one hour, followed by the PBS-T washing for three times, 10 minutes each. The membrane was then incubated with the HRP labeled secondary antibody for one hour, followed by three times PBS-T wash, 10 minutes each. The ECL Plus (GE Healthcare, Singapore) was used as detection reagent. The result was exposed on Kodak X-ray medical film and developed by Kodak X-OMAT 2000. For the S protein detection, the rabbit anti-HA (1:1000) or rabbit anti-SARS-CoV S protein serum (1:4000) was used as primary antibody. The HRP-linked Donkey anti-rabbit antibody (1:5000) (GE Healthcare, Singapore) was used as secondary antibody.

Chemical cross-linking

Chemical cross-linking was performed as described by Scianimanico et al.(Scianimanico, Schoehn et al. 2000). Briefly stated, purified HR1 or HR2 peptides were dissolved in cross-linking buffer (50 mM Hopes, pH 8.0) at concentration of 2 mg/ml and cross-linked through the addition of glutaraldehyde (Sigma-Aldrich, US). If the homo-oligomers of HR1 or HR2 peptides were formed, the cross-linking reaction should be able to protect those oligomers, so that the homo-oligomers of HR1 or HR2 can be detected in later Tricine / SDS-PAGE. The reaction is allowed to proceed for one hour on ice with the cross-linker at concentration of 0.1, 0.5, 1, 2.5, 5, 7.5, 10, 25 mM, and then quenched with 50 mM Tris-HCl (pH 8.0, five minutes, at

room temperature). The products were resolved by Tricine / SDS-PAGE. (Running buffer: 100 mM Tris, pH 8.3; 100 mM Tricine; 0.1% SDS. Tris-HCl / SDS: 3M Tris-Cl; 0.3% SDS, pH8.45. Separating Gel: 1.1 M Tris-HCl / SDS, pH 8.45; 29:1 Acrylamide/Bis-acrylamide, 12%; Glycerol 11.65%; 10% (w/v) ammonium persulfate, 1:540 added; TEMED 1:1800 added. Stacking Gel: 0.68 M Tris-HCl / SDS, pH 8.45; 29:1 Acrylamide/Bis-acrylamide, 2.67%; 10% (w/v) ammonium persulfate, 1:540 added; TEMED 1:1800 added. Tricine Gel Sample buffer (BIO-RAD, Singapore): 200 mM Tris-HCl, pH 6.8; 2% SDS; 40% glycerol; 0.04% Coomassie Blue G-250, 200 mM β-mercaptoethanol was added before sample loading.)

For native gel PAGE, the running method is according to "Molecular Cloning: A Laboratory Manual, Third Edition, Cold Spring Harbor Laboratory Press." (Running Buffer: 25 mM Tris, pH 8.3; 192 mM glycine. Separating Gel: 0.375 M Tris-HCl, pH 8.8; 29:1 Acrylamide / Bis-acrylamide, 12%; 10% (w/v) ammonium persulfate, 1:100 added; TEMED 1:2500 added. Stacking Gel: 0.125 M Tris-HCl, pH 6.8; 29:1 Acrylamide / Bis-acrylamide, 5.1%; 10% (w/v) ammonium persulfate, 1: 100 added; TEMED 1:1000 added. Native sample buffer (BIO-RAD, Singapore): 62.5 mM Tris-HCl, pH 6.8; 40% glycerol; 0.01% Bromophenol Blue.)

Chromatography

The HPLC analysis was performed using a Shimadzu chromatography system (LC-10ATVP, Shimadzu, Japan) equipped with a photodiode array detector (SPD-M10ATVP). The RP-HPLC used a Zorbax XDB-C8 column (4.6 × 150 mm, Agilent Technologies, US) with a flow rate of 1 mL/min and Buffer A (aqueous 0.04% TFA) and Buffer B (0.04% TFA in 90% acetonitrile). Factor Xa digested MBP-HR1 or MBP-HR2 solution mixture was loaded for analysis and purification. Cleaved HR1 or HR2 from MBP-HR1 or MBP-HR2 was the target for purification. A gradient of

10-60% Buffer B over 40 minutes was used to separate products with UV monitoring at 220 nm. Purified HR1 and HR2 peptide and their mixture in different pH solution were loaded for size exclusion chromatography. Size exclusion chromatography used a BioSuite 250 HR SEC column (7.8 × 300 mm, Waters, U.S.) with a flow rate of 1 mL/min and an isocratic gradient using phosphate buffered saline. The apparent molecular weights of peak fractions were calculated by comparing with protein standards (LMW gel filtration calibration kit, Amersham Biosciences, U.S.) consisting of ribonuclease A (13.7 kDa), chymotrypsinogen A (25 kDa), ovalbumin (43 kDa) and bovine serum albumin (67 kDa).

Circular Dichroism (CD) Spectroscopy

The CD spectra was obtained on a Chirascan CD spectropolarimeter (Applied Photophysics, Leatherhead, UK) using a quartz cuvette of 0.1 mm optical path length at room temperature. Purified HR1 and HR2 peptides and their equal molar mixture were dissolved PBS with different pH were loaded for CD scan. Scans were conducted from 190 to 240 nm, with an integration time of 1 second and a bandwidth of 1 nm. The protein concentration used for this study was 0.5 mg/ml.

Mass Spectrometry (MS)

Purified HR1 and HR2 peptides were submitted to the Proteins and Proteomics Center of the National University of Singapore for mass spectrometry analysis. Molecular weights were obtained using a time of flight (TOF) mass spectrometry instrument (Voyager-DE™ STR Biospectrometry™ Workstation) with a positive matrix-assisted laser desorption ionization (MALDI) mode. Peptides were identified by three cycles of protein sequencing (ABI Procise® 494 Protein Sequencer).

Results

Expression and Purification of the SARS-CoV S Protein HR1, HR2 and C-HR2

To produce substantial quantities of the S protein, the HR region and the HR1 and HR2 segments, each was expressed in *E. coli* as an MBP-fusion protein containing a Factor Xa cleavage site. The expressed fusion proteins were soluble and readily affinity-purified with amylose resin columns (Figure 2-2). Because of the relatively small size of the target peptides compared to the MBP, the inserted HR2 region was even smaller than the replaced multiple cloning site. The expressed MBP-HR2 and MBP-Cys-HR2, was even smaller than the expression control. For the great size difference of the MBP and HR2 regions, the post-cleavage RP-HPLC purification was impeded by the presence of the MBP (Figure 2-3.A) To overcome this limitation, a methanol precipitation protocol is developed to remove the MBP (please reference to material and methods). The big protein, MBP, is efficiently removed by a 50% methanol precipitation method, with no observed loss of HR2. This method substantially improves the RP-HPLC purification (Figure 2-3.B and 2-3.C). The methanol precipitation method was also applied to the purification of the HR1 with good results. Thus, the inclusion of the methanol precipitation step during purification simplified the preparation of the HR1, HR2 and Cys-HR2 peptides and results in good yields. The Cys-HR2 will be used for thiazolidine ligation. (Please see Chapter 4).

Figure 2-2

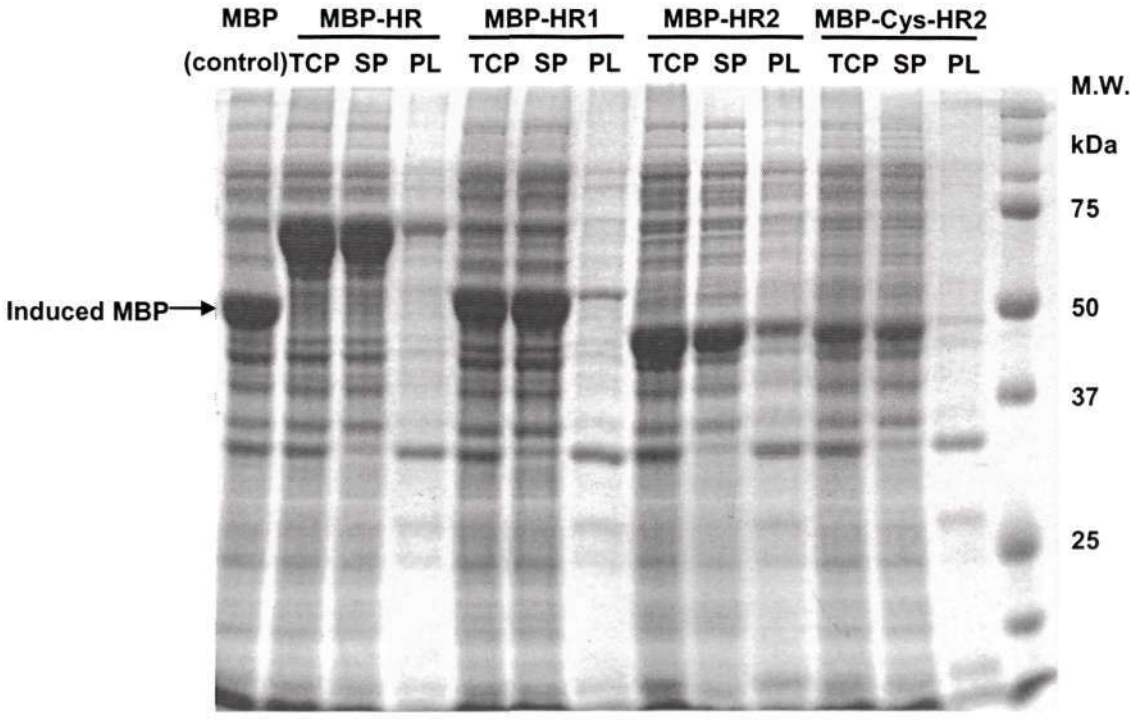


Figure 2-2. Expression of HR region in pMAL system.

The heptad repeat regions of the SARS-CoV S protein were fused with MBP and expressed as the MBP fusion protein in the pMAL system (New England BioLabs). The fusion proteins are indicated as MBP-HR, MBP-HR1, MBP-HR2 and MBP-Cys-HR2 respectively according to the HR segments it fused with. MBP-Cys-HR2 is the fusion protein which introduced a cysteine at the N-terminal of HR2 segment. The empty expressed vector MBP was used as a positive control. The lysate of the MBP fusion proteins expression cells are loaded for 10% SDS-PAGE. The induced protein expression band was shown after Coomassie blue stained on the SDS-PAGE gel. TCP: Total cell protein lysate, SP: Supernatant of the cell lysate, PL: Pellet of the cell lysate.

Figure 2-3

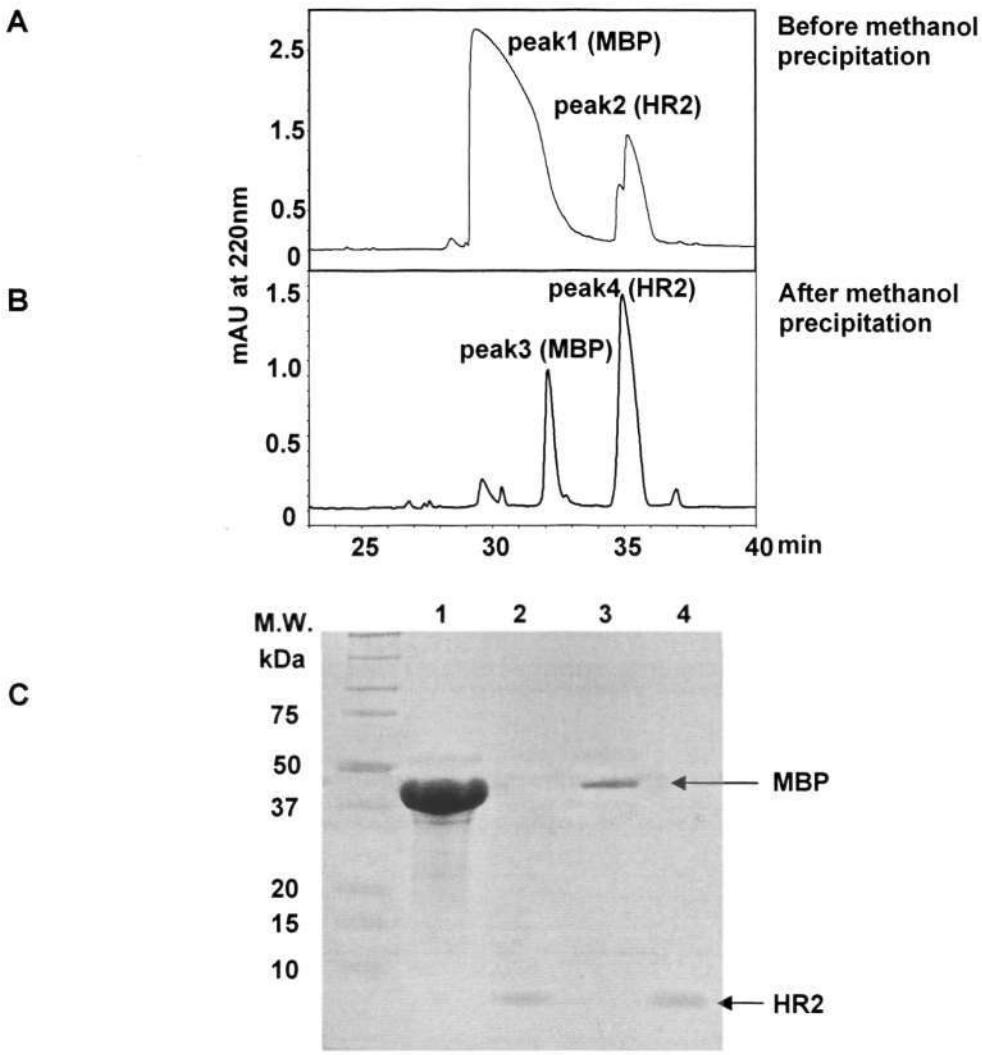


Figure 2-3. Purification of HR2 Using Methanol Precipitation followed by RP-HPLC.

A. The RP-HPLC running curve of the cleaved MBP-HR2 before methanol precipitation. Peak1 and peak2 were loaded to lane 1 and lane 2 for Tricine / SDS-PAGE (Figure 2-3.C) respectively. B. RP-HPLC running curve of cleaved MBP-HR2 after the method precipitation. Peak3 and peak4 were loaded to lane 3 and lane 4 for Tricine / SDS-PAGE (Figure 2-3.C) respectively. C. Tricine / SDS-PAGE of the collected peaks from the RP-HPLC of the digested MBP-HR2 samples.

Homo-oligomerization of HR1 and HR2

The molecular masses of the purified HR1 and HR2 were determined by the MALDI-TOF MS. The observed molecular masses of monomeric HR1 (12119.29 Da) and HR2 (4365.18 Da) agreed with their calculated molecular weights of 12116.13 Da and 4367.32 Da respectively (Figure 2-4.A and 2-4.B). Interestingly, the homogeneous solutions of both HR1 and HR2 also produce a series of peaks in the MS and those were found to be homo-oligomers. HR1 was found to form dimers and trimers, while the HR2 formed oligomers up to an octamer. It is possible that the higher orders of the HR1 and HR2 oligomers were formed, but were not detected within the scanning range of the MS analysis. The propensity of HR1 and HR2 spontaneously forming such multimeric structures hints that there exists an affinity between those peptide monomers. This self-affinity might support the native trimeric quaternary structure of the S protein.

Figure 2-4

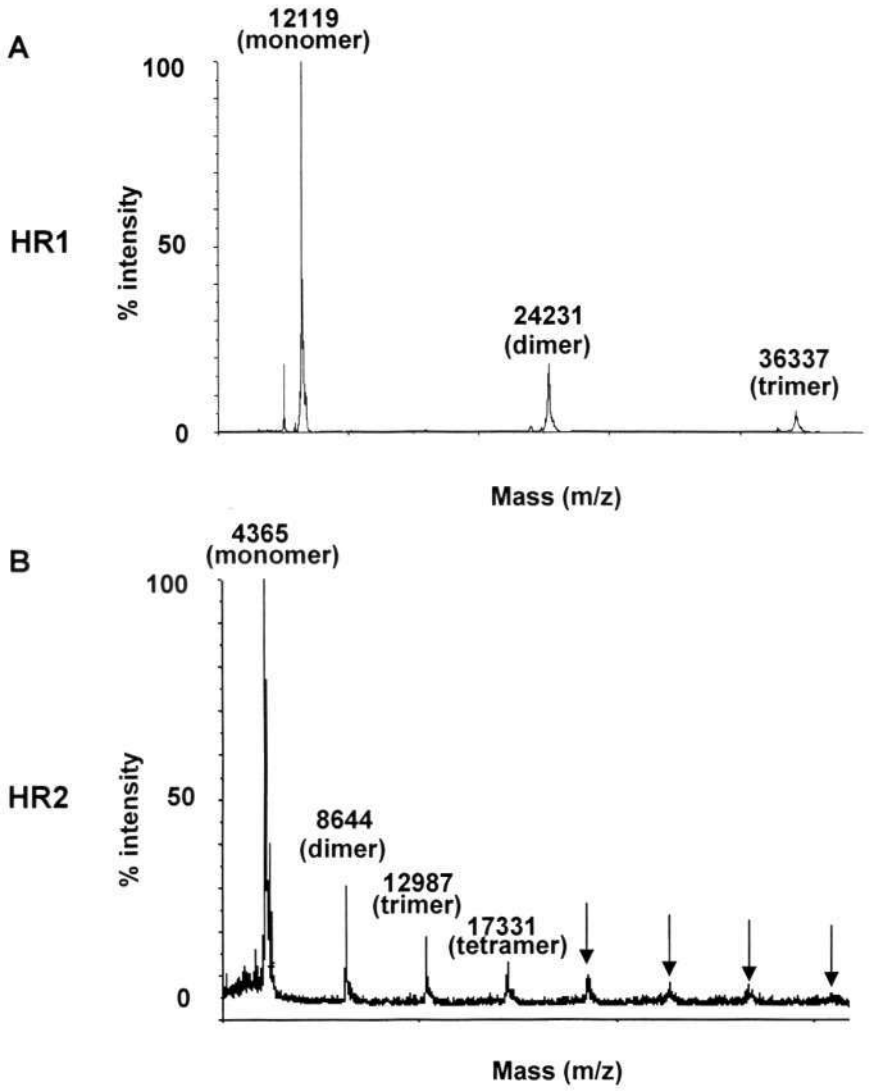


Figure 2-4. MALDI-TOF Mass Spectrum of HR1 and HR2

The molecular masses of the purified HR1 and HR2 were determined by the MALDI-TOF MS. A. Mass spectrum of the purified peptide HR1. The three peaks are believed to be the monomer, dimer and trimer of HR1. B. Mass spectrum of the purified peptide HR2. The serials of peaks are believed to be the monomer, dimer, trimer and even higher oligomers of HR2. The observed molecular weight is marked one each peak. The observed monomer mass agreed with the theoretical molecular weights of HR1 (12116.13 Da) and HR2 (4367.32 Da).

To confirm the homo-oligomeric states of HR1 and HR2, a chemical cross-linking of HR1 (12.1 kDa) and HR2 (4.3 kDa) was performed individually at pH 8.0. The addition of the cross-linking reagent glutaraldehyde to the HR1 solution resulted in stabilization of dimer (24 kDa ≈ 2 × 12.1 kDa) and trimer (36.3 kDa ≈ 3 × 12.1 kDa) and tetramer (48 kDa ≈ 4 × 12.1 kDa), as well as monomeric HR1 (Figure 2-5.A, Lane 1). Similarly, homo-oligomers of the HR2 were visualized after the addition of the cross-linking reagent. In the presence of the low concentration of glutaraldehyde, bands corresponding to HR2 monomers and dimers, even to pentamer, were observed (Figure 2-5.B, Lane 3). As the glutaraldehyde concentration increases, bands corresponding to the even higher oligomers of the HR2 can be observed (Figure 2-5.B, Lanes 4-7). The results indicated that the HR1 and HR2 homo-oligomers were preferentially formed. Interestingly, oligomers larger than tetramers of HR1 were not observed during the chemical cross-linking experiments, suggesting that the tetramer may be the saturated state for the HR1.

Figure 2-5

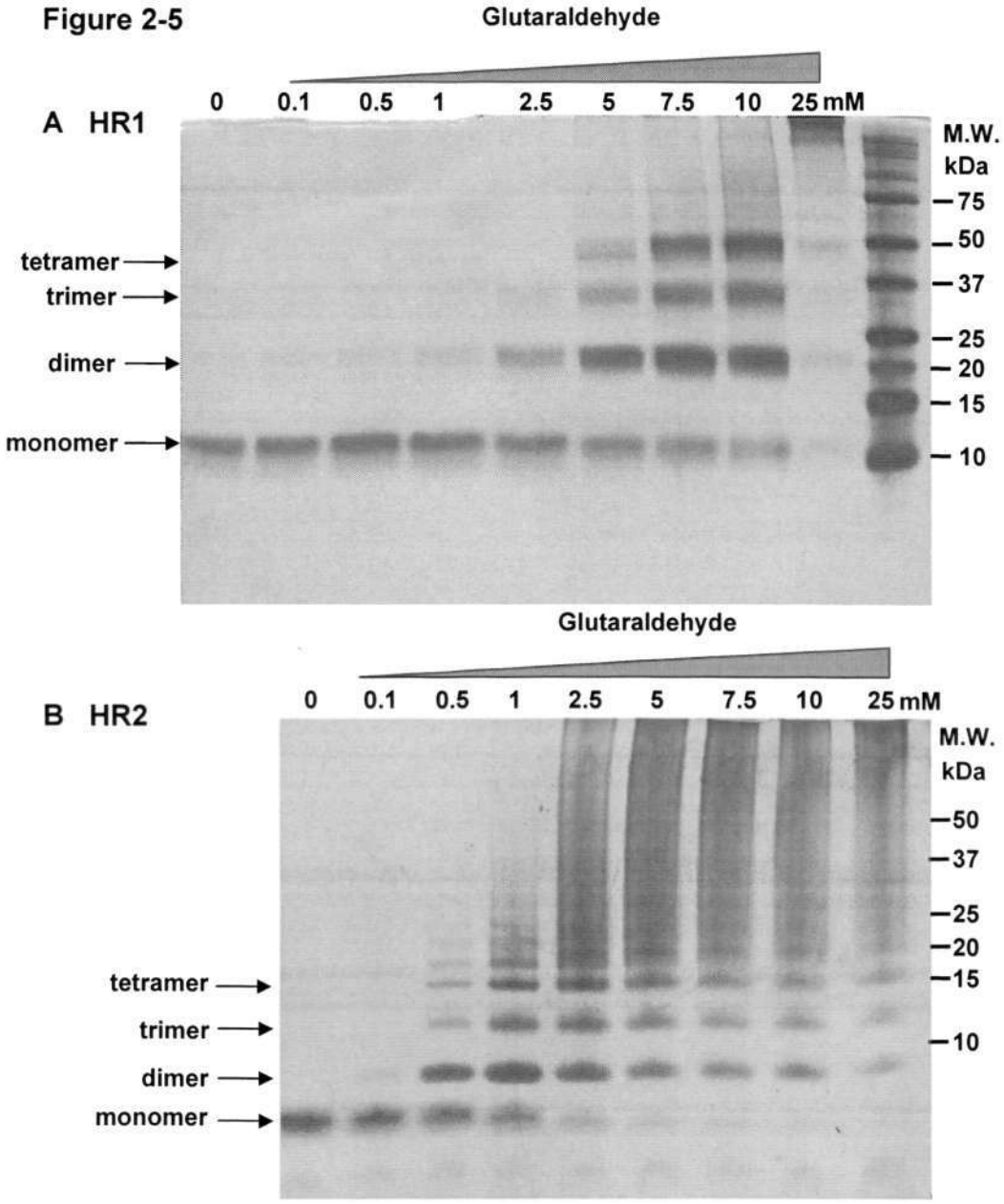


Figure 2-5. Cross-linking of HR1, HR2 at Different Concentration of Glutaraldehyde.

Monomer of HR1 and HR2 had the potential to form oligomers. The interactions between monomers were too weak so the oligomers formed were hardly detected in the stained SDS-PAGE gel. Here the cross-linker was used to link each of them to oligomers. The purified HR1 peptides (2 mg/ml) (Figure 2-5.**A**) and HR2 peptides (2 mg/ml) (Figure 2-5.**B**) were cross-linked by the glutaraldehyde at different concentration and then resolved by the 12% Tricine / SDS-PAGE, followed by silver staining. Oligomer bands of HR1 and HR2 were indicated by the arrows.

Interaction between the HR1 and HR2

Several groups have confirmed that peptides representing HR1 and HR2 interact in vitro (Bosch, Martina et al. 2004; Tripet, Howard et al. 2004), suggesting the interaction in vivo. To confirm this interaction, we first detected the interaction of HR1 and HR2 to the S protein.

Affinity purified MBP fusion protein MBP-HR1 and MBP-HR2, were used to test the HR1 and HR2 interaction with the S protein. The fusion proteins bind to the amylose column because of the existence of the MBP. The HR1 or HR2 were used to pull down the C-terminal HA-tagged SARS-CoV S protein, which was expressed in the insect cell system. The cell lysate flow through, the wash buffer and the amylose resin was collected respectively for the western blot to detect the existence of the S protein using anti-HA antibody (Figure 2-6.A). Comparing with the flow through, the S protein was enriched in the amylose resin. In the earlier experiment, we found the solely bound MBP would not be able to pull down the S protein. The result showed the MBP fused HR1 and HR2 can bind to the S protein, suggesting the interaction of HR1 and HR2 to the S protein. We believe the binding capacity originates from the binding potential between the heptad repeat region and its counterpart on the S protein.

Figure 2-6

Figure 2-6.A. S protein Pull Down Assay to Detect the Interaction of HR1 / HR2 with the S Protein.

The HR1 and HR2 segment are fused with MBP to form fusion proteins MBP-HR1 and MBP-HR2 respectively. The fusion proteins were used to pull down S protein to confirm the interaction between S protein and the HR regions. The MBP fusion proteins were loaded to amylose column first. Purified MBP was also loaded to one column as a negative control. The fusion protein can bind to the amylose resin by its MBP part. After pre-loading, the columns were washed and loaded with HA-tagged S protein, which was ectodomain of S protein (amino acids 1 to 1188) expressed from insect cell SF9 and it was provided as a gift by my colleague Lu Yanning (Lu, Liu et al. 2008). After loading, the columns were washed using column buffer. The exposed HR regions were expected to interact with S protein and retain S protein on the column. Flow through, washing samples and resin from each of the MBP fusion protein bound column were collected and examined by western blot for the existence of S protein. Rabbit anti-SARS-CoV S serum was used as primary antibody.

Lane 1 / 2: The S protein flow through from the column binding MBP-HR1 / MBP-HR2. Lane 3 / 4: The 5th wash sample from the column binding MBP-HR1 / MBP-HR2. Lane 5: Insect cell lysate, which contains the HA tagged SARS-CoV S protein (positive control). Lane 6 / 7: amylose resin from the column binding MBP-HR1 / MBP-HR2. Lane 8: amylase resin from the column binding MBP (negative control).

To confirm this, we used the RP-HPLC purified HR1 and HR2 peptides to test the interaction between them. HR1 and HR2 were mixed according to the molar ratio 1:1.2, 1:1, 1.2:1, and run on the native PAGE (Figure 2-6.B).

The length of HR1 is almost 3 times of HR2. In PAGE, more HR1 need to be loaded to the gel if it is equal mole to HR2. For example, 1.2 mg HR1 was approximately equal mole to 0.43 mg of HR2. So, the HR1 band looks darker than HR2 band on the stained PAGE Gel when equal mole of them is loaded. We did not select higher different ratio like 1:2, 1:1 or 2:1 to test in the experiment. Otherwise the HR1 band will appears quite dark and the HR2 band might be too faint comparing with HR1. So we select molar ratio 1:1.2, 1:1, 1.2:1 because it will gives better staining result. In the stained native PAGE gel, except for the hetero-oligomer bands in ratio 1:1.2, a weak band representing HR2 was shown. In ratio 1.2:1, a weak band corresponding to HR1 was shown. When they were mixed equal molar, only the hetero-oligomer bands can be detected. The results suggested that there exists an interaction between the HR1 and HR2. The composition of HR1 and HR2, in their hetero-oligomer, is 1:1. The results support the structure model of SARS-CoV fusion core that is composed of 3:3 HR1 and HR2 bundles.

My colleague, Dr. Ma Zengshuan, used the cross-linking reagent ethylene glycol-bis(succinimidyl succinate) (EGS) to protect the hetero-oligomers of HR1 and HR2. Other than the monomer bands, two bands corresponding to 32 and 49 kDa can be found in the Coomassie blue stained SDS-PAGE gels, which were supposed to be the hetero-dimer and hetero-trimer of the HR1 and HR2 complex. (32 kDa $\approx$ 2 $\times$ (12.1 + 4.3) kDa), (49 kDa $\approx$ 3 $\times$ (12.1 + 4.3) kDa). Similar results can be found in Rao's and Tien's groups' work (Xu, Zhu et al. 2004; Ni, Zhu et al. 2005), which supported that hetero-oligomers formed by HR1 and HR2 were 3:3 HR1 and HR2 bundles.

Figure 2-6.B

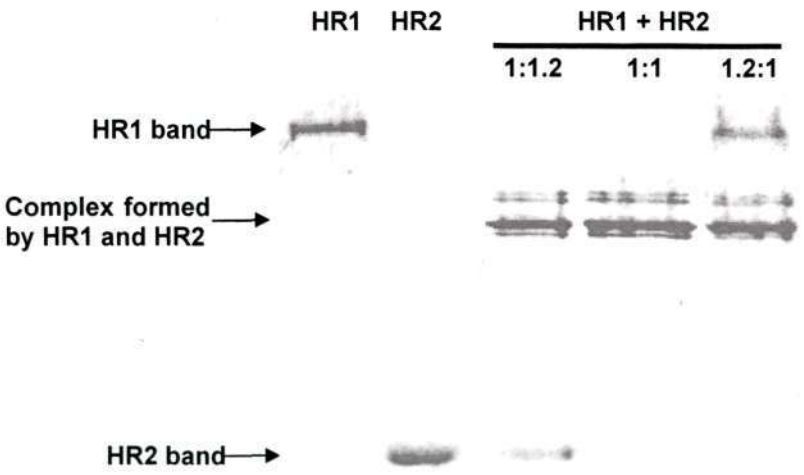


Figure 2-6.B. Native PAGE detection of HR1 and HR2 interaction.

Purified HR1 and HR2 peptides were mixed 1:1.2, 1:1, and 1.2:1 according to molar ratio and loaded for native PAGE. The HR1, HR2 bands and complex formed by HR1 and HR2 were shown after Coomassie blue stain.

pH-dependent Homo-oligomerization of HR1 and HR2

The multimeric forms of HR1 and HR2 were further examined using size exclusion chromatography (SEC), with varying pH: acidic, neutral and basic. Regardless of the pH, separate HR1 and HR2 solutions were found to contain only a single species, although the molecular weight of the species was pH-dependent (Figure 2-7.A and 2-7.B). Firstly, HR1 was found to elute at a position corresponding to 49 kDa (tetramer) in basic and neutral pH (pH 9 and 7), while at pH 5, the elution position corresponded to the monomeric HR1 peptide. Similarly, HR2 eluted at positions corresponding to 44 kDa (10-mer) and 26 kDa (6-mer) at the basic and neutral pH respectively. At pH 5, however, the HR2 eluted at a position corresponding to < 13 kDa, reflecting mono- or dimeric character. The data indicated that the large homo-oligomers of HR1 or HR2 were present under the basic and neutral conditions, but were unstable at the acidic pH and monomeric conformation is preferred.

The S protein exists in a trimeric conformation on the viral surface (Beniac, Andonov et al. 2006), which may suggest the innate affinity between each of the monomers. Thus, it is not surprising that the HR1 and HR2 oligomers are observed.

Figure 2-7

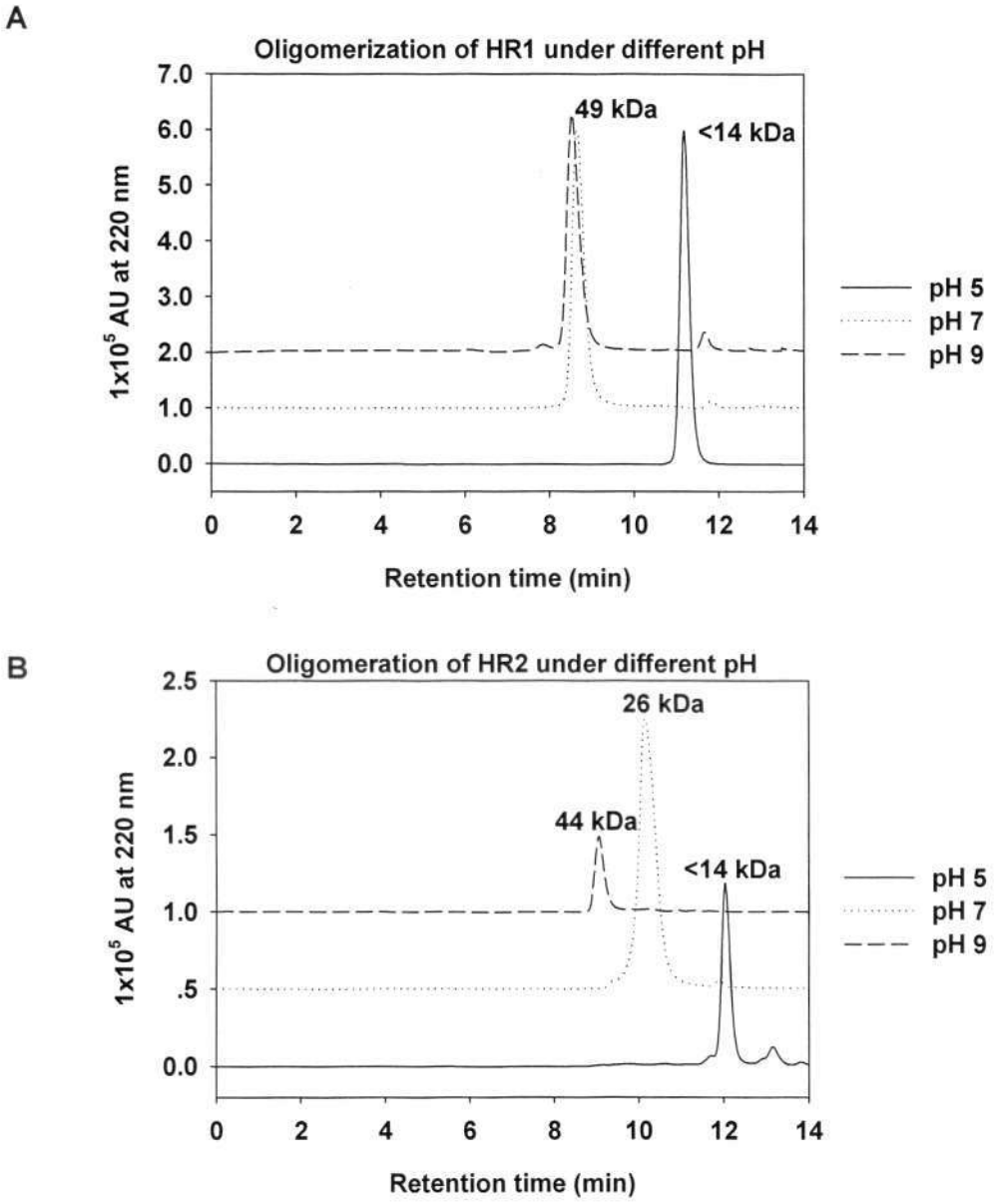


Figure 2-7. Homo-oligomerization of HR1 and HR2 at Different pH.

