Development of SARS-COV therapeutics using quaternary protein mimetics

Neo, Tuan Ling

2009

Neo, T. L. (2009). Development of SARS-COV therapeutics using quaternary protein mimetics. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/41875>

<https://doi.org/10.32657/10356/41875>

DEVELOPMENT OF SARS-COV THERAPEUTICS USING QUATERNARY PROTEIN MIMETICS

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

Neo Tuan Ling

School of Biological Sciences

**A thesis submitted to the Nanyang Technological University in partial
fulfillment of the requirement for the degree of Doctor of Philosophy
2009**

Acknowledgements

I would like to express my gratitude to all those who had given me the possibility to complete this thesis and its accompanying research work. My sincere gratitude goes out to my thesis supervisor, Professor James P. Tam, whom I am indebted to for his stimulating suggestions, discussions, encouragement and support through these years.

I would like to thank Assoc. Prof. Susana Shochat, Assoc. Prof. Liu Ding Xiang and Asst. Prof. Kristen Sadler for their expertise, advice and comments to my work.

Special thanks to Dr. Lu Yanning, for her expertise, encouragement and support in every aspect. I would also like to thank all the former and present members of Professor Tam's laboratory and all the staff and students from School of Biological Sciences for their help and support all this while. To the examiners, taking precious time off their busy schedules to read this thesis, thank you very much. Finally, I would like to thank my family and friends for their constant emotional and moral support through ups and downs during these times.

Table of Contents

Acknowledgements	1
Table of Contents	2
List of Tables	7
List of Figures	8
Abbreviations	10
Abstract	14
Chapter 1: Introduction	16
1. Quaternary Proteins	16
1.1 Quaternary proteins	16
1.2 Quaternary proteins in viruses	19
2. Virus Envelope Glycoprotein	21
2.1 Virus entry and fusion	21
2.2 Class I virus fusion proteins	22
3. Severe Acute Respiratory Syndrome Associated Coronavirus (SARS-CoV)	27
3.1 Background and SARS as an infectious disease	27
3.2 SARS Coronavirus (SARS-CoV)	28
3.3 Structural proteins of coronaviruses	28
3.4 Immune responses to SARS-CoV	29
4. SARS-CoV Spike Protein and its functional domains	33
4.1 SARS-CoV Spike (S) Protein	33
4.2 Structural features and functional domains of S protein	33
4.3 Receptor Binding Domain (RBD)	34
4.4 Putative Fusion Peptide (FP)	36
4.5 Heptad repeats 1 and Heptad repeats 2 (HR1 and HR2)	37

4.6 Pre-transmembrane or Membrane Proximal External Region (MPER)	37
4.7 Development of SARS-CoV vaccines and therapeutics	38
5. MPER in Class I virus fusion proteins	43
5.1 Conservation and importance of MPER in Class I viruses	43
5.2 MPER in SARS-CoV S protein	45
5.3 Proposed functions of MPER in virus fusion	45
6. Aims	49
6.1 Experimental design to study MPER of SARS-CoV	50
6.2 Experimental design and approach to a high affinity probe – “Synthetic antibody”	51
Chapter 2: Importance of MPER in SARS-CoV Viral Infectivity	53
1. Introduction	53
2. Materials and Methods	56
2.1 Site-directed mutagenesis of SARS-CoV S protein	56
2.2 Preparation of pseudotyped SARS-CoV containing different S protein mutants	57
2.3 Quantification of virus by reverse transcriptase (RT) assay	57
2.4 Single cycle infectivity assay	58
2.5 Solid phase peptide synthesis	58
2.6 Preparation of LUVs	59
2.7 Liposome leakage assay	59
2.8 Circular dichroism spectroscopy	60

3. Results	61
3.1 Site-directed mutagenesis of MPER of SARS S protein	61
3.2 Quantification of virus by RT assay	63
3.3 Infectivity of mutant pseudovirus on VeroE6 cells	
– Luciferase assay	63
3.4 Synthesis of mutant MPER peptides	65
3.5 Liposome leakage abilities of MPER mutant peptides	65
3.6 CD spectroscopy of MPER mutant peptides under different TFE conditions	69
4. Discussion	74

Chapter 3: MPER Quaternary Protein Mimetics of Intermediate States

During Fusion Event	79
1. Introduction	79
2. Materials and Methods	84
2.1 Solid phase peptide synthesis	84
2.2 Synthesis of N terminal linked and C terminal linked dimers	84
2.3 Preparation of LUVs	85
2.4 Liposome leakage assay	85
2.5 Ion conductance experiments	86
2.6 SARS CoV production	87
2.7 Toxicity tests and SARS CoV inhibition assay on Vero E6 cells	87
2.8 Circular dichroism spectroscopy	88

3. Results	89
3.1 Synthesis of MPER peptide and its quaternary protein mimetics as dimers	89
3.2 Liposome leakage abilities of MPER monomer and dimers	92
3.3 Ion conductance of MPER in lipid bilayer	94
3.4 CD spectroscopy of MPER and MPER dimers	95
3.5 Inhibitory activities of MPER monomer and dimers on SARS CoV infectivity	100
4. Discussion	102

Chapter 4: Quaternary Protein Mimetics via Self Interacting Sequences in SARS-CoV Spike Protein using a Novel Method – Implications in Antibody Mimetics	109
1. Introduction	109
2. Material and Methods	112
2.1 Expression and purification of SARS-CoV S protein ectodomain	112
2.2 SPOT synthesis and screening of SARS-CoV S protein peptide library	113
2.3 Peptide synthesis of biotinylated peptides	113
2.4 Pull down assay	114
2.5 Surface plasmon resonance	114
2.6 Immunofluorescence staining	115
3. Results	116
3.1 Synthesis and screening of S protein peptide library	116
3.2 Tabulation and analysis of results from screening	120

3.3 Validation of interaction by S protein pull down assay	123
3.4 Additive effects of peptides through noncovalent heterovalency	127
3.5 Homodimer increases pull down affinity to S protein	128
3.6 Determination of peptide binding affinity using surface plasmon resonance	132
3.7 C6 homodimers aggregates to form fibril-like structures	137
3.8 Functional application: Comparison of conventional Ab vs. synthetic Ab from their usage in immunofluorescence assay of expressed S protein	137
4. Discussion	141
 Chapter 5: Conclusions and Perspectives	 149
1. Studies on MPER of SARS-CoV S protein	149
2. Quaternary protein mimetics as ‘Synthetic Antibodies’	155
 Publications	 160
 References	 161

List of Tables

Table 1a.	Class I and Class II virus fusion proteins	25
Table 1b.	Comparison of Class I and Class II virus fusion proteins	25
Table 2.	Proposed functions of MPER within Class I viruses	47
Table 3.	Mutations introduced in Trp-rich region of S protein	62
Table 4.	Infectivity of different SARS-CoV pseudotyped with S protein mutants	64
Table 5.	Comparison of the infectivity of SARS-CoV pseudotyped with S protein Ala- and Phe-mutants	64
Table 6.	Synthetic MPER peptides from SARS-CoV and analogs	66
Table 7.	Comparison of structure, biophysical properties and activities of MPER monomer, NT dimer and CT dimer	104
Table 8.	Tabulation of peptide sequences from screening of S peptide library	121
Table 9.	Analysis of peptide sequences from screening of S peptide library	122
Table 10.	Peptides from SPOT screening results for synthesis	124

List of Figures

Figure 1.	Proposed model for Class I virus fusion	26
Figure 2.	Negative stained electron microscopy of SARS-CoV and phylogram of coronaviruses showing classification of SARS-CoV as a distinct group	31
Figure 3.	The genome structure of SARS-CoV	32
Figure 4.	Overall structure of SARS-CoV RBD in complex with neutralizing antibody Fab m396. Sequence and secondary structure assignment of RBD	41
Figure 5.	Ribbon representations of N3-C2 complex and model of the postfusion structure of SARS-CoV S2	42
Figure 6.	Conserved motifs in coronaviruses S protein	44
Figure 7.	Liposome leakage abilities of MPER peptides	68
Figure 8.	Circular dichroism of MPER mutant peptides	70
Figure 9.	Circular dichroism spectra of MPER mutant peptides under increasing TFE concentrations	72 73
Figure 10.	Surface-rendered model of averaged SIV Env spike tomograms	82
Figure 11.	Design of MPER protein mimetics	90
Figure 12.	Chemical structure of N terminal linked dimer and C terminal linked dimer	91
Figure 13.	Leakage of carboxyfluorescein from LUVs by MPER monomer, NT dimer and CT dimer	93
Figure 14.	Ion conductivity of MPER peptides across planar lipid bilayer	96
Figure 15.	Circular dichroism spectroscopy of MPER monomer and dimers under 0% and 80% TFE condition	98
Figure 16.	Circular dichroism spectroscopy of MPER monomer and dimers under different percentage of TFE	99
Figure 17.	Dose response curves showing inhibition of SARS-CoV activity on Vero E6 cells by MPER monomer, NT dimer and CT dimer	101
Figure 18.	Peptide library of SARS CoV S protein synthesized on the membrane for screening with S protein ectodomain (S1188HA)	117
Figure 19.	Screening of peptide library with S protein ectodomain (S1188HA)	118

Figure 20.	Immobilized peptide arrays with SARS CoV S protein ectodomain	119
Figure 21.	Schematic of S protein pull down assay	125
Figure 22.	Pull down of S ectodomain (S1188HA) using biotinylated peptides chosen from screening peptide array	126
Figure 23.	Pull down of S ectodomain (S1188HA) using combination of biotinylated peptides	129
Figure 24a.	Pull down of S ectodomain (S1188HA) using varying concentrations of G15 and H18 peptides	129
Figure 24b.	Pull down of S ectodomain (S1188HA) using a combination of biotinylated peptides showing additive effects of C6 and D4.	130
Figure 25.	Pull down of S1188HA using biotinylated C6 monomer and C6 dimer peptides	131
Figure 26.	C6 dimers aggregate to form fibril-like structures	131
Figure 27.	Surface plasmon resonance of binding between purified S protein and biotinylated A8 peptide	133
Figure 28.	Surface plasmon resonance of binding between purified S protein and biotinylated A8 peptide	134
Figure 29.	Surface plasmon resonance of binding between purified S protein and biotinylated C6 peptide	135
Figure 30.	Surface plasmon resonance of binding between purified S protein and biotinylated C6 dimer	136
Figure 31.	Confocal imaging of immunofluorescent stained 293T cells expressing full length SARS S protein	139
Figure 32.	Confocal imaging of immunofluorescent stained 293T cells expressing full length SARS S protein	140
Figure 33.	Interaction between NP-1 (HR1) and CP-1 (HR2) of SARS CoV spike protein as predicted by molecular modeling	143
Figure 34.	The eight classes of steric zippers	144
Figure 35.	Structure of linker for synthesis of peptide homodimer	147
Figure 36.	Conformational changes of S protein during SARS CoV fusion events	153

Abbreviations

ACE2: angiotensin-converting enzyme 2
Ag: silver
AIDS: acquired immuno-deficiency syndrome
Ala: alanine
APS: ammonium persulfate
ATCC: American type culture collection
BEVS: baculovirus expression vector system
BSA: bovine serum albumin
BSL3: Biosafety level 3
CD: circular dichroism
CDRs: complementarity determining regions
CF: carboxyfluorescein
CPE: cytopathic effects
cryoEM: cryo-electron microscopy
Cys: cysteine
DAPI: 4',6-diamidino-2-phenylindole
DLS: dynamic light scattering
DMEM: Dulbecco's Modified Eagle Media
DMF: N,N-dimethylformide
DMPC: dimyristoyl-phosphatidylcholine
DNA: deoxyribonucleic acid
dNTP: deoxyribonucleotide triphosphate
dpi: days post-infection
DRC: dose response curve
dsDNA: double stranded deoxyribonucleic acid
DTT: dithiothreitol
EboV: Ebola virus
ECL: enhanced chemiluminescence solution
EDTA : ethylenediaminetetraacetic acid
EM: electron microscopy

Env: Envelope protein
E protein: small envelope protein
ER: endoplasmic reticulum
FIV: feline immunodeficiency virus
Fmoc: 9H-(f)luoren-9-yl(m)eth(o)xy(c)arbonyl
FP: fusion peptide
gp160: glycoprotein 160
gp120: glycoprotein 120
gp41: glycoprotein 41
GPs: glycoproteins
HA: influenza hemagglutinin
HeBS: Hepes buffered saline
HEPES: 4-(2-hydroxyethyl)-1-piperazineethane-sulphonic acid
His: Histidine
HIV-1: human immunodeficiency virus type 1
HPLC: high-performance liquid chromatography
HR: heptad repeat
HRP: horse radish peroxidase
Hz: hertz
IBV: infectious bronchitis virus
IC₅₀: half maximal inhibitory concentration
Ig: immunoglobulin
Ile: isoleucine
IMCB: Institute of Molecular and Cellular Biology of Singapore
K_d: dissociation constant
kDa: kilo Dalton
LB: lysogeny broth
Leu: leucine
LUV: large unilamellar vesicle
Lys: lysine
L-SIGN: L-specific ICAM-grabbing non-integrin
mAb: monoclonal antibody
MALDI-TOF: Matrix-assisted laser desorption/ionization - Time of flight
moi: multiplicity of infection

MPER: membrane-proximal external region
M protein: membrane protein
N: asparagine
N protein: nucleocapsid protein
ORFs: open reading frames
PBS: phosphate-buffered saline
PCR: polymerase chain reaction
PEG: polyethylene glycol
Phe: phenylalanine
POPC: 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphatidylcholine
POPE: 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphatidylethanolamine
POPG: 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphatidylserine
preTM: pretransmembrane region
PVDF: polyvinylidene fluoride
RBD: receptor-binding domain
RBM: receptor binding motif
RNA: ribonucleic acid
RT: reverse transcriptase
rpm: revolutions per minute
RU: resonance unit
R18: octadecyl rhodamine B
SARS: severe acute respiratory syndrome
SARS-CoV: SARS associated coronavirus
SDS: sodium dodecyl sulfate
SDS-PAGE: SDS polyacrylamide gel electrophoresis
SIGN: specific intercellular adhesion molecule-3-grabbing non-integrin
SIV: simian immunodeficiency virus
SP: signal peptide
SPR: surface plasmon resonance
S protein: spike glycoprotein
TAE: Tris-acetate-EDTA
TCID₅₀: 50% tissue culture infectious dose
TEMED: N,N,N',N'-Tetramethylethylenediamine
TFA: Trifluoroacetic acid

TFE: 2,2,2-trifluoroethanol

TIS: triisopropylsilane

TM: transmembrane

Tris: tris(hydroxymethyl)aminomethane

Trp: tryptophan

Trp-rich: tryptophan-rich

TRSS: transcription-regulatory sequences

Tyr: tyrosine

UTR: untranslated region

VSV: vesicular stomatitis virus

VSV-G: vesicular stomatitis virus glycoprotein

WHO: World Health Organization

wt: wildtype

3D: three dimensional

6HB: six helix bundle

Abstract

SARS-CoV, a novel virus belonging to group IV coronavirus, was discovered in association with cases of severe acute respiratory syndrome (SARS) in 2003. The spike (S) protein of SARS-CoV is responsible for receptor binding and membrane fusion. The MPER, the Trp-rich membrane proximal external region in the S protein, is both enigmatic and therapeutically important. It is absolutely conserved in members of CoV family and highly conserved in other RNA viruses such as HIV and Ebola. In HIV, MPER contains the epitope essential for broadly neutralizing antibodies, and MPER-derived peptides are potent inhibitors of viral entry.

Structurally, S protein is trimeric and membrane-anchored as a transmembrane protein. Interestingly, MPER undergoes different structural states, from monomer, trimer and possibly hexamer during the fusion event. The self association and rearrangement of MPER in their quaternary states may provide important clues to the fusion event. My thesis focuses on understanding the physical states of MPER and their roles in the entry mechanism of SARS-CoV. Mutational and biophysical studies on MPER showed the importance of Trp in synchronization of SARS-CoV fusion event. My results support the importance of Trp in MPER function, both residually and positionally. The residual importance of Trp identity is due to the 'indole ring' effect, whereas positional importance reveals that increase in helicity may be the key conformational change in S protein-mediated fusion. Taken together, it is believed that the role of Trp in MPER is its contribution to the plasticity of the fusion protein structure, to take on different oligomeric intermediate quaternary structures to synchronize the fusion event.

Quaternary protein mimetics of MPER were designed to mimic and to understand the different oligomeric intermediate states of MPER during fusion. The results support current hypothesis that MPER takes an active role to perturb the apposed membrane during fusion and provides a low energy path for continuous lipid flow, thus enabling fusion. This study sheds light on the possible mechanism of the fusion process, that may involve the S protein undergoing a series of intermediate states prior to fusion pore formation. The key lies with the Trp of MPER contributing flexibility to S

protein structure to allow this series of conformational changes. Using the MPER protein mimetics, the conserved virus fusion mechanism was targeted for development of mutation resistant antivirals. We show that this strategy is promising and could provide potential antiviral candidates for SARS CoV.

The MPER quaternary protein mimetics inspired us to extend its use as a high-affinity probe for capturing proteins based on protein-protein interactions. If successful, the anticipated usage could be an alternative tool to antibodies – ‘synthetic antibody’. Using S protein ectodomain as a proof of concept, we have found self-interacting peptides to bind the target protein and able to pull down S protein from cell lysates. According to the nature of their interaction, they can be classified into four groups. Enhanced affinity of binding to S protein could be achieved via multivalency either by using homodimers or heterodimers. These peptides may act through self-interaction as implied from their fibril-forming abilities. Using immunofluorescence staining of cells expressing S protein, our prototypic ‘synthetic antibody’ was found to be comparable or better than monoclonal anti-S antibodies. The simplicity and potential extensive usage of this technology brings excitement and hope to research, diagnostics and therapeutics.

Chapter 1

Introduction

1. Quaternary Proteins

1.1 Quaternary proteins

Proteins, together with biological molecules like nucleic acids and polysaccharides, are the essential building blocks of cell. Not only do proteins make up most of the dry mass of a cell, they are also indispensable for executing nearly all the cell's functions, from enzymes, receptors, ion channels molecular machines to other specialized proteins like antibodies, hormones, and elastic fibers. Chemically, proteins are highly complex in their structures as well as in the diversity and sophistication of their functions.

According to the classical nomenclature of Linderström-Lang and Schellman (1959), there are four levels of organization in protein structure. Firstly, the primary structure which denotes the amino acid sequence of the protein. Stretches of amino acids within the protein that form α helices and β sheets constitutes the protein's secondary structure. The full three-dimensional organization of a protein is referred to as the tertiary structure, and if protein complexes of more than one protein molecule are formed, the complete structure is designated as the quaternary protein structure. Most proteins are made of more than one protein molecule and thus, they have a quaternary structure. Furthermore, nearly all, proteins interact to form dimers or higher order

oligomers to carry out their cellular function. Thus, the quaternary structure of proteins is almost invariably essential to its function.

The protein quaternary structure involves a variety of bonding interactions, which includes hydrogen bonding, salt bridges and disulphide bonds that helps cluster and maintain the several protein molecules into a specific three-dimensional geometry. The quaternary state of protein structure is very common in nature and vital to the protein's function. For instance, monomeric protein may be inactive, while dimerization or multimerization may activate the protein to carry out its intended function (eg. in receptor dimerization). Proteins that are composed of multimers, where their quaternary structures are critical for their function, are abundant in nature. Typical examples include growth factors (active dimers), cell membrane receptors (active dimers), insulin (inactive hexamer, active monomer) and virus proteins (multimers). Specific non-covalent interactions governing quaternary structure of proteins allow formation of quaternary state complexes that are important for catalytic activity of some enzymes (eg. HIV proteases). Quaternary structures of these proteins can exhibit a very wide range of stability and half-lives (eg. insulin). They may also confer additional stability to structures of macromolecular assemblies, which can be seen in oligomeric proteins (eg. virus proteins and aldolases). Quaternary structure in proteins may also serve other important roles in protein function. For example, association of proteins may be required to prevent build up of excessive osmotic pressure, for compartmentalization or aid in the control of degradation in protein turnover.

Hemoglobin has been the paradigm oligomeric protein, since the early discovery of its

quaternary structure by Svedberg in the mid 1920s. Svedberg's description of its quaternary structure, showing changes to quaternary structures during oxygen binding, was further confirmed by Perutz' X-ray crystallographic studies (1960). Further introduction of the concept of allostery by Monod *et al.* (1963, 1965), gives a central role to quaternary structure and the way it changes when ligand binds.

Another example of importance of quaternary structure is insulin. Insulin is a hormone in the body that regulates the blood glucose level. It is a small protein made up of 51 amino acids, which contains two protein chains connected by three disulphide bonds. Two disulphide bonds are used to hold the two chains together while the third linkage forms a loop within a single chain. It is a classic example of a protein with quaternary structure. Insulin molecule readily form dimers and in the presence of zinc ions (or other divalent metal ions) three dimers assemble together to form a hexamer, which is utilized in the biosynthesis, processing and for storage in B cells. However, the active form is the insulin monomer that binds to its receptor directly and is responsible for the expression regulatory role of insulin (Derewenda *et al.* 1989). Using rDNA-derived biosynthetic proteins, DiMarchi *et al.* was able to synthesize Humalog ® (LisPro-human insulin), an insulin analog by reversing the Pro-Lys sequence to Lys-Pro. Such a rearrangement changes the rate of dissociation from hexamer to monomer (DiMarchi *et al.* 1994, Anderson *et al.* 1998). This ability to engineer the quaternary states of insulin results in a rapidly-dissociating hexamer, optimized for pharmacological performance, without increased risk of abnormal immunological response, benefiting millions of diabetes patients.

As the advancement of genomics technology provided the ease of obtaining amino

acid sequences, the focus shifted to looking at the primary structure. However, the importance of quaternary structure of protein has regained attention in recent years with protein assemblies as targets of several recent structural genomics initiatives (Russell *et al.* 2004, Janin 2007).

1.2 Quaternary proteins in viruses

The most successful organisms that exploit quaternary proteins are the viruses. Viruses have very small genomes (eg. HBV genome ~3.2kb) with a large amount of coding information packed into it. The virus ensures the use of every nucleotide in the genome for protein coding and reads more than half the genome in two reading frames. Although viruses maximize the use of its small genome, it typically encodes less than twenty proteins (only seven proteins for HBV), depending heavily on host cell functions as well as the efficient use of the virus proteins.

Due to the limited number of proteins, virus proteins are exploited fully by forming multimers, using the different quaternary states to generate diversity and forms within their limited capacity. In addition, the quaternary states may also confer stability to protein structures. For example, many virus capsid proteins exhibit polymorphic quaternary structure, whereby the same protein subunit is found in different structural environments. Such variations are required as genetic economy limits the number of gene products. Quaternary proteins are also exemplified by virus enzymes, which are usually multimeric in nature for their function. For instance, the functional efficacy of the HIV-1 Rev protein is highly dependent on its ability to assemble onto its HIV-1

RNA target (the RRE) as a multimeric complex (Jain *et al.* 2001). The virus fusion protein is also made up of dimer or trimer associated together on virus membrane that changes conformation during fusion events. Understanding the importance of quaternary structure in virus proteins can aid in development of therapeutics, diagnostics and more in depth knowledge about the viruses.

2. Virus Envelope Glycoprotein

2.1 Virus entry and fusion

A virus is a sub-microscopic particle (ranging in size from 20 - 300 nm) capable of infecting cells of a biological organism. Ubiquitous and abundant in nature, they are composed of a small number of macromolecules, genetically-coded by virus but produced by the host organism. Viruses can replicate themselves only by infecting a host cell, and therefore cannot reproduce on their own. Their life cycle involves binding to receptors on host cells, fusion with host cell membrane, uncoating of viral capsid, replication, assembly and release of new viruses.

Receptor binding and the subsequent conformational changes leading to fusion of virus envelope to host cell membrane are events essential for virus entry and poses an attractive target for therapeutic intervention (Zwick *et al.* 2001, Hofmann *et al.* 2005, Chu *et al.* 2008, Liu *et al.* 2009, Du *et al.* 2009, Haim *et al.* 2009). Virus fusion glycoproteins are found on surfaces of virion structure and play the major role of binding to receptors (cell surface molecules) to gain entry into host cell. During the entry process, from receptor binding to membrane fusion, the virus fusion glycoprotein undergoes a multitude of substantial structural changes that eventually allow the energetically unfavorable mixing of lipid membranes to occur.

Virus membrane fusion can occur either at the plasma membrane or intracellularly after uptake by endocytosis (Earp *et al.* 2005). Direct fusion with plasma membrane is triggered by virus-receptor interactions at neutral pH while other viruses that depend

on internalization by receptor-mediated endocytosis (such as clathrin-dependent, non-caveolae-dependent) usually require exposure to mild acidic pH within organelles of the endocytic pathway.

For RNA viruses, virus fusion proteins have been characterized based on certain key structural features into two classes, namely, Class I (*Retroviridae* and *Coronaviridae*) and Class II (*Flaviviridae* and *Togaviridae*) (Table 1a) (Kielian and Rey 2006).

2.2 Class I virus fusion proteins

Class I virus fusion proteins mediate membrane fusion by refolding the fusion protein to a highly stable rod-like structure with a central trimeric α -helical coiled coil. Such coiled coil structures are symbolic of the Class I virus fusion proteins and can aid in defining Class I virus fusion proteins by computer prediction or characterization of such a structure (Carr and Kim 1993). Class I virus fusion proteins share several important features that can be exemplified by the influenza-virus haemagglutinin (HA).

HA is a trimeric protein that has both receptor-binding and fusion activities of influenza virus. It is synthesized as a fusion-inactive precursor, HA0, which is cleaved by host proteases into HA1 (receptor-binding) and HA2 (responsible for fusion), which are disulphide-bonded. HA2 is maintained in a 'metastable' state at the virus surface (Ruigrok *et al.* 1986, Carr *et al.* 1997) and has a conserved hydrophobic fusion peptide that lies in the N terminus of HA2, which inserts into the target

membrane during fusion. In the pre-fusion HA, the N-terminal fusion peptide in HA2 is buried at the trimer interface within the α -helical coiled-coil stalk region. Low pH triggers virus fusion, in which the HA1 trimer contacts at the head is destabilized, causing the globular head domains to dissociate. This allows a loop-to-helix transition of a segment of HA2 that was previously buried underneath HA1, projecting the fusion peptide towards the target membrane, where it inserts irreversibly (Bullough *et al.* 1994). This ‘pre-hairpin intermediate’ (Eckert and Kim 2001) then undergoes a drastic conformational change that involves the C terminal end of the long HA2 α -helix jackknifing back, reversing the direction of the viral-membrane-proximal segment of HA2, which then interacts in an anti-parallel fashion with the groove formed by the N-terminal trimeric coiled coil. The final post-fusion conformation of HA2 is therefore a highly stable rod with the TM and fusion-peptide segments together at the same end of the molecule. This structure is called a ‘trimer of hairpins’ (Eckert and Kim 2001), and is considerably more stable than the metastable pre-fusion form. A model for Class I virus fusion is shown in Figure 1.

Class II virus fusion protein (*Togaviridae*, *Flaviviridae*, etc) was found to have three-dimensional structure radically different from that of influenza HA (Lescar *et al.* 2001). A comparison of structural features of Class I and II virus fusion proteins is listed in Table 1b. Despite having distinctly different structural features, Class I and Class II virus fusion proteins refold during fusion to give very similar post-fusion hairpin structures, with fusion peptides or loops and the TM domains on the same end of the stable protein rod (Jardetsky and Lamb 2004, Kielian and Rey 2006). Both classes forms trimer of hairpins in post-fusion structures, with Class I having a central α -helical coiled-coil structure while Class II trimer of hairpins are characterized by β

structures (Skehel and Wiley 2000, Harrison 2005, Kielian and Rey 2006, Lamb *et al.* 2006, Stiasny and Heinz 2006). The striking resemblance in the formation of trimer of hairpins indicate a surprising convergence of Class I and Class II fusion mechanisms, and may point to a universal membrane fusion mechanism that applies to non-homologous virus fusion proteins (Jardetsky and Lamb 2004, Kielian and Rey 2006).

Understanding the structure of Class I and Class II virus fusion proteins allows the development of inhibitors at specific steps of fusion. Such inhibitors are useful for dissection of fusion mechanism and development of novel antiviral strategies (Kielian and Rey 2006).

Virus family	Examples
Class I	
<i>Orthomyxoviridae</i>	Influenza virus HA2 protein
<i>Paramyxoviridae</i>	Simian virus 5 F1 protein
<i>Filoviridae</i>	Ebola virus GP2 protein
<i>Retroviridae</i>	Moloney murine leukemia virus TM protein, HIV-1 gp41 protein
Coronaviridae	Mouse hepatitis virus and SARS virus S2 proteins
Class II	
<i>Flaviviridae</i>	E proteins of flaviviruses tick-borne encephalitis and dengue viruses
<i>Togaviridae</i>	E1 protein of alphavirus Semliki Forest virus

HA, haemagglutinin; SARS, severe respiratory syndrome; TM, transmembrane; gp/GP, glycoprotein.

Table 1a. Class I and Class II virus fusion proteins (Adapted from Kielian and Rey 2006)

Feature	Class I (influenza HA)	Class II (SFV E1)
Conformational change during fusion	Metastable fusion protein trimer to stable fusion protein trimer	Metastable dimer to stable fusion protein trimer
Predominant secondary structure of fusion protein	α -helix	β -sheet
Post-fusion structure	Trimer of hairpins with central α -helical coiled coil	Trimer of hairpins composed of β structure
Maturation to prefusion state through:	Proteolytic processing of fusion protein	Proteolytic processing of companion protein
Fusion peptide location in metastable structure	N-terminal peptide buried in trimer interface	Internal loop at fusion protein tip, capped by dimer interaction

HA, haemagglutinin; SFV, Semliki Forest virus

Table 1b. Comparison of Class I and Class II virus fusion proteins (Adapted from Kielian and Rey 2006)

Figure 1. Proposed model for Class I virus fusion. (a) The metastable conformation of a trimeric generalized fusion protein, with helical domain A in orange, helical domain B in pink, and the TM domain in purple. (b) After binding to a receptor on the cellular membrane, or on exposure to the low pH found in intracellular compartments (endosomes), the protein forms an extended conformation and the hydrophobic fusion peptide (red) inserts into the target membrane. (c) Several trimers are thought to be involved. (d) Protein refolding begins. The free energy thereby released causes the membranes to bend towards each other. (e) Formation of a restricted hemifusion stalk allows the lipids in the outer leaflets of the membranes to mix. (f) Protein refolding completes, forming the final, most stable form of the fusion protein, with the fusion peptide and transmembrane domain anti-parallel to each other but in the same membrane. Only (a) and (f) have been observed by crystallography, but biochemical data support many of the proposed steps. (Taken from Jardetsky and Lamb 2004)

3. Severe Acute Respiratory Syndrome Associated Coronavirus (SARS-CoV)

3.1 Background on SARS as an infectious disease

Severe acute respiratory syndrome (SARS) was a contagious atypical pneumonia that caused a major epidemic between November 2002 and July 2003, with a high mortality rate of 9.6% (8096 known cases and 774 deaths) as listed in the WHO's 21 April 2004 concluding report¹. The main symptoms are high fever ($>38^{\circ}\text{C}$), dry cough, shortness of breath or breathing difficulties. Changes in chest X-rays indicative of pneumonia also occur². SARS may be associated with other non-specific symptoms. The only common symptom to all patients seems to be a fever above 38°C . To date, there is no vaccine and effective treatment for SARS (Stockman *et al.* 2006). Antibiotics are not effective and existing treatments are mostly supportive with antipyretics, supplemental oxygen and ventilatory support as needed.

Drawing valuable experience from the SARS outbreak, there is an increase and continued global vigilance, surveillance and laboratory bio-safety practices. Due in part to these efforts and quarantining and euthanizing animals that may have had exposure to SARS-CoV, there has since been no new cases of SARS reported (Shi and Hu 2008, Liang *et al.* 2006). However, SARS should not be overlooked and is still a safety concern globally due to possible reintroduction of a SARS-like CoV (SL-CoV) into humans and risks of SARS escaping from laboratories (Normile 2004, Orellana 2004). Development of rapid and sensitive diagnostic tests to monitor the spread of this infectious disease is needed in view of potential outbreaks in future.