The SEC chromatography, which separates particles on the basis of size, was used here to analyze the homo-oligomerization of HR1 and HR2 at different pH. The purified HR1 and HR2 peptides were dissolved in a PBS pH 5.0, pH 7.0 and pH 9.0, respectively. The SEC chromatography was performed on a BioSuite 250 HR SECcolumn, with isocratic elution of the respective PBS solution at a flow rate of 1.0 ml/min. The SEC chromatography profiles of HR1 (**A**), HR2 (**B**) at pH 5.0, pH 7.0 and pH 9.0 were shown respectively. The molecular weights of peaks were calculated by a comparison with the protein standards.

pH-independent Hetero-oligomerization of HR1 and HR2

We also examined the size of the resultant HR1 + HR2 hetero-oligomer using size exclusion chromatography (SEC). An equal molar mixture of HR1 and HR2, is subjected to size exclusion chromatography at pH 5, 7 and 9 (Figure 2-8.A). At each pH level tested, the major peak corresponded to a molecular weight around 49 kDa, indicating 3 HR1 associated with 3 HR2.

RP-HPLC has a non-polar stationary phase and an aqueous, moderately polar mobile phase. Compounds stick to reverse phase HPLC columns in high aqueous mobile phase and are eluted from RP HPLC columns with high organic mobile phase. Peptides can be separated by running a linear gradient of the organic solvent. Although the HR1 + HR2 hetero-oligomer appeared a 49 kDa peak at all the pH we tested, which indicating a relatively stable structure, we believed the HR1 and HR2 will disassociated by the solvent change. When the polarity of the aqueous around the molecule changed to non-polar phase, the interaction between HR1 and HR2 might be disrupted. The linear gradient of the organic solvent could give better separation of HR1 and HR2. So, we used RP-HPLC to analyze the 49 kDa peak collected from Figure 2-8.A.

As shown in Figure 2-8.B, the peak collected in Figure 2-8.A was separated to two peaks in Figure 2-8.B. Each peak was confirmed to be HR1 and HR2 respectively. HR1 and HR2 were found to be in equal molar ratio. The results suggest that the hetero-oligomerization of HR1 and HR2 was stable at pH 5 during size exclusion chromatography. However, the hetero-oligomer dissociated to monomeric HR1 and HR2 when subjected to RP-HPLC.

Figure 2-8

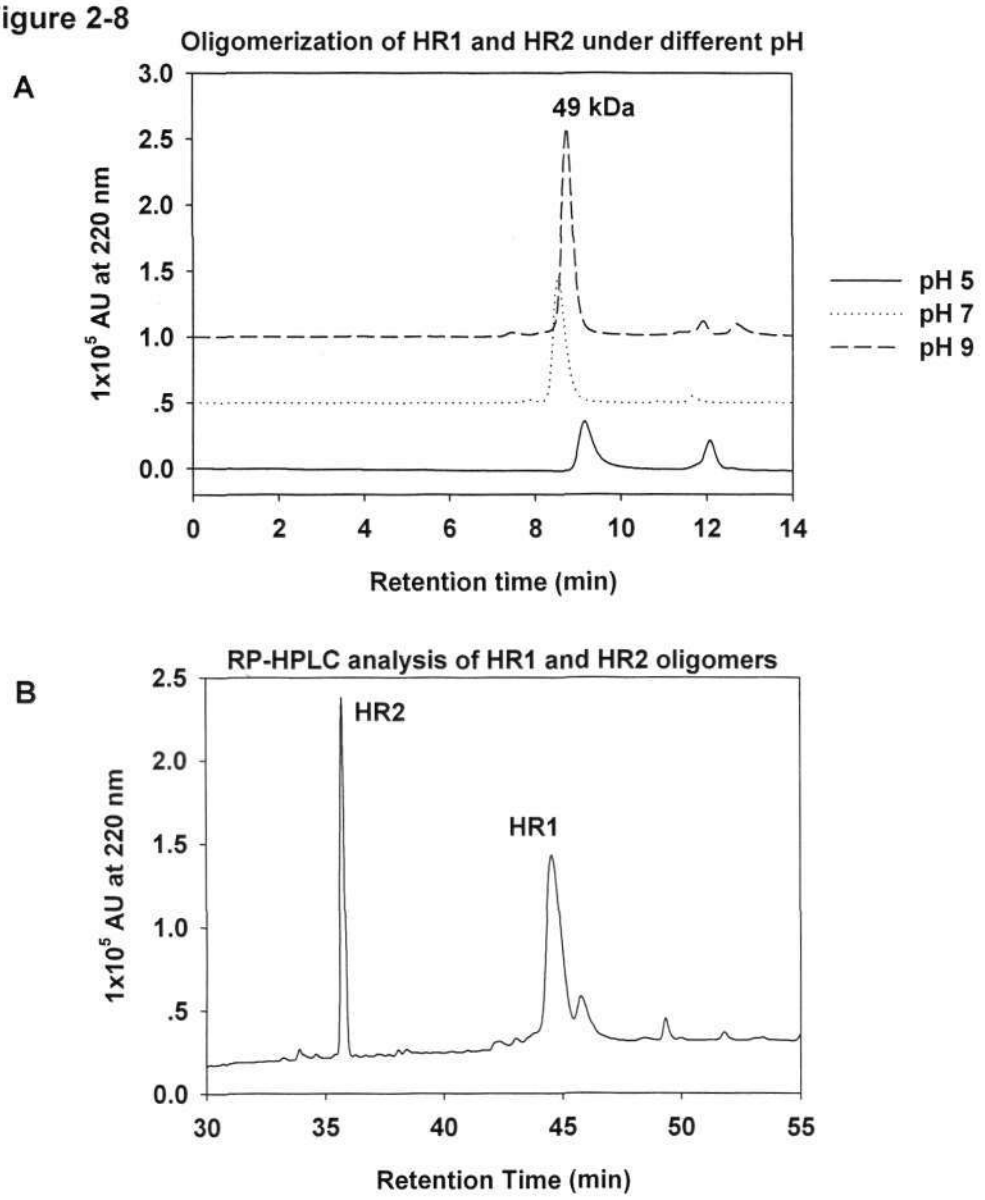


Figure 2-8. Hetero-oligomerization of HR1 + HR2 at Different pH.

The SEC chromatography was used here to analyze the hetero-oligomerization of HR1 + HR2 at different pH. The equal molar HR1 and HR2 mixture were dissolved in PBS at pH 5.0, pH 7.0 and pH 9.0 respectively. The SEC was performed on a BioSuite 250 HR SEC column with isocratic elution of respective PBS solution at a flow rate of 1.0 ml/min. The SEC chromatography profiles of the HR1 + HR2 mixture at pH 5.0, pH 7.0 and pH 9.0 are shown (**A**). The molecular weights of peaks were calculated by performing a comparison with the protein standards. The single peak in Figure 2-8. A. were collected and used for RP-HPLC analysis. The RP-HPLC profiles are shown (**B**). The eluted peaks were confirmed to be composed of HR1 and HR2.

Circular Dichroism (CD) Spectroscopy

The secondary structures present within the HR1, HR2 and the HR1 + HR2 equal molar mixture solutions, were examined by the CD. The CD spectra of HR1, HR2 and HR1 + HR2 mixture exhibited α-helical structures, with double minima at 208 nm and 222 nm (Figure 2-9). The HR2 displayed an unordered structure. The helical state of both HR1 and HR2 were affected by the pH. The helical content was found to significantly decrease as the pH became more acidic. Furthermore, a substantial increase of helicity was observed by the addition of HR1 to HR2. The helical state of HR1 + HR2 mixtures was not affected by the pH changing (Figure 2-9.C), which may suggest that the HR1 and HR2 mixture maintained a relative stable structure in the test condition.

Figure 2-9

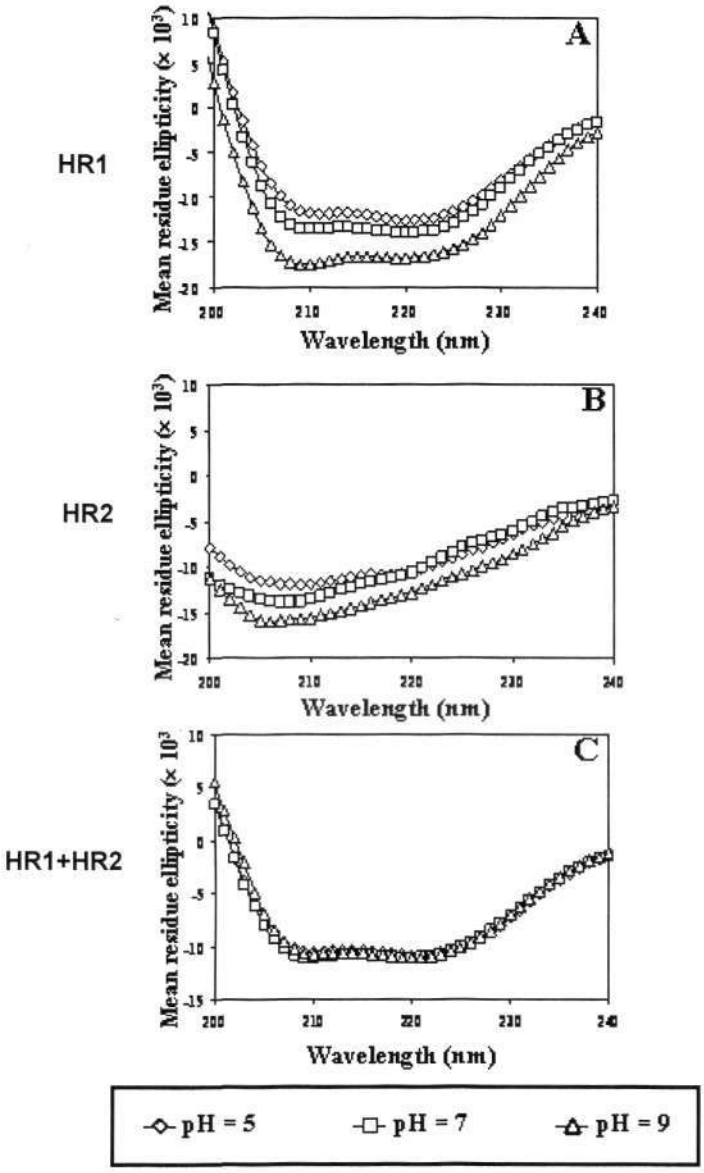


Figure 2-9. CD Spectra of HR1, HR2 and their Mixture.

CD was used here to examine the secondary structures present within the HR1, HR2 and the HR1 + HR2 equal molar mixture. HR1 (**A**), HR2 (**B**) and the HR1 + HR2 equal molar mixture (**C**) were dissolved in the PBS at pH 5, 7, 9 and examined by the CD. The CD spectra were obtained on a Chirascan CD spectropolarimeter using a quartz cuvette of 0.1 mm optical path length at room temperature. Scans were conducted from wavelengths of 200 to 240 nm, with an integration time of 1 second and a bandwidth of 1 nm.

Discussion

The fusion core of SARS-CoV was crystallized, and its structure was determined (Xu, Lou et al. 2004; Duquerroy, Vigouroux et al. 2005). It was a six-helix bundle structure with three HR2 helices packed against the hydrophobic grooves of the central coiled coils formed by three parallel HR1 helices in an oblique antiparallel manner. This complex is thought to represent the post-fusion conformation. It was a highly acid stable structure, which was resistant to proteolytic degradation and thermal denaturation in the SDS gel at 60°C (Bosch, Martina et al. 2004).

In our study, other than the trimer, the HR1 formed tetramer was determined by the chemical cross-linking and chromatographic analysis. The HR2 formed from dimer to pentamer and even higher oligomers through chemical cross-linking and chromatographic analysis. Our finding partially agreed with those of Tripet and his coworkers (Tripet, Howard et al. 2004), who have reported that the HR1 forms a stable tetramer by sedimentation analysis. In examining their data of the HR2 formed trimer, we found their sedimentation equilibrium experiments show a band greater than the theoretical mass of trimer at a concentration of 250 μM, which may also be interpreted as higher oligomerized peptides. The fusion core of the SARS-CoV S protein is formed by three HR2 docking on the groove of the trimerized HR1. Based on our understanding of the class I viral fusion protein, the helical state of HR2 is induced by binding to the helical HR1. The CD analysis of HR1 and HR2 mixture supported this theory. The HR2 displayed an unordered structure through CD analysis to HR2 peptides. We believed the HR2 exist in the unordered structural state as monomer, it is possible for it to associate together to form higher oligomers. So, the higher oligomer states were observed in the MALDI-TOF Mass analysis, chemical

cross-linking analysis, and size exclusion chromatography. The HR1, which centered in the fusion core as a trimerized coiled- coil, maintained a partial helical state in our study. The tetramer was the highest oligomer found in our study, suggesting that the stable existence form of HR1 might saturate at the tetramer. The HR1 provided the structural support for the fusion core and the hydrophobic grooves for the HR2 docking. The hydrophobic face between each HR1 strand was packed inside the coiled-coil, while exposing its hydrophilic side outside. So, this conformation may limit the strand numbers of the HR1. The most recent research resolved the tetramer crystal structure of the HR2, which indicated that the HR region could have another stable existence, other than the trimer form (Deng, Liu et al. 2006).

In our study, we also confirmed the function of the putative heptad repeat regions we expressed. Firstly, we tested and confirmed their interaction with the S protein in the S protein pull down assay. We believe the interaction between the HR regions with the S protein originated from the interaction between the HR regions and their counterparts on the S protein. Through native gel and SEC analysis, we confirmed that the interaction between HR1 and HR2. HR1 and HR2 formed stable trimeric complexes in a 1:1 ratio, which may be the six-helix bundle. We also explored the pH effects on the HR1 and HR2. The CD analysis applied in our study showed the secondary structure of the HR1 + HR2 mixture was maintained at pH 5, while the structure of HR1 or HR2 was affected by pH. The acidic pH decreased the helical state of HR1. Combining the SEC analysis results to the HR1 and HR2, it is believed that the homo-oligomer state of HR1 and HR2 was affected by the pH. Acidic conditions will disrupt the homo-oligomer state of HR1 and HR2. When HR1 and HR2 are formed as hetero-oligomers, their structure is relatively stable and the hetero-oligomer would not be disrupted by the acidic pH we tested.

The influenza HA is the best characterized pH-dependent viral membrane fusion glycoprotein. Many studies revealed that when exposed to the acidic pH, non-covalent interactions between the HA subunits are disrupted (Doms, Gething et al. 1986; Eckert and Kim 2001). The loss of even one single interchain salt bridge could destabilize the trimer formation of HA (Doms, Gething et al. 1986). The loss of interchain salt bridges at acidic pH may be the reason for the disassembly of the HR1 and HR2 homo-oligomers in our study.

Under the current class I viral envelop protein-mediated entry mechanism, there exists at least three successive conformations of the envelop fusion core (Eckert and Kim 2001). They are native prefusogenic state, pre-hairpin intermediate state, and post-fusion hairpin state. Studies on the structure of the fusion core suggest that this region of the S protein experienced a transient dissociation and re-association during the fusogenic conformational transition (Duquerroy, Vigouroux et al. 2005). The low stability of the HR1 and HR2 homo-oligomers fits the general mechanism of class I viral fusion proteins, where the HR1 and HR2 must dissociate and then re-associate in order to form the pre-hairpin intermediate. The acid-induced structural state change in HR1 and HR2 and oligomeric state of HR1 and HR2, can be affected by the pH, suggests that the pH plays a critical role in the dissociation and re-association. Low pH disrupted the interaction between the HR1 and HR2 homo-oligomers, which freed the HR1 and HR2 for heteromerization during the dissociation / re-association transition of S2 protein, which facilitates the formation of the six-helix bundle conformation.

Several group's work demonstrated that the viral entry process of SARS-CoV was affected by pH (Nie, Wang et al. 2004; Simmons, Reeves et al. 2004). Their work were all using S protein pseudotyped virus, the S protein plays a unique role in their

experiment. But the effect of pH to S protein, to its structure and conformation, were seldom mentioned. Our experimental data on the heptad repeat regions supported that the SARS-CoV takes the pH-dependent pathway for viral entry. We hypothesized that the neutral pH helps the HR region maintain its homo-trimer state in the viral membrane. When it was taken into the endocytosis vesicle, the acidic environment destablizes its homo-trimer state and assists the binding of HR1 and HR2.

Although the pH affects the oligomer state of HR1 and HR2, the viral fusion process is a complicated process and involves a series of conformational changes and is affected by many factors. Receptor binding, protease cleavage and low-pH environment inside the cell may all be involved. They may not only affect the oligomer states of the heptad repeat region, the conformational state of S2 domain of the S protein and evenly the S1 domain may be affected by those factors. Although our study in this chapter mainly focused on two heptad regions, our study provided a clue that the pH may be one of the critical factors affecting the SARS-CoV viral entry. Further intensive study is necessary to unveil the entry mechanism of SARS-CoV, especially for those experiments done in vivo, which will be the main focus in the next chapter.

Chapter 3

Characterization of the SARS-CoV S Protein Mediated Viral Entry

Characterization of the SARS-CoV S Protein Mediated Viral Entry

Introduction

Enveloped viruses use their envelope glycoprotein to invade the host cell. This process is initiated by attaching and binding their envelope glycoprotein to cellular receptors, and then it is followed by the conformational changes leading to membrane fusion and genome delivery (Dimitrov 2004). The SARS-CoV uses its spike glycoprotein to gain entry to the host cell. It is believed that the coronavirus spike protein mediated fusion shares many similarities with the other class I viral fusion proteins mediated viral-cell fusion (Bosch, van der Zee et al. 2003).

Both the viral entry of HIV-1 and the influenza virus are mediated by their envelop protein, HIV Env and influenza HA. As the model class I viral fusion proteins, they both require the host protease cleavage to render the protein fusion component. The viral fusion proteins are metastable and ready to be activated after the proteolytic events. In the HIV-1, the conformational change is activated by the receptor binding in neutral pH. But for the influenza virus, the conformational change of its HA protein was triggered by the proton binding, which provides the virus with a mechanism for detecting when it has arrived in the low pH milieu of the endosome. Thus, the virus membrane fusion takes place at different locations and stages in the viral entry process. For the HIV-1, the fusion occurs in the early entry stage and the viral membrane fuses directly with plasma membrane in a pH-independent manner. For the influenza virus, the fusion happens within the organelles of the endocytic pathway in a mildly acidic pH-dependent manner.

Comparing the viral entry pathway of both two viruses, several similarities and differences stand out and might enlighten our understanding of the SARS-CoV, which may use a similar invasion strategy of one of them, or both.

The proteolytic cleavage event and the pH are believed to be the two crucial factors that affect the viral entry process when comparing the viral entry pathway of HIV-1 and the influenza virus. SARS-CoV is the coronavirus bears class I viral fusion protein. How the proteolytic events and pH affect its viral entry process is an intriguing topic for further study. Although the receptor binding is another crucial factor which affects the viral entry process, and there is reports mentioned S protein enters cell through receptor mediated clathrin- and caveolae-independent pathway, cholesterol- and sphingolipid-rich lipid raft are involved (Wang, Yang et al. 2008), we chose to focus our study on proteolytic cleavage and pH effect in the SARS-CoV viral entry process.

Combining the result of our study on that pH affect the oligomer states of the heptad repeat regions, we extended our study to the effect of pH in the SARS-CoV viral entry and the proteolytic effect of spike protein of SARS-CoV. To analyze the function of SARS-CoV spike protein, we also have to increase the safety of studies on human pathogens. The pseudotyped viruses are adopted in our research. The heterogenic glycoprotein can be incorporated into the retroviral virions. The pseudotyped virus maintains host ranges of the original virus which glycoprotein is incorporated, while it can perform only one cycle infection (Landau, Page et al. 1991; Dong, Roth et al. 1992; Suomalainen and Garoff 1994). The pseudotyped virus particles are used as safe tools in the study of the viral entry mechanism, cell tropism, or receptor identification (Deng, Liu et al. 1996; Deng, Unutmaz et al. 1997; Wool-Lewis and Bates 1998; Chan, Empig et al. 2001).

Using the codon-optimized S protein, the S protein expression level was greatly improved through comparing it with the native sequence S protein in our tissue culture. The S protein was then incorporated into the HIV-1 based virus particle to form the S protein pseudotyped virus. To enhance the effect, the receptor of the SARS-CoV, angiotensin-converting enzyme 2 (ACE2) was expressed in 293 cell lines and the stable ACE2 expressing cell line was setup as an effector cell for study of the SARS-CoV viral infection. The entry process of the S protein pseudotyped virus showed the pH dependent manner. The cellular surface expresses the S protein and was treated by the trypsin protease. When mixed with the ACE2 expressing 293T cell, the obvious cell-cell fusion was observed. The following study showed that the S protein mediated cell-cell fusion is not affected by the pH.

Materials and Methods

Plasmid Construction

The full length of the S protein sequence containing the vector pJX40-S was kindly provided as a gift by Dr. Liu Dingxiang from the Institute of Molecular and Cellular Biology (IMCB), Singapore. The sequence is identical to the SARS-CoV Singapore strand 2774 (GeneBank Accession No. AY283798.) The full length of the S protein nucleotide was amplified by the PCR from the plasmid pJX40-S and then inserted into the pCDNA3.1 expression vector (Invitrogen, Singapore). Primers used were A (5'-CGC GGA TCC ATG TTT ATT TTC TTA-3') and B (5'- ATA AGA ATT CTT ATG TGT AAT GTA A-3') for the full length S protein. Primer C (5'-CCG GAA TTC ATG TTT ATT TTC TTA TTA TTT-3') and B were used to clone the full length S protein gene into the chicken beta-actin / rabbit beta-globin hybrid promoter

(AG) containing the vector pCAGGS (BCCM, Belgium) to form the pCAGGS-S expression vector. The full length S protein gene was also cloned into the human EF-1 promoter containing the pEF1 / V5His vector (Invitrogen, Singapore) to form the pEF1-SHis vector; to Human ubiquitin C promoter containing the vector pUB6 / V5-His (Invitrogen, Singapore) to form the pUB6-SHis vector; to the retroviral vectors pFB-Neo to form the pFB-Neo-S (Stratagene, Singapore). The codon-optimized S protein vector pTSh and the Luciferase gene containing the HIV-1 base pseudovirus production vector pNL4.3.Luc-R-E-pro, as well as the vector pAMLV are provided by Professor Deng Hongkui from the College of Life Sciences, Peking University, China. The VSV-G glycoprotein expression vector pHCMV-VSV-G was provided by Dr. Liu Dingxiang from IMCB, Singapore. The SARS-CoV receptor ACE2 sequence containing vector was provided as a gift by Professor M. Farzan from Harvard University, U.S. The ACE2 gene was cloned into pCDNA3.1 expression vector using primer D (5'-CGG GGT ACC ATG TCA AGC TCT TCC-3') and E (5'-CCG CTC GAG CTA AAA GGA GGT CTG-3') to produce a full length ACE2 expression vector pCDNA3.1ACE2. Pfu DNA Polymerases (Stratagene, Singapore) was used in all the PCR reactions.

Cell Lines

In the following, Dulbecco's Modified Eagle's Medium (DMEM) (Hyclone, Singapore) with 10% fetal bovine serum (Hyclone, Singapore) is mentioned as a normal culture media for all the cell culture. The 293T cells are cultured in normal culture media as the producer cell line for protein expression and pseudotyped virus production. The 293 cell was used to produce the ACE2 stable expression cell line. The pCDNA3.1-ACE2 expression vector 1 μg was transfected using Lipofectamine (Invitrogen, Singapore) to a 12 well tissue culture dish seed with 293 cells. 24 hours post infection,

the pCDNA3.1-ACE2 transected well was trypsinized for cell counting. 4000 cells are selected and seeded in the 10 cm tissue culture dish (TPP, Singapore) with DMEM, 10% FBS and 2 mg/ml Geneticin (Invitrogen, Singapore). The single cell clone in the 10 cm tissue culture dish was transferred to a 12 well tissue culture dish for further culture. The expression of the ACE2 was detected by the anti-ACE2 poly-clone antibody (Santa Cruz Biotechnology, U.S.). The stable ACE2 expressing cell line was named 293-ACE2.

Expression of S Protein

The 293T cells were transfected by Lipofectamine 2000 (Invitrogen, Singapore) according to the product manual. The 40 hours post transfection cell lysate were analyzed by the SDS-PAGE followed by western blot. For the T7 polymerase-driven expression, cells were exposed to the vaccinia virus vTF1.1 (Alexander, Moss et al. 1992) (Gift from Dr. Liu Dingxiang from IMCB, Singapore). One hour post infection, they were washed and further cultured for analysis.

Expression and purification of SARS-CoV S protein ectodomain.

The nucleotide fragment corresponding to 1-1188 amino acid residues of SARS-CoV S protein were amplified with a HA tag at the C-terminus (S1188HA). Primers used were A (5'-AGTCGAATTCCGAAC ATGTTTATTTTCTTA-3') and B (5'-GCCCTCTAGATTAAGCGTAATCTGGAACATCGTATGGGTACATCTCGA GATATTTTCCCAATTCTT-3'). The fragment was inserted into the transfer vector pVL1392 (BD Bioscience Pharmingen). The resulted plasmid was co-transfected with linearized baculovirus DNA (BD Biosciences) into Sf9 cells. Successful recombinant baculoviruses were selected and amplified. Sf9 cells were infected with the recombinant baculovirus at a multiplicity of infection (moi) of 3–10. At 4 days post-infection, cells were collected and lyzed in lysis buffer (50 mM Tris, pH 7.5; 150 mM

NaCl; 0.1% Nonidet P40; complete protease inhibitor cocktail). The protein was purified using the anti-HA affinity column (Roche applied Science).

Western Blot

The lysate of the transfected cell, stable cell line, or viral pellets were subjected to SDS–PAGE, which was followed by transferring to the polyvinylidene fluoride (PVDF) membrane (Millipore, Singapore). The membrane was blocked in 5% low fat milk containing PBS-T for one hour at room temperature. The membrane was then incubated with diluted anti-serum, or the antibody in the 5% low-fat milk containing PBS-T at room temperature for one hour, followed by PBS-T washing for three times, 10 minutes each. The membrane was then incubated with the HRP labeled secondary antibody for one hour, followed by three times PBS-T wash, 10 minutes each. The ECL Plus (GE Healthcare, Singapore) was used as detection reagents. The result is exposed on Kodak X-ray medical film and developed by the Kodak X-OMAT 2000.

For the S protein detection, the rabbit anti-SARS-CoV S protein serum (1:4000 dilutions) was used as primary antibody. The HRP-linked Donkey anti-rabbit antibody (1:5000) (GE Healthcare, Singapore) was used as a secondary antibody. For the ACE2 detection, a goat-anti ACE2 antibody (Santa Cruz Biotechnology, U.S) was used as the primary antibody. The HRP linked rabbit anti-goat antibody (Delta Biolabs, U.S) (1:1000) was used as a secondary antibody. For the HA tag detection, a rabbit anti-HA antibody (Delta Biolabs) (1:1000) is used as the primary antibody. The HRP-linked Donkey anti-rabbit antibody (1:5000) (GE Healthcare, Singapore) was used as the secondary antibody.

Immunofluorescence Detection of Surface Protein Expression

Decanted the culture media and washed the transfected cell or stable cell line by PBS for three times. Fixed the cell samples in 4% paraformaldehyde in the pH 7.4 PBS for 15 minutes at room temperature. Washed the samples twice with ice cold PBS. The cells were incubated with 1% BSA in PBS-T for 30 minutes to block unspecific binding of the antibodies. Incubated the cells in the diluted antibody in 1% BSA in PBS-T in a humidified chamber for one hour on ice. Decanted the solution and washed the cells three times in PBS, five minutes for each wash. The cells were incubated with the secondary antibody in 1% BSA for one hour on ice in dark. Decanted the secondary antibody solution and washed it three times with the PBS, five minutes each time in the dark, and then observed under the fluorescence microscope. For detection of the S protein, rabbit anti-SARS-COV S protein serum (1:400 dilutions) was used. The FITC-labeled swine anti-rabbit antibody (1:200) (DAKO Cytomation, Denmark) was used as a secondary antibody. For the ACE2 detection, a goat anti-ACE2 antibody (Santa Cruz Biotechnology, U.S) (1:100) was used as a primary antibody. The FITC-labeled mouse anti-goat antibody I (1:200) was used as a secondary antibody.

Pseudotyped Virus Production and Infection

The SARS-CoV pseudotyped virus, HIV / SARS, was produced by the standard calcium phosphate co-precipitation method (Deng, Liu et al. 1996). Briefly, the 10 μg pNL4.3.Luc-R-E-pro (Connor, Chen et al. 1995) and 10 μg pTSh were co-transfected into the 293T cells in 10 cm dishes. Supernatants, 48 to 72 hours post transfection, were harvested and spun down at 1000 g for 10 minutes. The supernatant was then filtered through 0.45 μM pore size membrane (Millipore, Singapore) and used in the infection assays. The pseudotyped virus was purified by

ultracentrifugation through a 20% sucrose cushion in the SW28 (Beckman Coulter, Singapore) rotor at 50,000 g for 90 minutes, and then resuspended in 100 μl PBS. The virus titer was determined and normalized by the EnzChek® Reverse Transcriptase Assay Kit (Invitrogen, Singapore). The supernatant containing pseudotyped virus equal to $5x10^4$ relative transcriptase activity unit was used to infect cells in the 24-well plates ($4–8 \times 10^4$ cells / well). The cells were lyzed at 48 hours post-infection. Twenty microliters of lysate was tested for luciferase activity by using the Luciferase Assay system (Promega, Singapore) according to the product manual and measured for 10 seconds in a TD-20 / 20 Luminometer (Turner biosystems, Singapore).

Lysosomotropic Agents Treatment

The 293-ACE2 cells were pre-incubated with serum free DMEM containing serials of dilutions of NH_4Cl one hour before pseudovirus infection. The HIV (SARS), the HIV (AMLV), or the HIV (VSV-G) pseudoviruses were added onto the 293-ACE2 cells in the presence of NH_4Cl at various concentrations. The culture media was replaced with fresh DMEM containing 10% FBS 12 hours later. The luciferase activity was measured 48 hours post infection.

Confocal Microscope

One day before transfection, the chamber slides (Iwaki, Singapore) were coated with poly-L-lysine (Sigma-aldrich, Singapore) for one hour at room temperature. The chamber slides are then washed with sterile H_2O (three times, five minutes each). The cover slides were allowed to dry completely and sterilize under the UV light for at least four hours. The cell or the cell lines were then grown in the chamber slides. Post transfection 24 hours later, the culture media was decanted and the transfected cell, or stable cell line, was washed by PBS for 3 times. Fixed the cell

samples in 4% paraformaldehyde in PBS pH 7.4 for 15 minutes at room temperature. The samples were washed twice with ice cold PBS. The cells were incubated with 1% BSA in the PBS-T for 30 minutes to block unspecific binding of the antibodies. Cells that were incubated with diluted antibody in 1% BSA were left in a humidified chamber for one hour on ice. The solution was decanted and the cells were washed three times in the PBS, five minutes each time. The incubated cells in the secondary antibody in 1% BSA were left in the dark for one hour on ice. Decanted the secondary antibody solution and washed it three times with the PBS, five minutes each, and leave it in the dark. Chamber slides were mounted with a drop of the mounting medium and observed under a confocal microscope.

To detect the S protein, the rabbit anti-SARS-COV S protein serum (1:400 dilutions) was used. The FITC-labeled swine anti-rabbit antibody (1:200) (DAKO Cytomation, Denmark) was used as secondary antibody. For the ACE2 detection, goat anti-ACE2 antibody (Santa Cruz Biotechnology, U.S) (1:100) was used as primary antibody. The FITC-labeled mouse anti-goat antibody (1:200) was used as secondary antibody. Pictures of the FITC fluorescent cells were taken by ZEISS LSM510 META Axiovert 200M confocal microscope.

Cell Fusion Assay

24 hours post transfection, the S protein transfected 293T cells, were detached by the EDTA or the TPCK trypsin (15 μg/ml) treatment for 15 minutes at 37°C. The EDTA and the TPCK trypsin were removed by PBS washing three times and mixed with the 293-ACE2 cells in pH 5, 7, 9 pre-adjusted DMEM, 10% FBS culture media. After 30 minutes, the culture media was changed back to the normal culture media for further culture at 37°C. 24 hours later, the cell was stained with Hoechst (Invitrogen) and observed under fluorescent microscopy.

For confocal microscopy observation, 24 hours after the S protein expression vector pTSh transfected the 293T, it is incubated with the fluorescent probe CMTMR (Molecular Probes, Singapore) according to the product manual for 30 minutes. The 293-ACE2 cell line was incubated with fluorescent probe CMFDA (Molecular Probes, Singapore) for 30 minutes. The cell is then washed by the PBS for three times to remove the fluorescent probe. The latter step was described in the previous paragraph. After the 293-ACE2 cell mixed with the EDTA or the TPCK Trypsin treated 293T cell, the cells were cultured on the poly-L-Lysine pre-coated chamber slides. 24 hours later, the cell was mounted, using the mounting media (DAKO Cytomation, Denmark) and observed under a confocal microscope.

Proteolytic Analysis of the S Protein

The ultracentrifuge concentrated S pseudotyped viral particles, or the insect SF9 cell expressed truncated S protein S1188HA (Provided by my colleague, Dr. Lu Yanning. The S protein was truncated at amino acid 1188 and HA tagged at the C-terminal) were treated with a series of dilutions using TPCK Trypsin for 15 minutes. The digestion samples were added to the SDS-PAGE loading buffer and boiled at 100°C immediately. For the PNGase treatment, the S1188HA was treated with the PNGase (New England BioLabs, U.S) for one hour before performing the TPCK Trypsin digestion.

Result

Expression of the S protein in Mammalian Cell

The full-length S protein gene (Singapore strand 2774), was first cloned into the pCDNA3.1. No obvious S protein expression was detected by the western blot of the lysate of pCDNA3.1-S transfected 293T cell. There exist two promoters on the pCDNA3.1 vector, which can drive the downstream gene expression. They are the CMV (cytomegalovirus) promoter and the T7 polymerase promoter respectively. To induce the S protein expression, the recombinant vaccinia virus vTF1.1, which expressed the T7 polymerase, was used to infect the pCDNA3.1-S vector transfected 293T cell. A protein band close to 200 kDa, was detected in the vTF1.1 infected cell lysate (Figure 3-1.A, lane 1), which was believed to be the glycosylated form of S.