¹ http://www.who.int/csr/sars/country/table2004_04_21/en/index.html

² <http://www.who.int/csr/sars/sarsfaq/en/>

3.2 SARS Coronavirus (SARS-CoV)

A previously uncharacterized virus, termed SARS-associated coronavirus (SARS-CoV), was isolated from SARS patients (Ksiazek *et al.* 2003, Peiris *et al.* 2003) and demonstrated to cause disease in infected nonhuman primates (Fouchier *et al.* 2003) (Figure 2A). The genome of SARS-CoV is a 29727 nucleotides long, polyadenylated RNA and the GC content is 41% (Rota *et al.* 2003). Its genome organization is typical of coronaviruses with characteristic gene order of 5'- replicase (rep), spike (S), envelope (E), membrane (M), nucleocapsid (N) -3' and short untranslated regions at both termini (Figure 3). Findings indicated that SARS-CoV is not closely related to any of the previously characterized coronaviruses and forms a distinct group within the genus *Coronavirus* (Figure 2B).

3.3 Structural proteins of coronaviruses

The structural proteins of coronaviruses, S, E, M and N function during host cell entry, virion morphogenesis and release. During virion assembly, N binds to a defined packaging signal on viral RNA, leading to formation of the helical nucleocapsid. M is localized at specialized intracellular membrane structures, and interactions between M and E proteins and nucleocapsids result in budding through the membrane. The S protein is incorporated into the viral envelope by interaction with M and mature virions are released from smooth vesicles (Garoff *et al.* 1998).

The S proteins of coronaviruses are large type I membrane glycoproteins responsible for receptor binding and membrane fusion. The S proteins of some coronaviruses are cleaved into S1 and S2 subunits by host proteases. S proteins also contain important virus-neutralizing epitopes and amino acid changes can dramatically affect the virulence and *in vitro* host cell tropism (Sanchez *et al.* 1999, Leparc-Goffart *et al.* 1998).

3.4 Immune responses to SARS-CoV

After coming into contact with SARS-CoV, infection can trigger humoral and cellular immune response in the human body. After approximately two weeks post-infection, specific immunoglobulin G and immunoglobulin M against SARS-CoV can be detected, this reaches a peak after sixty days post-infection and remains at a high level until 180 days post-infection (Mo *et al.* 2005). In patients who had recovered from SARS, high titres of neutralizing antibodies and cytotoxic T cell responses specific to SARS-CoV were found (Xu and Gao 2004, Zhong *et al.* 2005). The level of responses correlated well with the SARS disease outcome and this suggests that for clearance of SARS-CoV infection, both humoral and cellular immune responses are vital (Li *et al.* 2006).

Neutralizing antibodies and/or T cell immune responses mainly targeting SARS S protein can be raised directly (Li *et al.* 2006, Zhi *et al.* 2005, Buchholz *et al.* 2004, Zakhartchouk *et al.* 2007), although it can be raised against several SARS-CoV proteins as well (See *et al.* 2008, Dutta *et al.* 2008, Jin *et al.* 2005). This points to a

critical role played by specific immune responses induced by the S protein in combating against the SARS-CoV infection.

Figure 2. Negative stained electron microscopy of SARS-CoV (Panel A) and phylogram of coronaviruses showing classification of SARS-CoV as a distinct group (Panel B). (Taken from Ksiazek *et. al.*, N Engl J Med, 2003)

Figure 3. The genome structure of SARS-CoV. Replicase and structural regions are shown together with the predicted cleavage products in ORF1a and ORF1b. The position of the leader sequence (L), the 3' poly(A) tract and the ribosomal frameshift site between ORF1a and ORF1b are also indicated. Each box represents a protein product (NS, non-structural protein). Colours indicate the level of amino-acid identity with the best-matching protein of other coronaviruses. The SARS-CoV accessory genes are white. Filled circles indicate the positions of the nine transcription-regulatory sequences (TRSs) that are specific for SARS-CoV (5' ACGAAC3'). (Taken from Stadler *et al.* 2003)

4. SARS-CoV Spike protein and its functional domains

4.1 SARS-CoV Spike (S) protein

The importance of SARS S protein is obvious as it is one of the first viral proteins that are encountered by the host cell due to its localization on the virion surface. S protein is the main antigenic component that induces host immune responses among all the structural proteins of SARS-CoV. Its involvement in receptor recognition, virus attachment and entry makes it one of the most important and attractive target for development of SARS vaccine and therapeutics. The SARS-CoV S protein, like those of the other coronaviruses, mediates receptor binding and subsequent membrane fusion. They form characteristic corona of large, distinctive spikes on the viral envelopes (Holmes 2003, Lai and Cavanagh 1997) and are major targets for neutralizing antibodies. On SARS-CoV, S proteins are observed as 20-40 nm complex surface projections that surround the virion particles (Ksiazek *et al.* 2003) (Figure 2A). The overall similarity in sequence between the predicted amino acid sequence of SARS-CoV S protein and other coronaviruses is low (20-27%), except for some conserved sequences in the S2 subunit (Rota *et al.* 2003). This low level of similarity makes it difficult to draw conclusive parallels on structural and functional similarities with the other coronaviruses.

4.2 Structural features and functional domains of S protein

SARS-CoV S protein contains 23 potential N-linked glycosylation sites (Rota *et al.*

2003) and several conserved functional motifs at the N and C-termini. The S2 domain is more conserved than the S1 domain. The main functional domains includes the receptor binding domain (RBD), putative fusion peptide, heptad repeat 1 (HR1), heptad repeat 2 (HR2), pre-transmembrane region, and the cytoplasmic tail. The N terminal of SARS-CoV S protein contains a short type I signal sequence consisting of hydrophobic amino acids that are presumably removed during cotranslational transport through the endoplasmic reticulum. The C terminal, which contains a transmembrane (TM) domain and a cysteine-rich cytoplasmic tail, is highly conserved in SARS-CoV (Rota *et al.* 2003). Based on amino acid sequence, the SARS-CoV S protein lacks the basic amino acid cleavage site found in group 2 and 3 coronaviruses, suggesting that it is probably not cleaved into S1 and S2 subunits. However, similar to other coronaviruses, it was reported that SARS S protein can be cleaved into S1 and S2 subunits by proteases such as trypsin (Li *et al.* 2005), factor Xa (Du *et al.* 2007) and Cathepsin L (Bosch *et al.* 2008). The cleavage site for trypsin is mapped to R667-S668 (Lai *et al.* 2006), while the cleavage site for Cathepsin L is at T678-M679 on S protein (Bosch *et al.* 2008). Cathepsin L cleaves upstream of the fusion peptide and this cleavage is essential for the activation of membrane fusion domain following entry to progress through the fusion events. It was suggested that Cathepsin L proteolysis may occur within the endosome (Simmons *et al.* 2005).

4.3 Receptor binding domain (RBD)

SARS-CoV S protein attaches the virus to its cellular receptor, angiotensin-converting enzyme 2 (ACE2) (Li *et al.* 2003). This is a cell surface zinc peptidase which is also

utilized by another human coronavirus, HCoV-NL63, as receptor (Hofmann *et al.* 2005). The crystal structure of ACE2 ectodomain reveals a claw-like N-terminal peptidase domain, with the active site at the base of a deep groove, and a C-terminal “collectrin” domain (Towler *et al.* 2004). It was found that a fragment of S1, containing residues 318 to 510, is sufficient for tight binding to the peptidase domain of ACE2 (Xiao *et al.* 2003, Wong *et al.* 2004, Babcock *et al.* 2004). This fragment is termed the receptor binding domain (RBD), and is the critical determinant of virus-receptor interaction (Figure 4A). Changes in just a few residues in the RBD can lead to efficient cross-species transmission (Li *et al.* 2005, Song *et al.* 2005) and the RBD also contains important viral-neutralizing epitopes (Sui *et al.* 2004, van der Brink *et al.* 2005, He *et al.* 2005).

The RBD contains two subdomains: a core and an extended loop. The core is a five-stranded anti-parallel β sheets ($\beta 1$ to $\beta 4$ and $\beta 7$), with three short connecting α helices (αA to αC) (Figure 4B). Disulfide bonds connect cysteines 323 to 348, 366 to 419, and 467 to 474. The other remaining cysteines are disordered but two (378 and 511) are in the same neighborhood (Li *et al.* 2005). The extended loop subdomain lies at one edge of the core and presents a gently concave outer surface formed by a two-stranded β sheet ($\beta 5$ and $\beta 6$). The base of this concavity cradles the N terminal helix of ACE2 and residues 445 to 460 of the RBD anchor the entire receptor binding loop to the core of the RBD. This loop, which contains residues 424 to 494, makes all the contacts with ACE2 and is referred to as the receptor binding motif (RBM).

The RBM surface is complementary to the receptor tip with about $1700\text{\AA}^2$ of buried surface at the interface. Fourteen residues of viral S protein contact with eighteen

residues of the receptor and hydrophilic interactions among amino acid side chains predominate. Six RBM residues at this interface are tyrosines, which present both a polar hydroxyl group and a hydrophobic aromatic ring (Li *et al.* 2005). The significance of these structural features will become apparent in discussion.

It has also been reported that SARS-CoV has other alternative receptors to bind to host cells, such as DC SIGN (dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin) and/or L-SIGN (liver/lymph node-SIGN) (Jeffers *et al.* 2004, Yang *et al.* 2004). Crucial contact points on S protein for DC-SIGN and L-SIGN mediated virus entry are seven asparagine-linked glycosylation sites at amino acid positions 109, 118, 119, 158, 227, 589 and 699. These are distinctly different from RBD for ACE2, thus, S protein may be able to bind to these alternative receptors independently. However, the significance of having these alternative receptors and their actual function remains to be elucidated.

4.4 Putative fusion peptide (FP)

In general, fusion peptides are short (16 to 26 residues), hydrophobic sequences that are rich in alanine, glycine and phenylalanine. The presence of a canonical fusion tripeptide (YFG or FXG) is highly conserved among the fusion peptides of retroviruses, paramyxoviruses, influenza virus and filoviruses. A proline residue at or near the centre of many fusion peptides has been implicated as critical for the interaction of peptide with the target cell lipid membrane. A 19 amino acid putative SARS-CoV fusion peptide was identified using synthetic peptide and corresponds to

residues 770 to 788 on the SARS-CoV S2 region (Sainz *et al.* 2005). It was shown to partition strongly into large unilamellar vesicles (LUVs), induces fusion and leakage of LUVs and adopts a β -sheet structure.

4.5 Heptad repeats 1 and heptad repeats 2 (HR1 and HR2)

Class I virus fusion protein contains characteristically two heptad repeats known as HR-N and HR-C that formed coiled coil structures within the ectodomain of virus fusion proteins. Its corresponding counterparts in SAR-CoV S protein has been determined to be HR1 and HR2 respectively. In the pre-fusion state, the two HR regions oligomerize to form homotrimer coiled coil cores. Following insertion of fusion peptide into the target cell membrane, conformational changes induced by receptor binding causes HR interactions between residues 916 to 950 of HR1 and residues 1151 to 1185 of HR2 (Tripet *et al.* 2004). This mediates the formation of an anti-parallel six helix coiled coil structure, termed trimer of hairpins or 6 helix bundle (6HB) (Figure 5). It is believed to facilitate apposition of both viral and target cell membranes, causing fusion and subsequent virus entry. The complementarity of HR1 and HR2 to form a 6HB is exploited in my work to develop new tools.

4.6 Pre-transmembrane (PTM) or Membrane proximal external region (MPER)

Almost all Class I virus fusion proteins possess a tryptophan-rich (Trp-rich) membrane proximal external region (MPER). This region exists in all coronaviruses S protein S2 region and juxtaposes the transmembrane domain. They share a highly

conserved ten-residue sequence Y(V/I)KWPW(W/Y)VWL that is rich in aromatic amino acids with 3 to 4 Trp residues. This region is also conserved in SARS-CoV S protein and is the focus of my thesis. It will be discussed in more details in the following section.

4.7 Development of SARS-CoV vaccines and therapeutics

S protein presents an attractive and important target for development of vaccines and therapeutics due to its involvement in receptor binding and membrane fusion process. It is also the main antigenic component, amongst other structural proteins, that triggers host immune responses. Since the outbreak of SARS, there has been continual effort in the search of vaccines and therapeutics against this emerging infectious agent. Research focuses mainly on the S protein, which includes (i) vaccines based on S protein; (ii) S protein-based therapeutics; (iii) monoclonal antibodies against S protein and (iv) antiviral compounds and other small molecules.

Vaccines based on full length S protein (Yang *et al.* 2004, Bisht *et al.* 2004, Chen *et al.* 2005, Kam *et al.* 2007, He *et al.* 2006) and RBD can induce neutralizing antibodies (He *et al.* 2005, He *et al.* 2004, Bonavia *et al.* 2003, Du *et al.* 2007). However, they have only been tested in animal models which are usually less robust, providing virus replication but lacking clinical symptoms. Moreover, harmful immune responses that cause liver damage or enhanced infection after challenge with SARS-CoV also raises concerns regarding the safety and protection efficacy offered by vaccines containing full length S protein (Czub *et al.* 2005, Weingartl *et al.* 2004).

For S protein-based therapeutics, efforts had been focused on peptides that interrupt RBD-ACE2 interaction (Wong *et al.* 2004, Hu *et al.* 2005, Han *et al.* 2006), peptides that interfere with S protein cleavage (Zheng *et al.* 2005) and peptides that block HR1-HR2 interaction from forming fusion active core (Liu *et al.* 2004, Zheng *et al.* 2005, Yuan *et al.* 2004). These peptides have shown potent inhibitory activity and can be developed as novel therapeutics against SARS-CoV infection. However, studies had only been carried out *in vitro* and their *in vivo* inhibitory activity in animal models must be evaluated before considering further development. T20 (Jiang *et al.* 1993, Wild *et al.* 1994), a peptide that blocks the formation of fusion-active core in HIV at nanomolar level (Liu *et al.* 2007, Chan *et al.* 1998) has an equivalent counterpart in SARS but inhibits at micromolar range. However, the peptide in SARS should be more advantageous than T20 in HIV as SARS has a shorter incubation time which requires only a few injections at early stage of acute phase, compared to HIV patients which must be injected twice per day for the patient's lifetime. Nevertheless, such peptides face a major disadvantage of potential selection of escape mutants with altered host range phenotypes (McRoy and Baric 2008).

Neutralizing mouse (He *et al.* 2006, He *et al.* 2005, Zhou *et al.* 2004, Lai *et al.* 2005) and human monoclonal antibodies (mAbs) (Nie *et al.* 2004, Traggiai *et al.* 2004, Zhu *et al.* 2007, Coughlin *et al.* 2007, Rockx *et al.* 2008) are shown to have highly potent effects on neutralizing SARS-CoV. However, repeated use of mouse mAbs can cause a human-anti-mouse antibody response that can clear the mouse mAbs from blood, preventing its therapeutic effects and eliciting an allergic response in patients. Some human mAbs can induce broad protection against lethal homologous and heterologous

SARS-CoV challenge, minimizing possible emergence of mAb escape mutants (Rockx *et al.* 2008).

Some other efforts include antivirals like inhibitors of cathepsin L (Huang *et al.* 2006, Simmons *et al.* 2005, Simmons *et al.* 2006) and small molecules that target S protein (Stadler *et al.* 2008, Yi *et al.* 2004, Kao *et al.* 2004). However, these only showed inhibitory activity *in vitro* and require further studies in animal models. A promising hope would be gene targeting with small interfering RNA to inhibit SARS-CoV infection, replication and/or interruption of S gene expression (Wu *et al.* 2005, Akerstrom *et al.* 2007, He *et al.* 2006, Zheng *et al.* 2004, Qin *et al.* 2004, Zhang *et al.* 2004). This novel strategy had been demonstrated to work in *in vitro* and *in vivo* studies (Li *et al.* 2005).

Figure 4. Overall structure of the SARS-CoV RBD in complex with the neutralizing antibody Fab m396 (Panel A). Sequence and secondary structure assignment of RBD. (Panel B). (Taken from Prabakaran *et. al.*, JBC, 2006)

Figure 5. Ribbon representations of the N3-C2 complex and model of the postfusion structure of SARS-CoV S2. (A) The N3-C2 trimer. N3 is blue and C2 is red (B) An N3-C2 heterodimer, with highlighted “start and end” positions of secondary structure element. (C) Superposition between a HR1-HR2 heterodimer and a N3-C2 heterodimer. (D) Schematic representation of the proposed S2 postfusion structure. (Taken from Supekari *et. al.*, PNAS, 2004)

5. MPER in Class I virus fusion proteins

5.1 Conservation and importance of MPER in Class I viruses

The MPER is almost totally conserved in the *Coronavirus* family as revealed by sequence alignment (Rota *et al.* 2003) (Figure 6). It is also highly conserved in Class I viruses, such as HIV and Ebola virus. According to the trimer-of-hairpins model, this aromatic region would also align with the fusion peptide and TM domain during apposition of target cell and viral membranes, possibly adding to the overall hydrophobicity of the environment, and contributing to the distortion of lipid membranes necessary for fusion (Saez-Cirion *et.al.*, 2003, Salzwedel *et.al.*, 1999).

Peptides derived from MPER region have been demonstrated to possess a high tendency to partition into the membrane interface according to the Wimley and White (WW) interfacial hydrophobicity scale (Wimley and White 1996, Yau *et al.* 1998). Studies in HIV-1 gp41 and EboV glycoprotein 2, using synthetic peptides derived from region analogous to MPER domain, showed that these regions partition into and perturb the integrity of lipid vesicles (Suarez *et al.* 2000, Saez-Cirion *et al.* 2003, Suarez *et al.* 2000).

Figure 6. Conserved motifs in coronaviruses S protein (Adapted from Rota et al., 2003). Alignment of the C-terminal region of the SARS-CoV and reference coronavirus S proteins was generated with ClustalX 1.83. Residues that match the SARS-CoV sequence exactly are boxed. The red box indicates the amino acid sequence Y(V/I)KWRW(Y/W)VWL which is a conserved motif in all three coronavirus groups and SARS-CoV. This region is referred to as Trp-rich region in this study.

5.2 MPER in SARS-CoV S protein

Sequence alignment of coronavirus S proteins reveals a similar region in SARS-CoV S protein that is enriched in aromatic amino acids and extraordinarily conserved. It lies in an identical location to the MPER of HIV and EboV. It was recently reported that this conserved and highly aromatic MPER of CoV strongly partitions into the membranes of lipid vesicles with a preference for vesicles containing anionic lipids. It is also able to perturb membrane integrity and cause leakage of vesicle contents. These findings are consistent with the previous findings for HIV and EboV (Suarez *et al.* 2000, Saez-Cirion *et al.* 2003) and suggest a functional role of MPER in virus fusion and entry.

5.3 Proposed functions of MPER in virus fusion

Currently, there is no specific role and function assigned to the MPER and its mechanism of action to mediate viral entry is unclear. Research has largely focused on the putative roles of MPER during fusion and possible interactions with lipids and cholesterol. There are several hypotheses as to how the MPER works to aid in virus entry. Using structural data from HIV gp41, Weissenhorn *et.al.* proposed that the MPER is flexible (Weissenhorn *et al.* 1997). Similar to gp41, MPER is structurally disordered in the fusion pH-induced conformation of HA2 in influenza (Bullough *et al.* 1994). Together, they suggest that the MPER may act as a flexible extender that helps bring the virus and host cell membrane into close proximity for fusion to occur.

In HIV, MPER has similar characteristics to the fusion peptide and could interact with membranes. Salzwedel *et.al.* (1999) suggested that MPER may undergo a conformational change from β strand to a partly α helical structure when it contacts the lipid membrane, just like the fusion peptide.

Saez-Cirion *et.al.* (2003) subsequently suggested that the MPER may serve as an agent to promote membrane destabilization required for fusion. This may occur as the helical core of the fusion protein undergoes conformational changes to bring the two hydrophobic regions (FP and MPER) to close proximity, and the MPER taking the active role to perturb the apposed membranes during fusion. This may also involve interactions with membrane components such as cholesterol (Salzwedel *et al.* 1999).

It was further proposed that the MPER, in conjunction with the FP and the TM anchor, forms a continuous track of hydrophobic, membrane-interacting surfaces that provide a low energy (low barrier) path for lipid flow and membrane fusion during virus fusion (Sainz *et al.* 2005). Other possibilities also include the MPER acting as a contributor to trimer formation and maintenance of its stability (Saez-Cirion *et.al.*, 2003, Salzwedel *et.al.*, 1999). A summary of the proposed functions of MPER within the Class I viruses has been tabulated into Table 2.

In addition to these proposed roles, we hypothesize that Trp in the MPER is important in the synchronization of fusion events and functions like a ‘conformational keeper’ of fusion. Our rationale is that Trp in MPER may contribute to the plasticity of the fusion protein intermediate structures. Based on biophysical evidence, the virus fusion

Table 2. Proposed functions of MPER within Class I viruses

Virus	Proposed function	References
Influenza	Flexible extender that helps bring virus and host membrane in close proximity	Bullough <i>et al.</i> 1994
HIV	Flexible linker that helps orient fusion machinery for fusion to occur	Weissenhorn <i>et al.</i> 1997 Dimitrov <i>et al.</i> 2003
	Bind membrane possibly by interacting with membrane components such as cholesterol	Salzwedel <i>et al.</i> 1999 Vincent <i>et al.</i> 2002 Epand <i>et al.</i> 2003, 2005
	Membrane binding and destabilization facilitating lipid exchange between virus and host cell membrane	Suarez <i>et al.</i> 2000 Schibli <i>et al.</i> 2001 Saez-Cirion <i>et al.</i> 2002 Vishwanathan <i>et al.</i> 2008
	Forms a continuous track of hydrophobic membrane-interacting surfaces with FP and TM to provide a low energy path for lipid flow and membrane fusion during virus fusion	Sainz <i>et al.</i> 2005
	Contributor to trimer formation and maintenance of its stability	Saez-Cirion <i>et al.</i> 2003 Salzwedel <i>et al.</i> 1999
SIV	Extended coiled-coil – Post fusion bundle extends through MPER, which contributes to stability and thus, using the energy for fusion	Lay <i>et al.</i> 2004
SARS CoV	Promote membrane destabilization favoring fusion pore formation and fusion	Guillen <i>et al.</i> 2005

protein undergoes different structures during fusion as intermediate states. The unusually rich Trp cluster in MPER may facilitate the transition of different intermediate states in the fusion.

6. Aims

Quaternary protein mimetics are peptidic compounds that mimic the quaternary structure of a protein, its properties or function. The purpose of my study is to explore the use of quaternary protein mimetics for therapeutic uses, to inhibit virus infection and also as a biochemical tool for detection of protein-protein interactions. The importance of quaternary structure in protein function may be investigated by replacing critical residues important for quaternary structure and function of protein, which may alter the quaternary structure and in turn affecting protein functionality. In this thesis, we focus on the study of the importance of the quaternary structure of proteins, using virus fusion proteins.

Structurally, S protein is trimeric and membrane-anchored as a transmembrane protein. Interestingly, MPER undergoes different structural states, from monomer, trimer and possibly hexamer during the fusion event. The self association and rearrangement of MPER in their quaternary states may provide important clues to the fusion event. The main objective of my thesis is on the development of novel therapeutics and diagnostics for SARS-CoV using quaternary protein mimetics. Specifically, I focus on understanding the roles of the highly conserved Trp-rich MPER in SARS-CoV S protein, its conformational changes leading to fusion that could lead to its development as a novel therapeutic candidate against SARS-CoV. An extension of this work focuses on using a novel strategy to prepare high affinity probe against a target protein based on its own sequence. SARS-CoV S protein is used as an example for a proof of concept and such probes can have profound implications on the development of diagnostics for SARS-CoV and other Class I viruses.

6.1 Experimental design to study MPER of SARS-CoV

The MPER is highly conserved in the coronaviruses. It is exposed on the virion surface and hypothesized to lie on the water-lipid interface of the viral membrane. In HIV-infected sera, this exposed region is also shown to be the epitope for broadly neutralizing antibodies that could be raised (Suarez *et al.* 2000, Zwick *et al.* 2001, Schibli *et al.* 2001). These features of MPER make it an attractive and potential target for drug and vaccine development. However, the lack of a high resolution structure of the MPER on the virus fusion protein also makes it difficult to explore its exact mechanism of action. An objective of this thesis is to study the MPER of SARS-CoV and its role during virus fusion and entry.

The experimental design involves mutational studies on the MPER of SARS-CoV S protein on pseudotyped viruses. Pseudotyped viruses are chosen as they are safe and replication incompetent. They also closely mimic the SARS-CoV and allow virus entry into target cells to be assayed. Non-conservative and conservative mutations were employed to study the importance of MPER as well as the importance of aromaticity in the region. To complement the mutational studies, biophysical properties of model synthetic peptides of MPER as quaternary protein mimetics were studied to confirm the *in vivo* results and membrane remodeling properties of the peptides were studied. These quaternary protein mimetics were designed to emulate different fusion states of MPER and provide insights to their structure and mechanism in lipid environment.

6.2 Experimental design and approach to a high affinity probe – “Synthetic antibody”

This novel strategy is based on the hypothesis that proteins possess self complementary sequences within their amino acid primary structure. Examples of self-binding peptides are found in viral envelope glycoproteins in both Class I and Class II. They are usually located within ordered secondary structures such as β sheets and α helices. Other examples include prion proteins and many sera proteins that form fibrils. My hypothesis is that the active state of MPER is trimeric, involving self-binding of MPER peptides. However, these self-binding interactions are likely weak. Thus, my design is to improve the binding affinity by combining peptides with additive effects through a chemical linker.

The experimental design is to screen a library of overlapping peptides of the protein sequence of interest with the protein of interest itself to identify sequences that can interact with the protein. SARS S protein was used, in this case, as an example for proof of concept. This is followed by further confirmation of interaction by pull down assays and testing for possible additive effects of different combination of peptides and their dimeric form. Surface plasmon resonance was used to quantify the binding efficiency and kinetics of binding to S protein. Peptide candidates that demonstrate high binding affinity were synthesized with a biotin probe and used as reagent in downstream applications to compare its efficacy with conventional antibodies.

This strategy may aid in the development of inhibitors for SARS-CoV by blocking receptor binding or fusion events. It can also generate a panel of SARS-CoV specific

peptides that SARS-CoV can bind to, opening new opportunities in diagnostics. This platform technology can be extended to other Class I or even Class II viruses to aid their development of therapeutics and diagnostics.

Chapter 2

Importance of MPER in SARS-CoV Viral Infectivity

1. Introduction

The SARS-CoV has a 29727 nucleotides long, polyadenylated RNA genome with a GC content of 41% (Rota *et al.* 2003). Its genome organization is typical of coronaviruses with the characteristic gene order of 5'- replicase (rep), spike (S), envelope (E), membrane (M), nucleocapsid (N) -3' and short untranslated regions at both termini. To date, there is no vaccine and effective treatment for SARS. Antibiotics are not effective and existing treatments are mostly supportive with antipyretics, supplemental oxygen and ventilatory support as needed. Development of rapid and sensitive diagnostic tests to monitor the spread of this infectious disease is also needed in view of potential outbreaks in future.

Spike (S) protein, the outermost component of coronaviruses, is responsible for the viral attachment, fusion and entry of coronaviruses into host cells. S protein is a Class I virus fusion glycoprotein that mediates membrane fusion by refolding the fusion protein to a highly stable rod-like structure with a central trimeric α -helical coiled coil, characteristic of Class I virus fusion proteins. In the pre-fusion state, the two heptad repeat (HR) regions oligomerize to form homotrimer coiled coil cores. Following insertion of fusion peptide into the target cell membrane, conformational changes induced by receptor binding cause HR interactions between residues 916 to 950 of HR1 and residues 1151 to 1185 of HR2 (Tripet *et al.* 2004). This mediates the

formation of an anti-parallel six helix coiled coil structure during fusion, termed trimer of hairpins or 6 helix bundle (6HB). It is believed to facilitate apposition of both viral and target cell membranes, causing fusion and subsequent virus entry.

A highly conserved ten-residue sequence Y(V/I)KWPW(W/Y)VWL rich in aromatic amino acids with 3 to 4 tryptophan (Trp) residues resides just before the transmembrane (TM) domain, termed the membrane-proximal external region (MPER). Sequence alignment reveals that the MPER is almost totally conserved in the *Coronavirus* family (Rota *et al.* 2003). It is also highly conserved in Class I viruses, such as HIV and Ebola virus. According to the trimer-of-hairpins model, this aromatic region would also align with the fusion peptide and TM domain during apposition of target cell and viral membranes, possibly adding to the overall hydrophobicity of the environment and contribute to the distortion of lipid membranes necessary for fusion (Saez-Cirion *et.al.*, 2003, Salzwedel *et.al.*, 1999).

Peptides derived from MPER region have been demonstrated to possess a high tendency to partition into the membrane interface according to the Wimley and White (WW) interfacial hydrophobicity scale (Wimley and White 1996, Yau *et al.* 1998). Studies in HIV-1 gp41 and EboV glycoprotein 2, using synthetic peptides derived from the MPER domain, showed that these regions partition into and perturb the integrity of lipid vesicles (Suarez *et al.* 2000, Saez-Cirion *et al.* 2003, Suarez *et al.* 2000). It was recently reported that this conserved and highly aromatic MPER of CoV strongly partitions into the membranes of lipid vesicles with a preference for vesicles containing anionic lipids. It is also able to perturb membrane integrity and cause leakage of vesicle contents. These findings are consistent with the previous findings

for HIV and EboV (Suarez *et al.* 2000, Saez-Cirion *et al.* 2003) and suggest a functional role of MPER in virus fusion and entry. Currently, the specific mechanism of action of MPER remains unknown.

It has been shown that the MPER of SARS-CoV S protein is important in viral infectivity (Lu *et al.*, 2008). This section focuses on investigating the importance of aromatic residues in MPER of SARS-CoV and their structure-function relationship, as there is an unusually rich cluster of aromatic amino acids present within the short region. In this study, the Ala-scan mutational studies on SARS-CoV MPER is extended to replacing with Phe to partially recover aromaticity and the importance of aromatic residues for the function of MPER in membrane perturbation is further investigated. Structural changes that occur when aromatic residues are being substituted were also studied and how these changes relate to corresponding functional changes were also analyzed.

2. Materials and Methods

2.1 Site directed mutagenesis of SARS-CoV S protein

All mutants were constructed based on plasmid pcDNA3.1-OPT9-S, which contains a codon-optimized SARS-CoV S gene and was kindly provided by Prof. Zhang Linqi (Aaron Diamond AIDS Research Center, Rockefeller University, New York 10016). A series of plasmids containing mutated SARS-CoV S protein gene sequences were prepared, based on pcDNA3.1-OPT9-S, using QuickChange® II XL site-directed mutagenesis kit (Stratagene, La Jolla, CA), following the manufacturer's instructions.

The PCR reaction mixture was prepared as follows: 5µl of 10x reaction buffer, 3µl (10ng) of dsDNA template, 3µl (125ng) of forward and reverse primers each, 1µl of dNTP mix, 3µl of Quick Solution and adding ddH₂O to a final volume of 50µl. 1µl of PfuUltra HF DNA polymerase (2.5U/µl) was added before performing PCR. Cycling conditions were as follows: 95°C for 1min, followed by 18 cycles of 95°C for 50sec, 60°C for 50sec and 68°C for 12 min, and a final elongation step of 68°C for 7min.

The PCR products were digested using DpnI (1µl of 10U/µl DpnI in 50µl of products) at 37°C for 1h. The DpnI treated DNA were transformed into XL10-Gold® Ultracompetent cells and grown on LB ampicillin agar plates for >16h to produce clones. All clones were confirmed by sequencing and the respective plasmids were amplified using Qiagen MidiPrep kit (Qiagen, Valencia, CA) for transfections.