Although, the vTF1.1 induced the S protein expression under the T7 promoter, in the latter study, it's hard to remove the vTF1.1 from the tissue culture system. So, the S protein gene has been cloned into different expression vectors with different promoters, expecting an efficient S protein expression. The following vectors have been tried to express the S protein gene: chicken beta-actin / rabbit beta-globin hybrid promoter (AG) containing vector pCAGGS (BCCM, Belgium); human EF-1 promoter containing vector pEF1 / V5His (Invitrogen, Singapore); Human ubiquitin C promoter containing vector pUB6 / V5-His (Invitrogen, Singapore), retroviral vectors pFB-Neo (Stratagene, U.S). Those vectors transfected cell lysate were used for western blot analysis and the results were shown in Figure 3-1.A from lane 5 to 10. Only the chicken beta-actin / rabbit beta-globin hybrid promoter drove the detectable S protein expression, a weak band corresponding to the size of the vTF1.1 induced S protein expression band.

Figure 3-1

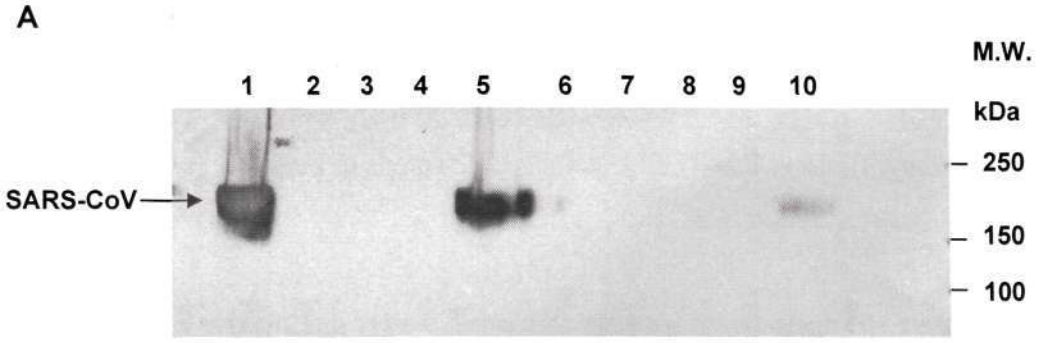


Figure 3-1.A. Western Blot Detection of the S Protein Expression in the 293T Cell.

Lane 1: pCDNA3.1-S transfected cell with vaccinia virus induction. Lane 2: Mock transfected cell with vaccinia virus induction. Lane 3: Used for maker running. Lane 4: Mock transfected cell. Lane 5: Codon-optimized S protein expression vector pTSh transfected cell. Lane 6: pCDNA3.1-S transfected cell. Lane 7: Human ubiquitin C promoter containing vector pUB6-SHis transfected cell. Lane 8: Human EF-1 promoter containing vector pEF1-SHis transfected cell. Lane 9: Retroviral vector pFB-Neo-S transfected cell. Lane 10: Chicken beta-actin / rabbit beta-globin hybrid promoter containing vector pCAGGS-S transfected cell. 293T cell was transfected by Lipofectamine 2000 (Invitrogen, Singapore). Lysate of the transfected cell was analyzed by the western blot. The expression of the S protein was detected using the rabbit anti-SARS-COV S protein serum (gift from Dr. Liu Dingxiang from the Institute of Molecular and Cellular Biology, Singapore).

Although the pCAGGS vector gave a detectable S protein expression, the expression level was not suitable for efficiently producing the S protein pseudotyped virus in the next step of our study. So, the codon optimized S protein expression vector pTSh (Gift from Professor Deng Hongkui, Peking University, China) was applied here. The native S protein nucleotide sequence was changed to the most prevalent codon used in human gene (Nie, Wang et al. 2004). The S protein expression was detected in the pTSh transfected 293T cell lysate (Figure 3-1.A, lane 5).

The S protein is a type I membrane protein. To confirm the cellular surface expression of S protein in pTSh transfected 293T cells, the rabbit anti-SARS-COV S protein serum was used in immunofluorescence assay. The 293T cell was transfected with pTSh vector and an empty vector. Diluted rabbit anti-SARS-COV S protein serum was incubated with the pTSh, or the mock transfected cell followed by FITC-labeled secondary antibody conjugation. The pTSh transfected 293T cell showed ring-like patterns under the fluorescent microscope. Observation under the light phase showed the lighting ring corresponding to the cellular membrane. The mock transfected cell has no obvious fluorescent lighting part (Figure 3-1.B). The immunofluorescence assay indicated the S protein was efficiently expressed on the 293T cell surface by transfecting the codon-optimized S protein expression vector pTSh.

Figure 3-1.B

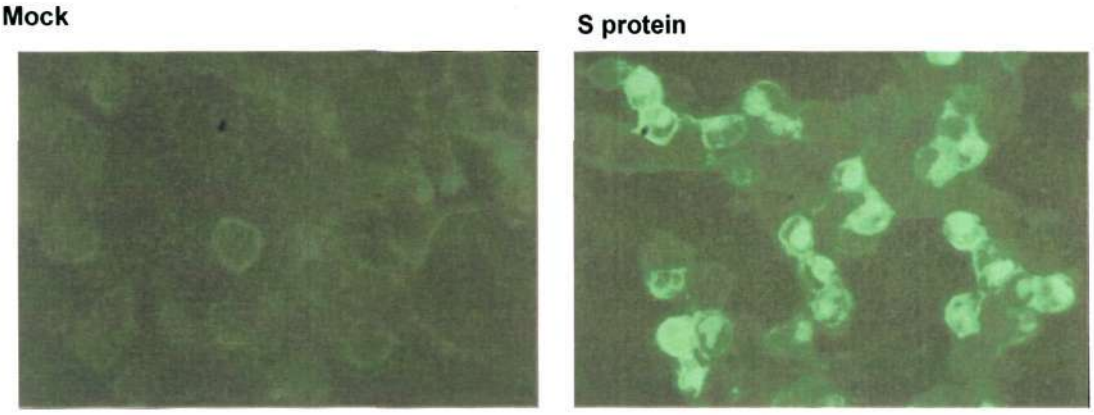


Figure 3-1.B. Immunofluorescence Detection of the S Protein Expression in the 293T Cell

Mock transfected 293T cell and pTSh vector transfected 293T cell were incubated with the rabbit anti-SARS-COV S protein serum. The FITC-labeled secondary antibody was used as a conjugate. The picture of the fluorescent cells was taken using an Olympus IX70 fluorescence microscope.

Setup ACE2 Expression Cell Line

The angiotensin-converting enzyme 2 (ACE2), was confirmed to be the major receptor of SARS coronavirus (Li, Moore et al. 2003). The exogenous ACE2 expression in the refractory cell line allowed the SARS infection (Mossel, Huang et al. 2005). The SARS-COV S protein pseudotyped retrovirus study showed the viral infectivity was affected by the expression level of the ACE2 on the cell surface (Moore, Dorfman et al. 2004). So, constructing the SARS-CoV receptor ACE2 expressing cell line will provide advantages to our viral entry study. In our previous study, the transfected 293 cell had higher expression level of the ACE2 compared with the Vero E6 cell and the transfected Hela cell. Besides, the 293 cell was a non-permissive cell line to SARS-CoV (Mossel, Huang et al. 2005). So, the 293 cells were selected to construct the stable ACE2 expressing cell line.

The ACE2 gene containing expression vector pCDNA3.1-ACE2 was transfected to the 293 cell. The pCDNA3.1 vector contained one neomycin resistant gene. The neomycin resistant cell clone might be the ACE2 expression clone. So, the cell clone in the G418 containing media was selected for further culture and examination of the ACE2 expression. The cell clone, which gives higher ACE2 expression level than the other clones, was further cultured for 20 generations. There was no obvious reduction of the ACE2 expression level. The cell line was named 293-ACE2. The expression level of the ACE2 was examined by the western blot. The 293-ACE2 cell line gave a band around 120 kDa, in the identical position of the band that showed in the pCDNA3.1-ACE2 transfected cell (Figure 3-2.A). The western blot result confirmed the expression of ACE2 in the 293-ACE2 cell line.

Figure 3-2.A

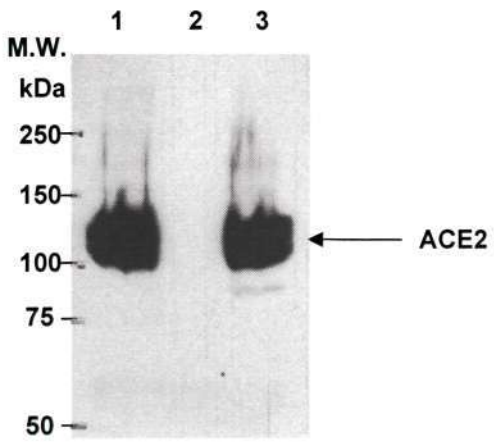


Figure 3-2.A. Western Blot Detection of the ACE2 Protein Expression in the transient transfected cell and stable cell line.

Lane 1: pCDNA3.1-ACE2 transfected cell. Lane 2: Mock transfected cell. Lane 3: ACE2 stable expressing cell line 293-ACE2. The 293 cell was mock transfected, or transfected by pCDNA3.1-ACE2 expression vector by Lipofectamine 2000 (Invitrogen, Singapore). Lysate of the transfected cell and the stable cell line 293-ACE2 is analyzed by the western blot. The expression of ACE2 protein was detected by goat anti-ACE2 antibody (Santa Cruz Biotechnology, U.S).

As the receptor of the SARS-CoV, the surface expression of ACE2 needs to be confirmed. Immunofluorescence technology was used here to observe the surface expression of ACE2 in both pCDNA3.1-ACE2 transfected 293 cell and the 293-ACE cell line. The transient expression in the transfected 293 cell and the stable expression in the 293-ACE2 cell line were both detected. The cell showed ring-like patterns under fluorescence microscope observation (Figure 3-2.B). The light phase observation showed the lighting ring corresponding to the cellular membrane. (Data not shown). The majority of the pCDNA3.1-ACE2 transfected cells showed green fluorescence. In the view field, all the 293-ACE cells appeared green fluorescence. Combining the western blot and immunofluorescence assay result, it indicated that the 293-ACE2 gave reasonable ACE2 expression on the cell surface. The 293-ACE2 cell line can be used as a target cell for the SARS-CoV viral entry.

Figure 3-2.B

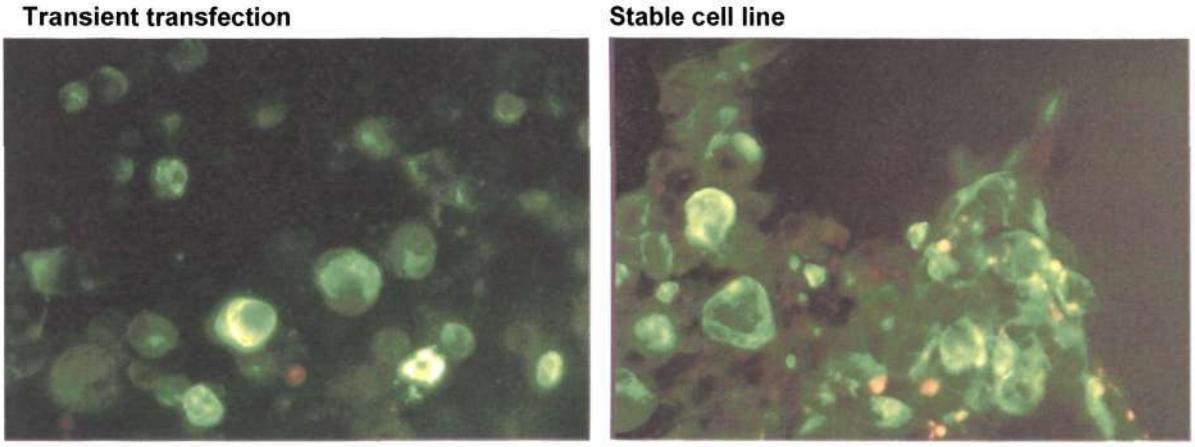


Figure 3-2.B. Immunofluorescence Detection of the ACE2 Expression in the Transient Transfected Cell and Stable Cell Line

The pCDNA3.1-ACE2 transfected the 293 cell and the ACE2 stable expressing cell line 293-ACE2 were incubated with the goat anti-ACE2 protein antibody. The FITC-labeled secondary antibody was used as a conjugate. The pictures of the fluorescent cells were taken using an Olympus IX70 fluorescence microscope.

Production of the SARS-CoV Spike Protein Pseudotyped Virus

The S protein pseudotyped virus HIV / SARS was produced by co-transfecting the 293T cells with the vector pTSh and pNL4.3.Luc.-R-E-pro (Gift from Dr. Deng Hongkui, Peking University, China) (Deng, Liu et al. 1996), using the calcium phosphate precipitation method. The culture supernatant was collected at 24, 48 and 72 hours post transfection. The virus particles were concentrated by ultracentrifugation through a 20% sucrose cushion. The concentrated viral particles were loaded to the SDS-PAGE for western blot analysis. The rabbit anti-SARS-COV S protein serum recognized a band around the 200 kDa corresponding to mature form of the S protein. This indicated that the expressed S glycoprotein was incorporated into pseudotyped particles (Figure 3-3.A). Judging by the S protein band detected in the western blot, less S protein pseudotyped viral particles are generated in the first 24 hours. As culturing time increasing, there was more pseudovirus particles generated from 48 to 72 hours. It was due to the transfected cells beginning to float out after three days culture, so the viral supernatant were not collected after 72 hours.

Figure 3-3.A

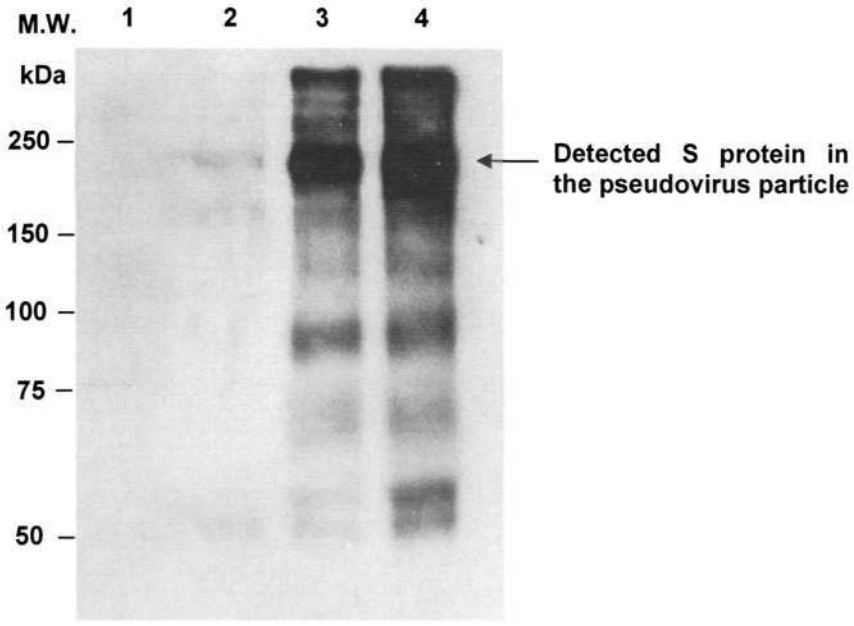


Figure 3-3.A. Western Blot Detections of Pseudovirus Particles Produced by the 293T Cell.

Lane 1. The S protein pseudotyped virus particle contain cultural media. Lane 2, 3, 4: 24, 48, 72 hours post transfection, concentrated pseudovirus particle. The 24, 48, 72 hours post transfection, pseudovirus culture media was spun down at 1000 g for 10 minutes to remove the cell debris. The supernatant was centrifuged through 20% sucrose cushion at 18000 g for 90 minutes. The pellet was resuspended in 100 µl PBS as concentrated pseudovirus samples. The unconcentrated and concentrated pseudovirus samples are loaded in the 8% SDS-PAGE for the western blot analysis. The S protein was detected by the rabbit anti-SARS-COV S protein serum.

The generated pseudovirus was based on the incorporation of the SARS-COV S protein into the retroviral particles. In our case, we use the HIV-1 based pseudovirus. The activity of reverse transcriptase carried by the HIV-1 virion can be determined by the reverse transcriptase assay, which is in proportion to the viral titer. The unconcentrated and concentrated 24, 48 and 72 hours viral supernatants were examined by the EnzCheck reverse transcriptase assay kit (Invitrogen, Singapore) to determine the viral titer. The results were shown in Figure 3-2.B. Corresponding to the western blot result in Figure 3-2.A, there were more pseudoviruses generated in 48 hours and 72 hours. The result suggested that the infectious S protein pseudotyped virus particle should be collected from 48 to 72 hours.

Figure 3-3.B

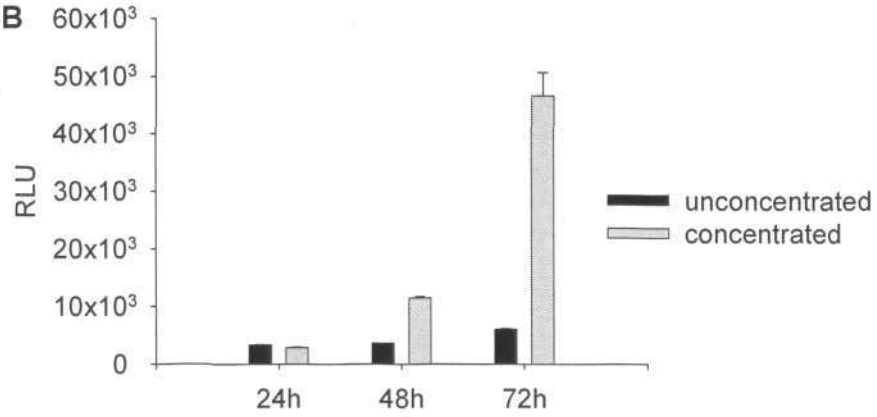


Figure 3-3.B. Reverse Transcriptase Based Virus Titer Determination Assay.

The S protein pseudotyped virus titer in culture media (black bar) and the concentrated pseudovirus containing culture media (gray bar), were examined by using the EnzChek® Reverse Transcriptase Assay Kit (Invitrogen, Singapore). The operation was done according to the product manual. The pseudovirus viral titer is shown as relative reverse transcriptase unit.

The HIV-1 based pseudovirus was generated with a luciferase gene as a reporter. The viral infectivity was reflected as the activity of luciferase expressed in the pseudovirus infected cell. The luciferase activity can be determined by a luciferase assay. There were several ways to enhance the viral infectivity. One way was to enhance the pseudovirus particles binding to the cellular surface. Here, we tried DEAE-Dextran and polybrene, the commonly used reagent to enhance viral binding. The results indicated that DEAE-Dextran and polybrene increases the viral infectivity of the pseudovirus (Figure 3-3.C). Using the same concentration, it seems that the polybrene had better effect comparing the DEAE-Dextran.

Figure 3-3.C

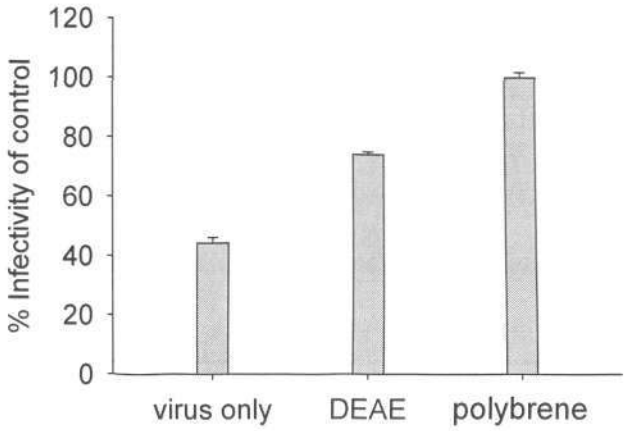


Figure 3-3.C. The Optimization of the S Protein Pseudotyped Virus Infection

The S protein pseudotyped virus containing media alone, pseudovirus containing media with polybrene (5 μg/ml), and pseudovirus containing media with DEAE-Dextran (5 μg/ml) was used to infect the 293T-ACE2 cell line. The infectivity of pseudovirus is proportion to the expression level of the luciferase gene it carries. The activity of luciferase was measured by the Luciferase Assay System (Promega, Singapore) according to the product manual.

The S protein Mediated Viral Entry Pathway

There are two main two pathways of enveloped viruses applied for viral entry. One is the direct fusion of the viral membrane and plasma membrane at neutral pH on the cellular surface. The other is receptor mediated endocytosis pathway, which required the low pH inside the endosome to trigger the membrane fusion. Ammonium chloride is one of the vacuolar acidification inhibitors. It can block the entry process of the virus which undergoes the endocytosis pathway (Sieczkarski and Whittaker 2002). Here, we used the SARS-COV S protein pseudotyped virus to infected ammonium chloride pre-treated 293-ACE2 cells. The vesicular stomatitis virus G (VSV-G) protein mediates viral fusion at mildly acidic pH. The mouse amphotropic murine leukemia virus (AMLV) envelope protein mediated the pH-independent viral entry process. So, VSV-G and AMLV Env pseudotyped HIV-1 based virus particles HIV (VSV-G) and HIV (AMLV) were also generated as controls to infect the ammonium chloride pre-treated 293-ACE2 cells.

As expected, the HIV (VSV-G) and HIV (AMLV) demonstrate pH-dependent and independent entry. The viral entry of SARS-COV S protein pseudotyped virus HIV (SARS) was affected by ammonium chloride (Figure 3-4). No obvious viral infectivity was detected when the target cell 293-ACE2 is treated with 25 mM ammonium chloride. This result was consistent with other researchers' work (Nie, Wang et al. 2004; Yang, Huang et al. 2004). The SARS-COV S protein pseudotyped virus displayed the pH dependent viral entry manner, suggesting that the SARS-CoV may use endocytosis pathway for viral entry.

Figure 3-4

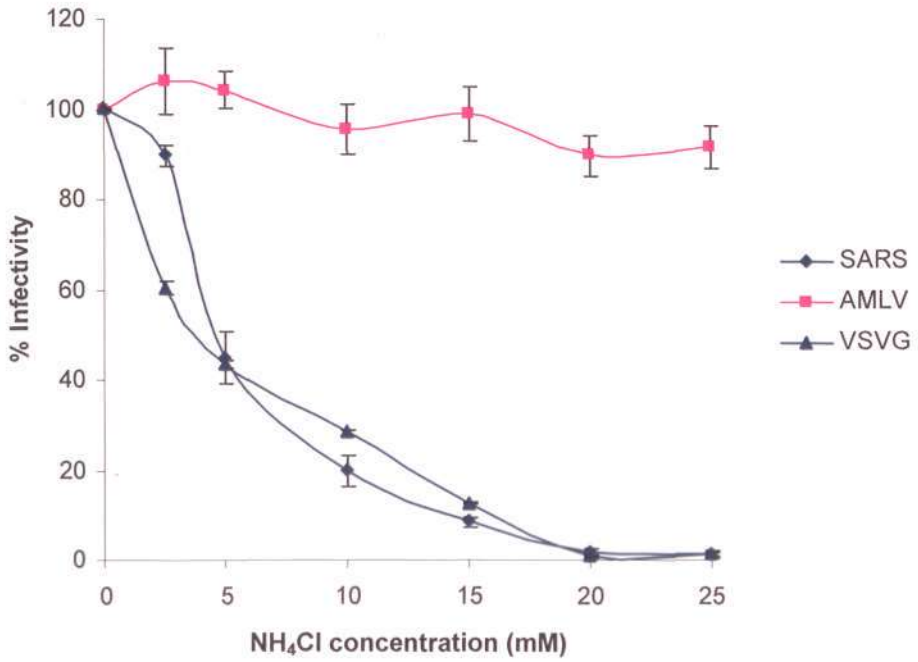


Figure 3-4. pH Dependent Viral Entry of the SARS-CoV S Pseudovirus.

The SARS-CoV S protein (SARS), Amphotropic murine leukemia virus 4070A envelop protein (AMLV) and vesicular stomatitis virus G protein (VSVG) pseudotyped HIV-1 virus, were used to infect the 293-ACE2 expressing cell under different concentration of the NH_4Cl. The viral infectivity was reflected by the activity of luciferase reporter gene and determined by luciferase assay. The results were shown in the form of mean ± SD, which represented the results from 3 different experiments.

S Protein Mediated Cell-Cell Fusion

The expression of the coronavirus spike protein may lead to cell-cell fusion in the host cell it expressed (Gallagher, Escarmis et al. 1991; Li and Cavanagh 1992; Gombold, Hingley et al. 1993; Bosch, van der Zee et al. 2003), while there is less syncytium observed in our study when we express the S protein in 293T cell. Western blot and immunofluorescence assay in previous experiment confirmed the expression of the S protein in 293T cell, but we desired to further confirm the cellular surface expression of the S protein in the 293T cell and observe its expression profile. To do this, confocal microscope was used to observe the immunofluorescence labeled S protein on the cellular surface. In brief, the 293T were transfected pTSh expression vector for S protein expression. The transfected cells were incubated with the rabbit anti-SARS-COV S protein serum and then conjugated with FITC-labeled anti-rabbit secondary antibody. Under confocal microscope (Figure 3-5), the S protein seems to be evenly distributed on the 293T cell surface. Observing under the light phase, the fluorescence portion co-localized with the plasma membrane. The result confirmed the cellular membrane integration of the S protein and its surface expression on the 293T cell. The cellular surface expression of ACE2 was also observed under confocal microscope. Goat anti-ACE2 antibody is commercial available. The 293-ACE2 cell line was incubated with the anti-ACE2 antibody followed by the FITC-labeled secondary antibody conjugation. Under confocal microscope, the fluorescence portion of the 293-ACE2 cell was co-localized with plasma membrane in light phase (Figure 3-5). The result confirmed that the receptor of SARS-CoV, ACE2 was expressed in the 293 cell surface.

Figure 3-5

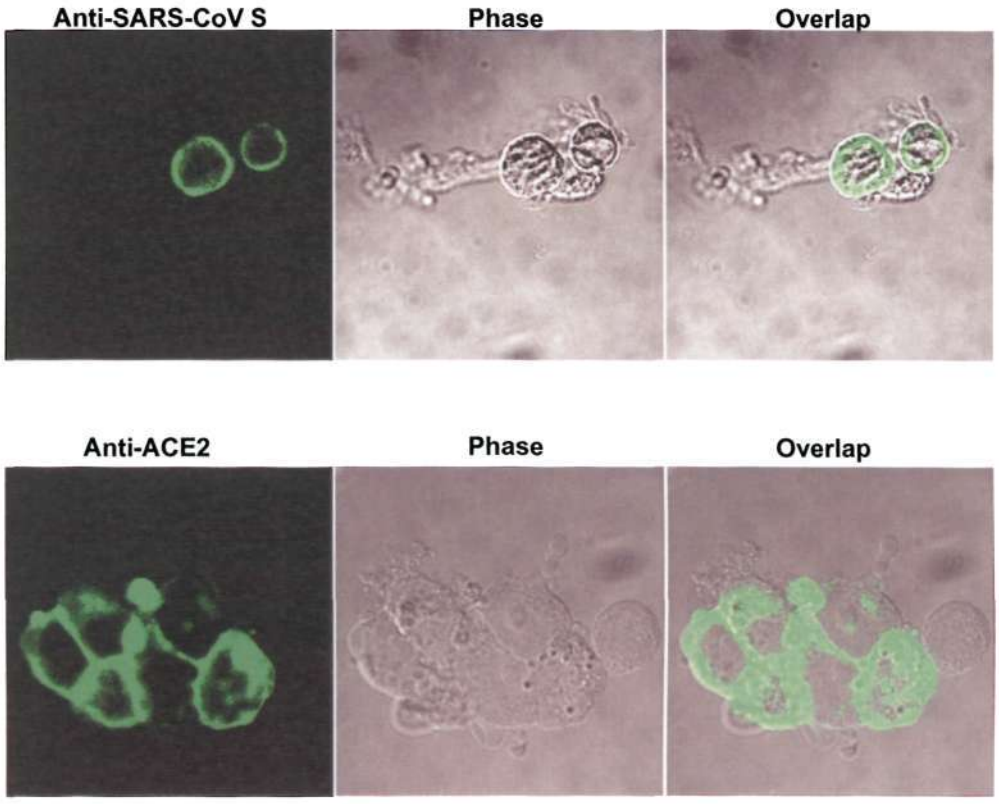


Figure 3-5. Confocal Microscope Analysis of the Surface Expression of the ACE2 and S Protein.

The 293T cell was transfected pTSh S protein expression vector using Lipofectamine 2000 (Invitrogen, Singapore) and grown on the poly-L-Lysine pre-treated chamber slides. 24 hours post transfection, the transfected cell was stained with the rabbit anti-SARS-COV S protein serum. The 293-ACE2 cell line was stained with the goat anti-ACE2 antibody. The FITC-labeled secondary antibody was used as a conjugate. The picture of fluorescent cells was taken using ZEISS LSM510 META Axiovert 200M confocal microscope.

After confirming the surface expression of both the S protein and ACE2, we further studied the membrane fusion activity mediated by the SARS-COV S protein. As mentioned previously, the proteolytic cleavage is important for priming the class I viral fusion protein. The HIV (SARS) pseudoviruses demonstrated the pH dependent manner for viral entry, suggesting the S protein mediated viral entry was affected by the pH. So, we focused mainly on how the proteolytic event and the pH affect the S protein mediated membrane fusion. We studied on these two factors through a developed cell–cell fusion assay.

The pTSh was transfected to the 293T cell and the cell was cultured 24 hours allowing the S protein to express on cell surface. The S protein expression cells were then detached from the culturing dish by EDTA treatment. The receptor expressing cell, 293-ACE2, were cultured in the pH pre-adjusted culture media at pH 5, 7 and 9. The detached cells were mixed with the 293-ACE2 cell in the pH pre-adjusted media expecting cell-cell fusion. Usually, the cell-cell fusion leads to the syncytia formation. To monitor syncytia formation, Hoechst was used to stain the nucleolus for facilitating observation. No cell–cell fusion occurred at pH 5 or pH 7 or pH 9 when the EDTA treated the S protein expressing cell was mixed with the 293-ACE2 cell line (Figure 3-6, group A). We then treated the S protein expression 293T cell with trypsin at concentration of 15 μg/ml. Similar concentration is used to activate the influenza HA (Klenk, Rott et al. 1975). The trypsin treated S protein expressing cells were then mixed with 293-ACE2 cells. As shown in Figure 3-6 group B, obvious cell–cell fusion appeared after trypsin treatment, no matter what the pH was in our test here. The fluorescence microscope observation showed that the nuclei seem clustered together in the syncytia induced by the trypsin treated S protein. The giant clustered nuclei were easy to recognize under fluorescence microscope. Judging by the giant

clustered nuclei, average 30-40 syncytia can be observed in one well in the 12-well culture dish. The results demonstrated that the S protein mediated cell–cell fusion was enhanced by the trypsin treatment, while the tested pH had no obvious effect on this process.

Figure 3-6

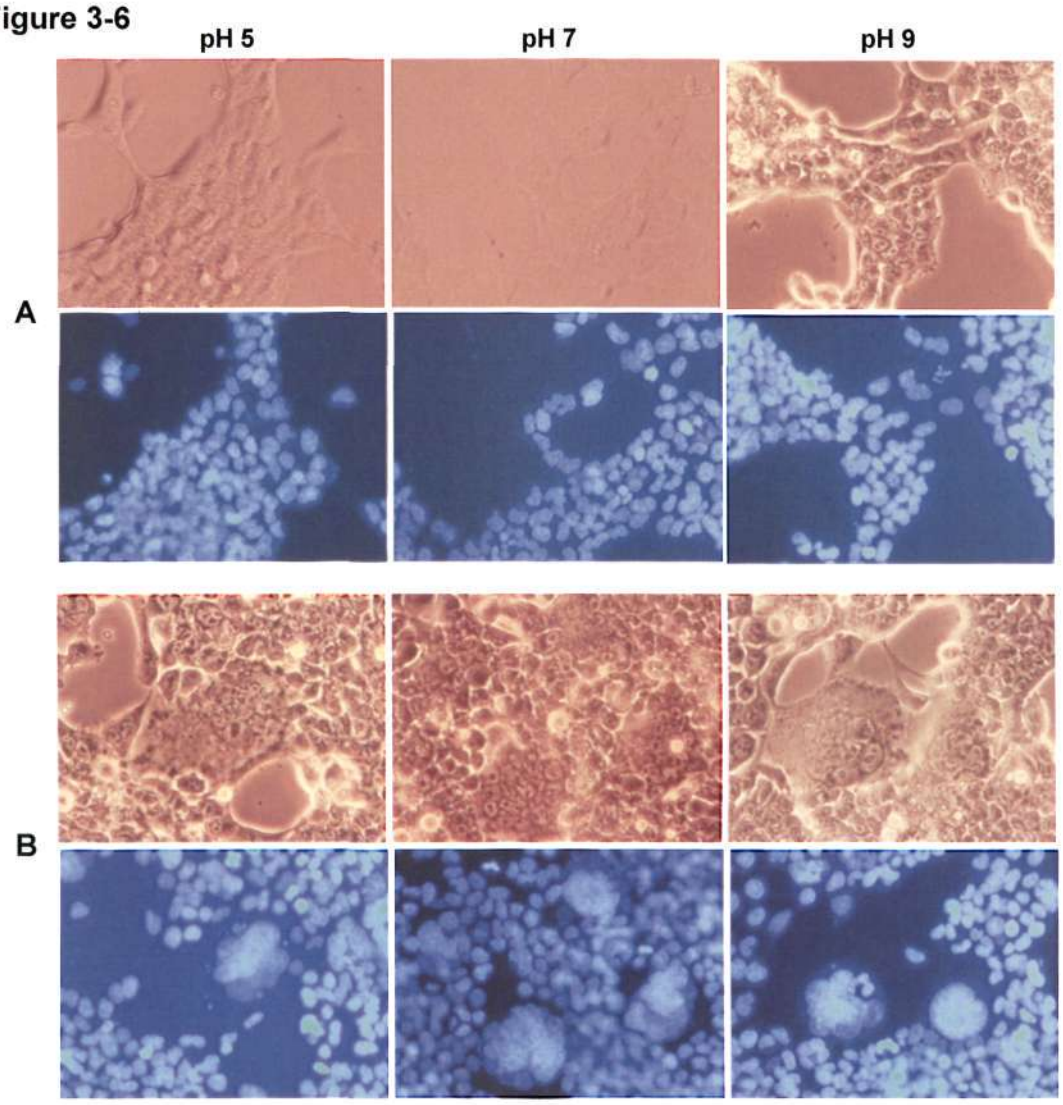


Figure 3-6. SARS-CoV S Protein Mediated Cell-Cell Fusion.

The pTSh SARS-COV S protein expression vector was transfected to 293T using Lipofectamine 2000 (Invitrogen, Singapore). 24 hours post transfection shows that the 293T cell was treated with EDTA (group **A**) or treated with TPCK treated trypsin protease (group **B**) for 15 minutes. The treated cells were mixed with the 293-ACE2 cell line in pre-adjusted pH 5, pH 7 and pH 9 DMEM media with 10% FBS and further cultured. 15 minutes later, the tissue culture media were changed back to normal DMEM media with 10% FBS. 24 hours later, the cell was stained by Hoechst (Invitrogen, Singapore) and observed under the fluorescence microscope. Photos were taken using an Olympus IX70 fluorescence microscope.