2.2 Preparation of pseudotyped SARS-CoV containing different S protein mutants

Pseudoviruses containing SARS CoV S protein were prepared by co-transfection of 293T cells using calcium phosphate transfection method with pNL4-3Luc⁺Env⁻Vpr⁻ and pcDNA3.1-OPT9-S mutant plasmids. pNL4-3Luc⁺Env⁻Vpr⁻ was kindly provided by Prof. Zhang Linqi (Aaron Diamond AIDS Research Center, Rockefeller University, New York 10016). The culture supernatant containing the pseudovirus was harvested 48h and 72h after transfection. Cell debris were cleared by filtration of the supernatant through a 0.45µm filter. The pseudovirus was concentrated by ultracentrifugation at 26,000 rpm for 3h (Zhang *et al.* 2004). The pelleted virus was resuspended in DMEM with 10% new born calf serum and stored as aliquots in -80°C.

2.3 Quantification of virus by reverse transcriptase (RT) assay

Concentrations of the different mutant pseudoviruses were quantified by their RT activity using EnzChek® reverse transcriptase assay kit (Molecular Probes, US), following the manufacturer's instructions. Briefly, reaction mixture was prepared by annealing the template and the primer provided and diluted 200-fold into the polymerization buffer. 20µl of the reaction mixture was aliquoted into each well of a 96-well microplate before 5µl of each standard or virus samples were added to each well. The reaction was incubated at 25°C for 30min before adding 2µl of 200mM EDTA to stop the reaction. Quantification of the samples was achieved by adding 173µl of PicoGreen working solution to each well containing samples and incubated for 5min at room temperature. Fluorescence of each sample was measured using

Tecan Safire Fluorescence Reader (MTX Lab Systems) and standard fluorescein wavelengths (excitation ~480nm, emission ~520nm). A standard curve of fluorescence versus RT activity was plotted and RT activity of experimental samples was then determined based on this standard curve.

2.4 Single cycle infectivity assay

Vero E6 cells (3×10^4 cells/well in 300 μ l) were plated in 48-well plates and cultured at 37°C, 5% CO₂, overnight. On the following day, Vero E6 cells were incubated 1h with standardized amounts of pseudotyped viruses in DMEM (based on RT assay, 0.5U of RT/well). After washing, the cultures were propagated for 48h. After 48h of incubation, the cells were assayed for their luciferase activity using Luciferase Assay System (Promega, Madison, WI). The cells were washed and lysed with 100 μ l of lysis buffer. Luciferase activity was determined after addition of 100 μ l of luciferase assay reagent into 20 μ l aliquot of cell lysate and counting the resultant scintillation for 15s using a TD-20/20 Luminometer.

2.5 Solid phase peptide synthesis

All amino acids and coupling reagents were purchased from Novabiochem (San Diego, CA, USA). The 0.34 mmol/g substituted rink amide resin was purchased from Advanced ChemTech (Louisville, KY, USA). All solvents used were of the highest commercial grade. Peptides were prepared by Fmoc chemistry using 20% piperidine

in DMF for 30 min at the deprotection step. Amino acids used during the coupling step were four times in excess. Final cleavage of all side chain protecting groups and the peptide from the resin was performed with cleavage mixture of 95% TFA, 2.5% water and 2.5% TIS. All peptides were HPLC purified and molecular mass determined by MALDI-TOF mass spectrometry.

2.6 Preparation of LUVs

Large unilamellar vesicles consisting of 90% POPC with 10% POPG (Avanti Polar Lipids) were prepared according to the extrusion method of Nayar *et al.*. Briefly, lipids were dried from chloroform/methanol solution with nitrogen gas stream and high vacuum overnight. Lipid vesicles were resuspended in the respective buffers for the different assays to bring the lipid concentration to 100mM. Samples were subjected to 15 freeze and thaw cycles followed by extrusion through 0.1 μ m polycarbonate membranes in an extruder.

2.7 Liposome leakage assay

Carboxyfluorescein encapsulated LUVs were prepared by resuspending dried lipid film in solution containing 50mM CF. The unencapsulated CF was removed from the CF encapsulated LUVs by gel filtration using Sephadex G-25 (PD-10 column from Amersham). The sizes of LUVs were examined by determination of the hydrodynamic radius obtained with dynamic light scattering. Peptides were added to

the CF-LUVs and measured immediately and after one hour incubation at 25°C. All measurements were done on a spectrofluorometer with excitation at 480nm and emission at 518nm and usage of a cut-off filter at 495nm. All readings were corrected and expressed as percentage leakage.

2.8 Circular dichroism spectroscopy

CD measurements were made on Chirascan CD spectrometer (Applied Photophysics). Cell of 0.1mm pathlength (Hellma Uk Ltd.) were used for all the measurements at 25°C. Samples were measured between 190nm and 240nm, with a 0.1nm step resolution, a measurement speed of 60nm/min and a 1nm bandwidth. Baselines were either water or the respective buffer solutions in which the peptides were dissolved in. At least four repeat scans were obtained for each sample and its respective baseline. The average baseline spectrum was subtracted from the average sample spectrum before the net spectrum was smoothed with a Savitsky-Golay filter. Secondary structure analyses were performed using CDNN program with the respective spectrum as input. The secondary structure contents were then plotted against TFE percentage used in the experiments.

3. Results

3.1 Site directed mutagenesis of MPER of SARS-CoV S protein

To investigate the importance of aromatic amino acids of MPER in SARS-CoV viral entry, mutational studies of MPER of S proteins were performed. Alanine scan of the aromatic residues in MPER was first used to determine the positional importance of each Trp, Phe or Tyr residues. Multiple mutations of the MPER aromatic residues were then performed. The Ala-mutants were then replaced with a Phe residue to determine the importance of aromaticity contributed by each aromatic residue to the region. The different constructs of S protein mutants are shown in Table 3. This part of work was performed in collaboration with Dr. Lu Yanning. Site-directed mutagenesis strategy was used to generate plasmids with mutant S proteins. All plasmids containing different mutated S genes were selected and confirmed via sequencing.

In order to determine S protein-mediated infectivity in VeroE6 cells, a system using an HIV-1 plasmid expressing luciferase reporter gene was pseudotyped with SARS-CoV S protein and its various mutants to generate the pseudotyped SARS virus. Pseudotyped retroviruses containing S protein mutants were generated in 293T cells by co-transfecting the HIV-1 luciferase reporter plasmid together with the mutated S protein-encoding plasmid.

Table 3. Mutations introduced in Trp-rich region of S protein

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

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

3.2 Quantification of virus by RT assay

Production of pseudotyped virus was followed by concentration of the virus obtained and its quantification in order to ensure that comparable amounts of viruses were used for the infectivity assay. Quantification was performed by determining the levels of reverse transcriptase (RT) activity in the samples and comparing to that from a standard calibration curve of known RT activity. All pseudotyped virus samples were quantified and comparable amounts were used in the infectivity assay that followed.

3.3 Infectivity of mutant pseudovirus on VeroE6 cells – Luciferase assay

Pseudoviruses were used to infect VeroE6 cells, which had been shown to be a permissive cell line for SARS-CoV. The infectivity of the pseudoviruses in VeroE6 cells were determined by measuring the luciferase activity in the cell after infection. It was shown that tri-, tetra- and penta-substitution with Ala (WYF5A, WY4A, W3A) completely abrogated infectivity, while single- and double-substitution with Ala (W2A-4/6, W2A-4/9, W2A-6/9, W1194A, W1196A, Y1197A, W1199A and F1202A) substantially decrease infectivity by >90% (Table 4). In contrast, Phe-substituted mutants are able to restore 10-25% infectivity comparing to the wildtype (Table 5). The percentage infectivity restored by Phe mutants at amino acid position of SARS-CoV pseudotypes in descending order is as follows: W1196 > W1199 > Y1197 > W1194. These results suggest that maintenance of the aromaticity of MPER in S protein is essential for SARS-CoV infectivity. The varying degree to which the level of infectivity is being restored may also provide clues as to which aromatic residues

Table 4. Infectivity of different SARS-CoV pseudotyped with S protein mutants

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

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

Table 5. Comparison of the infectivity of SARS-CoV pseudotyped with S protein Ala- and Phe-mutants

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

contribute more to the important role of MPER in viral infectivity.

3.4 Synthesis of mutant MPER peptides

The yield of S protein by expression was low, thus, to further investigate the biophysical properties of MPER and its mutants, peptides corresponding to the MPER region were chemically synthesized. The use of linear peptides provides a convenient and direct approach for our studies. It also solves the problem of obtaining native S protein to carry out the studies. However, it should be noted that such peptides often are conformationally flexible and may not be a true representation of the native conformation. As MPER's main feature is its extraordinarily rich aromatic sequence, it is predicted to be hydrophobic. It should be noted that in this case, amino acids upstream of MPER were included to improve solubility. MPER, in this report, is defined to be between amino acid positions 1187 to 1202 on the wildtype SARS CoV Spike protein. Wildtype MPER, KYEQYIKWPWYVWLGF (wt MPER), was synthesized by solid phase peptide synthesis using Fmoc chemistry. Aromatic residues within the MPER were replaced with Ala accordingly during synthesis, using Fmoc chemistry as well (Table 6). All peptides were HPLC purified and their respective molecular weights were verified by MALDI-TOF mass spectrometry.

3.5 Liposome leakage abilities of MPER mutant peptides

To test the potential of MPER mutant peptides in perturbing membrane integrity,

Table 6. Synthetic MPER Peptides from SARS-CoV and analogs

Name	Sequence Position	Amino Acid Sequence
wt MPER	1187 – 1202	K Y E Q Y I K W P W Y V W L G F
W1194A	1187 – 1202	K Y E Q Y I K A P W Y V W L G F
W1196A	1187 – 1202	K Y E Q Y I K W P A Y V W L G F
Y1197A	1187 – 1202	K Y E Q Y I K W P W A V W L G F
W1199A	1187 – 1202	K Y E Q Y I K W P W Y V A L G F
F1202A	1187 – 1202	K Y E Q Y I K W P W Y V W L G A
3W3A	1187 – 1202	K Y E Q Y I K A P A Y V A L G F

Aromatic amino acids of MPER mutant analogs were substituted with alanine during synthesis (highlighted in blue).

carboxyfluorescein (CF) leakage assay was used. CF was encapsulated in LUVs at self-quenching concentration of 50mM. Any leakage of CF into the external medium results in dilution and thus, the relief of self-quenching and an increase in fluorescence. Peptides in solution were added to LUVs and dequenching occurs within a few seconds after addition. The amount of fluorescence measured is directly proportional to the extent of leakage of CF from the LUVs.

A pronounced membrane destabilizing effect of the wildtype MPER peptide (wt MPER) was observed by the leakage assay. Wildtype MPER induced 65% leakage (with reference to 100% leakage by Triton X-100) upon addition to the CF-encapsulated LUVs, and caused 100% leakage after 1h incubation (Figure 7). Y1197A has the same leakage assay profile as wt MPER, while W1199A showed only a slight decrease of about 10% in leakage compared to wt MPER. Thus, from my results, Ala mutation on tyrosine at position 1197 has no significant effect and Ala mutation on tryptophan at position 1199 has minimal effect on liposome leakage ability of MPER on the whole. The Ala mutations showed greatest effects with W1194A, W1196A, F1202A and 3W3A, diminishing the leakage abilities of the mutated MPER peptides to less than 20%, and totally abrogated leakage ability of MPER in the triple mutant even after 1h incubation. Amongst the single mutants, W1194A exhibits the lowest potential to cause leakage at 5% only. Two time points at 0h and 1h were taken and the extent of leakage exhibited by the various MPER mutant peptides follows a general trend that remained consistent even after an hour incubation of the reaction mixture, albeit with higher level of fluorescence recorded (Figure 7).

From my results, mutating aromatic residues nearing both ends of the MPER (i.e.

Figure 7. Liposome leakage abilities of MPER peptides.

Peptides were added to the CF-LUVs and measured immediately (pink) and after one hour incubation (blue) at 25°C. Recordings were done using a spectrofluorometer with excitation at 480nm and emission at 518nm and usage of a cut-off filter at 495nm. All readings were corrected and expressed as percentage leakage.

W1194A, W1196A and F1202A) has significantly greater effect in decreasing the membrane destabilizing ability of MPER than mutating aromatic residues at the centre of the MPER (Y1197A and W1199A), which retains almost all or if not all of its membrane destabilizing ability. No enhancement of leakage abilities were observed in the experiments, only diminished activities recorded. The percentage decrease in liposome leakage by Ala mutants at S protein amino acid position of synthetic peptides in descending order is as follows: W1194 > F1202 > W1196 > W1199 > Y1197. It is interesting to note the substantial difference in activities of W1196A and Y1197A in response to Ala mutation. Residing side by side, mutation of tyrosine at position 1197 has no significant effect compared to wt MPER while mutation of tryptophan at position 1196 reduces the leakage from 65% in wt MPER to 24% in W1196A, effectively reducing the membrane destabilizing activity by more than half.

3.6 CD spectroscopy of MPER mutant peptides under different TFE conditions

To investigate how the different Ala mutations may change the structure of MPER and affect their infectivity and membrane destabilizing potential, CD spectroscopy was used to analyze the secondary structures of the mutant peptides under varying TFE concentrations. From the results obtained, under 0% TFE, wildtype MPER adopts a less helical and more β sheet conformation (Figure 8a), while some mutants are more helical than the wt MPER and others do not show significant ordered secondary structures (more random). It is interesting to note that Y1197A and W1199A (dashed blue and red lines in Figure 8a), which had none or little effect on their membrane destabilizing activities due to mutation, adopts a more α -helical conformation than

Figure 8. Circular dichroism of MPER mutant peptides.

(a) CD spectra of wt MPER, Y1197A, W1199A and (b) other MPER mutant peptides in water at 25°C were overlaid in the respective diagrams. (c) CD spectra of wt MPER, Y1197A, W1199A and (d) other MPER mutant peptides in 80% TFE at 25°C were overlaid in the respective diagrams.

wt MPER. Mutants with more profound effects on their activities due to Ala mutation, W1194A, W1196A, F1202A and 3W3A (solid lines in Figure 8b), exhibits CD spectra that lacked characteristic secondary structure signatures (eg. α -helix and β -sheets) and lacking distinct maxima at $\sim 200\text{nm}$, indicating a predominantly random coil conformation compared to wt MPER.

Under 80% TFE (Figure 8c and 8d), all MPER peptides generally adopted a more ordered structure. All mutant peptides adopted more helical structure than wt MPER. Interestingly, under 80% TFE, spectra for W1194A and 3W3A exhibited the most pronounced α -helical content and this correlates with the liposome leakage data whereby both of these mutants had the most diminished membrane destabilizing activities, with 3W3A having totally abrogated activity.

Due to the chemical nature of TFE, it is used in this study to simulate a lipidic environment. From the CD spectra obtained (Figure 9), it can be seen that all mutants unanimously adopt a more significant and pronounced α -helical structure, with characteristic α -helix double minima at 208 and 222nm, under increasing TFE concentrations. From these results, the mutant peptides transit to more α -helical structure in an increasing lipid-like environment.

Figure 9. Circular dichroism spectra of MPER mutant peptides under increasing TFE concentrations.

(a) CD spectra of MPER mutant peptides W1194A, W1196A and Y1197A under increasing TFE concentrations at 25°C were overlaid in the same diagram respectively.

Figure 9. Circular dichroism spectra of MPER mutant peptides under increasing TFE concentrations.
 (b) CD spectra of MPER mutant peptides W1199A, F1202A and 3W3A under increasing TFE concentrations at 25°C were overlaid in the same diagram respectively.

4. Discussion

Fusion process is dictated by the protein-protein interaction of S protein and host receptors, which trigger specific and discrete conformational changes of the virus fusion protein. This fusion process is stepwise, precise and synchronized. It must occur in the correct sequence triggered by various conformational intermediates of the S protein. If conformational changes that cause fusion are activated too early, union of the virus and host membrane will not occur and fusion will be abrogated. In this highly synchronized process of virus fusion with host membrane, conformation intermediates play a key role in synchronization.

The MPER of Class I virus fusion glycoprotein is absolutely conserved in members of coronaviruses and highly conserved in other RNA viruses such as HIV, FIV and EboV (Sainz *et al.*, 2005). The importance of MPER in vaccine and therapeutics development had been demonstrated in HIV research as its epitopes are recognized by broadly neutralizing antibodies from human sera (Zwick *et.al.*, 2001). We believe that Trp in MPER plays an important role in these series of synchronization events. Thus far, research on MPER has largely focused on the putative roles that MPER takes on during fusion and possible interactions with lipids and cholesterol. My study on MPER aims to provide a mechanistic explanation and new insights to the role and importance of MPER during fusion events.

In this section of my thesis, I will address the importance of Trp in MPER for synchronization in the fusion event. And out of all Trp, W1194 appears to be the crucial player. This conclusion is deduced from the results of infectivity and liposome

leakage data. To investigate the importance of MPER in viral entry of SARS-CoV, pseudotyped SARS-CoV containing S protein mutants were used to determine S-protein mediated infectivity. Replacing aromatic amino acids in MPER by multiple-substituted Ala mutants completely abrogated infectivity while double- and single-substituted Ala mutants substantially decrease infectivity by >90%, suggesting that the aromatic amino acids of MPER are important in SARS-CoV viral entry.

To ascertain the contribution of aromaticity of MPER during viral fusion, Phe mutants were used to restore the conserved Trp residues. Our results show that these mutants with partial restoration of aromaticity exhibit 10-25% infectivity comparing to the wildtype, suggesting that aromaticity of the MPER plays a role in mediating viral fusion. However, the varying levels to which infectivity is restored for individual residues reveals stronger contribution by certain residues, particularly the location of Trp. For example in W1194F, infectivity cannot be restored by restoring aromaticity using Phe.

We hypothesize that Trp in MPER is the conformational keeper of the virus fusion event. When Trp in MPER is replaced by mutation to Ala, the ability of S protein to mediate fusion is lost, and its aromaticity cannot be totally restored by replacing with Phe mutants. As such, the conformational keeper is gone, resulting in diminished ability to cause synchronized fusion. Such importance and critical role of Trp residue in fusion protein mediating virus fusion is not without precedent. In the influenza hemagglutinin, Trp¹⁴ is a critical residue that stabilizes the kink in the boomerang structure of the HA fusion domain (Lai *et al.* 2006). Mutation to Ala removed its ability to cause fusion due to a more flexible structure as a result of the loss of the

kink structure (Lai *et al.* 2007). These observations illustrate the importance of Trp in contributing to the kink region of influenza HA fusion domain to form a boomerang structure for fusion function (Lai *et al.* 2007).

Trp is different from Phe because of its indole ring. At physiological pH (pH 7), indole N is not protonated. The indole N is a weak base that cannot be protonated under aqueous conditions but may be protonated in lipidic environment. Upon endosome acidification after the virus is endocytosed, pH-induced conformational changes may result in increased lipidic environment surrounding the protein, causing the N atom to become charged by protonation and its potential to interact with carboxyl side chains or charged side chains of lipids is greatly enhanced. The charged N in indole ring of Trp may also be important in allowing the oligomeric state of MPER to dissociate by charge-charge repulsion, as well as enabling dissolution of MPER. However, in the case of Phe substitution, this indole ring effect is absent.

Apart from the Trp residue importance in MPER function, we postulate positional importance of Trp in MPER function as well. In particular, Trp at position 1194 is vital for the function of MPER in infectivity and destabilization of lipid membrane. Substituting Trp at 1194 to Ala renders the structure more helical as determined by CD spectroscopy. This changes the MPER from one state to another. Increase in helicity in MPER seems to be the important conformational change that allows fusion to proceed. The premature increase in helical structure of the W1194A mutant causes the MPER to progress past the fusion stage even before meeting the host membrane, losing the only chance to exert its fusion effect. We hypothesize that the importance of Trp in MPER lies not only in the precise guarding and synchronization of fusion

events, but also to provide plasticity to the fusion protein structure, such that it can adopt different states during fusion.

To conclude the findings in this section, we postulate that MPER has a role in the synchronization of precisely timed fusion events. The importance of the MPER and its unusually rich Trp content has been widely reported and studied in a few Class I glycoprotein viruses. From my studies, it is believed that the specific and precise changes in conformation that results in the different intermediate states during fusion can be attributed to Trp in the MPER. The key to the importance of Trp in MPER is its contribution to the plasticity of the fusion protein structure. The virus fusion protein must have the flexibility and plasticity to take on different structures during fusion due to the different intermediate states of fusion. Mutations to Ala increase helix content, causing the pre-fusion structure to be in a 'locked' state, thus forgoing the plasticity to adopt conformational changes during fusion. From my studies, two major contributing factors determining the importance of Trp in MPER function are (i) residual importance (importance of identity of residue) and (ii) positional importance (importance of position of residue).

Initially termed the aromatic-rich region by some researchers, we believe that it is more than aromaticity that is important in MPER's role in mediating fusion. The importance of Trp in MPER is exhibited not just in the unusually high occurrence in MPER but also supported by mutation and Phe rescue experiments. Our results show that the position to which Trp reside is also vital. The positional importance of Trp in MPER is seen in particular on position 1194, where mutation to Ala results in total loss of activity that cannot be restored by replacing with an aromatic Phe residue. The

resulting increase in helicity due to Ala mutation could prematurely progress the fusion protein conformation to post-fusion state, thereby inactivating its fusion capability.

Following the investigation of Trp and its role in MPER, my study moved on in the next section to look at the importance of oligomeric state and how to exploit the oligomeric state and conformational changes of MPER during fusion.

Chapter 3

MPER Quaternary Protein Mimetics of Intermediate States During Fusion Event

1. Introduction

Fusion of virus with host cell membrane is an essential process for SARS CoV entry and successful infection to occur. This is mediated by S protein, a Class I virus fusion glycoprotein, which undergoes many structural changes and different intermediate states during the fusion event. Briefly, at pre-fusion state, the two HR regions oligomerize to form homotrimer coiled coil cores. Upon receptor binding, induced conformational changes cause HR1 to fold back to HR2, via interactions between residues 916 to 950 of HR1 and residues 1151 to 1185 of HR2 (Tripet *et al.* 2004). This results in the formation of an anti-parallel six helix bundle (6HB), a coiled coil structure during fusion state, which is also known as a trimer of hairpins (Park *et al.* 2003, Liu *et al.* 2004, Chu *et al.* 2008). It is believed that the 6HB facilitates apposition of both viral and target cell membranes, causing fusion and subsequent virus entry.

MPER is highly conserved in Class I viruses, such as HIV, Ebola virus as well as SARS CoV. Peptides derived from this region have been demonstrated to possess a high tendency to partition into the membrane interface, according to the Wimley and White (WW) interfacial hydrophobicity scale (Wimley and White 1996, Yau *et al.* 1998), as well as empirically in HIV-1 and EboV studies (Suarez *et al.* 2000, Saez-

Cirion *et al.* 2003, Suarez *et al.* 2000). This conserved and highly aromatic MPER of CoV also strongly partitions into the membranes of lipid vesicles with a preference for vesicles containing anionic lipids. It is also able to perturb membrane integrity and cause leakage of vesicle contents. These findings are consistent with the previous findings for HIV and EboV (Suarez *et al.* 2000, Saez-Cirion *et al.* 2003) and suggest a functional role of MPER in virus fusion and entry. Currently, the specific mechanism of action of MPER remains unknown.

There are several hypotheses as to how the MPER works to aid virus entry. Structural data from HIV gp41 suggested that the MPER is flexible (Weissenhorn *et al.* 1997). In influenza HA2, MPER is structurally disordered in the fusion pH-induced conformation (Bullough *et al.* 1994). It was suggested that MPER may undergo a conformational change from β strand to a partly α helical structure when in contacts the lipid membrane, just like the fusion peptide (Salzwedel *et al.* 1999). MPER may serve as an agent to promote membrane destabilization required for fusion (Saez-Cirion *et al.* 2003). This may occur as the helical core of the fusion protein undergoes conformational changes to bring the two hydrophobic regions (FP and MPER) to close proximity, and the MPER taking the active role to perturb the apposed membranes during fusion. This may also involve interactions with membrane components such as cholesterol (Salzwedel *et al.* 1999).

It was further proposed that the MPER, in conjunction with the FP and the TM anchor, forms a continuous track of hydrophobic, membrane-interacting surfaces that provide a low energy (low barrier) path for lipid flow and membrane fusion during virus fusion (Sainz *et al.* 2005). Other possibilities also include the MPER acting as a

contributor to trimer formation and maintenance of its stability (Saez-Cirion *et.al.*, 2003, Salzwedel *et.al.*, 1999). A summary of the proposed function of MPER based on studies in different Class I viruses has been tabulated into Table 2.

We hypothesize that initially the prefusion aromatic MPER may lie on the membrane interface together as a trimer extending from the S protein trimer, similar to that observed in HIV gp41 (Figure 10). The prefusion state MPER then undergoes structural changes to become fusion-active MPER, initiated by receptor binding and acidification of endosome. The MPER then dissociates from the bundle of trimer, like legs splaying apart, and this fusion state structure serves as a destabilization motif to interact with and disrupt the lipid bilayer of host cell membrane.

In the previous chapter, we introduced the hypothesis of the importance of Trp in MPER due to its contribution to ‘plasticity’ of MPER structure required during fusion. The MPER adopts different forms during fusion that allows its function during the fusion events. In this study, we wish to provide evidence to support membrane interacting properties of MPER in Class I virus fusion glycoprotein using SARS CoV MPER, and to construct quaternary protein mimetics of MPER intermediate fusion states for further studies. The key design of this novel protein mimetic approach for antivirals against MPER is the multivalent parallel peptide chains which mimic the multivalent interactions of MPER. Such quaternary protein mimetics have protein-like properties and are druggable. Two types of quaternary protein mimetics were designed that likely mimic ‘active’ intermediate states of MPER during fusion events. The two types of quaternary protein mimetics are made up of the same MPER sequence but either linked at the N terminal or C terminal. As a result, these two

Figure 10. Surface-rendered model of averaged SIV Env spike tomograms. Env spike (blue) and membrane (grey) in (a) side, (b) intermediate and (c) top views and (d) an intermediate view at a higher density threshold. The spike and membrane surfaces (in a–d) were rendered at different thresholds, the latter adjusted to conform to the known thickness of the lipid bilayer (4 nm). (Taken from Zhu *et al.* 2006).

dimers should adopt and mimic different intermediate states, with difference in properties expected. It can be foreseen that due to the close resemblance in nature to antimicrobial peptides, which disrupt membranes indiscriminately, the potential membrane leaking and destabilizing activity of CT dimer predicted may not represent an ideal therapeutic candidate. In contrast, NT dimer is expected to inhibit virus fusion by self associating to MPER, preventing fusion from proceeding, without damaging effects to cell membrane. We investigated their membrane perturbing potencies and the secondary structures that they adopt under different environments.

2. Materials and Methods

2.1 Solid phase peptide synthesis

All amino acids and coupling reagents were purchased from Novabiochem (San Diego, CA, USA). The 0.34 mmol/g substituted rink amide resin was purchased from Advanced ChemTech (Louisville, KY, USA). All solvents used were of the highest commercial grade. Peptides were prepared by Fmoc chemistry using 20% piperidine in DMF for 30 min at the deprotection step. Amino acids used during the coupling step were four times in excess. Final cleavage of all side chain protecting groups and the peptide from the resin was performed with cleavage mixture of 95% TFA, 2.5% water and 2.5% TIS. All peptides were HPLC purified and molecular mass determined by MALDI-TOF mass spectrometry.

2.2 Synthesis of N terminal linked and C terminal linked dimers

For C terminal-linked MPER dimer, a linker [NH_2 - β -Ala-Lys(NH_2)-Lys(Biotin)-MBHA resin] was first made on resin to create a double amino group per functional amino group on the MBHA resin. The β -alanine was used as a spacer to allow equal distance between the amino groups and the point of branching. This linker is then used to extend the MPER chain on the amino groups as with the monomer. For N terminal-linked MPER dimer, MPER monomers with an additional Cys at the N terminal were synthesized. A linker that is used to ligate two MPERs N terminally was synthesized. The linker was oxidized with sodium periodate (NaIO_4) and HPLC

purified. The aldehyde linker was then used immediately for ligation with Cys-MPER to yield the N terminal-linked MPER dimer. All peptides were HPLC purified and molecular mass determined by MALDI-TOF mass spectrometry.

2.3 Preparation of LUVs

Large unilamellar vesicles consisting of 90% POPC with 10% POPG (Avanti Polar Lipids) were prepared according to the extrusion method of Nayar *et al.*. Briefly, lipids were dried from chloroform/methanol solution with nitrogen gas stream and high vacuum overnight. Lipid vesicles were resuspended in the respective buffers for the different assays to bring the lipid concentration to 100mM. Samples were subjected to 15 freeze and thaw cycles followed by extrusion through 0.1µm polycarbonate membranes in an extruder.

2.4 Liposome leakage assay

Carboxyfluorescein encapsulated LUVs were prepared by resuspending dried lipid film in solution containing 50mM CF. The unencapsulated CF were removed from the CF encapsulated LUVs by gel filtration using Sephadex G-25 (PD-10 column from Amersham). The sizes of LUVs were examined by determination of the hydrodynamic radius obtained with dynamic light scattering. Peptides were added to the CF-LUVs and measured immediately and after one hour incubation at 25°C. All measurements were done on a spectrofluorometer with excitation at 480nm and

emission at 518nm and usage of a cut-off filter at 495nm.

2.5 Ion conductance experiments

Synthetic 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphatidylethanolamine (POPE), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphatidylserine (POPS), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphatidylcholine (POPC), and dimyristoyl-phosphatidylcholine (DMPC) were purchased from Avanti Polar Lipids. Decane (>99%) and (N-[2-hydroxyethyl]piperazine-N9-[2-ethanesulfonic acid]) were obtained from Sigma-Aldrich, chloroform from Sigma-Aldrich, and ethanol (>99%) and sodium chloride (NaCl) were obtained from Fluka. Ion conductance across lipid bilayer was measured (Wilson et al. 2004) by forming planar lipid bilayers across a 250- μ m aperture in the wall of a Delrin cup separating the aqueous solutions in the cis and trans chambers. The working volume was of 1 mL in each chamber. A lipid mixture of 5:3:2, POPE:POPS:POPC (50 mg/mL) was dissolved in chloroform and dried under N₂ gas. The lipid cake was then resuspended in decane. A small amount of lipid mixture was painted across the aperture of the Delrin cup. About 1 mg of peptide was added to the cis chamber under continuous stirring, until ion channel activity was detected, typically after 15–30 min. The chambers contained 500mM NaCl, 5 mM HEPES, pH 7.2 (cis), and 5 mM NaCl and 5 mM of HEPES, pH 7.2 (trans). Recordings were performed in asymmetric ionic conditions using a different buffer. The cis and trans solutions were connected via an Ag/AgCl electrode and 1 M NaCl, 2% agar bridge. Currents were amplified using a Warner Instrument amplifier and digitized at 5 kHz and filtered at 10 kHz using pCLAMP 9.2 software (Axon Instruments, Inc.). The

trans chamber was set as reference and the cis chamber was held at 100 mV potential. Stirring was stopped after detection of channel activity.

2.6 SARS CoV production

Vero E6 cells were maintained at monolayer in 75cm² cell culture flasks. The cells were inoculated with viral stock and incubated at 37°C under 5% CO₂ for 48 hours. Viruses were harvested from the supernatant when cytopathic effects (CPE) were observed in more than 75% of the cells. The supernatant was clarified at 2500rpm for 10min and stored at -80°C in aliquots. The viruses obtained were titrated to get TCID₅₀.

2.7 Toxicity tests and SARS CoV inhibition assay on Vero E6 cells

For SARS CoV inhibition assays, virus dose of 100 TCID₅₀ was used. From the selected dilution of 100 TCID₅₀, serial dilutions were made. Different concentrations of peptides to be tested were prepared and added to the appropriate virus dilutions in a 96 well plate. The virus-peptide mixtures together with the controls were incubated for 1hour under 5% CO₂ and 37°C. Trypsinized Vero E6 cells were then added to all the wells and incubated for another 4-5 days. The experimental plate is examined daily for cytopathic effects (CPE) in the form of plaque formation and the inhibitory activities were recorded. The toxicity tests for the peptides on Vero E6 cells were carried out using the same protocol without the addition of the virus. All experiments

were performed in the BSL3 certified laboratory in Singapore General Hospital (SGH).