Through the trypsin treated S protein induced cell-cell fusion, we wanted to confirm that the cell-cell fusion was originated from the S protein expressing cell with the 293-ACE2 expressing cell. Another cell-cell fusion assay was performed here. Firstly, the pTSh expression vector was transfected to the 293T cell allowing the S protein expression. 24 hours post transfection, the cell was labeled with the fluorescent probe CMFDA, which stains the cell green color under the fluorescence microscope. The 293-ACE2 cell was labeled by fluorescent probe CMTMR, which gives the stained cell red color under the fluorescence microscope. The CMFDA stained S protein expressing cells were treated by EDTA and trypsin respectively and mixed with CMTMR stained 293-ACE2 cells in different pH pre-adjusted culture media as described before. The syncytia were observed under confocal microscope after the Hoechst nucleus staining. Consistent with previous fluorescence microscope observation results, there were no obvious syncytia formed after the EDTA treatment. In contrast, after TPCK trypsin treatment, the syncytia formed by the S protein expressing cell and 293-ACE2 cell can be easily observed under confocal microscope in all the pH we tested here (pH 5, 7 and 9) (Figure 3-7). The non-fusion cells were distinctly stained with a red or green color. The syncytia were stained with both red and green colors. So, they displayed yellow color in the overlap images from different channels. The nuclei staining showed that there exist clustered nuclei inside the syncytia. The giant fusion cell can also be observed under the light phase. The results confirmed that the syncytia originated from the cellular membrane fusion of the S protein expressing 293T cell and 293-ACE2 cell. The result also showed that the S protein mediated membrane fusion occurred at acidic, neutral, or basic pH, once it was cleaved by the trypsin. The results indicated that the proteolytic events were a crucial factor for the SARS-COV S protein mediated membrane fusion process,

suggesting that the proteolytic treatment was important for priming the SARS-CoV S protein, as it does in the other class I viral fusion protein. The pH had no notable effect on the S protein mediated cell-cell membrane fusion.

The advantage for the cell-cell fusion assay here was obvious. The fluorescent probe stain was fast and stable. The staining process can be shortened to 30 min while the fluorescent can maintain more than 2 days, which also avoid the multiple steps of washing and antibody labeling in immunofluorescent assay. This method allowed us to observe the cell-cell fusion 1 or 2 hours after the experiment when the cells were attached to the plate. Besides, the cell-cell fusion can be measured using flow cytometry and provides quantitative results. So this method was also used by the other research group to study the S protein mediated cell-cell fusion (Simmons, Reeves et al. 2004; Sha, Wu et al. 2006). Another cell-cell fusion assay might provide even more accurate quantitative results, which involved using the cells expressing T7 polymerase (Bosch, van der Zee et al. 2003; Belouzard, Chu et al. 2009; Madu, Roth et al. 2009). The cell-cell fusion can be measured by testing the activity of luciferase expressed under T7 promoter. The inhibitory effect of peptide to the cell-cell fusion can also be measured in this system. We tried to setup stable cell line using 293 cell and we could not get the suitable cell clone which can give high and stable level of T7 polymerase expression. So, we did not use this cell-cell fusion assay in our experiment.

Figure 3-7

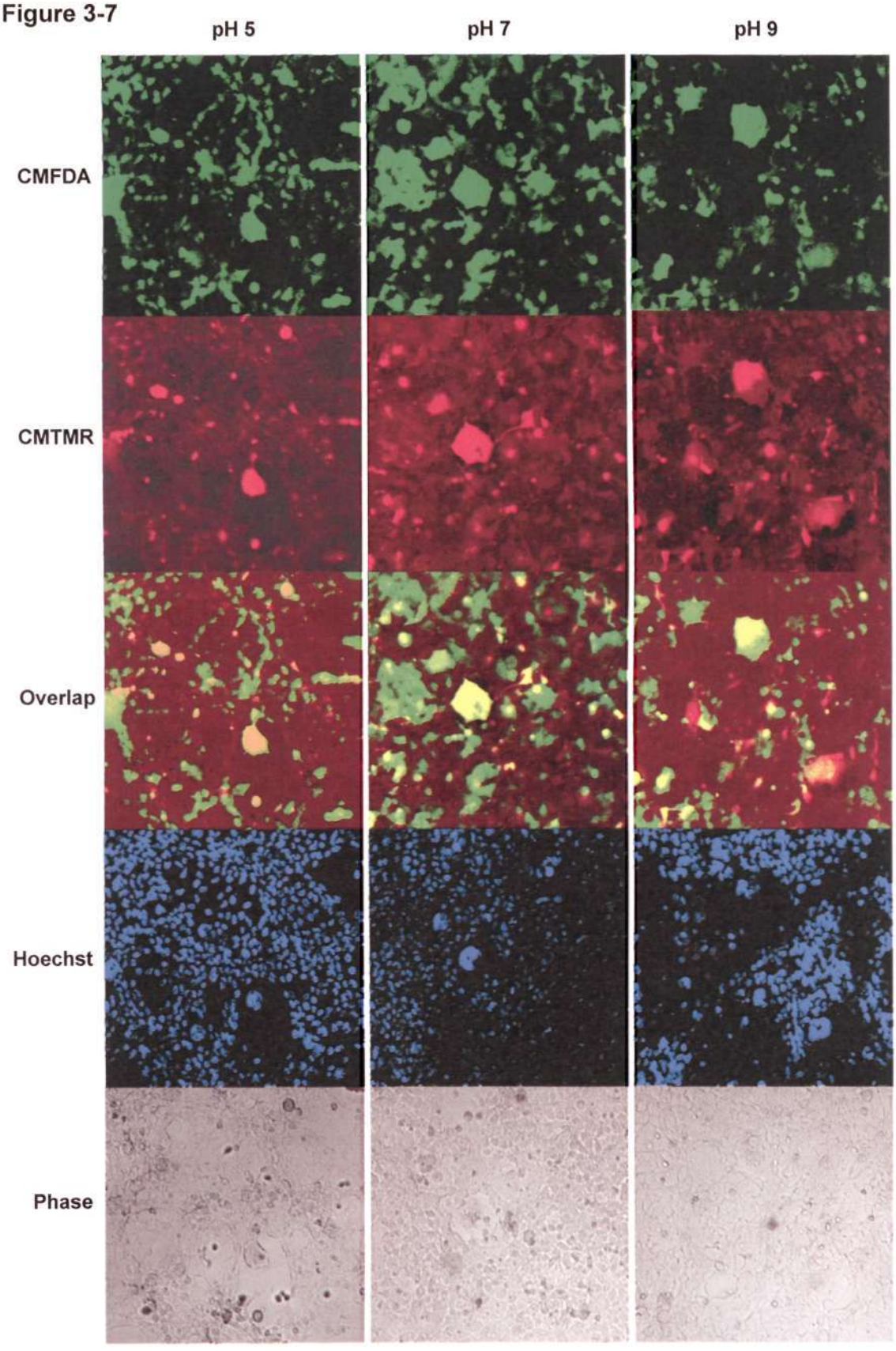


Figure 3-7. Confocal Microscope Analysis S Protein Mediated Cell-Cell Fusion

The pTSh SARS-COV S protein expression vector was transfected to 293T using Lipofectamine 2000 (Invitrogen, Singapore). 24 hours post transfection, the 293T cell was stained with the CMFDA probes (Molecular Probes, Singapore) according to the product manual. CMFDA are fluorescent chloromethyl derivatives that freely diffuse through the membranes of live cells. It will exhibit bright, green fluorescence in the cytoplasm at all physiological pH levels. The 293-ACE cell line was stained with the CMTMR probes (Molecular Probes, Singapore). CMTMR are also fluorescent chloromethyl derivatives which can exhibit bright, red fluorescence in the cytoplasm at all physiological pH levels. The transfected 293T cell was treated with TPCK treated trypsin protease for 15 minutes and then mixed with the 293-ACE2 cell line in the pre-adjusted pH 5, pH 7 and pH 9 DMEM media with 10% FBS and further cultured for 30 minutes. After that, the mixed cells were cultured in poly-L-lysine pre-coated chamber slides using normal DMEM media with 10% FBS. 24 hours later, the cell was stained with Hoechst (Invitrogen), covered by mounting media (DAKO Cytomation, Danmark) and observed under the confocal microscope. Photos were taken using a ZEISS LSM510 META Axiovert 200M confocal microscope.

Proteolytic Cleavage of the SARS-CoV S Protein by Trypsin

As described in previous experiment, the trypsin treatment enhanced the S protein mediated membrane fusion. We wondered what kind of changes happen on trypsin treated S protein. So, the trypsin treated S protein was examined by the western blot using the rabbit anti-SARS-COV S protein. To avoid the intervention of the immature form of S protein expressed in the transfected cell, we chose the S protein pseudotyped virus particle to treat with trypsin. The S protein integrated into the pseudovirus particle should be the mature form and it should stretch out like a spike on the pseudovirus membrane, which was similar to the stretched out S protein expressed in the cellular surface. The S protein pseudotyped virus was treated with trypsin at different concentration level for the same time period, as we treated the S protein expressing cell. The trypsin digestion was quenched by adding the SDS-PAGE loading buffer and boiling at 100°C for five minutes. The digested S protein samples were loaded to the SDS-PAGE for the western blot analysis. The rabbit anti-SARS-COV S protein serum was used for detection. The anti-serum was raised by injecting the rabbit with the SARS-COV S protein fragment from the amino acid 842 to 1175. It should recognize the S protein C-terminal half portion, which used to be named the S2 part of the S protein. The untreated S protein appeared as a 200 kDa band in the SDS-PAGE gel. After the trypsin treatment for 15 minutes, the 200 kDa full length S protein band and a band around 87 kDa appeared (Figure 3-8.A). It suggests that the S protein was first digested into two main parts, the 113 kDa portion, and the other was the 87 kDa portion, which corresponds to the S2 part of SARS-CoV S protein. Since the anti-serum used here only recognizes the S2 part, the 113 kDa was estimated by calculating the observed full length S protein and the released S2 part. The 113 kDa band was not detected on the western blot. The result also showed

that the S2 part can be further degraded into two parts, which appeared like two bands around 55 to 60 kDa. With the same digestion condition as we treated the S protein expressing cell, the S protein integrated in the pseudovirus was digested and released the 87 kDa S2 part, which suggested that the expressed S protein on the 293T cell surface was also digested and led to the exposure of S2.

Figure 3-8. A

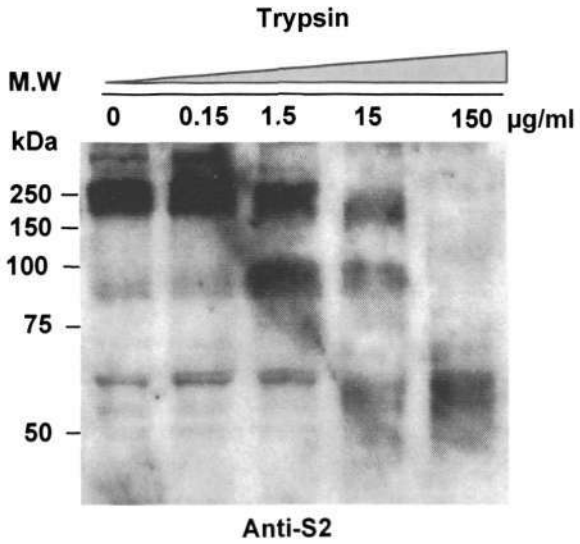


Figure 3-8.A Western Blot of TPCK Trypsin Treated S Protein on Pseudovirus Particles.

Ultracentrifugation concentrated S protein pseudovirus particles were treated by TPCK treated trypsin at different concentrations for 15 minutes. The digested samples were loaded to the SDS-PAGE for western blot analysis. The S protein fragments were detected by the rabbit anti-SARS-COV S protein antibody.

The SARS-COV S protein was one heavily glycosylated membrane protein. The glycosylation encumbered us to estimate the possible cleavage site by calculating the digested fragment size. The insect cell SF9 expressed S proteins were applied in our study, expecting less post translational modification. The SF9 cell expressed S protein is a HA tagged truncated form of the full length S protein. It was truncated at the amino acid 1188 to remove the membrane anchored portion and helped to release the expressed S protein. The SF9 expressed S proteins were treated by PNGase to remove the glycan chains first. The PNGase treated S protein samples were then digested by trypsin, using the same conditions as we used to digest S protein pseudotyped virus particles. The treated samples were then loaded to the SDS-PAGE for the western blot analysis. The rabbit anti-HA antibody was used for the detection of the C-terminal HA tagged S protein. The SF9 cell expressed S protein appeared a band around the 200 kDa (Figure 3-8.B). After PNGase treatment, the S protein showed a band near to the 160 kDa. The PNGase treated S protein was digested by trypsin and releases a fragment that appeared as a band around the 75 kDa. That fragment can be further degraded into several smaller fragments below 50 kDa. The SF9 cell expressed S protein was C-terminal HA tagged. The HA antibody recognized 75 kDa band was corresponding to the S2 portion of the SARS-CoV S protein. Estimating by using the observed fragment sizes at the same enzyme concentration as treated the S protein expressing cell, the SF9 expressed S protein should be digested into two fragments. One is the 75 kDa portion, the other is the 85 kDa portion corresponding to the S1 part, while that S1 part would not be detected by anti-HA antibody. Calculated by the observed size, combined with the sequence alignment result with the other coronavirus S protein, we believed the trypsin first digests the S protein around amino acid 660 to 672 of the full length S protein.

Figure 3-8.B

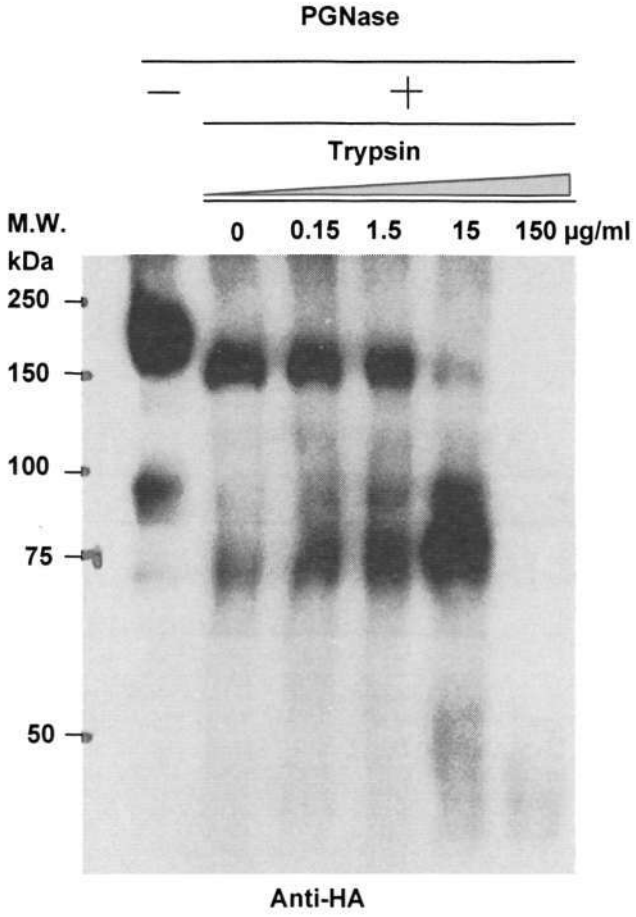


Figure 3-8.B Western Blot of TPCK Trypsin Treated Unglycosylated S Protein.

The insect cell expressed HA tagged SARS-COV S protein sample was treated by the PNGase to remove the glycans on S protein and then digested by the TPCK treated trypsin at different concentration for 15 minutes. The digested samples were loaded to the 8% SDS-PAGE for the western blot analysis. The HA tagged S protein fragments were detected by the rabbit anti-HA antibody (Delta Biolabs, U.S).

Discussion

The SARS-CoV is a newly emerged pathogen, which has brought damages to human health worldwide. How this new virus may invade the host cell and how to prevent this process becomes a hot topic for people to study. In our study, we found the suitable level expression of the S protein was achieved by optimizing the expression codon. The surface expression of the S protein was confirmed both by immunofluorescence assay and confocal microscope observation. Based on this, we produced a HIV base S protein pseudotyped virus particle. The refractory cell line 293 was changed to a susceptible cell line 293-ACE2 by integrating the ACE2 gene. The 293-ACE2 cell line was able to be infected by S protein pseudotyped virus. Thus, it served as the target cell for infection in the following studies. The SARS-CoV pseudotyped virus and the 293-ACE2 cell line were used as an efficient and safe system in our research for analyzing the character of the SARS-CoV infection. In our study, we mainly focused on how the pH and proteolytic events affect the SARS-CoV S protein mediated viral fusion process.

To mediate the fusion of viral and cellular membrane, the viral envelope proteins need to go through a series of conformational changes (Eckert and Kim 2001). The receptor binding is one of the events used to trigger the conformational changes of viral envelope protein. In HIV, the receptor triggers the rearrangement of its envelope protein subunit and leads to direct membrane fusion on the cell surface in pH-independent manner. In the other virus, like influenza, the receptor binding is not sufficient to trigger membrane fusion. The low pH is another crucial factor for triggering fusion. The low pH environment required by viral fusion exists in the acidified endosome in endocytosis process (Marsh and Pelchen-Matthews 2000).

Inhibitors to endosomal acidification can block viral entry process of those viruses (Nash and Buchmeier 1997). In our study, the S protein mediated pseudovirus entry was inhibited by one of the lysosomotropic agents NH_4Cl in the target cell lines 293-ACE2. The result indicated that the SARS-CoV S glycoprotein mediated viral entry requires endosome acidification, suggesting that the SARS-CoV may go through an endocytosis pathway for viral entry.

The NH_4Cl inhibit the acidification of the endosome. The inhibitory effect of NH_4Cl to SARS-COV S protein mediated viral entry suggests that there is a correlation between the pH changing and the conformational change of the S protein. In the following study (Fig. 3-6 and Fig. 3-7), the S protein mediated membrane fusion seems not affected by pre-treatment at different pH, while the trypsin treatment greatly enhanced the S protein mediated cell-cell fusion. This result suggested that the protease cleavage may be sufficient to trigger the conformational change of S protein. Since the cell-cell fusion also occurred in the neutral pH, it indicated the S protein mediated membrane fusion could happen at the cellular surface in neutral pH if it is activated by protease. It also indicated that direct fusion pathway might be taken instead of the endosome pathway if proper protease exists.

Although trypsin does not exist in lung and there are many other types of protease existence in lung of the SARS-CoV infected patients. The trypsin here is used to mimic those cellular proteases which involved in the viral entry process. Recently research development supported our study (Matsuyama, Ujike et al. 2005). We found in those Japanese scientists' experiment, the degradation profiles of SARS-CoV S protein looked similar when S protein was treated by trypsin, dispase, elastase and thermolysin. So, the trypsin here can mimic some of the cellular proteases. More recent study on the cleavage of SARS-CoV S protein showed that the trypsin could

recognize both of the functional basic motifs on S protein, which is believed to be important for the activation of S protein fusogenicity as well as the virion incorporation (Belouzard, Chu et al. 2009). Our cell-cell fusion results suggest that protease treatment seems to be the crucial factor for the S protein mediated viral entry, but the pH is not. One possible explanation is that the conformational change of S protein induced by the pH might be reversible like VSV-G (Doms, Keller et al. 1987). Considering that our study is based on the S pseudotyped virus and the SARS-CoV receptor expressing cell 293-ACE2, we have to emphasize that the S protein binding with receptor the ACE2 will be the pre-condition for our study. No obvious syncytium is observed neither in trypsin treated S protein expressing cell nor ACE2 expressing cell alone. We are prone to propose another possible explanation for our experiment data here: The receptor binding triggers the conformational change of the S protein, which might help the S protein to expose its cleavage site. The cellular proteases activate the S protein's membrane-fusion ability by cleavage. The low pH will facilitate the conformational change induced by proteolytic event.

We considered three major factors during viral entry: the receptor binding, proteolytic cleavage and the pH. We believed the first two will be the dominant forces for the SARS-COV S protein mediated viral entry. There are accumulated experiment evidences to support this explanation. Through a cry-electron microscopy study, people find the ACE2 binding results in a structure change of the SARS-CoV S protein (Beniac, Devarennes et al. 2007). The proteolytic treatment to the S protein will bypass lysosomotropic agent mediated inhibition of the viral entry (Simmons, Gosalia et al. 2005). Different from Furin, there exists evidence to support that the cathepsin L might be an important activating protease for the SARS-CoV infection (Huang, Bosch et al. 2006). Although the pH has a mirror effect during the SARS-

COV S protein mediated viral entry, there also exists evidence that the low pH will affect the conformational state of the S protein ectodomain. This is consistent with our previous study results on the heptad repeat region. The pH will also affect the activity of the cellular protease that might affect of the cleavage of the S protein. This explains why the SARS-CoV S protein-mediated entry is sensitive to lysosomotropic agents. For example, the cathepsin L is more stable and efficient in acidic pH (Simmons, Gosalia et al. 2005). Further specific study is required to clarify these possibilities.

Many viral enveloped glycoproteins are cleaved by protease into two parts. In the influenza virus, the cleavage of the HA protein is necessary for its viral infectivity. In coronaviruses, like MHV, BCoV, and HCoV-OC43, the spike proteins are cleaved into two parts, S1 and S2. There usually exists a furin consensus sequence motif, RXR / KR in the S1–S2 junction region (Molloy, Anderson et al. 1999; Rockwell, Krysan et al. 2002) in coronavirus. Mutation on the furin protease site prevents cleavage and also diminishes the cell-cell fusion activity of MHV (Sturman, Ricard et al. 1985; Taguchi 1993; Bos, Heijnen et al. 1995; Yamada, Takimoto et al. 1998; de Haan, Stadler et al. 2004).

But in the SARS-COV S protein, the furin-like protease cleavage site is missing. Sequence alignment of SARS-COV S protein with the other coronavirus S protein traced down the corresponding sequence to the amino acid 660 to 672. Two basic amino acids were included in that area, which are the potential targets for trypsin cleavage. One of the S protein sample used in the trypsin cleavage assay comes from the insect cell SF9 expressed S protein, which is ectodomain of S protein truncated at amino acid 1188 with an HA tag. Recent research work about conformational states of the SARS-CoV S ectodomain also used SF9 cell expressed S protein(Li, Berardi et al. 2006), which is from amino acid 12 to 1190 with N-terminal

honeybee melittin signal sequence and a C-terminal histidine tag. Our S protein trypsin digestion profile looks similar to the trypsin cleavage profile in their research work. The trypsin cleavage site was mapped to amino acid 667 on the S protein. This result agrees with our study on the location of trypsin cleavage site.

Although we did not detect the cleaved form of the S protein from the transfected 293T cell, a published study indicates that the S protein of severe acute respiratory syndrome (SARS) is cleaved in the virus infected Vero-E6 cells (Wu, Shang et al. 2004). The 85 kDa S2 portion of the S protein detected in a virus infected cells is quite similar to the 87 kDa S2 band detected in our S protein pseudotyped virion trypsin digestion assay. There also exists evidence that the S protein might be cleaved by cellular protease. For example, the S protein was cleaved when the S protein pseudotyped virus was incubated with the ACE2 expressing 293T cells (Du, Kao et al. 2007). Here we might provide an explanation about why no cleaved form of S protein was observed in the S protein expressing cell. The putative cleavage site is not efficiently used, or it is not fully exposed before the receptor binding triggers the conformational change of the S protein. The S protein relies on cellular proteolytic treatment to fulfill its membrane fusion potential. Pre-treatment of the S protein before infection of the target cell reduces viral infectivity (Simmons, Reeves et al. 2004; Simmons, Gosalia et al. 2005). So, the S protein is barely cleaved before the receptor engagement. The cleavage may happen after virion release, either in the cell supernatant or inside the target cell.

Based on the SARS-COV S protein pseudotyped virus and the 293-ACE2 expressing cell, we characterized the SARS-CoV S protein-mediated entry process. It is believed that the receptor and the cleavage are crucial for the viral entry process, but there still exists the possibility that the SARS-CoV might use the other pathway

for viral entry. Further study is required to clarify the entry mechanism of the SARS-CoV.

Chapter 4

Design and Develop Trimeric Protein Mimetics derived from the SARS-CoV S Protein Heptad Repeat Region as a Viral Entry Inhibitor

Design and Develop Trimeric Protein Mimetics derived from the SARS-CoV S Protein Heptad Repeat Region as a Viral Entry Inhibitor

Introduction

Severe acute respiratory syndrome (SARS) is a lethal disease (Poutanen, Low et al. 2003; Tsang, Ho et al. 2003), which is caused by a novel coronavirus—SARS coronavirus (SARS-CoV) (Drosten, Gunther et al. 2003; Rota, Oberste et al. 2003). Although the outbreak of SARS is over, animal reservoirs of the SARS-CoV appear to exist. The possibility of its future reemergence remains a threat to public health. To counter this, efficacious drugs are required for prevention and treatment of this disease.

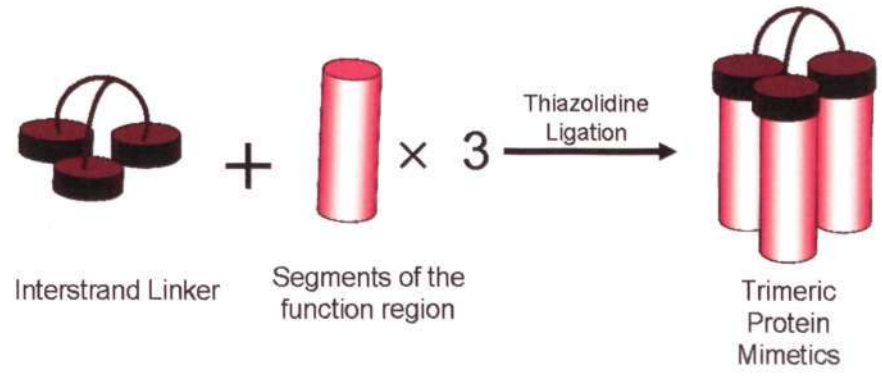
The spike protein of the coronavirus is a multifunctional protein that plays a central role in the pathogenesis of the coronavirus infections. The SARS-CoV S protein has been classified to a class I viral fusion protein. It is supposed to use the similar entry mechanism as the other viruses, which bear the class I viral fusion proteins. The initial stage of viral entry involves exposure of the fusion peptide. Exposed fusion peptide is inserted into the cellular membrane. With the fusion peptide inserted in the cellular membrane, while the transmembrane domain is still anchored in the viral membrane, the stretched HR1 and HR2 are like an opened hairpin. This intermediate used to be named the "pre-hairpin" structure. Subsequently, the HR2 is folded back to surround the HR1 in an anti-parallel manner to form a trimer of the hairpin-like structure (Bullough, Hughson et al. 1994; Lu, Blacklow et al. 1995; Chan,

Fass et al. 1997; Tan, Liu et al. 1997; Weissenhorn, Dessen et al. 1997; Caffrey, Cai et al. 1998; Weissenhorn, Calder et al. 1998; Weissenhorn, Carfi et al. 1998; Baker, Dutch et al. 1999). Considering the helical state of HR1 and HR2, the folded trimer structures are sometimes called the six-helix bundle, fusion core, or coiled-coil. This conformational change and the free energy released from this process are sufficient to facilitate juxtaposition of the viral and cellular membranes. Followed by the virus-cell membrane fusion, the viruses penetrate the cellular membrane and gain viral entry.

In vitro introduced HR peptides will serve as a competitor of the original HR counterparts. They will prevent the formation of the six-helix bundle, thus inhibiting the consequent membrane fusion. For example, the peptides analogs of the HR regions from the viral fusion protein in HIV-1 (Qureshi, Coy et al. 1990; Gallaher, Segrest et al. 1992; Wild, Oas et al. 1992; Wild, Greenwell et al. 1993), the Ebola virus (Watanabe, Takada et al. 2000), or paramyxovirus (Lambert, Barney et al. 1996; Young, Li et al. 1999), show an inhibitory effect on those virion infectivity. As a result, the HR region is a potential target for antiviral inhibitor or drug development. In fact, a peptide overlaps the HR2 and the aromatic domain of gp41 shows an inhibitory effect on the HIV-1 viral fusion down to the nanomolar. Thus, it became the peptide drug Fuzeon™ (DP178, T-20 or enfuvirtide), which is the first FDA approved HIV-1 entry inhibitor. These results provide insights for developing a new class of antivirus and stimulate scientists to continue their research interests to design peptide fusion inhibitors targeting other viruses, such as SARS-CoV (Garry 2003; Kliger and Levanon 2003).

Several groups show that analogs from the HR1 or HR2 region can inhibit the SARS-CoV entry (Bosch, Martina et al. 2004; Liu, Xiao et al. 2004; Yuan, Yi et al. 2004; Zhu, Xiao et al. 2004). However, the inhibition concentrations for those peptides are in the micromolar range (Bosch, Martina et al. 2004; Liu, Xiao et al. 2004). A possible explanation for the high viral inhibition dosage of those inhibitory peptides is due to the weak interaction between those inhibitory peptides to their targeted HR regions, or their proteolytic instability and propensity for aggregation in aqueous solutions for their amphiphilic nature. The design and development of the inhibitory peptides, with higher binding affinity / stability, might be a solution. Refer to the experience of the other researchers in the design of the HIV-1 entry inhibitors, protein mimetics might improve on those deficiencies of the monomers peptide inhibitors. Several laboratories have designed synthetic protein constructs of both HR regions of GP41 in HIV (Delwart, Mosialos et al. 1990; Fields and Noble 1990; Eckert and Kim 2001; Louis, Bewley et al. 2001; Root, Kay et al. 2001; Bewley, Louis et al. 2002; Louis, Nesheiwat et al. 2003; Sia and Kim 2003). Successful design of the quaternary protein mimetics will largely exploit the Leu / Ile heptad repeat motif and the propensity of the HR regions to associate (Eckert and Kim 2001; Louis, Bewley et al. 2001; Root, Kay et al. 2001; Bewley, Louis et al. 2002; Louis, Nesheiwat et al. 2003; Sia and Kim 2003). We plan to design the trimeric protein mimetics based on the HR2 region, expecting higher helicity and stability. The protein mimetics are supposed to bind with the pre-hairpin intermediate structure and block the viral fusion process (Figure 4-1).

Figure 4-1



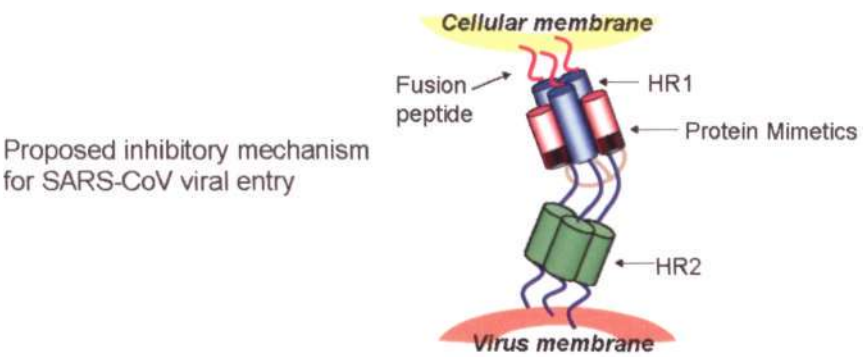


Figure 4-1. Design of the Trimeric Protein Mimetics and its Proposed Inhibitory Mechanism for the SARS-CoV Viral Entry

The aldehyde groups on the trivalent interstrand linker molecular will react with the N-terminal cysteine on the HR region derived peptides to form trimeric protein mimetics through thiazolidine ligation. The trimeric protein mimetics are supposed to bind with the pre-hairpin structure of the SARS-CoV S protein and block the conformational change of it. Thus, it may block the subsequent membrane fusion process and serve as the viral entry inhibitor.

A general and mild method for the conjugation of peptides through site-specific thiazolidine formation is developed in our laboratory. This method is based on the specific reaction between the aldehyde and aminothiol of the cysteine under acidic conditions to form a stable thiazolidine product. This ligation method is used to prepare the trimeric N- and C-peptides of gp41 as fusion-state protein mimetics in our laboratory.

Here, we report that mimetics design of the quaternary structures of the HR2 regions of SARS-CoV S protein by trimerization of the HR2 peptides with a scaffold. The quaternary mimetics show an increase in ordered structure, proteolytic stability, and inhibitory potency against the SARS-CoV S protein pseudotyped virus. The trimerized HR2 mimetics also mimic the trimeric fusogenic state of the HR2 in the S protein. In vitro study of the mimetics also provides clues for the study of the SARS-CoV viral entry mechanism.

Materials and Methods

Plasmid Construction

The pET24a expression vector is modified by inserting one Nde I and one Xho I digested short sequence including the N-terminal His tag to the Xho I endonuclease site from a modified pET32 expression vector, pET4517 (gift from Dr. Wang Wei, Institute of Materia Medica, China), to create the modified expression vector pET24am. The modified pET24am contains the N-terminal His Tag, thrombin cleavage nucleotide sequence, S tag, enterokinase cleavage nucleotide sequence, multiple cloning sites and the C-terminal His tag. The SARS-CoV (Singapore strain 2774) S protein heptad repeat region derived fragment sequences are subcloned into the pET24am expression vector using the following primers, with one Factor Xa cleavage site at the N-terminal. HR1C (amino acid 900-949, amino acid sequence: ENQKQIANQFNKAISQIQESLTTTSTALGKLQDVVNQNAQALNTLVKQLI) uses primer A (5'-CCG GGA TCC ATC GAG GGA AGG GAG AAC CAA AAA CAA ATC-3') and primer B (5'-TGG TGC TCG AGT TAG CTA AGT TGT TTA AC-3'). HR2 (amino acid 1151-1185, amino acid sequence: ISGINASVVNIQKEIDR LNEVAKNLNESLIDLQEL) use primer C (5'-ATC GGA TCC ATC GAG GGA AGG ATT TCA GGC ATT AAC-3') and primer D (5'-GTG CTC GAG TTA CAA TTC TTG AAG GTC-3'). The CL36 (amino acid 1151-1185 with one additional N-terminal cysteine, amino acid sequence: CISGINASVVNIQKEIDRLNEVAKNLNES LIDLQEL) use primer E (5'-ATC GGA TCC ATC GAG GGA AGG CGC ATT TCA GGC ATT AAC-3') and primer D. IL29 (amino acid 1151-1179) use primer C and primer F (5'-GTG CTC GAG TTA TGT GTT TAA TGC TTG-3'). The pfu DNA Polymerases (Fermentas, Singapore) are used in all PCR reactions.

Cell Lines

In the following, Dulbecco's Modified Eagle's Medium (DMEM) (Hyclone, Singapore), with 10% fetal bovine serum (Hyclone, Singapore), is mentioned as a normal culture media for all the cell culture. The 293T cells are cultured in normal culture media as the producer cell line for protein expression and pseudotyped virus production. The SARS-CoV receptor ACE2 expressing cell line 293-ACE2 is cultured in normal culture media, with 2 mg/ml geneticin.