2.8 Circular dichroism spectroscopy

CD measurements were made on Chirascan CD spectrometer (Applied Photophysics). Cell of 0.1mm pathlength (Hellma Uk Ltd.) was used for all the measurements at 25°C. Samples were measured between 190nm and 240nm, with a 0.1nm step resolution, a measurement speed of 60nm/min and a 1nm bandwidth. Baselines were either water or the respective buffer solutions in which the peptides were dissolved in. At least four repeat scans were obtained for each sample and its respective baseline. The average baseline spectrum was subtracted from the averaged sample spectrum before the net spectrum was smoothed with a Savitsky-Golay filter. Secondary structure analyses were performed using CDNN program with the respective spectrum as input. The secondary structure contents were then plotted against TFE percentage used in the experiments.

3. Results

3.1 Synthesis of MPER peptide and its quaternary protein mimetics as dimers

MPER peptide and its quaternary protein mimetics were chemically synthesized to study their different putative conformational states that occur during fusion and how these “intermediate” states may contribute to the functional role of MPER during virus fusion. The aim is to design two types of MPER quaternary protein mimetics: an ‘opened’ structure of dimer and a ‘closed’ structure of dimer. Linking the N terminal of MPER peptide dimers is predicted to form a ‘closed’ structure due to hydrophobic interaction at the free C terminal ends. While linking the C terminal of MPER peptide dimers can result in an ‘opened’ structure, under acidified conditions, due to charge-charge repulsion at the free N terminal ends (Figure 11).

Due to MPER’s extraordinarily rich aromatic sequence and predicted hydrophobicity, hydrophilic amino acids upstream of MPER were included to improve solubility. MPER, in this context, is defined to be between amino acid positions 1187 to 1202 on the wildtype SARS CoV Spike protein. MPER monomer, KYEQYIKWPWYVWLGF (KF16), was synthesized by solid phase peptide synthesis using Fmoc chemistry (Carpino and Han 1972). Two versions of MPER dimers were synthesized, which is linked at either the N terminal (NT dimer) or C terminal (CT dimer). For the NT dimer, MPER monomer with an additional Cys at the N terminal was synthesized, and the purified and unprotected monomer was then ligated to an aldehyde-Lys linker via a thiazolidine bond (Figure 12a). For the CT dimer, a Lys-based linker was coupled to the resin before synthesizing the MPER sequence which yielded a CT dimer after

Figure 11. Design of MPER protein mimetics. The schematic shows the different location of the functional elements of SARS CoV S protein. MPER (green), in our context, is defined to be between amino acids 1187 to 1202. The MPER monomer, KYEQYIKWPWYVWLGF, its N terminal linked and C terminal linked dimer were designed and proposed in our studies as mimetics of MPER.

Figure 12. Chemical synthesis of N terminal linked dimer and C terminal linked dimer. (a) Linker molecule made up of two Ser branching from Lys was synthesized by solid phase peptide synthesis. The linker was then ligated to Cys-MPER by formation of thiazolidine linkage. The final product was N terminal linked dimer consisting of two MPER₁₁₈₇₋₁₂₀₂ linked together at their amino terminal via thiazolidine. (b) C terminal linked dimer was synthesized by solid phase peptide synthesis directly followed by final deprotection and cleavage from MBHA resin. It consists of two MPER₁₁₈₇₋₁₂₀₂ linked together at their carboxy terminal to amino functional groups on the linker.

final cleavage from resin (Figure 12b). All peptides were HPLC purified and their molecular mass confirmed using MALDI-TOF mass spectrometry.

3.2 Liposome leakage abilities of MPER monomer and dimers

To test the membrane-active property of MPER monomer and dimers in perturbing membrane integrity, carboxyfluorescein (CF) leakage assay was used (Isenberg *et al.*, 2002). CF was encapsulated in LUVs at self-quenching concentration of 50mM. Any leakage of CF into the external medium results in dilution and thus, the relief of self-quenching and an increase in fluorescence. Peptides in solution were added to LUVs and dequenching occurs within a few seconds after addition. The amount of fluorescence measured is directly proportional to the extent of leakage of CF from the LUVs.

Similar to previous results, a pronounced membrane destabilizing effect of the MPER monomer was observed as determined by the leakage assay. MPER monomer induced 65% leakage (with reference to 100% leakage by Triton X-100) upon addition to the CF-encapsulated LUVs, and caused 100% leakage after 1h incubation. CT dimer exhibited a similar leakage profile as MPER monomer, causing 74% leakage. In contrast, the NT dimer was unable to produce leakage, yielding only 1% leakage. In general, the MPER dimers showed contrasting results, with CT dimer exhibiting about 10% more leakage than the MPER monomer, while the NT dimer recorded only 1% leakage.

Figure 13. Leakage of carboxyfluorescein from LUVs by MPER monomer, KYEQYIKWPWYVWLGE, NT dimer and CT dimer. Peptides were added to the CF-encapsulated LUVs and measured immediately (blue) and after one hour incubation (purple) at 25°C.

Prolonged incubation after one hour did not alter the overall trend of CF leakage exhibited by the MPER monomer and dimers, remaining consistent albeit with higher level of fluorescence recorded (Figure 13). From the results, the enhanced leakage ability of CT dimer compared to MPER monomer is significant. More importantly, there is substantial difference in the behavior of the two forms of dimers with almost total abrogation of membrane destabilizing ability of the NT dimer.

To provide additional insight to the contrasting effect of the dimers, the reaction mixtures were subjected to dynamic light scattering (DLS) before and after experiment, to determine any particle size change. NT dimer, which was unable to cause significant leakage, recorded an increase in liposome size from the initial diameter of 149nm to 209nm (N terminal linked dimer) whereas the CT dimer that could cause significant and enhanced leakage had widely distributed particle size under DLS, with a majority particle size below 15nm in diameter. Calculations showed that the increase in diameter could correspond to the final diameter after two liposomes of the same initial size fused to form a larger one. This result suggests that there might be lipid mixing at monolayer level without disrupting the bilayer and causing leakage of contents.

3.3 Ion conductance of MPER in lipid bilayer

To further examine the perturbation effects of MPER monomer and dimers on lipid membrane, the planar lipid bilayer technique was used. In this technique, an artificial planar lipid bilayer is formed across a hole connecting two macroscopic

compartments. Peptide was then added to one of the compartment and the current flowing through the membrane was monitored under voltage-clamp conditions. Spikes of current were recorded across the membrane when peptides were tested. From the 20s frame shown, after addition of MPER monomer, occasional spikes of current across the membrane with maximum amplitude of $\sim 5\text{pA}$ were recorded (Figure 14). For NT dimer, no significant spikes of current were recorded, with only minute amplitudes of disturbances (Figure 14). When CT dimer was added, frequent spikes of current were registered through the 20s time frame, with maximum amplitude of $\sim 15\text{pA}$.

In general, the results showed that CT dimer was able to cause significantly more frequent current leaks across lipid bilayer and at higher current amplitudes (about three times higher) than that recorded by MPER monomer and NT dimer. NT dimer exhibited the lowest inclination to cause membrane perturbation, thus, its disability to cause significant leakage of currents across the lipid bilayer. These correlate well with the liposome leakage data.

3.4 CD spectroscopy of MPER and MPER dimers

As mentioned in the previous chapter, where the importance of the Trp residues of MPER to fusion and its contribution to the role of MPER was examined, I would like to look at how the different states of conformational change during fusion contributes to the functional role of MPER during fusion. To investigate the structures of MPER peptides under lipidic mimetic media, CD spectroscopy was used to determine the

Figure 14. Ion conductivity of MPER peptides across planar lipid bilayer. MPER peptides were added and stirred till activity was observed. Ion conductances across the lipid bilayer were recorded as current amplitude against time. The *trans* chamber was earthed and the *cis* chamber was held at 60mV.

secondary structures of the MPER monomer and dimers under varying TFE concentrations. In water, MPER monomer and CT dimer do not display significant ordered secondary structures, whereas the CD spectra for NT dimer showed presence of some α -helical structure (Figure 15a). Due to the chemical nature of TFE, it is used in this study to simulate a lipidic environment. Upon introduction to 80% TFE, MPER monomer and dimers attained more ordered α -helical structures, as seen from the increase in negative bands at ~ 222 and 210nm (Figure 15b). The amplitude for both minima at ~ 222 and 210nm of NT and CT dimers are very similar but the amplitude for both minima of MPER monomer is about two times less than NT and CT dimers.

To examine the specific structural change profile of MPER monomer and dimers in increasing TFE, further investigation of the structural contents of the peptides with step-wise increment of TFE (from 0% to 40%) was performed. The presence of increasing concentration of TFE induced a progressive increase in the α -helical content for MPER monomer (Figure 16a) and NT dimer (Figure 16b), but the amplitude of both minima of the monomer is less than NT dimer for more than two-fold. For CT dimer, a similar increase in α -helical content as that in NT dimer is observed, however, this final structure in TFE was attained regardless of TFE percentage (Figure 17c). From the results obtained, NT and CT dimers progress towards an almost similar final α -helical structure in a lipidic environment, albeit with different profiles of structural change. In general, by using a secondary structure prediction program CDNN, there is a substantial increase in α -helical content for MPER monomer (18 – 21%), NT dimer (20 – 27%) and CT dimer (20 – 25%).

Figure 15. Circular dichroism spectroscopy of MPER monomer and dimers under 0% and 80% TFE condition. (a) CD spectra of MPER monomer (blue), NT dimer (green) and CT dimer (red) in 0% TFE. (b) CD spectra of MPER monomer (blue), NT dimer (green) and CT dimer (red) in 80% TFE.

Figure 16. Circular dichroism spectroscopy of MPER monomer and dimers under different percentage of TFE. (a) CD spectra of MPER monomer under 0% (blue), 20% (green) and 40% (red) TFE. (c) CD spectra of N terminal linked dimer under 0% (blue), 20% (green) and 40% (red) TFE. (d) CD spectra of C terminal linked dimer under 0% (blue), 20% (green) and 40% (red) TFE.

3.5 Inhibitory activities of MPER monomer and dimers on SARS CoV infectivity

To evaluate the inhibitory activities of MPER monomer and dimers on SARS CoV, virus neutralization assays were performed. SARS CoV virus were recovered and used to prepare virus stock for the neutralization assay. The neutralization assays were performed using peptide concentration of 3.125 μ M to 50 μ M. After the virus and peptide were mixed and added to Vero E6 cells, the samples were incubated for 72 hr at 37°C. Cytopathic effects (CPE) were observed under microscope and recorded. The data were tabulated and IC₅₀ was obtained where possible. Toxicity of the peptides was tested on Vero E6 cells from 3.125 μ M to 50 μ M. These were performed in exact condition as the virus neutralization assay. After 72 hours incubation in 37°C, monomeric MPER and NT dimer showed no cell toxicity to Vero E6. However, CT dimer displayed toxicity, killing cells in the tested concentration range.

For MPER monomer, inhibitory activity on SARS CoV infectivity was observed with an IC₅₀ of ~15 μ M, as obtained from the dose-response curve (Figure 17). Inhibitory activity by NT dimer on SARS CoV infectivity was also observed with an IC₅₀ of ~7 μ M, a two fold increase in inhibitory potency compared to the MPER monomer. For CT dimer, despite showing cell toxicity, inhibitory activity was also observed after discounting contribution of CPE due to cell toxicity. IC₅₀ was too large to be obtained from the dose-response curve of CT dimer (Figure 17). It is possible that CT dimer might be more potent than NT dimer, thus causing cell toxicity. Thus, the actual effect of CT dimer could have been partially masked by CPE due to toxicity on Vero E6 cells.

Figure 17. Dose response curves showing inhibition of SARS-CoV activity on Vero E6 cells by MPER monomer, NT dimer and CT dimer. Different concentrations of peptides were added to the appropriate virus dilution in a 96 well plate. The virus-peptide mixtures were incubated for 1hour under 5% CO₂ and 37°C, before adding trypsinized Vero E6 cells and incubated for another 4-5 days. Cytopathic effects (CPE) in the form of plaque formation and inhibitory activities were recorded daily. DRCs of MPER monomer (blue), NT dimer (green) and CT dimer (red) were used to determine the IC₅₀ of the peptides.

4. Discussion

The MPER is absolutely conserved in members of coronaviruses and highly conserved in other RNA viruses such as HIV, FIV and EboV (Sainz *et al.* 2005). The importance of MPER in HIV vaccine and therapeutic development had been demonstrated as its epitopes are recognized by broadly neutralizing antibodies from human sera (Zwick *et al.* 2001). This region exists in S protein S2 region of all coronaviruses and juxtaposes the transmembrane domain. They share a highly conserved ten-residue sequence Y(V/I)KWPW(W/Y)VWL that is rich in aromatic amino acids with 3 to 4 Trp residues. Peptides derived from MPER region have been demonstrated to possess a high tendency to partition into the membrane interface according to the Wimley and White (WW) interfacial hydrophobicity scale (Wimley and White 1996, Yau *et al.* 1998). Studies in HIV-1 gp41 and EboV glycoprotein 2, using synthetic peptides derived from region analogous to MPER domain, showed that these regions partition into and perturb the integrity of lipid vesicles (Suarez *et al.* 2000, Saez-Cirion *et al.* 2003).

Our study suggests that the MPER region of SARS CoV S protein, spanning residues 1187-1202, has the ability to perturb phospholipid bilayer and may cause lipid mixing. Our previous work showed the essential role played by MPER in SARS CoV infectivity by mutational studies in pseudoviruses (Lu *et al.* 2008) and its inhibitory effects on SARS CoV infection of Vero E6 cells. These data indicate that the MPER may have a functional role in virus fusion by aiding membrane destabilization. In this section, we designed peptides to mimic the various fusion states of MPER and our results showed that these peptides show properties that agree with our design. This could represent the first effort whereby protein mimetics of prefusion and fusion state

intermediates of MPER has been synthesized.

In general, results from liposome leakage and ion conductance experiments showed membrane destabilizing and perturbing potential of MPER. A summary of the results illustrating the different properties and structural data of the MPER monomer and dimers are tabulated in Table 7. It is interesting to note that such membrane destabilizing activities exist only in certain states of the fusion intermediates. From the chemical structure and structural data obtained, it is used to predict the structures of MPER peptides (Table 7). NT dimer resembles a 'closed' U-shaped structure due to hydrophobic interactions at the free C termini. This only exposes partially the surfaces of MPER available for interactions and mimics one of the fusion-active intermediate (where all MPERs are close together in trimeric state). CT dimer resembles an 'opened' V-shaped structure due to charge repulsion under acidic conditions at the free N termini. This allows the exposure of all the MPER available surfaces for interactions and thus, mimics a later intermediate state during fusion. Similarly in MPER monomer, all the available surfaces are exposed to exert its effect but with decreased activity as seen in all the experiments. This could be due to the next neighboring MPER being spatially far apart and not as closely linked together like in the CT dimer.

In this section, I would like to put forth my hypothesis of the virus fusion mechanism. Before receptor binding, the pre-fusion state virus fusion protein is a trimer with its MPER held close together via hydrophobic interactions. This is a fusion-inactive structure in which, unless fusion is activated and triggers conformational changes that permits it to progress to the next intermediate state, it remains in a 'locked' state that

Table 7. Comparison of structure, biophysical properties and activities of MPER monomer, NT dimer and CT dimer.

	MPER monomer	NT dimer	CT dimer
Diagrammatic Representation	<p>Distance between monomers are far apart without linker</p>	<p>Hydrophobic interaction</p>	<p>Charge repulsion under acidic conditions</p>
Structure	'Open-end' monomer (All surfaces exposed)	'Closed' U-shaped structure (Partially exposed)	'Opened' V-shaped structure (More surfaces exposed)
Leakage of liposome	Active (<70%)	Inactive (<5%)	Active (>70%)
Ion conductance across lipid bilayer	Occasional spikes Max amplitude ~5pA	No spikes observed Only small amplitude of disturbances recorded	Frequent spikes Max amplitude ~15pA
SARS CoV inhibition	Active inhibition (IC ₅₀ ~25μM) No cell toxicity	Active inhibition (IC ₅₀ ~7μM) No cell toxicity	N.A. Toxic to cells
Circular dichroism	Attain ordered α-helix content under increasing TFE but lower content	Progressively attain more ordered α-helix content under increasing TFE	Attain ordered α-helix content regardless of %TFE (i.e. helix content remains unchanged with increasing TFE)

does not allow MPER to exert its function for virus fusion to occur. As the virus binds to host receptor, it is at the same time internalized via endocytosis. When acidification of the endosome occurs, MPER becomes positively charged causing charge repulsion which dissociates the trimeric structure. The Trps in MPER becomes critical in this step by contributing flexibility to the structure to allow conformational changes. This fusion-active intermediate (described as 'opened' state here), having all surfaces exposed for interaction, then allows MPER to exert its membrane destabilizing potential on the host membrane, before going into a non-fusogenic form that is more α -helical than pre-fusion, with a conformation that is 'locked' from reverting back to the fusion-active form again.