Expression of SARS-CoV S Protein HR Fragment

Freshly transformed *E. coli* BL21 cells (Novagen, CA, U.S.) were grown in Luria Broth (LB) medium to log phase, (OD_{600}=0.6) and subsequently induced by adding the IPTG (isopropyl-β-D-thiogalactopyranoside, GIBCO BRL, U.S.) to a final concentration of 0.3 mM. Eight hours later, the bacteria were spun down using a JA-10 rotor in Avanti^{TM} centrifuge (Beckman Coulter, U.S). They are resuspended in 1/20 volume of the column buffer (50 mM NaH_2PO_4, 300 mM NaCl, pH 8.0) and homogenized by the Microfluidizer processor (Microfluidics, U.S). The cell debris was removed by centrifugation at 20000 rpm for 15 minutes in a JA-22.5 rotor in the Avanti^{TM} centrifuge. The supernatant was loaded onto the ProBond resin (Invitrogen, Singapore) preloaded column for affinity purification. The protein-loaded column was then washed with 20 column volume of column buffer and eluted with the column buffer containing imidazole with increasing concentration. Fractions containing the His-tagged protein were identified using the SDS-PAGE.

These fractions were concentrated using the Amicon Ultra filter devices (Millipore, Singapore) and changed the buffer to protease Factor Xa digestion buffer (20 mM Tris·Cl, 50 mM NaCl, 1 mM CaCl2, pH 6.5). The affinity purified His-

tagged HR fragments were digested by the Factor Xa at room temperature. The digested HR2 fragments were further purified using reversed-phase high pressure liquid chromatography (RP-HPLC).

Solid-Phase Peptide Synthesis

All amino acids and coupling reagents were purchased from Novabiochem (San Diego, CA). Two synthetic versions of unprotected peptides, IL25 and CL26, were prepared. They are peptides from amino acid 1161-1185 on SARS-COV S protein without and with an additional cysteine at their N-terminals. By solid phase chemistry on a trifluoroacetic-acid (TFA)-labile Fmoc-Rink-amide-resin using 9-fluorenylmethoxycarbonyl (Fmoc) chemistry, 95% TFA / 2.5% Water / 2.5% TIS (triisopropylsilane) was used to cleavage and remove all of the protecting groups of synthesized IL25. 94.5% TFA / 2.5% Water / 2% EDT (1, 2-ethanedithiol) / 1% TIS was used to cleavage and remove all of the protecting groups of synthesized CL26. After being cleaved for four hours, the cleavage supernatant was precipitated by ice-cold diethyl ether in 1:10 volume. The precipitates were spun down and freeze dried. The crude peptides were further purified by the C18-reverse phase HPLC.

Scaffold Production

The scaffold precursor Ser^{1}-βAla^{2}-Lys^{3}(Ser)-Lys^{4}(Ser)-Ser^{5}-Ser^{6}-Ala^{7} was prepared by Fmoc chemistry. The tripeptide C-terminal tail, Ser^{5}-Ser^{6}-Ala^{7}, was added to the C-terminal of the branched tripeptide scaffold βAla^{2}-Lys^{3}(Ser)-Lys^{4}(Ser), to increase solubility and facilitate the HPLC purification. The synthesis of this C-terminal tail Ser(tBu)-Ser(tBu)-Ala was assembled stepwise by Fmoc-chemistry on the Rink resin. The scaffold was then assembled by Lys^{4} with Fmoc-Lys(Mtt). It was continued further with Fmoc-chemistry, sequential couplings of the Fmoc-Lys(Mtt),

and Fmoc-βAla and Fmoc-Ser(tBu) that protect the hexapeptide Fmoc-Ser(tBu)^1-Lys(Mtt)^2-Lys(Mtt)^3-Ser(tBu)^4-Ser(tBu)^5-Ala^6. The Mtt protecting groups on the Lys were moved by 2% TFA / 1% TIS in DCM for 20 minutes and followed by 1% TFA / 1% TIS in DCM for 20 minutes, two times. The deprotected Lys was coupled with the Fmoc-Ser(tBU). The protecting groups Fmoc on all the three N-terminal residues were removed by the 20% piperidine / DMF. The 95% TFA / 2.5% Water / 2.5% TIS treatment cleaved the scaffold precursor from the resin. The cleavage supernatant was precipitated by ice-cold diethyl ether in 1:10 volume. The precipitates were spun down and freeze dried. The crude peptides were further purified by the C18-reverse phase HPLC.

Sodium Periodate Oxidation

The aldehyde-functionized scaffold was generated efficiently by the oxidation of the N-terminal Ser with five molar equivalents of sodium periodate added to the precursor in an aqueous solution (100 μl/mg) buffered at pH 5.5 by 0.2 M acetate (Geoghegan and Stroh 1992). After 20 minutes, the reaction mixture was purified by the HPLC to remove the side product formaldehyde formed in the reaction to afford a 90% yield of the aldehyde-scaffold, which was used immediately for the thiazolidine ligation to form the trimeric protein mimetics.

Thiazolidine Ligation

Thiazolidine ligation forming the trimeric protein mimetics was performed using a 3.3:1 molar of the Cys-peptides, with the aldehyde-scaffold in a deaerated mixture of 10% acetonitrile and 90% water containing 0.05% TFA at pH 2.5. The reaction was kept under nitrogen and monitored by the HPLC. The reaction was

completed at 48 hours. The major product was in the range of 40-60% yield after purification by the HPLC.

Chromatography

The HPLC analysis was performed using Shimadzu chromatography system (LC-10ATVP, Shimadzu, Japan) equipped with a photodiode array detector (SPD-M10ATVP). The RP-HPLC uses Vydac C4 and C18 column (4.6 × 250 mm, 10 × 250 mm, Grace, U.S) with a low rate of 1 or 2 mL/min and Buffer A (aqueous 0.05% TFA) and Buffer B (0.05% TFA in 90% acetonitrile). A gradient of 0-20% Buffer B, over 20 minutes, is used to separate the scaffold and its precursor. A 20-70% Buffer B, over 40 minutes, is used to separate the other products with UV monitoring at 220 nm.

CD measurement

The CD spectra was obtained on a Chirascan CD spectropolarimeter (Applied Photophysics, Leatherhead, UK) using a quartz cuvette of 0.1 mm optical path length at room temperature. Scans were conducted from 185 to 260 nm, with an integration time of 0.2 second and a bandwidth of 1 nm. The protein concentration used for this study was 1 mg/ml. A quantitative evaluation of fractional helicities was based on the mean residue ellipticities according to the method proposed by Wu (Wu, Ikeda et al. 1981).

Proteinase K Treatment

The proteinase K (Fermentas, Singapore) at concentration 0.5 μg/ml and 0.125 μg/ml were used to treat peptides and protein mimetics. The incubation time was 20 minutes. The digestion monitored by the HPLC was quenched by loading the digestion samples for the HPLC analysis with buffer A containing 0.05% TFA.

Viral Inhibition

To produce the HIV-luc (SARS) pseudotyped virus, 2×10^6 293T cells were cotransfected with 10 μg of pNL-4-3E-R-Luc (HIV-luc) and 10 μg codon optimized SARS-CoV S protein expression vector using the calcium phosphate method. Culture media was replaced 12 hours after transfection. After 48-72 hours, the culture supernatant was collected and filtered through a 0.45 μm membrane. For the inhibition assay, the HIV-Luc (SARS) pseudotyped virus was incubated with serials of diluted peptides or protein mimetics at 37°C for 30 minutes. The virus and peptide mixture was transferred to 96 well plates pre-seeded with 293-ACE2 cells (3×10^3 cells per well). Each concentration was tested in 5 wells. The media was replaced 12 hours after infection. The cells were lyzed for luciferase assay 48 hours after infection. The luciferase activity was measured by Luciferase Assay Kit (Promega, Singapore) on the TD-20 / 20 Luminometer (Turner biosystems, Singapore).

Hemolytic Assays of Peptides and Protein Mimetics

The peptide and protein mimetics were dissolved in water and serially diluted in phosphate buffered saline (PBS). The assay was performed by adding 20 μl of sample to 80 μl of 1% suspension of the washed human type O red blood cells (RBCs) in PBS. The plate was incubated at 37°C for one hour and centrifuged at 1,000 rpm for six minutes. Aliquots of 75 μl were transferred to a 96 well U bottomed microtiter plate. (Nunc, Singapore). The absorbance was measured at 415 nm with an automatic plate reader Safire2™ (Tecan, Singapore). The level of hemolysis was calculated as the percentage of maximum lysis (1% Triton X-100 control) after adjusting for minimum lysis (PBS control). Synthetic melittin (Sigma) was used for comparison.

Results

Synthesis of SARS-CoV S Protein HR Fragments

The SARS-CoV S protein HR fragments (Figure 4-2) were prepared by the chemical method and biological method. Among them, the IL25 (amino acid 1161-1185) and the CL26 (IL25 with N-terminal cysteine) were prepared by solid phase peptide synthesis. The HR1C (amino acid 900-949), the HR2 (amino acid 1151-1185), the CL36 (HR2 with N-terminal cysteine), and the IL29 (amino acid 1151-1179) were generated by the protein expression in *E. coli*.

For large scale preparation of peptides by protein expression in *E. coli*, efficient expression systems, which can give higher expression level of soluble protein, are in demand. Also, for the peptides which will be used for ligation next step, the first amino acid at N-terminal must be Cys, other than the Met. For these two major concerns, the expression system with the N-terminal fusion tag for affinity purification might be a choice. The MBP fusion protein system (pMAL system, New England BioLabs), the GST fusion protein system (pGEX-5X, Amersham), the Intein system (pTwin-1 system, New England BioLabs) as well as the pET system (pET24a, Novagen) were tested. Only the MBP fusion protein system gave a higher soluble protein expression level.

Figure 4-2

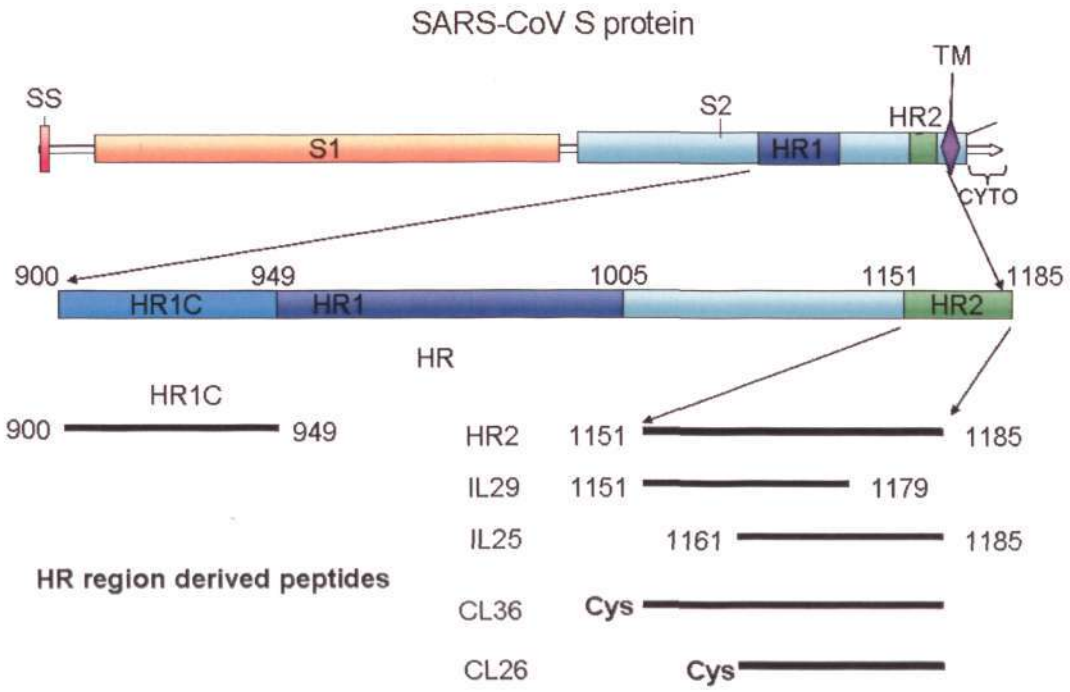


Figure 4-2. A Schematic View of the S Protein of the SARS-CoV and the HR Region Derived Fragments.

The location of the heptad repeats 1 and 2 (HR1 and HR2), the transmembrane domain (TM) and the cytoplasmic tail (CYTO) in the SARS-CoV S protein are shown. There is a signal sequence (SS) followed by a cleavage site at the start of the sequence. S1 and S2 stand for the homologic regions of SARS-COV S protein to the other coronavirus S protein S1 and S2 domain. Among the HR region, the HR1 starts from amino acid 900 to 1005. HR2 starts from amino acid 1151-1185. One region within HR1, is the binding site of HR2, and it is named HR1C (amino acid 900-949). Amino acid 1161-1179 is confirmed to be the helical region of HR2 by crystal structure. Several segments spanning that region of HR2 and HR1C are listed and prepared by chemical method or biological method. The CL36 and CL26 are the N-terminal cysteine version of the HR2 and IL25.

Although the corresponding fragments can be expressed by the pMAL system we used in chapter 2, pMAL system is not an ideal system for preparing large amount of peptides. MBP was 11 fold bigger than the peptide we expected. In theory, for every 100 mg of MBP (42.5 kDa), one gets about only 8.6 mg of the 35 amino acids HR2 peptide (4 kDa). The yields were even lower in practical operation. Because of this, we modified pET24am system. The previous HR fragments were expressed in the pET24a vector (Novagen, U.S) and failed to get a higher or more soluble protein expression. It is also hard to remove the C terminal tag and add the N-terminal cysteine. However, its short expression open reading frame can be exploited for the peptide expression. The pET24a vector was modified to insert the N-terminal His Tag, the thrombin cleavage nucleotide sequence, the S tag, and the enterokinase cleavage nucleotide sequence. To obtain the expected peptide, without introducing any unwanted amino acid, we added the Factor Xa cleavage at the N-terminal of each peptide. The factor cleavage site is only four amino acids and it cleaved at the last Arg. Thus, no additional amino acid was left on the peptides and it was easy to generate the N-terminal cysteine peptides. The whole expression frame for the HR2 inserted expression vector was only 90 amino acids (Figure 4-3.A). In theory, for every 100 mg purified protein, one could get about 39 mg of the 35 amino acids peptide HR2 (4 kDa). In addition, the high affinity of the His tag to the Ni-Chelat resin, provided a lot of convenience for the large scale expression.

Figure 4-3.A

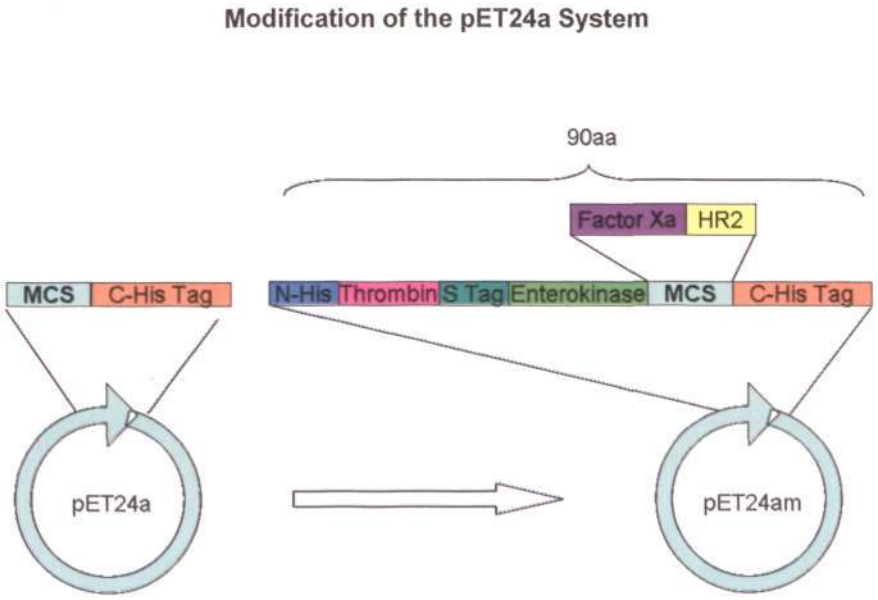


Figure 4-3. Modification of the pET24a Expression Vector and its Protein Expression

A. pET24am Vector is Modified to insert the N-terminal His Tag

The N-terminal His tag (N-His), Thrombin cleavage site (Thrombin), S tag, Enterokinase cleavage site (Enterokinase) nucleotide sequences were inserted into the pET24a expression vector. To avoid introducing any additional amino acid, the Factor Xa cleavage is added at the N-terminal of each peptide. The expression open reading frame of the modified pET24a vector was only 90 amino acids with the HR2 insertion.

HR fragment are obtained from the SARS-CoV (Singapore strand 2774) S protein containing vector pJX40-S by PCR and subcloned into the modified pET24a vector (pET24am) for expression. The pET24am vector gave efficient expression of the HR region derived fragments. Using the HR2 insertion vector pET24am-HR2 as an example, the pET24am-HR2 expressed the N-terminal His tagged HR2 fragment and gave a 10 kDa band in the Tricine / SDS-PAGE gel (Figure 4-3.B), which corresponded to the expected size of the 90 amino acids expression frame. Most of the induced proteins were in the supernatant. The other HR region derived fragments expressed by the pET24am vector were also in the soluble portion of the expressed protein. The purity of the eluted protein was suitable for protease digestion and further purification. For the short expression frame, the protease digestion was quite efficient compared with the pMAL system. In the pMAL system, the existence of a huge MBP in the N-terminal might cover the cleavage site and make the cleavage time several folds longer.

Figure 4-3.B

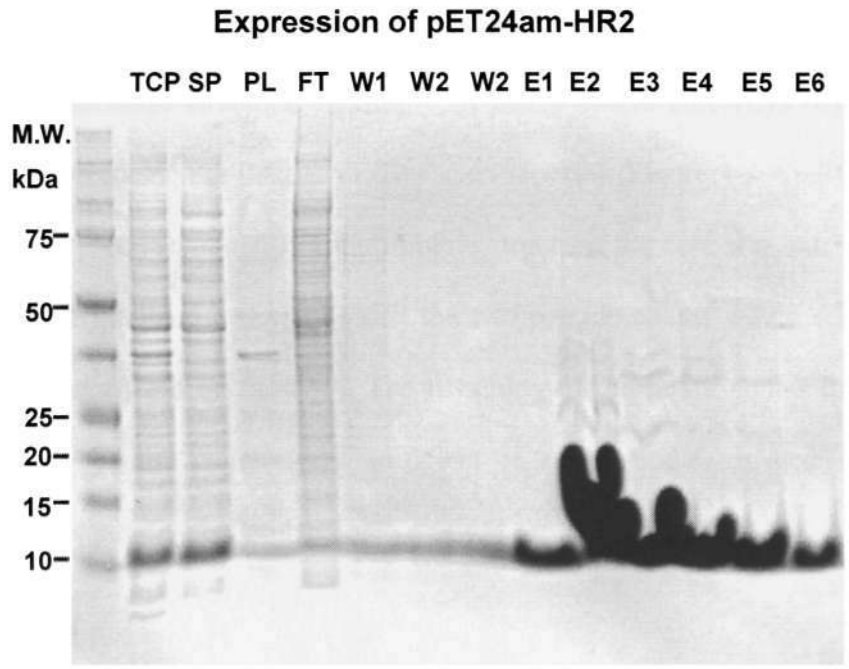


Figure 4-3.B. Expression and Purification Profile of the pET24am-HR2 in *E. coli*.

The HR2 insertion vector pET24am-HR2 was expressed in the *E. coli* BL21 and the lysate of bacteria is loaded onto the Tricine-SDS PAGE for analysis. TCP: Total cell protein lysate, SP: Supernatant, PL: Pellet, FT: flow through of the nickel column, W1 to W3: Washing fraction is 1 to 3. E1 to E6: Elution fractions 1 to 6 from the nickel column.

Design of the Trimeric Protein Mimetics

The design of trimeric protein mimetics was composed of two synthetic elements: a small peptidyl scaffold and monomeric Cys-peptides. The scaffold was designed to maximize structural flexibility. It had a di-Lys core (Lys^3-Lys^4 of the heptapeptide precursor Ser-βAla–Lys^3-Lys^4-Ser-Ser-Ala) (Figure 4-4.A). To provide flexibility and pseudosymmetry of the trimeric structure, the core was extended with the addition of the βAla at Lys^3, so that the two peptide chains would each be ten atoms from the α-atom on the Lys^3. The flexibility of the linker comes from many rotatable CH_2 bonds, which were sufficient to accommodate reverse turns for interlocking the putative coiled HR2.

Figure 4-4 A

Scaffold precursor sequence: Ser^1-βAla^2-Lys^3(Ser)-Lys^4(Ser)-Ser^5-Ser^6-Ala^7

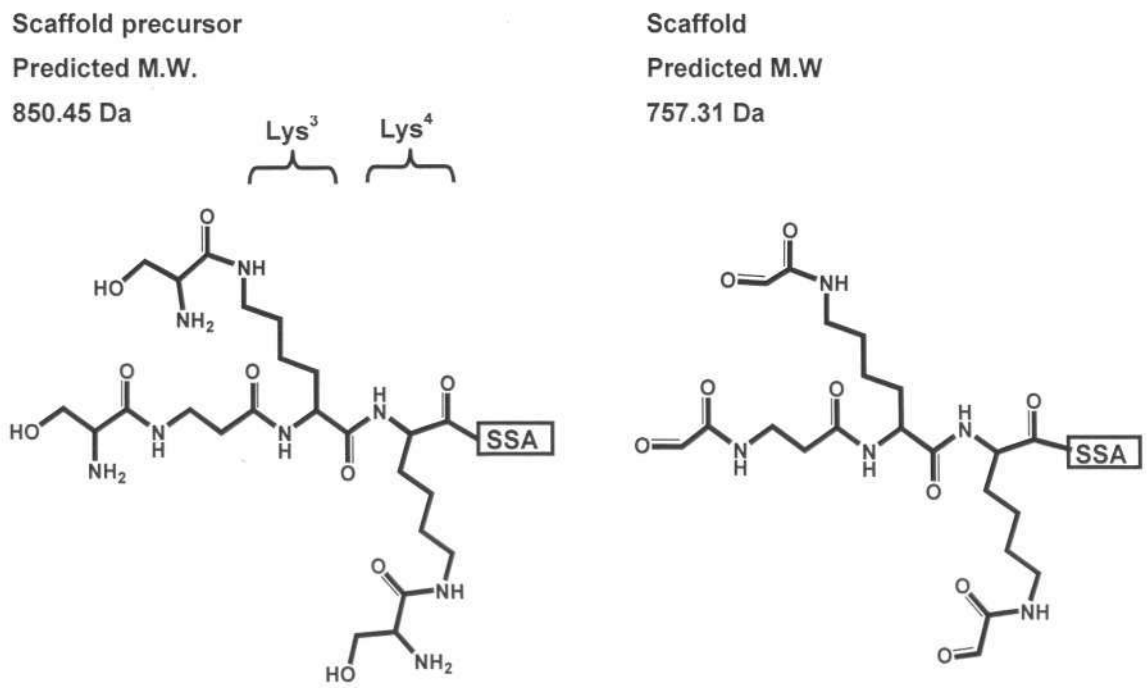


Figure 4-4. Generate Scaffold for Thiazolidine Ligation by $NaIO_4$ Oxidation.

A. Scaffold and its precursor.

The scaffold precursor Ser^1-βAla^2-Lys^3 (Ser)-Lys^4 (Ser)-Ser^5-Ser^6-Ala^7 was prepared by Fmoc chemistry. The tripeptide C-terminal tail, Ser^5-Ser^6-Ala^7 was added to the C-terminal of the branched tripeptide scaffold βAla^2-Lys^3 (Ser)-Lys^4 (Ser) to increase solubility and facilitate the HPLC purification. The molecule had a di-Lys core (Lys^3-Lys^4). An additional βAla is added at Lys^3, so that the two peptide chains will each be nine atoms from the α-atom on the Lys^3. The precursor has Ser at its amino terminus. The Ser is converted to aldehyde by periodate oxidation at pH 5.5 to form the scaffold. The tri-aldehyde-scaffold is used for ligation with the appropriate Cys-peptide monomers.

The scaffold precursor contained Ser at its amino terminus. It was converted to aldehyde by periodate oxidation at pH 5.5. The tri-aldehyde-scaffold was used for ligation with the appropriate Cys-peptide monomers. The tripeptide Ser-Ser-Ala was added at the C-terminal of the scaffold to increase aqueous solubility, to ease the HPLC purification, and to facilitate mass-spectrometry verification. The trimeric protein mimetics were artificial proteins consisting of three parallel strands. So it could not be prepared directly by recombinant methods, but it was well suited to be prepared by coupling the unprotected peptides through chemoselective ligation. Coupling the unprotected peptide, either at the amino or the carboxyl terminal to a functionalized scaffold, was a single step reaction without additional chemical transformation (Jun Shao 1995 ; Tam and Yu 2002). In our experiment, we chose to use the N-terminal ligation method through thiazolidine ligation to couple the peptidyl scaffold with the unprotected monomeric Cys-peptides (Figure 4-5.A). To fulfill this ligation scheme, we prepared two versions of unprotected peptides: HR2 and IL25 were used as the monomeric controls and their corresponding N-terminal cysteine analogs, CL36 and CL26, were used for chemoselective thiazolidine ligation to form trimeric protein mimetics.

Figure 4-5

A

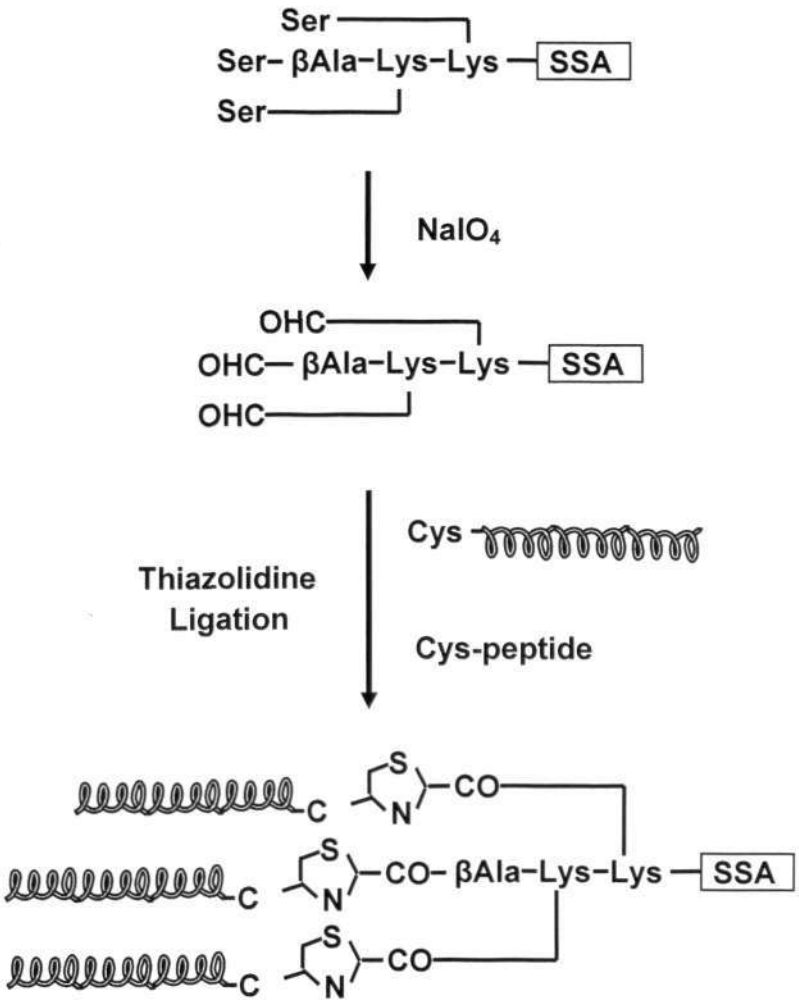


Figure 4-5. Synthesize Trimeric Protein Mimetics through Thiazolidine Ligation.

A. Synthetic Scheme of Trimeric Protein Mimetics.

The scaffold precursor was oxidized by the NaIO4 in acetonitrile containing 0.2 M acetate buffer (pH 5.5) 50:50 (v:v) and the oxidization was completed in 20 minutes. The reaction mixture was injected into the HPLC to remove the formaldehyde formed in the reaction. The HPLC fractions containing the aldehyde-scaffold are used directly for ligation with the Cys-peptides. Thiazolidine ligation is performed by mixing 3.3 molar excess of the Cys-peptide with aldehyde-scaffold in the HPLC eluent. The thiazolidine ligation was complete after 48 hours.

Periodate oxidation of the precursor to generate the scaffold

The HPLC purified scaffold precursor was oxidized by the sodium periodate to convert the N-terminal Ser to the aldehyde-function groups (Figure 4-4.A). The oxidation processes was monitored by the HPLC. The precursor molecule contained six Ser and was highly water soluble (>100 mg/ml). It was used to be eluted at 7-8.5 minutes in the 0-20% Buffer B 20 minutes gradient (Figure 4-4.B). The collected peak was examined by the ESI-Mass and the observed mass 851 Da was identical to its theoretical molecule weight. After oxidation, its retention time was retarded to 11 minutes. The eluted peaks were examined by the ESI-Mass and the observed mass was 758 Da, which was identical to the theoretical mass of the scaffold.

Figure 4-4.B

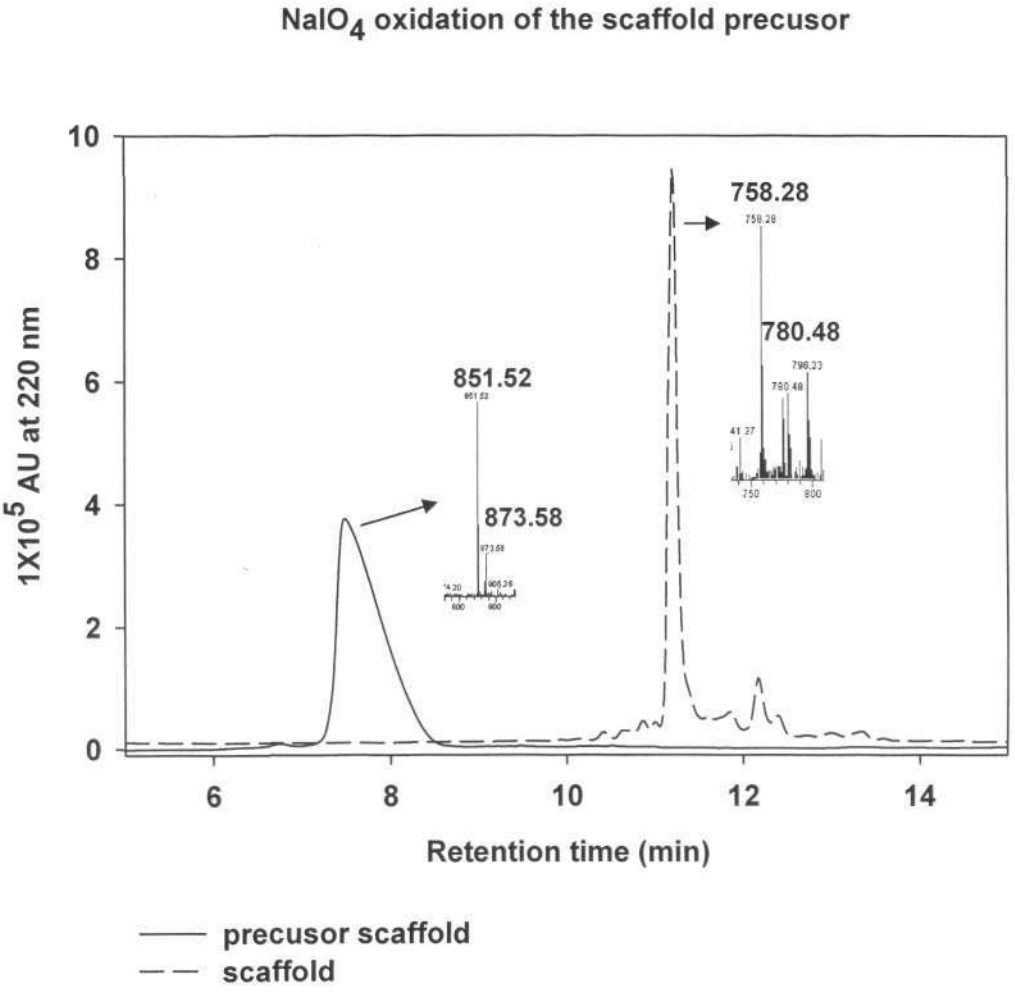


Figure 4-4.B. Purification of the Scaffold and its precursor.

The HPLC purification profiles of the scaffold precursor and scaffold are shown. The ESI-MS results of the elution fraction overlapped with the HPLC profile. The elution fraction of precursor comes out at 7-9 minutes. The precursor containing the fraction is examined by the ESI-MS to confirm the molecule weight. The synthesized precursor molecular gives the expected mass of 851 Da. The 873 Da is the Na+ attached form of the precursors. The scaffold retention time is 11-11.5 minutes. The ESI-MS results give the expected mass of 758 Da of the scaffold.

Thiazolidine Ligation

The thiazolidine ligation of the purified scaffold and Cys-peptides was performed in aqueous conditions at pH 2.5. The ligation was initially performed using 3:1 molar N-terminal cysteine peptide with aldehyde-scaffold. The ligation processes were monitored by the HPLC under wavelength 220 nm (Figure 4-5.B). The collected peaks were examined by the MALDI-TOF mass analysis. Peak A was confirmed to be the monomer of the CL36. Peak B was confirmed to be the oxidized dimer form of the CL26. Peak C was confirmed to be the side products, which was the scaffold dimerized cysteine peptides. Peak D was the expected trimeric CL36 protein mimetics. The side products were detected in the 48 hours ligation mixture as an incomplete ligation product, though it may be a result from the insufficient reactant, CL36 peptides. The ratio of the N-terminal cysteine peptide to aldehyde scaffold was then increased to 5:1. The scaffold dimerized cysteine peptides were still detected in the 48 hours ligation mixture. Even the 10:1 cysteine peptide to the aldehyde scaffold was used and the side products peak was not diminished.

Figure 4-5.B

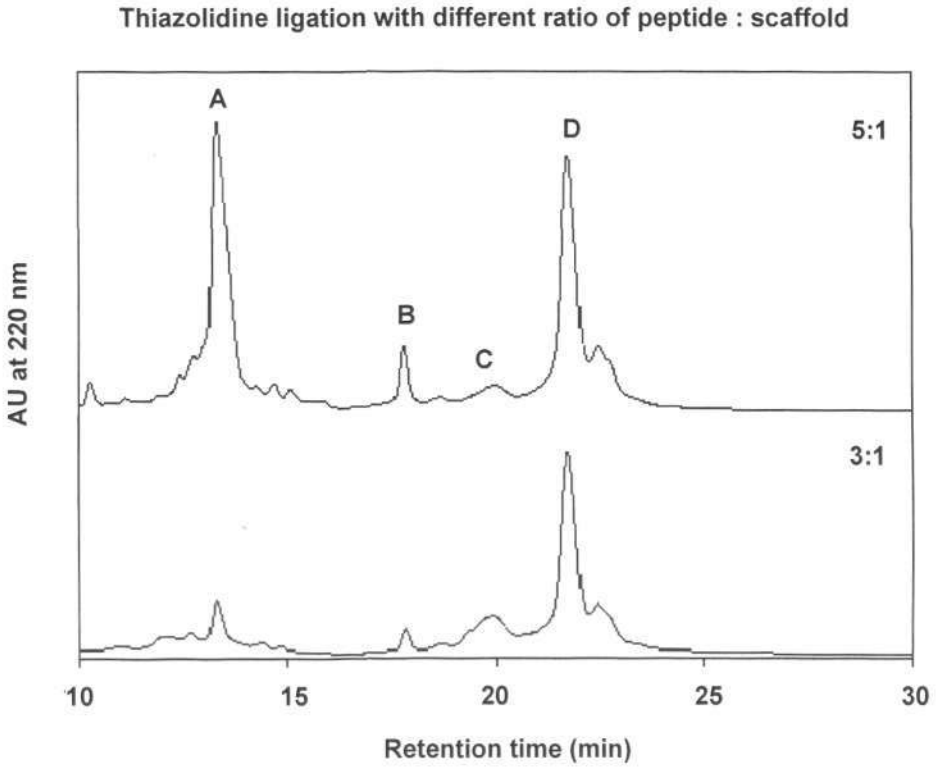


Figure 4-5.B. Thiazolidine Ligation at Different Peptides: Scaffold Ratio

Thiazolidine ligation with different peptides : scaffold ratios are monitored by the HPLC. The figure shows peptide : scaffold molar ratio 5:1 and 3:1 HPLC purification profile. The ligation 48 hours samples are injected to the HPLC and the elution fractions are examined by the MALDI-TOF Mass. Peak A, B, C, D are confirmed to be CL36 monomer, CL36 oxidized dimer, two CL36 attached incomplete ligation product and protein mimetics CL36 trimer respectively.

The appearance of side products may also be the result of incomplete ligation time. To prove our conclusion, we performed a time course for the thiazolidine ligation (Figure 4-5.C). The ligation mixtures were analyzed at different time points by the HPLC. In the first two hours, the incomplete ligation product gave an obvious peak at 33-34 minutes in the HPLC purification curve. With the ligation time extended, the peak was lower than before, while the protein mimetics peak was trending higher and higher, indicating the accumulation of the final products and the conversion of the incomplete products to the final products. Judging by the relative height of the reagent and products, there was no obvious yield increase after the ligation for 48 hours. The ligation time was even extended to 96 hours and peak C, the incomplete ligation products, still existed. A similar situation occurred when the CL26 was used to produce the trimeric protein mimetics. Considering convenience for purification, we chose 3.3:1 as the ratio to perform the thiazolidine ligation. Also, we accepted 48 hours, as the time for ligation completion.

Figure 4-5.C

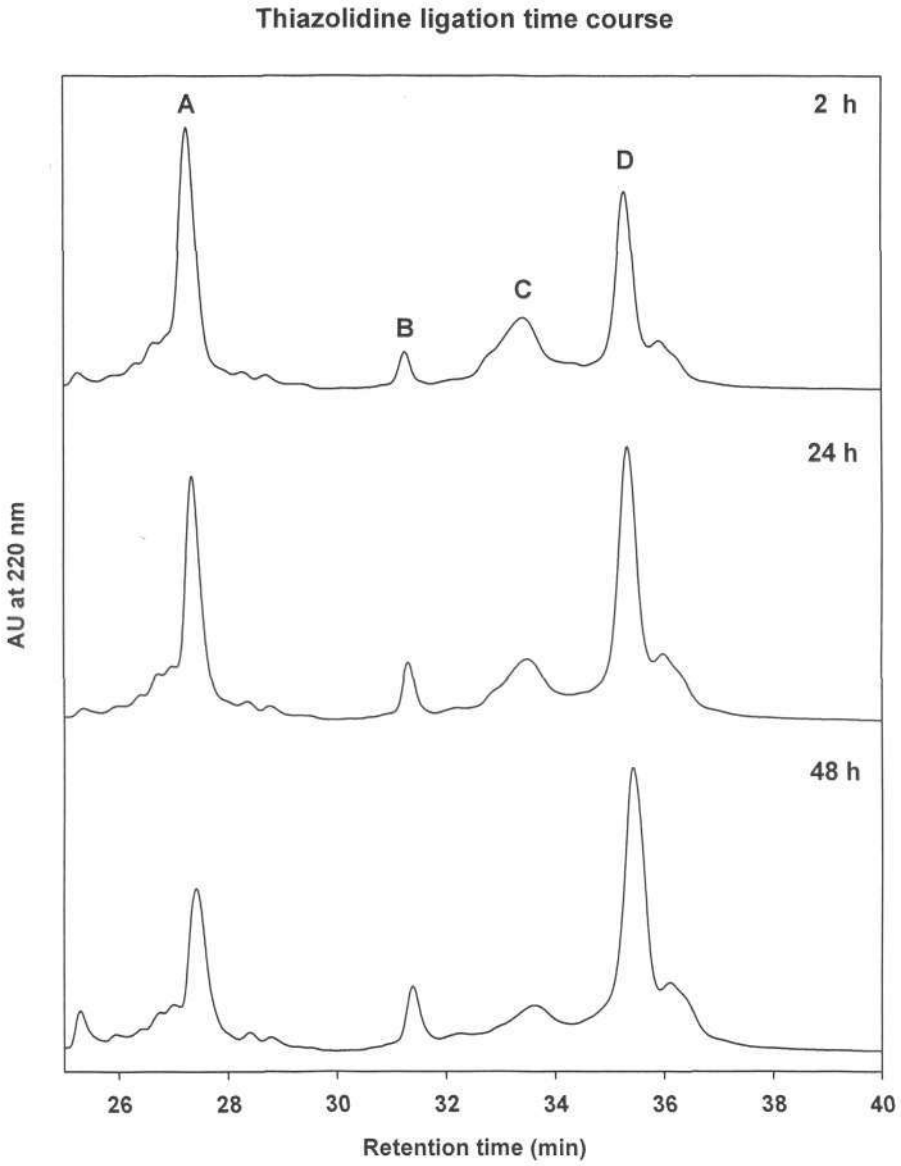


Figure 4-5.C. Time Course of the Thiazolidine Ligation.

The ligation process is monitored by the HPLC. Figure 4-5.C shows the ligation samples collected at the 2, 24 and 48 hours HPLC purification profiles. The elution fractions were examined by the MALD-TOF Mass. Peak A was identified to be the CL36 monomer. Peak B was identified as the oxidized CL36 dimer. Peak C was identified to be the incomplete ligation product. The scaffold is attached only to two CL36. Peak D is the trimeric protein mimetics.

CD Analysis of the Trimeric Protein Mimetics and the Peptides

The produced trimeric protein and the HR region derived fragments are analyzed by the CD to determine their helical states. Through CD examination, the HR2 region displays unordered structure in all the pH conditions we tested (Figure 4-6). As expected, the CL36 trimeric protein mimetics display an ordered helical structure in the neutral and basic pH. Interestingly, the HR2 trimeric protein mimetics, CL36 displays unordered structure in the acidic pH. The similar situation happens on the IL25 and CL26 trimeric protein mimetics. The IL25 corresponds to the helical region of the HR2. The IL25 displays unordered structure in all the pH we tested, while its ligated trimer form, CL26 protein mimetics, displayed an ordered helical state in neutral and basic pH, but failed to display any ordered state when exposed to acidic pH. That means the trimer state of the HR2 was disassembled by the low pH. In contrast to the HR2, the HR1C, which is the binding region of HR2 on HR1, displayed the helical structural state in all the pH we tested. Besides, the helical state of the HR1 did not change when the pH is changed. When the HR1 and HR2 peptides are equally molar mixed, the helical state of the mixture was greatly improved. All of the three peptides derived from the HR2 region were tested and the results were found to be similar. The entire HR2 region derived peptides displayed an unordered structure while they were induced to form higher helical states. The helical state of the mixtures would not change by the pH changing. The results suggested that the HR1 could induce the formation of the helical state of the HR2 region. Besides the HR2, the helical state of the HR1 is also improved, which means the HR1 and HR2 will react with each other to form a stable higher helical state oligomer. We suppose that the stable mixtures are the formed six-helix bundle.

Figure 4-6

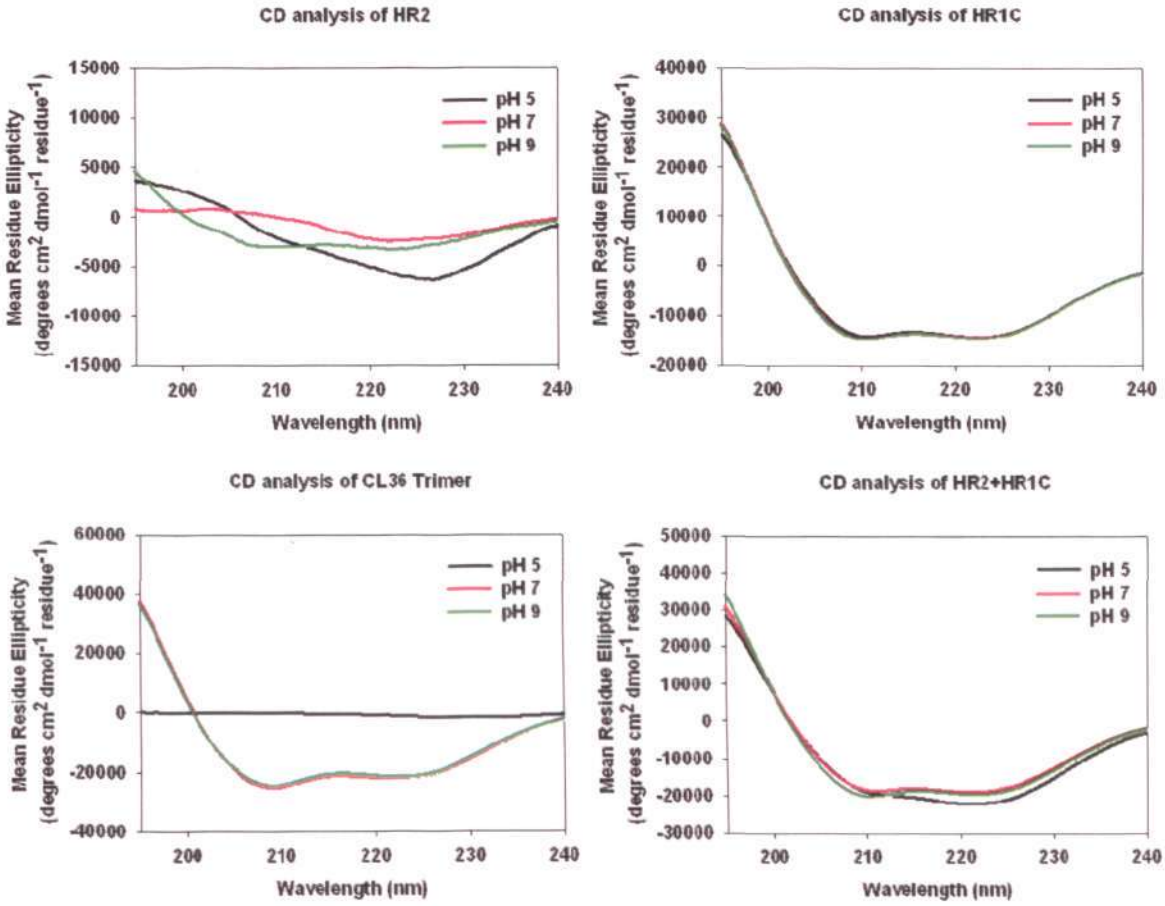


Figure 4-6. CD Analysis of the HR Region Derived Protein Mimetics and Peptides.

The peptide HR1C, HR2, CL36 trimeric protein mimetics (the ligated trimer form of the HR2) and the equal molar mixture of HR1C and HR2 were dissolved in PBS in pH 5, 7, 9 and examined by the CD. The CD spectra are obtained on Chirascan CD spectropolarimeter, using a quartz cuvette of 0.1 mm optical path length, at room temperature. Scans are conducted from wavelengths of 195 to 240 nm, with an integration time of 0.2 second and a bandwidth of 0.1 nm.

Proteinase K Treatment of Protein Mimetics and the Peptides

As covalently linked trimers, the protein mimetics were expected to be more stable and resistant to proteolytic degradation than their corresponding monomers. Proteinase K, a broad-spectrum endoprotease, was used here as the protease for degradation. Two concentrations (0.5 μg/ml and 0.125 μg/ml) of the proteinase K was used to digest the protein mimetics and the peptides in the PSB at 37°C, followed by the HPLC analysis (Figure 4-7). The HR2 monomer was treated by 0.5 μg/ml proteinase K initially. After 15 minutes digestion, the peak of the HR2 had completely disappeared, while the trimeric HR2 protein mimetics demonstrated proteolytic resistance. The peak of the HR2 trimeric protein mimetics still can be examined by the HPLC, with some degradation peaks. The IL25 was digested by the proteinase K at 0.125 μg/ml for 15 minutes, its peak was shrunk to half of its original height. The trimer of the CL26 still maintained 91% of its original height after the proteinase K treatment. The results indicated that the protein mimetics has a stronger proteolytic resistance as expected. The stability may come from their ordered structure compared with their monomer peptides.

Figure 4-7

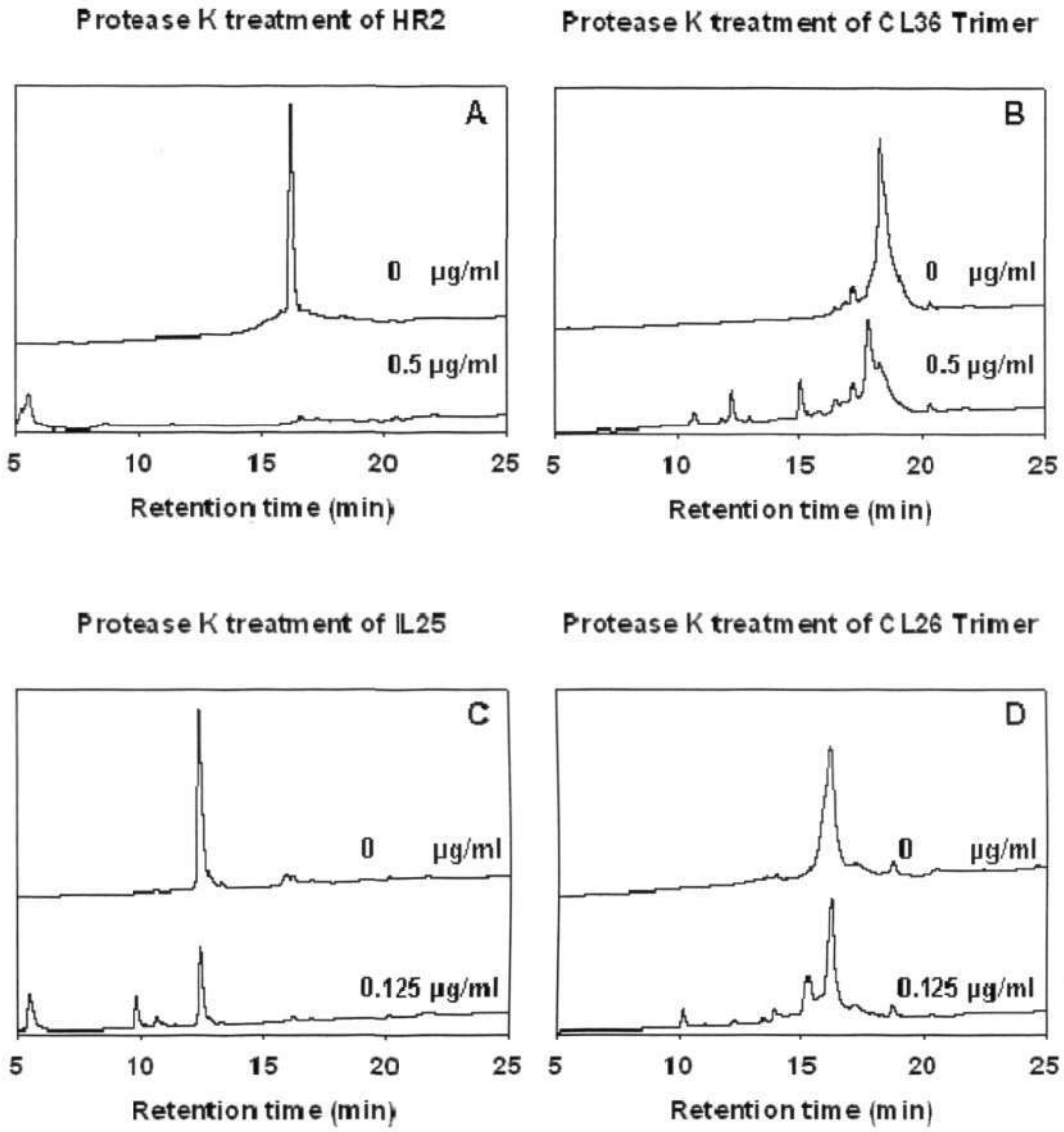


Figure 4-7. Protease K Treatment of Trimeric Protein Mimetics and Peptides.

Figure 4-7, A to D, shows the HPLC purification profiles of proteinase K treated synthesized protein mimetics and the peptides they are based on. The peptide and protein mimetics were digested by protease K at 15 minutes at 37°C. The observing wavelength is 220 nm. **A**. HR2, **B**. CL36 trimeric protein mimetics, **C**. IL25, **D**. CL26 trimeric protein mimetics. The peptides or protein mimetics were prepared in 1 mg/ml in the PBS. The proteinase K concentrations are indicated on the figures.

Protein Mimetics / Peptides Viral Inhibition Assay

The inhibitory activity of the peptides and protein mimetics were tested in the pseudovirus system mentioned in the previous chapter. The 293-ACE2 cell was infected with the SARS-CoV pseudotyped virus in the absence or presence of peptides or protein mimetics at 50 μM. The viral infectivity was determined by the luciferase assay. The peptides or protein, which showed the inhibitory effect, was selected and tested in the next round to determine the effective concentration. The HR2 peptide and CL36 trimeric protein mimetics showed the inhibitory effect on virus infection. The other peptides, HR1C and IL29 etc, had no obvious inhibitory effect on the first round of the viral inhibition assay. Different concentrations of the HR2 and CL36 protein mimetics were mixed with the SARS-CoV pseudotyped virus and then used to infect the 293-ACE2 cell. The viral infectivity was determined by the luciferase assay 48 hours after infection. As shown in Figure 4-8, the HR2 and CL36 protein mimetics demonstrated a dosage dependent inhibition effect on the pseudovirus. The CL36 trimeric protein mimetics demonstrated a stronger inhibitory effect compared with the HR2 peptides. The EC_{50} value of the HR2, calculated by software, is 7.07 μM, while the EC_{50} of the CL36 trimeric protein mimetics is 3.13 μM. The results suggested that the ordered structure may increase the affinity of the inhibitor to the target, thus increasing the inhibitory effect as we expected.

Figure 4-8

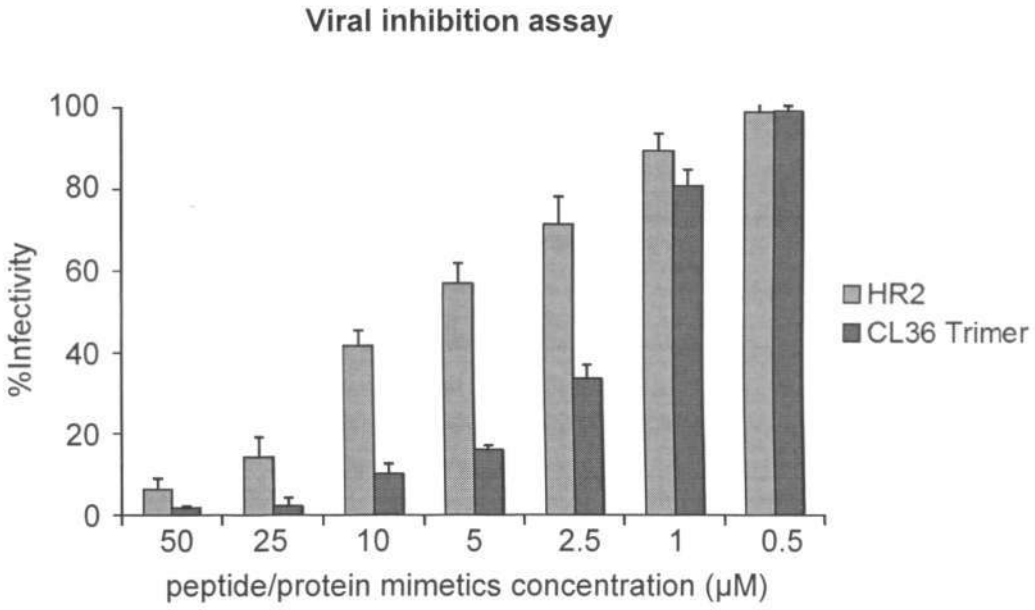


Figure 4-8. Protein Mimetics / Peptides Viral Inhibition Assay.

The SARS-CoV S protein pseudotyped HIV virus is used to infect the 293-ACE2 cell. The viral infectivity is reflected by the activity of luciferase carried by the pseudovirus, which is determined by the luciferase assay. The viral infectivity is represented as the means ± SD of the five separate experiments. The EC_{50} of the HR2 is 7.07 μM and the EC_{50} of the CL36 trimeric protein mimetics is 3.13 μM.

Hemolytic Assays of Peptides and Protein Mimetics

The toxicity of the produced peptides and protein mimetics was determined by the hemolytic assay. Fresh human O type red blood cells (RBC) were mixed with the peptides or protein mimetics to determine the hemolytic activity. The incubation time was one hour at 37°C. PBS was used as a negative control and the 1% Triton X-100 treated cell was set as 100% lysis. Melittin, which was known to bind to red blood cells and cause an increase in permeability, was used as a peptide control for the hemolysis. The results show that none of the peptides or protein mimetics results in hemolysis at the maximum tested concentration in the assay. Figure 4-9 showed the hemolytic activity of the HR2 and the CL36 trimeric protein mimetics as an example. The peptides or protein mimetics concentrations were increased to 125 μM or 250 μM and there was still no obvious hemolysis occurrence. The results indicated that the HR region derived peptides have less toxicity and it was a good target for design sequence specific peptide or protein inhibitors.

Figure 4-9

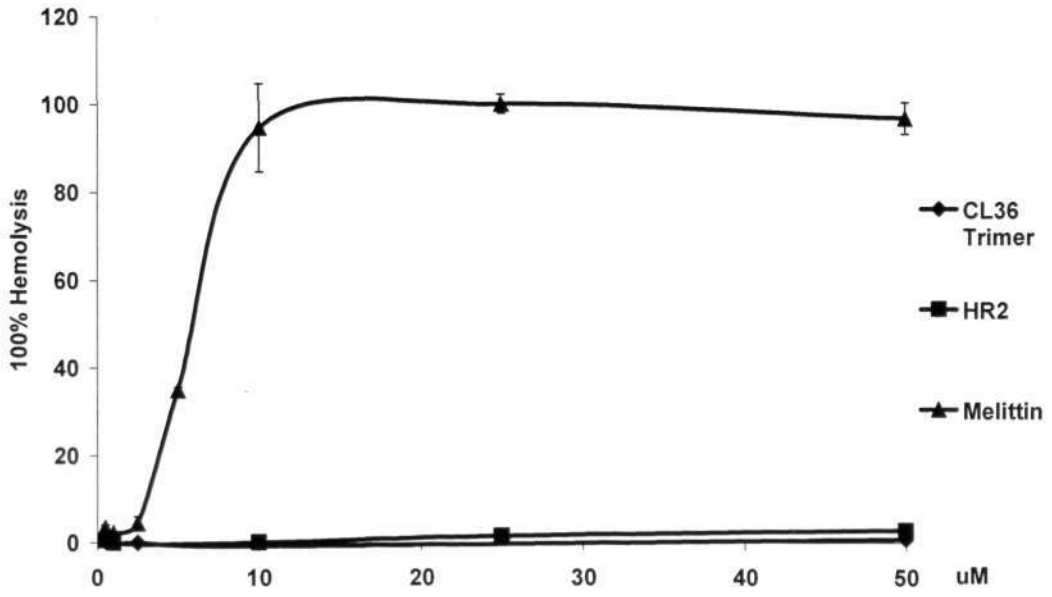


Figure 4-9. Hemolytic Assays of Peptides and Protein Mimetics.

Fresh human O type red blood cells (RBC) were mixed with the peptides or protein mimetics to determine the hemolytic activity. The incubation time was one hour at 37°C. The PBS was used as a negative control and the 1% Triton X-100 treated cell is set as 100% lysis. Melittin is used as a peptide control for hemolysis. The absorbance was measured at 415 nm with an automatic plate reader Safire2^{TM} . The results were shown as percentage lysis to the 1% Triton X-100 treated cell. The data represented the results from three separated experiments.

Discussion

With increased understanding of the entry mechanism of class I viral enveloped protein and the structure of the fusion core, different modalities have evolved to target this process, particularly the pre-hairpin structure, as pre-entry inhibitors. Among these modalities, inhibitors based on the engineered peptides or proteins, as inhibitors with enhanced helicity to mimic the pre-hairpin conformation, have found successes (Wild, Shugars et al. 1994; Judice, Tom et al. 1997; Chan, Chutkowski et al. 1998; Ferrer, Kapoor et al. 1999; Eckert and Kim 2001; Louis, Bewley et al. 2001; Root, Kay et al. 2001; Bewley, Louis et al. 2002; Sia, Carr et al. 2002; Louis, Nesheiwat et al. 2003). Most engineering approaches use a chimeric protein design and noncovalently-associated proteins, as models. Here, we adapted a scaffold approach to engineer the quaternary structure mimetics of the trimeric SARS-CoV S protein HR2 peptides, without significantly increasing their peptide lengths. In our approach, the protein mimetics were prepared conveniently and efficiently by a chemoselective ligation using unprotected peptide monomers.

Protein mimetics stabilize quaternary protein structures through interchain linkages. The HR1 and HR2 pre-hairpin structure is a quaternary structure formed by three monomers with helical propensity. To overcome the energy barrier of trimerization, the mimetics contained a scaffold, as an interstrand crosslinker, which constrained three monomers as a covalent-linked trimer to mimic a truncated version of the pre-hairpin structure. This design allowed for the interstrand stabilization of their hydrophobic faces, exposing only their hydrophilic faces to solvents that increase solubility and minimizes aggregation, which overcame the amphipathic nature of the heptad repeat region.

Previous work in the HIV-1 inhibitors study showed that recombinant methods were possible to produce either five or six-helix bundle containing both the HR1 and HR2 in an up and down fashion similar to the fusion core (Eckert and Kim 2001). Recombinant protein-based SARS-CoV entry inhibitors were also reported (Ni, Zhu et al. 2005). However, no matter what region was adopted in its design, the recombinant protein-based entry inhibitor was a folded polypeptide chain. The folding force originated from the propensity of the HR1 and HR2 association. The unavoidable usage of both HR1 and HR2 reduced the active site of the inhibitor. Some of the recombinant protein-based inhibitors worked like one single peptide (Ni, Zhu et al. 2005). To generate the inhibitor containing three-parallel-stranded compounds was a difficult task for the biosynthetic approach, due to the existence of three pairs of carboxylic and amino termini in the structure.

A site-specific chemoselective ligation is well suited to prepare such a kind of mimetics. Experimentally, the reported inhibitory peptides have been ligated to trimeric mimetics in aqueous solutions without a coupling reagent. This method also maximizes flexibility and minimizes chemical steps. Because elaborate chemical manipulations (protection and deprotection steps) were not required in the ligation step, it was the most convenient and direct method to achieve protein mimetics without side reactions.

Consistent with the previous finding, the HR2, the IL29, and the IL25 did not display significant helical structures in aqueous solutions. The CD measurements of the trimeric CL36 and trimeric CL26 in aqueous solutions provided support that they largely adopted conformations displaying both helical and extended structure, but were more ordered than their corresponding monomers, which were essentially disordered. The ordered structure might come from the increasing interstrand

interaction of the hydrophobic residues, after covalent linkage, thereby exposing the hydrophilic residues to solvents to provide a higher order of stability. The ordered structure may also enhance the bioactive conformations, which contributed to the increased inhibitory potency.

In addition, we might find that the trimerized HR2 was still affected by the pH. The associations between each coiled strands were diminished by the acidic pH. The CD analysis of the CL26 protein mimetics showed that the helical region of the HR2 was still affected by the pH. The acidic pH disassembled the trimer state of the HR2 region. The corresponding helical region on the HR1, which was the HR2 binding region, remained intact at the pH we tested. The results suggested that the pH might affect the homo-oligomer state of the HR2 region on the trimerized S protein in the viral membrane. The dissociation of the trimer state of the HR2 by the acidic pH helped to lower the energy barrier allowing for the completion of the conformational change. However, the free energy released by the formation of the six-helix bundle was sufficient to fulfill the conversion even without the low pH environment. As a result, the pH would facilitate the conformational change of the SARS-CoV S protein, but was not an essential factor in this process.

Consistent with the previous report, the HR2, but not the HR1, was able to inhibit the SARS-CoV S protein pseudotyped virus infection. The trimeric CL36 protein mimetics also demonstrated the viral inhibitory effect on a dosage dependent manner. The dosage dependent manner of inhibition might be explained due to their competitive binding to the HR1 region of the SARS-CoV S protein.

The HR2 peptide and the CL36 protein mimetics were both moderate inhibitors compared with the corresponding region derived peptides in the HIV-1, which showed the inhibitory effect at a nano molar grade. Even in coronavirus, the

EC_{50} of the HR2 region to the MHV infection was down to 0.9 μM. Although several research group reported peptides from the S2 domain HR2 region that had viral fusion inhibitory activity at the micromolar level (Bosch, Martina et al. 2004; Zheng, Guan et al. 2005), like 0.14 and 1.19 μM (Yuan, Yi et al. 2004), there were also several group reported peptide from HR2 region has poor antiviral activity (Liu, Xiao et al. 2004). After we compare the exact peptide sequences of those research groups, we found that maybe the interaction between HR1 and HR2 of SARS-CoV S protein was sequence specific. We believed the sequence difference leads to the variance between experiment results between different research groups.

For example, in Deng's group's results (Yuan, Yi et al. 2004), the peptide HR2-18 (amino acids 1161 to 1187) showed IC_{50} at 1.19 μM, while HR2-17 (amino acids 1158 to 1187) showed little anti-viral activity. The CP-1 (amino acids 1153 to 1189) in Jiang's group (Liu, Xiao et al. 2004) showed IC_{50} at 19 μM, the sequence HR2-11 (amino acids 1153 to 1187) in Deng's group showed little anti-viral activity (Yuan, Yi et al. 2004). The peptide HR2-38 (amino acids 1149 to 1186) in Gao's paper showed IC_{50} of 0.5-5 nM (Zhu, Xiao et al. 2004), the same peptide sequence showed IC_{50} of 1.02 μM in Tien's group (Ni, Zhu et al. 2005). In those research groups' work, except for Gao's group, the peptides centered from amino acids 1151 to 1185 showed the anti-viral activity from IC_{50} 19 to 1.02 μM. The HR2 sequence selected by our group is from amino acids 1151 to 1185, its IC_{50} in our result is 7.07 μM. The result was consistent with other research group's experiment data. We believe the conformation of the peptides may have been altered by adding or deleting a few amino acids, thus the consequent inhibitory effects would be totally different. Meanwhile, comparing with those peptides from MHV S protein HR2 region, the tendency of the SARS-CoV S protein HR2 region binding to HR1 region to form the

trimeric coiled-coil was rather low (Bosch, van der Zee et al. 2003). This also helped to explain the low antiviral activity of HR2 peptide and the CL36 trimeric protein mimetics in our experiment. The sequence we selected may not be the optimal sequence for HR1 binding.

Another explanation for the low inhibitory effect of the HR2 and its trimeric protein mimetics might be found from the entry pathway of the SARS-CoV. In previous reports, the heptad repeat region derived peptides inhibited the SARS-CoV entry at the micro molar grade, regardless of a different strand of the SARS-CoV or pseudovirus or host cell type. The inefficient inhibition of those peptides may be due to the insufficient peptide was trafficked into the endosome. So, our result might be an experiment evidence to support the endocytosis pathway of SARS-CoV viral entry. Recent research works from Taguchi's group supported this assumption (Ujike, Nishikawa et al. 2008). The peptide SR9 (amino acid 1151 to 1185) show no inhibitory effect on SARS-CoV infection. When the endocytosis pathway was blocked and the virus entered cell via cell surface, the SR9 showed IC_{50} to 100 nM. This result could help to explain why the HR2 peptide and CL36 trimeric protein mimetics had moderate anti-viral effect in our experiment. It also helped to explain that peptide with the same sequence (HR2-10) show little anti-viral activity in Deng's group's result (Yuan, Yi et al. 2004).

Recent research finding furthered our understanding of the SARS-CoV entry mechanism and provided insight into the entry inhibitor design. Consistent with our previous finding, the trypsin treatment enhanced the S protein mediated membrane fusion. The trypsin used in our experiment simulated the cellular proteases that may be produced by the inflammation cell in the lungs of SARS patients. The proteases produced by the SARS-CoV infected patients (Kawabata, Hagio et al. 2002) may

enhance the SARS-CoV infection in their lungs through direct cellular membrane fusion, leading to rapid infection and severe damage to their major organs.

Since the HR region derived peptide demonstrated inhibitory effect on the S protein mediated direct membrane fusion process, the HR region was still a good target for inhibitor development. The CL36 trimeric protein mimetics showed 2-3 times more potent than HR2 monomer in our experiment. Considering the CL36 trimeric protein mimetics contained more active sites for binding per molecule compare to HR2 monomer, it seems that trimerization of HR2 did not greatly improve its potency. But the trimeric protein mimetics demonstrated higher stability and stronger proteolytic resistance compared with the HR2 monomer when treated with proteinase K. Meanwhile, insufficient protein mimetics being incorporated into the endosome may be the reason for CL36 trimeric protein mimetic did not exhibit higher potency as we expected. This suggested that the design of the trimeric protein mimetics may be a good solution to prevent the severe damage of the patient's organ by the SARS-CoV infection, and it may be an effective drug to be used in SARS.

So, the future direction for the SARS-CoV viral entry inhibitor in our group can be classified to the following to directions. One is to seek the peptides with higher affinity to the fusion core, especially to HR1. Replacement of the key residues in the HR2 peptide to increase its affinity of binding with the HR1 region, to form the six-helix bundle could lead to improvement of its antiviral efficacy. In Taguchi's work, the EK substitution peptide, SR9EK13, showed similar anti-viral effect as the original HR2 peptide. That means the replacement of the key residues is a workable direction. To identify the correlation between the sequence and inhibitory effect, more refined mapping of the peptides needed to be done in this field.

Another direction is to tackle the problem of low antiviral activity of our peptides because of virus using endocytosis pathway. The anti-viral peptides need to be designed to be efficiently transferred into the endosome and to be stable in the environment. Adding the cell-permeable sequence, which allow peptides to penetrate the cellular membrane and go inside the cell, could be a possible solution. In this way, even the SARS-CoV take endocytosis pathway, the anti-peptides or protein mimetics are still be able to target the virus inside the cytoplasma, which make it to be new antiviral candidates to SARS-CoV.

The other research group also reported that the sequences outside HR region might have anti-viral effect. For example, peptides analogous to the SARS-CoV loop region had IC_{50} values 2–4 μM in viral plaque formation inhibition (Sainz, Mossel et al. 2006). Two peptides on ACE2 (amino acids 22 to 44 and 351 to 357) showed potent antiviral activity with IC_{50} about 0.1 μM when artificially linked together by glycine (Han, Penn-Nicholson et al. 2006). Peptides from receptor binding region of S protein or the S protein binding region on ACE2 which could block the receptor binding process can be the promising targets for viral entry inhibitor design. When linked with the peptides from HR region, they might be able to inhibit the viral entry from more than one step. So, they could block the viral entry no matter what entry pathway it takes.

Other than being used directly as entry inhibitor, the HR2 and CL36 protein mimetics can also be used to raise antibody, especially for CL36 protein mimetics. There were some progresses reported in this direction (Lai, Chong et al. 2005; Sadler, Zhang et al. 2008) and it showed quaternary protein approach to mimic conserved and functional domains of viral envelope proteins is an interesting topic in vaccine development. But we will mainly focus on the first two directions first in the future.

Chapter 5

Summarizing Discussion

Discussion

In this thesis, we studied the entry mechanism of the SARS-CoV and its inhibition strategy. For the SARS-CoV viral entry, we focused mainly on the effect of the pH and protease cleavage. For its inhibition, (this design concept is based on the previous research work on the HIV-1 entry inhibitor design in our laboratory.) we designed and developed the trimeric protein mimetics based on the heptad repeat region of the SARS-CoV S protein to inhibit the viral fusion of the SARS-CoV. In Chapter 2, we analyzed the heptad repeat region of the SARS-CoV S protein and confirmed its interaction with the S protein. We also found that the oligomerization states of the HR regions were affected by the pH, suggesting changing the pH may affect the conformation of S protein. This study was further detailed in Chapter 3. There, we studied the pH effect on the S protein mediated viral entry and membrane fusion. We found that the SARS-CoV viral entry may go through the pH-dependent endocytosis pathway and protease treatment enhances the S protein-mediated membrane fusion. The protease cleavage site is believed to be located around amino acid 667. In Chapter 4, we described our design and the development of the trimeric protein mimetics and its viral inhibition effect. Biophysical study of the protein mimetics also demonstrated that the conformational state of the trimerized HR2 region was affected by the pH. The implications of our findings on the SARS-CoV entry process and the strategy of its inhibition will be discussed below.

Receptor Binding Initiate the SARS-COV S protein Viral Entry Process

The receptor of SARS-CoV was identified first as Angiotensin-converting enzyme 2 (ACE2) (Li, Moore et al. 2003). Later, CD209L or L-SIGN or DC-SIGN2 was discovered to be an additional receptor of SARS-CoV. The ACE2 receptor binding region of S protein was mapped and amino acid 318 to 510 of S protein shown to be a critical site for recognition (Xiao, Chakraborti et al. 2003; Babcock, Esshaki et al. 2004; Wang, Chen et al. 2004; Wong, Li et al. 2004).

Cryo-EM study confirmed that there is a structural transition of the S protein upon receptor binding. The S protein was observed to be able to bind to three ACE2 receptors at once (Beniac, Devarennes et al. 2007). Treatment of cells with either SARS-CoV S protein or S protein pseudotyped viruses resulted in the translocation of ACE2 from the cell surface to endosomes (Wang, Yang et al. 2008). Interestingly, the purified receptor binding domain of S protein alone was able to bind ACE2 and induced internalization (Wang, Guo et al. 2008). This suggests the receptor binding domain of S protein will trigger ACE2 mediated cellular endocytosis signal pathway, by which SARS-CoV enters the susceptible cells.

In our study, the S protein pseudotyped virus was unable to infect the 293T. Its viral infectivity was detected after ACE2 was expressed on the cellular surface. During the S protein mediated cell-cell fusion assay, the syncytia formation was seldom observed in the S protein expressing 293T cell, even after trypsin treatment. But the TPCK trypsin-treated S protein can induce strong cell-cell fusion when mixing with ACE2 expressing 293 cell. Our data also demonstrate that the receptor binding is important for viral entry and S protein mediated membrane fusion, suggesting the receptor binding induces signal pathway and conformational change of S protein.

Cleavage Is Necessary for Activation of the Fusogenecity of SARS-COV S Protein

Among the three groups of coronaviruses, group 2 and 3 are proteolytically processed S protein into the S1 and S2 parts, roughly equal in size. Their cleavage site occurs at the last residue in a highly basic motif: RRFRR in the S protein of IBV (Cavanagh, Davis et al. 1986), RRAHR in the S protein of the MHV strain A59 (Luytjes, Sturman et al. 1987) or RRARR in the S protein of MHV strain JHM, and KRRSRR in the S protein of BCoV. Similar cleavage sites are found in the other group 2 S proteins, except for the S protein of SARS-CoV (Figure 5-1).

The cleavage sites on the S proteins of many coronavirus, such as those of MHV A59 (de Haan, Stadler et al. 2004), are cleaved by furin-like proteases, generating a receptor binding unit and membrane-anchored fusion unit. For SARS-CoV, furin-mediated cleavage of the spikes appeared not to occur (Simmons, Reeves et al. 2004). The sequence alignment shows that the most likely cleavage site on SARS-CoV is around amino acid 667, corresponding to the furin cleavage site on the S protein of MHV A59. This site is believed to be recognized by trypsin. Both trypsin and cathepsin L treatment activate the S protein fusion activity (Simmons, Reeves et al. 2004; Matsuyama, Ujike et al. 2005; Simmons, Gosalia et al. 2005; Huang, Bosch et al. 2006). Endosomal cathepsin L protease is believed to activate the S protein during endocytosis pathway. The cathepsin L's cleavage site has recently been identified to be amino acid 678 on S protein, 11 amino acids downstream of the trypsin cleavage site (Bosch, Bartelink et al. 2008). More recent report shows the existence of another cleavage site at amino acid 798-801, which could also be recogniezed by trypsin (Kam, Okumura et al. 2009). Furin-like protease cleavage site introduced at that position results in the cleavage of S protein and cell-cell fusion

without trypsin treatment when expressing on the cell surface (Watanabe, Matsyama et al. 2008).

These data based on the works of our laboratory and many others demonstrate that there is no stringency in the cleavage site of the SARS-CoV S protein for the fusion activity. It is believed that the S protein in the SARS-CoV spike exposes a protease-accessible loop between the S1 and S2 subunits that can be targeted by different enzymes, such as cathepsin L, trypsin or elastase under physiological conditions. Consistent with the presence of such a loop, the syncytium-inducing capacity of expressed SARS-CoV S protein was dramatically enhanced after the introduction of a functional furin cleavage site in the S1 / S2 junction region (Follis, York et al. 2006; Qiu, Hingley et al. 2006). No such loop has been detected in spikes of HCoV-NL63 or of most other subgroup 1 coronaviruses. The existence of multiple cleavage sites which are recognized by different proteases might also hint that the activation of the SARS-CoV S protein is a sequential process (Belouzard, Chu et al. 2009). This is a unique feature for coronavirus S protein and it has not been reported in the envelope protein of the other virus bearing class I viral fusion protein.

Figure 5-1

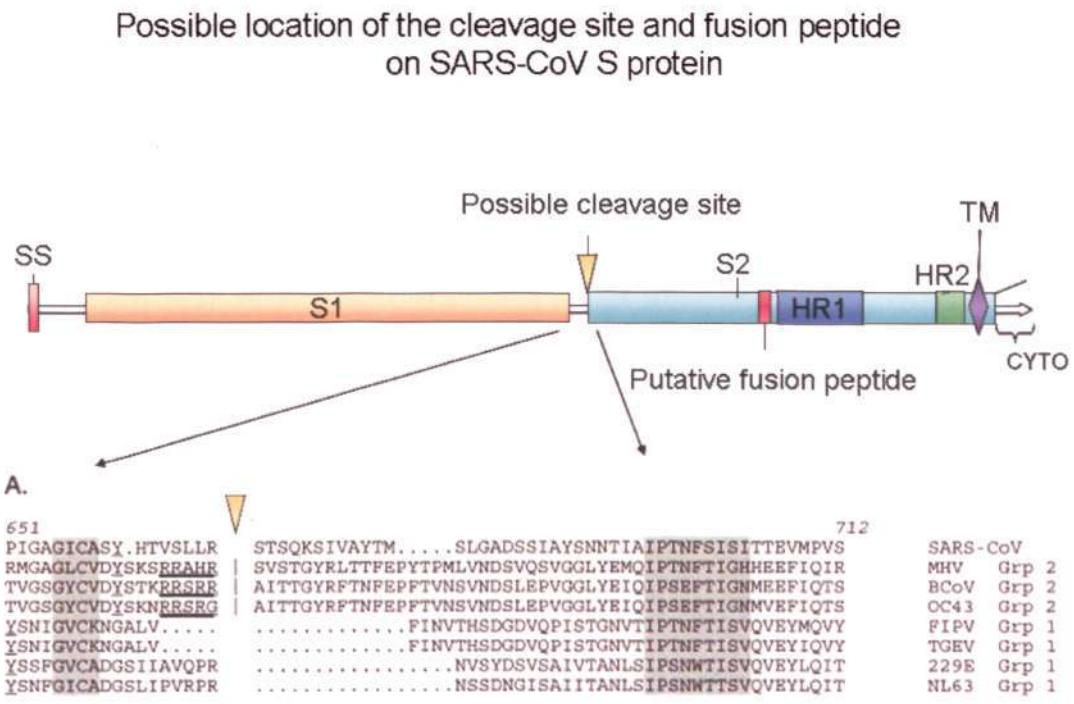


Figure 5-1. Location of the Possible Cleavage Site and the Putative Fusion Peptide.

Based on the sequence alignment result of the SARS-CoV S protein with the S protein of the other coronavirus, the possible cleavage site of S is shown on the diagram. The cleavage motifs on the other coronavirus S protein are shown, as the underline sequences in the sequence alignment result. The sequence alignment results come from (Follis, York et al. 2006). The location of the putative fusion peptide is drawn as the red box on the S protein according to the experiment result from (Yao, Ren et al. 2004; Sainz, Rausch et al. 2005).

In Chapter 3, we demonstrated that the cellular expressed S protein induced membrane fusion after trypsin treatment. The S protein mediates fusion between the protease activated S protein expressing cell and SARS-CoV receptor ACE2 expressing cell. This result was consistent with other researchers' findings. Several reports indicated that the protease treatment enhanced the SARS-CoV infection (Matsuyama, Ujike et al. 2005; Simmons, Gosalia et al. 2005; Du, Kao et al. 2007; Ujike, Nishikawa et al. 2008).

Interestingly, similar to what we found in the trypsin-treated S protein in Chapter 3, the S protein was treated with thermolysin, dispase, elastase or trypsin (Matsuyama, Ujike et al. 2005), all generating two fragments. The sizes of the two fragments generated by different protease cleavage are all similar. The cleavage pattern of the protease trypsin and cathepsin L treated S proteins was also similar, which is like the cleaved form of the S protein identified from the SARS-CoV infected Vero E6 cell (Wu, Shang et al. 2004). Our results support that there might be a protease-accessible loop in the middle of the SARS-CoV S protein, which is cleaved by the cellular protease and helped to render the fusion competent of the S protein. Although the basic motif for furin-like protease cleavage is not found in the SARS-CoV S protein, the other cellular proteases can also fulfill the cleavage of the S protein, such as the cathepsin L or trypsin. This helps to explain why the cell-cell fusion is enhanced after protease treatment.

The Location of the Fusion Peptide

The cleavages of the SARS-COV S protein are consistent with our present knowledge of bioprocessing of the calss I viral fusion protein. In class I viral fusion proteins, the cleavage usually occurs at the N-terminus or the upstream of the fusion

peptide. In the coronavirus, the cleavage is at the upstream amino acids and produces an internal fusion peptide. However, not all the S proteins in coronavirus are known to present to yield cleaved subunits (Vennema, Heijnen et al. 1990).

The SARS-CoV S protein can be activated by more than one protease and their cleavage sites are not identical (Huang, Bosch et al. 2006). The finding of the secondary protease cleavage site at aa 797 activating fusion activity implies that S protein has a fusion peptide together with two different HR regions (Watanabe, Matsyama et al. 2008). This suggests that the fusion peptide on SARS-CoV S protein is located downstream at the amino acid position 798.

Amino acid 858 to 886 on SARS-CoV S protein was predicted to be the fusion peptide because it is highly hydrophobic to be a membrane anchoring segment (Bosch, Martina et al. 2004; Guillen, Perez-Berna et al. 2005). Recent studies have shown that the region of amino acid 770 to 788 has a much more potent fusogenic activity than the former sequence (aa 858 to 886) to constitute the fusion peptide of the S protein (Sainz, Rausch et al. 2005; Madu, Roth et al. 2009). More recent studies suggest that the sequence (amino acid 798 to 815) following the secondary cleavage site might be a novel fusion peptide of SARS-CoV (Madu, Roth et al. 2009).

It is also possible that more than one region has membrane anchoring capabilities on the SARS-CoV S protein. Region amino acid 798 to 815 may be the first one acquiring the capability to bind to and insert into the membrane. Region amino acid 858 to 886 might serve as another internal fusion peptide which is essential for the subsequent steps in the fusion process. Its location at immediately to of the HR1 makes it adjacent to the transmembrane domain after the 6HB rearrangements of S protein during the fusion process (Supekar, Bruckmann et al.

2004). However, the precise location of the fusion peptide of the SARS-CoV S protein remains undetermined.

pH Effect during the SARS-CoV Viral Entry

An important factor that affecting the entry process of several enveloped viruses is pH, especially those internalized via endocytosis pathway. The acidic pH of the endosome has been used by viruses, at a suitable time and at a specific location to escape to cytoplasm, without being delivered to hydrolytic lysosomes (Helenius, Kartenbeck et al. 1980).

The pH also serves as a trigger for the conformational change for the viral fusion protein, leading to the fusion between the viral and the cellular membrane. For example, the protonation of the HA protein in the low pH environment triggered its irreversible conformational change.

In our study (Chapter 2), we found that the pH affected the oligomeric state of the peptides derived from HR region of SARS-CoV S protein and acidic conditions led to the dissociation of the homo-oligomers of those peptides. Through the biophysical analysis of the protein mimetics of the HR2 region (Chapter 4), we found that the trimerized HR2 was also affected by the pH. The association between each coiled strand was diminished by acidic pH, while the HR2 binding region on the HR1, remained intact at all pH tested. The results suggest that the pH affects the homo-oligomer state of the HR2 region on the trimerized S protein in the viral membrane. It is reported that not only for the HR region, but also the whole ectodomain of S protein are affected by pH (Li, Berardi et al. 2006). The ectodomain of S protein adopt different shape at different pH under electron microscope observation, suggesting the conformational change under different pH.

In Chapter 3, we found that the viral entry mediated by the SARS-CoV S protein was affected by the NH_4Cl, an inhibitor for the endosome acidification, suggesting that the SARS-CoV undertakes the endocytosis pathway for viral entry. The low pH environment for viral entry was thought to be necessary for conformational change of S protein via protonation. Other coronaviruses, such as TGEV (Hansen, Delmas et al. 1998), and the HCoV-229E (Nomura, Kiyota et al. 2004), also enter cells via endocytosis and the viral fusion occurs at the acidified endosomes. This conclusion is consistent with reports that the SARS-CoV viral entry is sensitive to several lysosomotropic agents, like NH_4Cl or Bafilomycin A, suggesting the low pH environment inside endosomes is necessary for conformational change of S protein (Simmons, Reeves et al. 2004; Yang, Huang et al. 2004).

In general, receptor-binding of the ACE2 initiates the viral entry process. The lysosomotropic agents, functioning to elevate the pH of acidic compartments, can inhibit the dissociation of the ligand from the receptor and trapping the receptor in the endosome (Ray, Basu et al. 2004), thus block the recycling of receptor. ACE2 can be trapped within perinuclear vacuoles for more than 14 hours after treatment by ammonium chloride, bafilomycin A1, or chloroquine (Wang, Yang et al. 2008).

Cathepsin L is one of the cellular proteases involved in the endocytosis pathway for SARS-CoV viral entry. It activates the S protein by a specific cleavage in the acidic environment inside the endosome. The cleavage of cathepsin L is more efficient at acidic pH (Huang, Bosch et al. 2006). Increasing pH from acidic to neutral results the decrease of cathepsin L activity. No obvious cleavage is observed under neutral pH (Bosch, Bartelink et al. 2008).

In chapter 3, we found that the S protein-mediated cell-cell fusion is not affected by pH after trypsin treatment, consistent with findings by other researchers.

For example, Simmons et al. have reported that the S protein-mediated cell-cell fusion is not affected by pH after protease treatment (Simmons, Reeves et al. 2004; Simmons, Gosalia et al. 2005; Bosch, Bartelink et al. 2008; Kawase, Shirato et al. 2008). The pH-dependency of the SARS-CoV viral entry can also be bypassed by protease treatment (Matsuyama, Ujike et al. 2005; Watanabe, Matsyama et al. 2008), which is believed to be another viral entry pathway being adopted by SARS-CoV through the direct membrane fusion.

Thus, pH seems to play multiple roles in the endocytosis pathway, affecting the viral entry directly or indirectly. It affect the recycling of the receptor, the activities of the endosomal protease which is required by S protein fusogenicity activation. It also affects the conformational change of the S protein, especially for the HR2 region. But, it is not the determining factor for the SARS-CoV viral entry process. Receptor binding and protease activation are the critical factors during the whole viral entry process. They play important roles in both direct membrane fusion pathway and pH dependent endocytosis pathway.

The role of HR region

In the formation of the fusion core, the HR2 needs to fold back along the groove of the central coil composed of the HR1. Therein exists a three-step conformational change process: 1) The HR2 stayed at the metastable homo-trimer on the S protein, 2) The HR2 disassembled, and 3) The HR2 packed against the groove of the central coiled HR1 in an anti-parallel manner. The low pH is believed to favor this process by dissociating the trimer state of HR2, which helps to lower the energy barrier needed to complete the conformational change.

Recent structural and dynamic properties of HR2 have been characterized (McReynolds, Jiang et al. 2008). These results demonstrate there exists an equilibrium between a structured trimer thought to represent a pre-fusion state and an ensemble of unstructured monomers thought to represent a novel transition state. Their experiment data support the three-step model from pre-fusion state to post-fusion state. Comparing with HR2 monomer, the trimerized HR2 protein mimetics shows higher order of secondary structure, the α-helix. The trimerized HR2 protein mimetics can be seen as three HR2 peptides constrained at the N-terminal. The increase of helicity in each of the HR2 peptide in the protein mimetics agrees with the CD analysis result in those researchers' work. The disordered states of HR2 decreasing with the increasing concentration of HR2 were also observed in our experiment, suggesting the pH is a factor in the oligomerization process of HR2.

The interactions between HR1 and HR2 have been further studied recently. Two different sites of interactions have been proposed by two individual research teams: Amino acid 916 to 950 of HR1 and amino acid 1151 to 1185 of HR2 (Tripet, Howard et al. 2004), and amino acid 902 to 952 of HR1 and amino acid 1145 to 1184 (Xu, Zhu et al. 2004). A recent study shows that amino acid 899 to 916 at N-terminal of HR1 form half of the relatively deep groove region for the binding of HR2 helical regions and make extensive interactions with amino acid 1161 to 1192 (Duquerroy, Vigouroux et al. 2005). The N-terminal of HR1 and C-terminal of HR2 may be important for the formation of a stable fusion core structure. The dimer of HR1 and tetramer of HR2 (four-helix bundle) in the crystal structure were also reported (Deng, Liu et al. 2006). This four-helix bundle state is considered one of the intermediate states of folding / unfolding of the fusion core structure during fusogenesis. In our

study in Chapter 2, both HR1 and HR2 were found to exist other than the trimer forms, which indicate the complicated interaction between HR1 and HR2.

The interaction between HR1 and HR2 is relatively strong. The HR1 sequence used in our study in Chapter 2 and Chapter 5 is different, through CD analysis, we found the stable structure formed by HR1 and HR2 are not affected by pH change, at least at the secondary structure level. This finding agrees with our understanding of the viral fusion process. In the early stage, from receptor binding to the pre-fusion stage, the S protein might need the low pH to facilitate its conformational change. But from pre-fusion stage to post fusion stage, even without the low pH environment, once the conformational change started, the free energy released by the formation of six-helix bundle is sufficient to fulfill the conversion. The low pH might only help to dissociate the trimerized HR2 region to facilitate the conformational change of the SARS-CoV S protein from pre-fusion state to the post-fusion stage, but it was not an essential factor in this process.

Possible SARS-CoV Coronavirus Viral Entry Model

Combining the results and literature precedences, we believe the SARS-CoV may use the following essential steps for its viral entry in the early stage: 1) receptor binding, 2) conformational change after receptor binding, leading to the exposure of the cleavage sites, and 3) cellular protease cleavage induced a conformational change of the S protein, leading to the viral and cellular membrane fusion.

It is generally agreed that multiple entry mechanisms are exploited by the SARS CoV to establish infection, indicating both pH-dependent endocytosis pathway and receptor-dependent direct membrane fusion pathways could be used to gain host cell entry. However, receptor binding and protease cleavage are the two critical

factors in both pathways (Simmons, Reeves et al. 2004; Matsuyama, Ujike et al. 2005; Simmons, Gosalia et al. 2005).

In the endosomal pathway, the virus particles that attach to its receptor, angiotensin-converting enzyme 2 (ACE2) (Li, Moore et al. 2003), is trafficked into the endosome. Then the S protein is cleaved at amino acid position 678 by an endosomal protease, like cathepsin L, which activates the S protein fusion activity. The cleavage event might also happen at amino acid position 797 by other cellular protease like furin. The decreasing pH inside the endosome disassembled the trimerized HR2 region on the SARS-CoV S protein, decreasing the energy barrier required for forming the six- helix bundle. Also, the decreasing pH was favored for cleavage of the S protein by the endosome protease. The exposed cleavage site of the S protein induced by the receptor binding was attacked by the endosome protease. The constraints preventing the necessary conformational changes required for fusion peptide insertion and fusion core formation would be removed. Once cleaved, the activated S protein would induce the fusion of viral membrane with the endosome membrane. In this pathway, the acidic pH alone is not sufficient to trigger fusion. Proteolytic cleavage on the viral fusion protein, by the acidic dependent endosome protease, was required for viral penetration. Similar cases were found in the Ebola virus (Ebert, Deussing et al. 2002; Chandran, Sullivan et al. 2005; Simmons, Gosalia et al. 2005).

In direct membrane fusion pathway, the virion S protein attached to cell surface ACE2 is activated for fusion by airway proteases such as trypsin or plasmin, which recognize two functionally important basic motif on S protein and cleaves at amino acid position 667 and amino acid position 797 of the S protein, (Li, Berardi et al. 2006; Bosch, Bartelink et al. 2008; Kam, Okumura et al. 2009). It leads to the

fusion of the viral membrane and plasma membrane and the subsequent viral entry from the cell surface (Matsuyama, Ujike et al. 2005). Entry from the cell surface is therefore possible only when there exist exogenous proteases which activate the S protein. In both pathways, the cleavage and fusogenic activation of SARS-CoV S protein are a critical step for virus entry into cells. However, the receptor binding is a pre-requisite.

Here, a model for the SARS-CoV viral entry to host cell is proposed. The receptor binding triggered the conformational change of the S protein as well as the signal pathway for ACE2 dependent endocytosis. The conformational change of S protein could lead to the exposure of the cleavage site on the S protein. The protease cleavage triggered further conformational changes of the S protein and led to the dissociation of S1 (the receptor binding subunit) and the exposure of fusion peptide. The insertion of fusion peptide to the cellular membrane and formation of the six-helix bundle by the HR region lead to the S protein mediated membrane fusion. If the cleavage occurs at the surface of the cell by the cellular protease or exogenous protease, the viral fusion will happened at the cellular surface, which will be the direct membrane entry pathway. If the cleavage happened in the endosome, by the pH-dependent endosome protease, such as the cathepsin L, the viral fusion occurred in the endosome, which will be the receptor mediated endocytosis pathway (Figure 5-3).

Figure 5-2

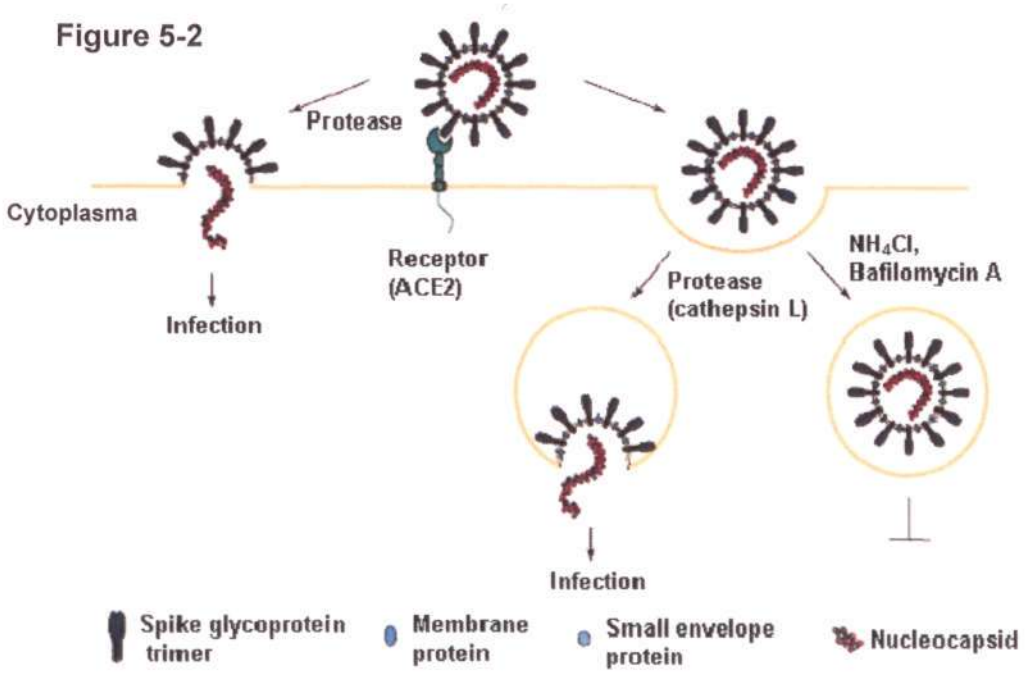


Figure 5-2. Possible SARS-CoV Entry Model.

The possible SARS-CoV entry pathways are shown in the diagram. The SARS-CoV can infect cells through direct membrane fusion and the endocytosis pathway. But, both of the pathways involve receptor binding and protease cleavage. Binding of the spike glycoprotein to the receptor (ACE2) can induce conformational alterations, which might induce the expose of the cleavage site on the S protein. After cellular protease cleavage activation, the S proteins perform further conformational change and facilitate viral fusion. If the cellular protease cleavage occurs at the cellular membrane after receptor binding, it leads to the direct membrane fusion. The S protein may also be cleaved in the endosome by cellular protease, such as the cathepsin L in the receptor mediated endocytosis pathway. It results in the fusion between the viral membrane and the endosome membrane. Acidification of the endosome will facilitate the conformational change of S protein and the cleavage of cellular protease, but it will not be the determining force to induce the conformational change. Inhibitors of endosome acidification, such as ammonium chloride (NH_4Cl) or bafilomycin A can block the endocytosis pathway, because they are believed to interfere with the cleavage of the S protein inside the endosome. The figure is adapted from (Hofmann and Pohlmann 2004).

HR Region is the Target for the SARS-CoV Viral Entry Study and Inhibitor Design

In our study, as well as others (Bosch, Martina et al. 2004), peptide of HR2 region have demonstrated a moderate inhibitory effects on the SARS-CoV viral entry, normally at the micromolar grade, while the HR1 region had no inhibitory effect. In contrast, the peptides of the HR1 region derived from the gp41 in the HIV-1 showed a potent inhibitory effect on the HIV infection. Recently, it was reported that the short peptide derived from the N-terminal region of the HR1, was able to inhibit the SARS-CoV infection (Yuan, Yi et al. 2004; Liu, Kao et al. 2009). As we discussed in Chapter 2, there existed an interaction between the HR1 and the HR2. The peptides derived from these regions were supposed to bind to the pre-hairpin intermediate during the S protein-mediated viral fusion, suggesting that peptides of the HR region could be a good target for designing entry inhibitors.

Previous works on the HIV-1 entry inhibitor design in our laboratory have successfully engineered HR2 region derived peptides that are ligated to form the protein mimetics. These protein mimetics showed an increase in the helical structure and structural stability. Trimeric protein mimetics are more potent inhibitors than their corresponding monomer's peptides. My results on SARS-CoV, however, showed that the viral inhibitory effect of the SARS-CoV HR2 trimeric mimetics did not increase substantially as we expected. Its low inhibition efficiency may be due to the sequence specificity of the inhibitory peptide. It is reported that the inhibition of the SARS-CoV infection are highly sequence specific. A few amino acids changes in the HR2 region may result in total loss of the inhibitory effect of HR2 peptides (Yuan, Yi et al. 2004).

Another plausible reason is the difference of the viral entry pathway. HIV-1 uses a direct membrane fusion pathway, whereas the SARS-CoV uses the receptor-mediated endocytosis pathway for viral entry. Limited peptide or protein mimetics are recruited into the endosome, which could be insufficient to block the viral entry. Also, the low pH inside the endosome may inhibit tight binding of the HR2 protein mimetics to the HR1 region (as what is indicated in our CD analysis in vitro in Chapter 4).

It was reported that the direct membrane fusion was 100 to 1000 fold more efficient than the endocytosis pathway in SARS-CoV viral entry. In major target organs such as bronchus, lungs and intestines (Nicholls, Poon et al. 2003; Peiris, Chu et al. 2003; To and Lo 2004; Tse, To et al. 2004), the protease secreted facilitated the SARS-CoV infection through direct membrane fusion and caused the high multiplication of the virus in those organs. Although the HR2 derived-peptide showed a low inhibitory effect to the endocytic viral entry, the HR2 derived peptide was reported to block the direct membrane fusion of the SARS-CoV at the nanomolar grade (Ujike, Nishikawa et al. 2008). Our protein mimetics demonstrated a higher stable structure and a resistance to proteolytic treatment (Chapter 4). The stronger proteolytic resistance and low hemolytic activity of the HR region derived protein mimetics, made them promising inhibitors for the direct membrane fusion adopted by the SARS-CoV in those target organs with existing secreted protease. Further study needs to be done in this field to determine the entry mechanism of the SARS-CoV and design potent inhibitors which could block the SARS-CoV viral entry in both pathways.

References

Alexander, W. A., B. Moss, et al. (1992). "Regulated expression of foreign genes in vaccinia virus under the control of bacteriophage T7 RNA polymerase and the Escherichia coli lac repressor." J Virol **66**(5): 2934-42.

Armstrong, J., H. Niemann, et al. (1984). "Sequence and topology of a model intracellular membrane protein, E1 glycoprotein, from a coronavirus." Nature **308**(5961): 751-2.

Babcock, G. J., D. J. Esshaki, et al. (2004). "Amino acids 270 to 510 of the severe acute respiratory syndrome coronavirus spike protein are required for interaction with receptor." J Virol **78**(9): 4552-60.

Baker, K. A., R. E. Dutch, et al. (1999). "Structural basis for paramyxovirus-mediated membrane fusion." Mol Cell **3**(3): 309-19.

Belouzard, S., V. C. Chu, et al. (2009). "Activation of the SARS coronavirus spike protein via sequential proteolytic cleavage at two distinct sites." Proc Natl Acad Sci U S A **106**(14): 5871-6.

Beniac, D. R., A. Andonov, et al. (2006). "Architecture of the SARS coronavirus prefusion spike." Nat Struct Mol Biol **13**(8): 751-2.

Beniac, D. R., S. L. Devarennes, et al. (2007). "Conformational Reorganization of the SARS Coronavirus Spike Following Receptor Binding: Implications for Membrane Fusion." PLoS ONE **2**(10): e1082.

Bewley, C. A., J. M. Louis, et al. (2002). "Design of a novel peptide inhibitor of HIV fusion that disrupts the internal trimeric coiled-coil of gp41." J Biol Chem **277**(16): 14238-45.

Bonavia, A., B. D. Zelus, et al. (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., L. Heijnen, et al. (1995). "Mutational analysis of the murine coronavirus spike protein: effect on cell-to-cell fusion." Virology **214**(2): 453-63.

Bosch, B. J., W. Bartelink, et al. (2008). "Cathepsin L functionally cleaves the severe acute respiratory syndrome coronavirus class I fusion protein upstream of rather than adjacent to the fusion peptide." J Virol **82**(17): 8887-90.

Bosch, B. J., B. E. Martina, et al. (2004). "Severe acute respiratory syndrome coronavirus (SARS-CoV) infection inhibition using spike protein heptad repeat-derived peptides." Proc Natl Acad Sci U S A **101**(22): 8455-60.

Bosch, B. J., R. van der Zee, et al. (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.

Bullough, P. A., F. M. Hughson, et al. (1994). "Structure of influenza haemagglutinin at the pH of membrane fusion." Nature **371**(6492): 37-43.

Caffrey, M., M. Cai, et al. (1998). "Three-dimensional solution structure of the 44 kDa ectodomain of SIV gp41." EMBO J **17**(16): 4572-84.

Carr, C. M. and P. S. Kim (1993). "A spring-loaded mechanism for the conformational change of influenza hemagglutinin." Cell **73**(4): 823-32.

Cavanagh, D. (1995). "The Coronavirus Surface Glycoprotein. ." Plenum Press,New York.

Cavanagh, D., P. J. Davis, et al. (1986). "Coronavirus IBV: partial amino terminal sequencing of spike polypeptide S2 identifies the sequence Arg-Arg-Phe-Arg-Arg at the cleavage site of the spike precursor propolypeptide of IBV strains Beaudette and M41." Virus Res **4**(2): 133-43.

Chambers, P., C. R. Pringle, et al. (1990). "Heptad repeat sequences are located adjacent to hydrophobic regions in several types of virus fusion glycoproteins." J Gen Virol **71 (Pt 12)**: 3075-80.

Chan, D. C., C. T. Chutkowski, et al. (1998). "Evidence that a prominent cavity in the coiled coil of HIV type 1 gp41 is an attractive drug target." Proc Natl Acad Sci U S A **95**(26): 15613-7.

Chan, D. C., D. Fass, et al. (1997). "Core structure of gp41 from the HIV envelope glycoprotein." Cell **89**(2): 263-73.

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

Chan, S. Y., C. J. Empig, et al. (2001). "Folate receptor-alpha is a cofactor for cellular entry by Marburg and Ebola viruses." Cell **106**(1): 117-26.

Chandran, K., N. J. Sullivan, et al. (2005). "Endosomal proteolysis of the Ebola virus glycoprotein is necessary for infection." Science **308**(5728): 1643-5.

Colman, P. M. and M. C. Lawrence (2003). "The structural biology of type I viral membrane fusion." Nat Rev Mol Cell Biol **4**(4): 309-19.

Connor, R. I., B. K. Chen, et al. (1995). "Vpr is required for efficient replication of human immunodeficiency virus type-1 in mononuclear phagocytes." Virology **206**(2): 935-44.

D.A.J.Tyrrell., J. D. A. D. M. B. C. H. C. D. H. M. S. H. L. M. K. M. (1965). "Coronavirus." Nature **220**(5168): 650.

de Groot, R. J., W. Luytjes, et al. (1987). "Evidence for a coiled-coil structure in the spike proteins of coronaviruses." J Mol Biol **196**(4): 963-6.

de Haan, C. A., L. Kuo, et al. (1998). "Coronavirus particle assembly: primary structure requirements of the membrane protein." J Virol **72**(8): 6838-50.

de Haan, C. A., P. S. Masters, et al. (2002). "The group-specific murine coronavirus genes are not essential, but their deletion, by reverse genetics, is attenuating in the natural host." Virology **296**(1): 177-89.

de Haan, C. A., K. Stadler, et al. (2004). "Cleavage inhibition of the murine coronavirus spike protein by a furin-like enzyme affects cell-cell but not virus-cell fusion." J Virol **78**(11): 6048-54.

de Haan, C. A., H. Vennema, et al. (2000). "Assembly of the coronavirus envelope: homotypic interactions between the M proteins." J Virol **74**(11): 4967-78.

Delwart, E. L., G. Mosialos, et al. (1990). "Retroviral envelope glycoproteins contain a "leucine zipper"-like repeat." AIDS Res Hum Retroviruses **6**(6): 703-6.

Deng, H., R. Liu, et al. (1996). "Identification of a major co-receptor for primary isolates of HIV-1." Nature **381**(6584): 661-6.

Deng, H. K., D. Unutmaz, et al. (1997). "Expression cloning of new receptors used by simian and human immunodeficiency viruses." Nature **388**(6639): 296-300.

Deng, Y., J. Liu, et al. (2006). "Structures and polymorphic interactions of two heptad-repeat regions of the SARS virus S2 protein." Structure **14**(5): 889-99.

Deregt, D., M. Sabara, et al. (1987). "Structural proteins of bovine coronavirus and their intracellular processing." J Gen Virol **68 (Pt 11)**: 2863-77.

Dimitrov, D. S. (2004). "Virus entry: molecular mechanisms and biomedical applications." Nat Rev Microbiol **2**(2): 109-22.

Doms, R. W., M. J. Gething, et al. (1986). "Variant influenza virus hemagglutinin that induces fusion at elevated pH." J Virol **57**(2): 603-13.

Doms, R. W., D. S. Keller, et al. (1987). "Role for adenosine triphosphate in regulating the assembly and transport of vesicular stomatitis virus G protein trimers." J Cell Biol **105**(5): 1957-69.

Dong, J., M. G. Roth, et al. (1992). "A chimeric avian retrovirus containing the influenza virus hemagglutinin gene has an expanded host range." J Virol **66**(12): 7374-82.

Drosten, C., S. Gunther, et al. (2003). "Identification of a novel coronavirus in patients with severe acute respiratory syndrome." N Engl J Med **348**(20): 1967-76.

Du, L., R. Y. Kao, et al. (2007). "Cleavage of spike protein of SARS coronavirus by protease factor Xa is associated with viral infectivity." Biochem Biophys Res Commun **359**(1): 174-9.

Duquerroy, S., A. Vigouroux, et al. (2005). "Central ions and lateral asparagine/glutamine zippers stabilize the post-fusion hairpin conformation of the SARS coronavirus spike glycoprotein." Virology **335**(2): 276-85.

Ebert, D. H., J. Deussing, et al. (2002). "Cathepsin L and cathepsin B mediate reovirus disassembly in murine fibroblast cells." J Biol Chem **277**(27): 24609-17.

Eckert, D. M. and P. S. Kim (2001). "Design of potent inhibitors of HIV-1 entry from the gp41 N-peptide region." Proc Natl Acad Sci U S A **98**(20): 11187-92.

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

Escors, D., E. Camafeita, et al. (2001). "Organization of two transmissible gastroenteritis coronavirus membrane protein topologies within the virion and core." J Virol **75**(24): 12228-40.

Escors, D., J. Ortego, et al. (2001). "The membrane M protein carboxy terminus binds to transmissible gastroenteritis coronavirus core and contributes to core stability." J Virol **75**(3): 1312-24.

F.A. Murphy, C. M. F., D.H.L. Bishop, S.A. Ghabrial, A.W. Jarvis, G.P. Martelli, M.A. Mayo, M.D. Summers. (1995). "Virus Taxonomy: The Sixth Report of the International Committee on Taxonomy of Viruses." Virus Taxonomy.

Ferrer, M., T. M. Kapoor, et al. (1999). "Selection of gp41-mediated HIV-1 cell entry inhibitors from biased combinatorial libraries of non-natural binding elements." Nat Struct Biol **6**(10): 953-60.

Fields, G. B. and R. L. Noble (1990). "Solid phase peptide synthesis utilizing 9-fluorenylmethoxycarbonyl amino acids." Int J Pept Protein Res **35**(3): 161-214.

Fischer, F., D. Peng, et al. (1997). "The internal open reading frame within the nucleocapsid gene of mouse hepatitis virus encodes a structural protein that is not essential for viral replication." J Virol **71**(2): 996-1003.

Follis, K. E., J. York, et al. (2006). "Furin cleavage of the SARS coronavirus spike glycoprotein enhances cell-cell fusion but does not affect virion entry." Virology **350**(2): 358-69.

Fosmire, J. A., K. Hwang, et al. (1992). "Identification and characterization of a coronavirus packaging signal." J Virol **66**(6): 3522-30.

Gallagher, T. M., C. Escarmis, et al. (1991). "Alteration of the pH dependence of coronavirus-induced cell fusion: effect of mutations in the spike glycoprotein." J Virol **65**(4): 1916-28.

Gallaher, W. R., J. P. Segrest, et al. (1992). "Are fusion peptides really "sided" insertional helices?" Cell **70**(4): 531-2.

Garry, R. F. and S. Dash (2003). "Proteomics computational analyses suggest that hepatitis C virus E1 and pestivirus E2 envelope glycoproteins are truncated class II fusion proteins." Virology **307**(2): 255-65.

Garry, W. R. G. a. R. F. (2003). "Model of the pre-insertion region of the spike (S2) fusion glycoprotein of the human SARS coronavirus: implications for antiviral therapeutics." [Online] http://www.virology.net/Articles/sars/s2model.html.

Geoghegan, K. F. and J. G. Stroh (1992). "Site-directed conjugation of nonpeptide groups to peptides and proteins via periodate oxidation of a 2-amino alcohol. Application to modification at N-terminal serine." Bioconjug Chem **3**(2): 138-46.

Goldsmith, C. S., K. M. Tatti, et al. (2004). "Ultrastructural characterization of SARS coronavirus." Emerg Infect Dis **10**(2): 320-6.

Gombold, J. L., S. T. Hingley, et al. (1993). "Fusion-defective mutants of mouse hepatitis virus A59 contain a mutation in the spike protein cleavage signal." J Virol **67**(8): 4504-12.

Gorbalenya, A. E., E. J. Snijder, et al. (2004). "Severe acute respiratory syndrome coronavirus phylogeny: toward consensus." J Virol **78**(15): 7863-6.

Guan, Y., B. J. Zheng, et al. (2003). "Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China." Science **302**(5643): 276-8.

Guillen, J., A. J. Perez-Berna, et al. (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.

Haijema, B. J., H. Volders, et al. (2004). "Live, attenuated coronavirus vaccines through the directed deletion of group-specific genes provide protection against feline infectious peritonitis." J Virol **78**(8): 3863-71.

Han, D. P., A. Penn-Nicholson, et al. (2006). "Identification of critical determinants on ACE2 for SARS-CoV entry and development of a potent entry inhibitor." Virology **350**(1): 15-25.

Hansen, G. H., B. Delmas, et al. (1998). "The coronavirus transmissible gastroenteritis virus causes infection after receptor-mediated endocytosis and acid-dependent fusion with an intracellular compartment." J Virol **72**(1): 527-34.

Harrison, S. C. (2005). "Mechanism of membrane fusion by viral envelope proteins." Adv Virus Res **64**: 231-61.

Heinz, F. X. and S. L. Allison (2001). "The machinery for flavivirus fusion with host cell membranes." Curr Opin Microbiol **4**(4): 450-5.

Heinz, F. X. and S. L. Allison (2003). "Flavivirus structure and membrane fusion." Adv Virus Res **59**: 63-97.

Helenius, A., J. Kartenbeck, et al. (1980). "On the entry of Semliki forest virus into BHK-21 cells." J Cell Biol **84**(2): 404-20.

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

Holmes, K. V., E. W. Doller, et al. (1981). "Tunicamycin resistant glycosylation of coronavirus glycoprotein: demonstration of a novel type of viral glycoprotein." Virology **115**(2): 334-44.

Holmes, K. V. and L. Enjuanes (2003). "Virology. The SARS coronavirus: a postgenomic era." Science **300**(5624): 1377-8.

Huang, I. C., B. J. Bosch, et al. (2006). "SARS coronavirus, but not human coronavirus NL63, utilizes cathepsin L to infect ACE2-expressing cells." J Biol Chem **281**(6): 3198-203.

Jeffers, S. A., S. M. Tusell, et al. (2004). "CD209L (L-SIGN) is a receptor for severe acute respiratory syndrome coronavirus." Proc Natl Acad Sci U S A **101**(44): 15748-53.

Judice, J. K., J. Y. Tom, et al. (1997). "Inhibition of HIV type 1 infectivity by constrained alpha-helical peptides: implications for the viral fusion mechanism." Proc Natl Acad Sci U S A **94**(25): 13426-30.

Jun Shao, J. P. T. (1995). "Unprotected Peptides as Building Blocks for the Synthesis of Peptide Dendrimers with Oxime, Hydrazone, and Thiazolidine Linkages." Journal of the American Chemical Society **117**(14): 3893-3899.

Kam, Y. W., Y. Okumura, et al. (2009). "Cleavage of the SARS coronavirus spike glycoprotein by airway proteases enhances virus entry into human bronchial epithelial cells in vitro." PLoS One **4**(11): e7870.

Kawabata, K., T. Hagio, et al. (2002). "The role of neutrophil elastase in acute lung injury." Eur J Pharmacol **451**(1): 1-10.

Kawase, M., K. Shirato, et al. (2008). "Protease-mediated entry via endosome of human coronavirus 229E." J Virol.

Kilby, J. M., S. Hopkins, et al. (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.

Klenk, H. D., R. Rott, et al. (1975). "Activation of influenza A viruses by trypsin treatment." Virology **68**(2): 426-39.

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

Ksiazek, T. G., D. Erdman, et al. (2003). "A novel coronavirus associated with severe acute respiratory syndrome." N Engl J Med **348**(20): 1953-66.

Kubo, H., Y. K. Yamada, et al. (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.

Kunkel, F. and G. Herrler (1993). "Structural and functional analysis of the surface protein of human coronavirus OC43." Virology **195**(1): 195-202.

Lai, M. M. C., K.V.Holmes. (2001). "Coronaviridae: the viruses and heir replication." Fundamental virology, 4th edition. Lippincott Williams & Wilkins, Philadelphia.

Lai, S. C., P. C. Chong, et al. (2005). "Characterization of neutralizing monoclonal antibodies recognizing a 15-residues epitope on the spike protein HR2 region of severe acute respiratory syndrome coronavirus (SARS-CoV)." J Biomed Sci **12**(5): 711-27.

Lamb, R. A., Kolakofsky, D (2001). "Paramyxoviridae: The viruses and their replication." Fields Virology. 4th Edition, Vol. 2 Lippincott-Raven, Philadelphia.: 1305-1340.

Lambert, D. M., S. Barney, et al. (1996). "Peptides from conserved regions of paramyxovirus fusion (F) proteins are potent inhibitors of viral fusion." Proc Natl Acad Sci U S A **93**(5): 2186-91.

Landau, N. R., K. A. Page, et al. (1991). "Pseudotyping with human T-cell leukemia virus type I broadens the human immunodeficiency virus host range." J Virol **65**(1): 162-9.

Laude, H., Masters, P.S. (1995). "The Coronavirus Nucleocapsid Protein." Plenum Press, New York.

Laude, H., D. Rasschaert, et al. (1987). "Sequence and N-terminal processing of the transmembrane protein E1 of the coronavirus transmissible gastroenteritis virus." J Gen Virol **68 (Pt 6)**: 1687-93.

Lescar, J., A. Roussel, et al. (2001). "The Fusion glycoprotein shell of Semliki Forest virus: an icosahedral assembly primed for fusogenic activation at endosomal pH." Cell **105**(1): 137-48.

Li, D. and D. Cavanagh (1992). "Coronavirus IBV-induced membrane fusion occurs at near-neutral pH." Arch Virol **122**(3-4): 307-16.

Li, F., M. Berardi, et al. (2006). "Conformational states of the severe acute respiratory syndrome coronavirus spike protein ectodomain." J Virol **80**(14): 6794-800.

Li, F., W. Li, et al. (2005). "Structure of SARS coronavirus spike receptor-binding domain complexed with receptor." Science **309**(5742): 1864-8.

Li, W., M. J. Moore, et al. (2003). "Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus." Nature **426**(6965): 450-4.

Liao, C. L., X. Zhang, et al. (1995). "Coronavirus defective-interfering RNA as an expression vector: the generation of a pseudorecombinant mouse hepatitis virus expressing hemagglutinin-esterase." Virology **208**(1): 319-27.

Liu, I. J., C. L. Kao, et al. (2009). "Identification of a minimal peptide derived from heptad repeat (HR) 2 of spike protein of SARS-CoV and combination of HR1-derived peptides as fusion inhibitors." Antiviral Res **81**(1): 82-7.

Liu, S., G. Xiao, et al. (2004). "Interaction between heptad repeat 1 and 2 regions in spike protein of SARS-associated coronavirus: implications for virus fusogenic mechanism and identification of fusion inhibitors." Lancet **363**(9413): 938-47.

Locker, J. K., J. K. Rose, et al. (1992). "Membrane assembly of the triple-spanning coronavirus M protein. Individual transmembrane domains show preferred orientation." J Biol Chem **267**(30): 21911-8.

Louis, J. M., C. A. Bewley, et al. (2001). "Design and properties of N(CCG)-gp41, a chimeric gp41 molecule with nanomolar HIV fusion inhibitory activity." J Biol Chem **276**(31): 29485-9.

Louis, J. M., I. Nesheiwat, et al. (2003). "Covalent trimers of the internal N-terminal trimeric coiled-coil of gp41 and antibodies directed against them are potent inhibitors of HIV envelope-mediated cell fusion." J Biol Chem **278**(22): 20278-85.

Lu, M., S. C. Blacklow, et al. (1995). "A trimeric structural domain of the HIV-1 transmembrane glycoprotein." Nat Struct Biol **2**(12): 1075-82.

Lu, Y., D. X. Liu, et al. (2008). "Lipid rafts are involved in SARS-CoV entry into Vero E6 cells." Biochem Biophys Res Commun **369**(2): 344-9.

Luo, Z., A. M. Matthews, et al. (1999). "Amino acid substitutions within the leucine zipper domain of the murine coronavirus spike protein cause defects in oligomerization and the ability to induce cell-to-cell fusion." J Virol **73**(10): 8152-9.

Luytjes, W., L. S. Sturman, et al. (1987). "Primary structure of the glycoprotein E2 of coronavirus MHV-A59 and identification of the trypsin cleavage site." Virology **161**(2): 479-87.

Machamer, C. E., M. G. Grim, et al. (1993). "Retention of a cis Golgi protein requires polar residues on one face of a predicted alpha-helix in the transmembrane domain." Mol Biol Cell **4**(7): 695-704.

Madu, I. G., S. L. Roth, et al. (2009). "Characterization of a highly conserved domain within the severe acute respiratory syndrome coronavirus spike protein S2 domain with characteristics of a viral fusion peptide." J Virol **83**(15): 7411-21.

Magiorkinis, G., E. Magiorkinis, et al. (2004). "Phylogenetic analysis of the full-length SARS-CoV sequences: evidence for phylogenetic discordance in three genomic regions." J Med Virol **74**(3): 369-72.

Marsh, M. and A. Pelchen-Matthews (2000). "Endocytosis in viral replication." Traffic **1**(7): 525-32.

Martina, B. E., B. L. Haagmans, et al. (2003). "Virology: SARS virus infection of cats and ferrets." Nature **425**(6961): 915.

Matsuyama, S., M. Ujike, et al. (2005). "Protease-mediated enhancement of severe acute respiratory syndrome coronavirus infection." Proc Natl Acad Sci U S A **102**(35): 12543-7.

McReynolds, S., S. Jiang, et al. (2008). "Characterization of the prefusion and transition states of severe acute respiratory syndrome coronavirus S2-HR2." Biochemistry **47**(26): 6802-8.

Michael M. C. Lai and, K. V. H. (1996). "Coronaviridae: The Viruses and Their Replication." Fields Virology. Lippincott-Raven, Philadelphia.: 1075-1103.

Molloy, S. S., E. D. Anderson, et al. (1999). "Bi-cycling the furin pathway: from TGN localization to pathogen activation and embryogenesis." Trends Cell Biol **9**(1): 28-35.

Moore, M. J., T. Dorfman, et al. (2004). "Retroviruses pseudotyped with the severe acute respiratory syndrome coronavirus spike protein efficiently infect cells expressing angiotensin-converting enzyme 2." J Virol **78**(19): 10628-35.

Mossel, E. C., C. Huang, et al. (2005). "Exogenous ACE2 expression allows refractory cell lines to support severe acute respiratory syndrome coronavirus replication." J Virol **79**(6): 3846-50.

Nash, T. C. and M. J. Buchmeier (1997). "Entry of mouse hepatitis virus into cells by endosomal and nonendosomal pathways." Virology **233**(1): 1-8.

Navas-Martin, S. R. and S. Weiss (2004). "Coronavirus replication and pathogenesis: Implications for the recent outbreak of severe acute respiratory syndrome (SARS), and the challenge for vaccine development." J Neurovirol **10**(2): 75-85.

Ni, L., J. Zhu, et al. (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.

Nicholls, J. M., L. L. Poon, et al. (2003). "Lung pathology of fatal severe acute respiratory syndrome." Lancet **361**(9371): 1773-8.

Nie, Y., P. Wang, et al. (2004). "Highly infectious SARS-CoV pseudotyped virus reveals the cell tropism and its correlation with receptor expression." Biochem Biophys Res Commun **321**(4): 994-1000.

Niemann, H., R. Geyer, et al. (1984). "The carbohydrates of mouse hepatitis virus (MHV) A59: structures of the O-glycosidically linked oligosaccharides of glycoprotein E1." EMBO J **3**(3): 665-70.

Nomura, R., A. Kiyota, et al. (2004). "Human coronavirus 229E binds to CD13 in rafts and enters the cell through caveolae." J Virol **78**(16): 8701-8.

Opstelten, D. J., M. J. Raamsman, et al. (1995). "Envelope glycoprotein interactions in coronavirus assembly." J Cell Biol **131**(2): 339-49.

Ortego, J., I. Sola, et al. (2003). "Transmissible gastroenteritis coronavirus gene 7 is not essential but influences in vivo virus replication and virulence." Virology **308**(1): 13-22.

Peiris, J. S., C. M. Chu, et al. (2003). "Clinical progression and viral load in a community outbreak of coronavirus-associated SARS pneumonia: a prospective study." Lancet **361**(9371): 1767-72.

Peiris, J. S., S. T. Lai, et al. (2003). "Coronavirus as a possible cause of severe acute respiratory syndrome." Lancet **361**(9366): 1319-25.

Poutanen, S. M., D. E. Low, et al. (2003). "Identification of severe acute respiratory syndrome in Canada." N Engl J Med **348**(20): 1995-2005.

Qiu, Z., S. T. Hingley, et al. (2006). "Endosomal proteolysis by cathepsins is necessary for murine coronavirus mouse hepatitis virus type 2 spike-mediated entry." J Virol **80**(12): 5768-76.

Qureshi, N. M., D. H. Coy, et al. (1990). "Characterization of a putative cellular receptor for HIV-1 transmembrane glycoprotein using synthetic peptides." Aids **4**(6): 553-8.

Raamsman, M. J., J. K. Locker, et al. (2000). "Characterization of the coronavirus mouse hepatitis virus strain A59 small membrane protein E." J Virol **74**(5): 2333-42.

Rapaport, D., M. Ovadia, et al. (1995). "A synthetic peptide corresponding to a conserved heptad repeat domain is a potent inhibitor of Sendai virus-cell fusion: an emerging similarity with functional domains of other viruses." EMBO J **14**(22): 5524-31.

Ray, R. B., A. Basu, et al. (2004). "Ebola virus glycoprotein-mediated anoikis of primary human cardiac microvascular endothelial cells." Virology **321**(2): 181-8.

Regl, G., A. Kaser, et al. (1999). "The hemagglutinin-esterase of mouse hepatitis virus strain S is a sialate-4-O-acetylesterase." J Virol **73**(6): 4721-7.

Rockwell, N. C., D. J. Krysan, et al. (2002). "Precursor processing by kex2/furin proteases." Chem Rev **102**(12): 4525-48.

Root, M. J., M. S. Kay, et al. (2001). "Protein design of an HIV-1 entry inhibitor." Science **291**(5505): 884-8.

Rota, P. A., M. S. Oberste, et al. (2003). "Characterization of a novel coronavirus associated with severe acute respiratory syndrome." Science **300**(5624): 1394-9.

Routledge, E., R. Stauber, et al. (1991). "Analysis of murine coronavirus surface glycoprotein functions by using monoclonal antibodies." J Virol **65**(1): 254-62.

Sadler, K., Y. Zhang, et al. (2008). "Quaternary protein mimetics of gp41 elicit neutralizing antibodies against HIV fusion-active intermediate state." Biopolymers **90**(3): 320-9.

Sainz, B., Jr., E. C. Mossel, et al. (2006). "Inhibition of severe acute respiratory syndrome-associated coronavirus (SARS-CoV) infectivity by peptides analogous to the viral spike protein." Virus Res **120**(1-2): 146-55.

Sainz, B., Jr., J. M. Rausch, et al. (2005). "Identification and characterization of the putative fusion peptide of the severe acute respiratory syndrome-associated coronavirus spike protein." J Virol **79**(11): 7195-206.

Schultze, B., D. Cavanagh, et al. (1992). "Neuraminidase treatment of avian infectious bronchitis coronavirus reveals a hemagglutinating activity that is dependent on sialic acid-containing receptors on erythrocytes." Virology **189**(2): 792-4.

Schultze, B., H. J. Gross, et al. (1991). "The S protein of bovine coronavirus is a hemagglutinin recognizing 9-O-acetylated sialic acid as a receptor determinant." J Virol **65**(11): 6232-7.

Schultze, B., C. Krempl, et al. (1996). "Transmissible gastroenteritis coronavirus, but not the related porcine respiratory coronavirus, has a sialic acid (N-glycolylneuraminic acid) binding activity." J Virol **70**(8): 5634-7.

Schwegmann-Wessels, C. and G. Herrler (2006). "Sialic acids as receptor determinants for coronaviruses." Glycoconj J **23**(1-2): 51-8.

Scianimanico, S., G. Schoehn, et al. (2000). "Membrane association induces a conformational change in the Ebola virus matrix protein." EMBO J **19**(24): 6732-41.

Senanayake, S. D., M. A. Hofmann, et al. (1992). "The nucleocapsid protein gene of bovine coronavirus is bicistronic." J Virol **66**(9): 5277-83.

Sha, Y., Y. Wu, et al. (2006). "A convenient cell fusion assay for the study of SARS-CoV entry and inhibition." IUBMB Life **58**(8): 480-6.

Sia, S. K., P. A. Carr, et al. (2002). "Short constrained peptides that inhibit HIV-1 entry." Proc Natl Acad Sci U S A **99**(23): 14664-9.

Sia, S. K. and P. S. Kim (2003). "Protein grafting of an HIV-1-inhibiting epitope." Proc Natl Acad Sci U S A **100**(17): 9756-61.

Siddell, S. G. (1995). "The *Coronaviridae*: an introduction. ." Plenum Press,New York. NY.

Siddell, S. G. (1995). "The small-membrane protein." The *Coronaviridae*I. Plenum Press, New York: 181-189.

Sieczkarski, S. B. and G. R. Whittaker (2002). "Dissecting virus entry via endocytosis." J Gen Virol **83**(Pt 7): 1535-45.

Simmons, G., D. N. Gosalia, et al. (2005). "Inhibitors of cathepsin L prevent severe acute respiratory syndrome coronavirus entry." Proc Natl Acad Sci U S A **102**(33): 11876-81.

Simmons, G., J. D. Reeves, et al. (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.

Singh, M., B. Berger, et al. (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 D. C. Wiley (1998). "Coiled coils in both intracellular vesicle and viral membrane fusion." Cell **95**(7): 871-4.

Spaan, W., D. Cavanagh, et al. (1988). "Coronaviruses: structure and genome expression." J Gen Virol **69 (Pt 12)**: 2939-52.

Sturman, L. S., C. S. Ricard, et al. (1985). "Proteolytic cleavage of the E2 glycoprotein of murine coronavirus: activation of cell-fusing activity of virions by trypsin and separation of two different 90K cleavage fragments." J Virol **56**(3): 904-11.

Sturman, L. S., C. S. Ricard, et al. (1990). "Conformational change of the coronavirus peplomer glycoprotein at pH 8.0 and 37 degrees C correlates with virus aggregation and virus-induced cell fusion." J Virol **64**(6): 3042-50.

Suarez, T., W. R. Gallaher, et al. (2000). "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.

Suomalainen, M. and H. Garoff (1994). "Incorporation of homologous and heterologous proteins into the envelope of Moloney murine leukemia virus." J Virol **68**(8): 4879-89.

Supekar, V. M., C. Bruckmann, et al. (2004). "Structure of a proteolytically resistant core from the severe acute respiratory syndrome coronavirus S2 fusion protein." Proc Natl Acad Sci U S A **101**(52): 17958-63.

Taguchi, F. (1993). "Fusion formation by the uncleaved spike protein of murine coronavirus JHMV variant cl-2." J Virol **67**(3): 1195-202.

Taguchi, F. (1995). "The S2 subunit of the murine coronavirus spike protein is not involved in receptor binding." J Virol **69**(11): 7260-3.

Tam, J. P. and Q. Yu (2002). "A facile ligation approach to prepare three-helix bundles of HIV fusion-state protein mimetics." Org Lett **4**(23): 4167-70.

Tan, K., J. Liu, et al. (1997). "Atomic structure of a thermostable subdomain of HIV-1 gp41." Proc Natl Acad Sci U S A **94**(23): 12303-8.

To, K. F. and A. W. Lo (2004). "Exploring the pathogenesis of severe acute respiratory syndrome (SARS): the tissue distribution of the coronavirus (SARS-CoV) and its putative receptor, angiotensin-converting enzyme 2 (ACE2)." J Pathol **203**(3): 740-3.

Tripet, B., M. W. Howard, et al. (2004). "Structural characterization of the SARS-coronavirus spike S fusion protein core." J Biol Chem **279**(20): 20836-49.

Tsang, K. W., P. L. Ho, et al. (2003). "A cluster of cases of severe acute respiratory syndrome in Hong Kong." N Engl J Med **348**(20): 1977-85.

Tse, G. M., K. F. To, et al. (2004). "Pulmonary pathological features in coronavirus associated severe acute respiratory syndrome (SARS)." J Clin Pathol **57**(3): 260-5.

Tsui, P. T., M. L. Kwok, et al. (2003). "Severe acute respiratory syndrome: clinical outcome and prognostic correlates." Emerg Infect Dis **9**(9): 1064-9.

Ujike, M., H. Nishikawa, et al. (2008). "Heptad repeat-derived peptides block protease-mediated direct entry from the cell surface of severe acute respiratory syndrome coronavirus but not entry via the endosomal pathway." J Virol **82**(1): 588-92.

Ujike, M., H. Nishikawa, et al. (2008). "Heptad repeat-derived peptides block protease-mediated direct entry from the cell surface of severe acute respiratory syndrome coronavirus but not entry via the endosomal pathway." J Virol **82**(1): 588-92.

Vennema, H., G. J. Godeke, et al. (1996). "Nucleocapsid-independent assembly of coronavirus-like particles by co-expression of viral envelope protein genes." EMBO J **15**(8): 2020-8.

Vennema, H., L. Heijnen, et al. (1990). "Intracellular transport of recombinant coronavirus spike proteins: implications for virus assembly." J Virol **64**(1): 339-46.

Vennema, H., P. J. Rottier, et al. (1990). "Biosynthesis and function of the coronavirus spike protein." Adv Exp Med Biol **276**: 9-19.

Wang, H., P. Yang, et al. (2008). "SARS coronavirus entry into host cells through a novel clathrin- and caveolae-independent endocytic pathway." Cell Res **18**(2): 290-301.

Wang, P., J. Chen, et al. (2004). "Expression cloning of functional receptor used by SARS coronavirus." Biochem Biophys Res Commun **315**(2): 439-44.

Wang, S., F. Guo, et al. (2008). "Endocytosis of the receptor-binding domain of SARS-CoV spike protein together with virus receptor ACE2." Virus Res **136**(1-2): 8-15.

Watanabe, R., S. Matsyama, et al. (2008). "Entry from cell surface of SARS coronavirus with cleaved S protein as revealed by pseudotype virus bearing cleaved S protein." J Virol.

Watanabe, S., A. Takada, et al. (2000). "Functional importance of the coiled-coil of the Ebola virus glycoprotein." J Virol **74**(21): 10194-201.

Weissenhorn, W., L. J. Calder, et al. (1998). "The central structural feature of the membrane fusion protein subunit from the Ebola virus glycoprotein is a long triple-stranded coiled coil." Proc Natl Acad Sci U S A **95**(11): 6032-6.

Weissenhorn, W., A. Carfi, et al. (1998). "Crystal structure of the Ebola virus membrane fusion subunit, GP2, from the envelope glycoprotein ectodomain." Mol Cell **2**(5): 605-16.

Weissenhorn, W., A. Dessen, et al. (1999). "Structural basis for membrane fusion by enveloped viruses." Mol Membr Biol **16**(1): 3-9.

Weissenhorn, W., A. Dessen, et al. (1997). "Atomic structure of the ectodomain from HIV-1 gp41." Nature **387**(6631): 426-30.

Wild, C., T. Greenwell, et al. (1993). "A synthetic peptide from HIV-1 gp41 is a potent inhibitor of virus-mediated cell-cell fusion." AIDS Res Hum Retroviruses **9**(11): 1051-3.

Wild, C., T. Oas, et al. (1992). "A synthetic peptide inhibitor of human immunodeficiency virus replication: correlation between solution structure and viral inhibition." Proc Natl Acad Sci U S A **89**(21): 10537-41.

Wild, C. T., D. C. Shugars, et al. (1994). "Peptides corresponding to a predictive alpha-helical domain of human immunodeficiency virus type 1 gp41 are potent inhibitors of virus infection." Proc Natl Acad Sci U S A **91**(21): 9770-4.

Wilson, I. A., J. J. Skehel, et al. (1981). "Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 A resolution." Nature **289**(5796): 366-73.

Winter, C., C. Schwegmann-Wessels, et al. (2006). "Sialic acid is a receptor determinant for infection of cells by avian Infectious bronchitis virus." J Gen Virol **87**(Pt 5): 1209-16.

Wong, S. K., W. Li, et al. (2004). "A 193-amino acid fragment of the SARS coronavirus S protein efficiently binds angiotensin-converting enzyme 2." J Biol Chem **279**(5): 3197-201.

Wool-Lewis, R. J. and P. Bates (1998). "Characterization of Ebola virus entry by using pseudotyped viruses: identification of receptor-deficient cell lines." J Virol **72**(4): 3155-60.

Wu, C. S., K. Ikeda, et al. (1981). "Ordered conformation of polypeptides and proteins in acidic dodecyl sulfate solution." Biochemistry **20**(3): 566-70.

Wu, X. D., B. Shang, et al. (2004). "The spike protein of severe acute respiratory syndrome (SARS) is cleaved in virus infected Vero-E6 cells." Cell Res **14**(5): 400-6.

Xiao, X., S. Chakraborti, et al. (2003). "The SARS-CoV S glycoprotein: expression and functional characterization." Biochem Biophys Res Commun **312**(4): 1159-64.

Xu, Y., Y. Liu, et al. (2004). "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., Z. Lou, et al. (2004). "Crystal structure of severe acute respiratory syndrome coronavirus spike protein fusion core." J Biol Chem **279**(47): 49414-9.

Xu, Y., J. Zhu, et al. (2004). "Characterization of the heptad repeat regions, HR1 and HR2, and design of a fusion core structure model of the spike protein from severe acute respiratory syndrome (SARS) coronavirus." Biochemistry **43**(44): 14064-71.

Yamada, Y. K., K. Takimoto, et al. (1998). "Requirement of proteolytic cleavage of the murine coronavirus MHV-2 spike protein for fusion activity." Adv Exp Med Biol **440**: 89-93.

Yang, Z. Y., Y. Huang, et al. (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.

Yao, Q. and R. W. Compans (1996). "Peptides corresponding to the heptad repeat sequence of human parainfluenza virus fusion protein are potent inhibitors of virus infection." Virology **223**(1): 103-12.

Yao, Y. X., J. Ren, et al. (2004). "Cleavage and serum reactivity of the severe acute respiratory syndrome coronavirus spike protein." J Infect Dis **190**(1): 91-8.

Yokomori, K., S. C. Baker, et al. (1992). "Hemagglutinin-esterase-specific monoclonal antibodies alter the neuropathogenicity of mouse hepatitis virus." J Virol **66**(5): 2865-74.

Yoo, D. W., M. D. Parker, et al. (1991). "The S2 subunit of the spike glycoprotein of bovine coronavirus mediates membrane fusion in insect cells." Virology **180**(1): 395-9.

Young, J. K., D. Li, et al. (1999). "Interaction of peptides with sequences from the Newcastle disease virus fusion protein heptad repeat regions." J Virol **73**(7): 5945-56.

Yuan, K., L. Yi, et al. (2004). "Suppression of SARS-CoV entry by peptides corresponding to heptad regions on spike glycoprotein." Biochem Biophys Res Commun **319**(3): 746-52.

Zheng, B. J., Y. Guan, et al. (2005). "Synthetic peptides outside the spike protein heptad repeat regions as potent inhibitors of SARS-associated coronavirus." Antivir Ther **10**(3): 393-403.

Zhu, J., G. Xiao, et al. (2004). "Following the rule: formation of the 6-helix bundle of the fusion core from severe acute respiratory syndrome coronavirus spike protein and identification of potent peptide inhibitors." Biochem Biophys Res Commun **319**(1): 283-8.