From the results of liposome leakage experiments, CT dimer exhibits much stronger membrane destabilizing activity than MPER monomer. The 'opened' V-shaped structure of CT dimer exposes all available surfaces for interactions. Two MPER being held together in close proximity, with a continuous stretch of hydrophobic segments linked, thus, allows for deeper penetration into lipid bilayer to cause greater scale of destabilization compared to action by a single MPER monomer. For MPER monomer, activity is weaker as monomers are considered molecularly far apart. NT dimer lacks such activity as the hydrophobic segments are disrupted and not linked but may cause lipid mixing to a certain extent but without leaking contents from LUVs. These results correlate with the predicted activities of the corresponding structures. Being linked, MPERs of CT dimer are close together at molecular distance while in the monomeric forms, MPERs are far apart and difficult to exert synergistic membrane perturbing effects at molecular distance. This illustrates the synergistic effects of MPER in causing membrane destabilization during fusion and this effect

could be even greater for a trimer and multiple copies of trimeric S protein, which are believed to be involved in the virus fusion..

As shown by the ion conductance data, CT dimer can cause leakage of higher current amplitude than monomer and more frequently, but NT dimer cannot cause significant current leakage. This is probably because of NT dimer's structure that does not allow it to insert deeply into the membrane bilayer.

Intriguing it may seem at first, the trend of the virus neutralization data was reversed in these experiments. The cell toxicity of CT dimer could be due to its membrane perturbing effects (which is the greatest), which could have lysed the cell's membrane causing CPE. Thus, CT dimer's actual inhibitory effect on SARS CoV could have been masked by its toxicity to cells. Inhibitory activity of NT dimer is twice as strong as MPER monomer, as indicated by their IC_{50} value. This is as predicted because MPER monomer may interact with MPER to prevent fusion but the complex formed is not as strong and stable, with complex dissociation easier compared to NT dimer. In contrast, NT dimer may interact with MPER much stronger to form a very stable trimeric complex that is energetically favorable and not easily dissociated. In addition, the trimeric complex formed has all surfaces available for interaction buried, and thus there is little chance that it will participate in other interactions, causing instability to the complex formed.

In summary, the two different synthetic states contain the same MPER sequences but linked up differently, either at the N terminal or C terminal. This results in different arrangements of hydrophilic and hydrophobic segments. For CT dimer, this results in

a continuous hydrophobic segment that allows stable insertion into membrane giving it membranolytic activity as shown in my studies. However, this is not seen in NT dimer as the hydrophobic stretches are disrupted and not linked, thus, resulting in no significant membranolytic activity. My studies showed that the two different synthetic states have very diverse properties and may mimic different intermediate states of MPER during fusion. These differences in physical and biological properties are due to the ‘plasticity’ of MPER structure, which is required and essential for fusion events. It supports our hypothesis that MPER is able to adopt different forms during fusion that allows its function.

Our studies on SARS CoV MPER agree with work by others on MPER of other Class I virus fusion glycoprotein (eg. HIV, EboV, etc). This could represent a consensus functional element in Class I virus fusion glycoprotein essential for virus fusion, akin to that of fusion peptide (FP), given its high level of sequence conservation and its essential function during fusion. However, the MPER has an added advantage of being more water soluble than the hydrophobic FP, which presents positive implications in drug design, making it a better target for further drug development. The results in this section clearly point to a membrane perturbation role by MPER and that its peptide equivalent is able to inhibit virus infection.

The exact role of MPER during viral fusion and its mechanism remains unknown. Previous work found MPER to be flexible in HIV (Weissenhorn *et al.* 1997) and structurally disordered in influenza HA (Bullough *et al.* 1994), suggesting that MPER may act as a flexible extender that helps bring the virus and host cell membrane into close proximity for fusion to occur. It was suggested that MPER may undergo a

conformational change from β strand to a partly α helical structure when in contacts the lipid membrane, just like the fusion peptide (Salzwedel *et al.* 1999). Other work suggested that MPER may serve as an agent to promote membrane destabilization required for fusion (Saez-Cirion *et al.* 2003), which may also involve interactions with membrane components such as cholesterol (Salzwedel *et al.* 1999). It was further proposed that MPER, in conjunction with FP and TM anchor, forms a continuous track of hydrophobic, membrane-interacting surfaces that provide a low energy (low barrier) path for lipid flow and membrane fusion during virus fusion (Sainz *et al.* 2005). Other possibilities also include the MPER acting as a contributor to trimer formation and maintenance of its stability (Saez-Cirion *et al.*, 2003, Salzwedel *et al.*, 1999).

In conclusion, we believe that after fusion is activated, the S protein undergoes conformational changes whereby HR2 folds back to HR1 to form the 6 helix bundle. This not only pulls both membranes close to each other in preparation for fusion, but also brings the two hydrophobic regions of FP and MPER to close proximity, sandwiched between the virus and host membrane. This close proximity of MPER and FP may form a continuous track of hydrophobic surfaces that provide a low energy barrier path for disrupted lipids to flow, allowing the fusion pore to form. This will then allow genetic material of the virus to transfer to the host cell for replication and successful infection via the fusion pore.

Chapter 4

Quaternary Protein Mimetics via Self Interacting Sequences in SARS CoV Spike Protein using a Novel Method – Implications in Antibody Mimetics

1. Introduction

In the previous chapters, the use of quaternary protein mimetics for therapeutic uses and to inhibit virus infection was explored. In this chapter, how quaternary protein mimetics may be used as a tool for detection of protein-protein interactions will be addressed. The anticipated usage and properties of such agents may have implications as an alternative to the functions of antibodies.

Antibodies have proven to be an invaluable tool for research, diagnostics, therapeutics and in the human body for immune defense. Conventional antibodies bind to their epitopes (ligands) with high affinity ($K_d < 10^{-9}$ M for a good antibody) and specificity, making them useful for a wide range of applications. They are often used as capture agents in research, for purifications and detecting tags, also in diagnostics for detection of antigens. Despite their applications in a multitude of procedures and techniques, antibodies are proteins in nature and this results in many limitations.

Antibodies are largely limited to animal sources and biological production, which makes antibodies expensive and difficult to produce and purify in bulk quantities. Quality control of such biologics is also complicated due to the many factors involved,

as such, batch to batch variations exists. Antibodies are also temperature sensitive and require stringent storage, which makes field applications hard. They can lose their folded structure and therefore, their activity if they are required to be immobilized at surface-solution interfaces. Furthermore, neutralizing antibodies can only be obtained if the antigen is immunogenic.

Many researchers have tried using macromolecular biomolecules (such as nucleic acid aptamers and protein aptamers) as capture agents. However, there seems to be greater potential in developing smaller synthetic molecules as mimics of conventional antibodies. Small synthetic molecules are generally much cheaper and easier to synthesize and purify in substantial quantities with efficient quality control due to their defined reaction conditions and excellent yield. Small synthetic molecules are also easy to store with minimal destruction to their active structures, giving them the possibilities of wider field applications. They can be immobilized on surfaces without much worry on diminished activities. Moreover, it may be possible to develop synthetic molecules that bind to antigens that do not generate antibodies.

In this study, we propose a novel method to prepare synthetic antibody mimics for a given protein (SARS CoV S protein in this case). Our approach is based on the concept of molten globulin in protein folding, whereby ordered secondary structures fold first and form microdomains in molten globule state (Ikeguchi *et al.* 1997, Mizuguchi *et al.* 1999). This partially folded state primes further folding of tertiary and quaternary protein structures under permissive conditions. Ordered secondary structures show high propensity for strong interactions in proteins as molten globule state of protein folding harbors significant stability. We hope to make use of these

consistent secondary structural elements of interaction in proteins to identify regions of short peptides that can self interact with the target protein.

The proposed development of a novel peptide-based platform technology that aims to produce ‘synthetic antibodies’ for a target protein, involves fishing self-binding sequences from a given protein by a peptide library of overlapping peptides derived from the given protein sequence, and probing for interactions using the same protein. Positive candidate peptides are to be synthesized with a probe to validate the interactions via pull down assay. Different formats of the peptides are also to be synthesized to increase the affinity of the peptides in pulling down the protein. The peptide probes can then be used for downstream functional applications or proteomics (Figure 16). To establish the technology of the proposed strategy, we use the SARS-CoV S protein as an example to discover rules that govern self interactions that could further our understanding of predictive rules for use in other systems. Here, we report the discovery of a series of self-binding peptides in SARS-CoV S protein.

2. Materials and Methods

2.1 Expression and purification of SARS-CoV S protein ectodomain (S1188HA)

The nucleotide fragments corresponding to amino acid residues 1 to 1188 of SARS-CoV S protein were amplified by PCR from the plasmid pJX40-S which contains the SARS-CoV spike gene (gift from Institute of Molecular Cell Biology, Singapore) and a HA tag was added to the C-terminus of S protein (S1188HA). Primers used were A (5'-AGTCGAATTCCGAACATGTTTATTTTCTTA-3') and B (5'-GCCCTCTAGATTAAGCGTAATCTGGAACATCGTATGGGTACATCTCGAGATA TTTTCCCAATTCTT-3'). An upstream EcoR I site and a downstream BamH I site as well as a stop codon preceding the BamH I site were designed to insert the fragments into transfer vector pVL1392 (BD Bioscience Pharmingen), under the control of a strong viral promoter, polyhedrin promoter. To generate recombinant baculovirus, the plasmid containing S1188HA was co-transfected with linearized baculovirus DNA (BaculoGold™, BD Biosciences) using Transfer Buffer A & B (BD Bioscience Pharmingen) into insect cells following the manufacturer's instructions. Successful recombinants were identified by detecting the S1188HA expression using Western blot. The virus was amplified for three to four rounds to prepare the viral stock. Insect cells were infected with the recombinant baculovirus at a multiplicity of infection (moi) of 3-10. At 4 days post-infection (dpi), cells were collected by a cell scraper (Costar) and then resuspended in lysis buffer (50mM Tris, pH7.5; 150mM NaCl; 0.1% Nonidet P40), containing complete protease inhibitor cocktail tablets (Roche Diagnostics GmbH, Roche Applied Science), lysed on ice for 30 min, then centrifuged at 4°C, 30000g for 30 min.

2.2 SPOT synthesis and screening of SARS-CoV S protein peptide library

MultiPep from Intavis AG was used to synthesize the peptide array on Amino-PEG500-UC540 membrane from Intavis AG. SPOT synthesis protocol was used as previously described (Frank and Overwin 1996, Frank 2002). Overlapping peptide sequences derived from SARS-CoV S protein were synthesized on the membrane, whereby each peptide has a constant number of amino acid shift from the previous. Prepared membrane was blocked for 2 hrs with SuperBlock® Blocking Buffer (Pierce, Rockford, IL). Buffer was removed, and non-purified S1188HA (diluted in the blocking buffer to achieve a buffer to S1188HA ratio of 1:100), was incubated with membrane for 1hr. Membrane was washed 3 times with 2M urea for 10min each followed by the incubation of rabbit anti-HA antibody for 1hr. Subsequently, the membrane was washed with 2M urea as described above, followed by the incubation of goat anti-rabbit antibody for 1hr. 2M urea wash was performed, and positive interaction of peptides to S1188HA was detected by ECL Western blotting detection kit.

2.3 Peptide synthesis of biotinylated peptides

All amino acids and coupling reagents were purchased from Novabiochem (San Diego, CA, USA). The 0.34 mmol/g substituted rink amide resin was purchased from Advanced ChemTech (Louisville, KY, USA). All solvents used were of the highest

commercial grade. Biotinylated peptides were prepared by Fmoc chemistry using 20% piperidine in DMF for 30 min at the deprotection step. Amino acids used during the coupling step were four times in excess. Final cleavage of all side chain protecting groups and the peptide from the resin is performed with cleavage mixture of 95% TFA, 2.5% water and 2.5% TIS.

2.4 Pull down assay

UltraLink® Immobilized NeutrAvidin™ Protein (Pierce, Rockford, IL) was loaded with 0.05mM of biotinylated peptides by incubating for 20 minutes. The beads were then washed with PBS to remove residual peptides. This was followed by blocking of the beads for 20 minutes using SuperBlock® Blocking Buffer (Pierce, Rockford, IL). The beads were washed with PBS before incubation with S ectodomain in 2M urea for 1 hour. After removal of unbound S protein, the beads were washed with PBST and boiled in SDS sample buffer before loading onto 8% SDS PAGE to run. The resulting gel was transferred onto PVDF membrane for Western Blotting.

2.5 Surface Plasmon Resonance

The binding kinetics and affinity of the candidate peptides to HA-purified SARS S protein S1188HA were analyzed by surface plasmon resonance (Biacore 3000, Uppsala, Sweden). The purified S1188HA was covalently immobilized to a CM5 sensor chip via amine group using the amine coupling kit (Biacore) in 10mM sodium

acetate buffer, pH 4.5. Experiments were run at a flow rate of 10 μ l/min in HBS-EP buffer (Biacore). The surface was regenerated with 10mM glycine-HCl, pH 2.0. Binding kinetics parameters were measured with different molar concentrations of analytes and evaluated with BIA-EVALUATION software (Biacore).

2.6 Immunofluorescence staining

293T cells were transfected with plasmid coding for SARS full-length S protein using Lipofectamine 2000. Cells were collected and fixed with 4% paraformaldehyde in phosphate-buffered saline (PBS) and permeabilized with 0.5% Triton X-100 for 5 minutes at room temperature. Primary antibodies in 4% bovine serum albumin in Tris-buffered saline (TBS) with 0.05% Tween 20 were incubated for 1 hour at room temperature or overnight at 4°C. Cells were washed three times and incubated with appropriate secondary antibodies for 1 hour at room temperature. Slides were mounted in Prolong Gold Anti-Fade medium with DAPI (Invitrogen). Images were collected and analyzed on an Axiovert 200M inverted fluorescence compound microscope (Carl Zeiss, Germany)

3. Results

3.1 Synthesis and screening of S protein peptide library

Peptide libraries consisting of 310 overlapping sequences of SARS CoV S protein was synthesized on a PEG-derivatized membrane using Fmoc chemistry. The peptides cover the full length of 1255 amino acids of S protein, with each peptide 20 amino acids in length and moving 4 amino acids downstream with each peptide (Figure 18). These peptides were anchored by their carboxyl terminal to the PEG membrane.

The peptide library was used to screen for self-binding peptides that bind to S protein in a procedure similar to Western blot detection. The screening was performed first by blocking the PEG membrane with SuperBlock® Blocking Buffer to prevent non-specific reactions. The membrane was subsequently incubated with S protein ectodomain (S1188HA) with an HA affinity tag at the carboxyl terminus. The peptides binding S protein were detected by rabbit anti-HA antibody followed by anti-rabbit HRP secondary antibody, and finally detection by chemiluminescence (Figure 19).

The result of the screening is shown in Figure 20, where the spots indicate that peptide sequence of that spot shows self-binding to S protein ectodomain. The peptide sequence of the spots can be retrieved based on their coordinates on the membrane. The result is further tabulated and analyzed for their secondary structures, hydrophobicity and presence of glycosylation sites.

Overlapping peptides of full length S protein

Figure 18. Peptide library of SARS CoV S protein synthesized on the membrane for screening with S protein ectodomain (S1188HA). Peptides covering the full 1255 amino acid length of S protein, with each peptide 20 amino acids in length and moving 4 amino acids downstream with each peptide were synthesized on PEG derivatized membrane with their C terminal anchored to the membrane.

Figure 19. Screening of peptide library with S protein ectodomain (S1188HA)

Peptide library on the membrane was made by SPOT synthesis. Screening of the peptide library was performed by first incubating with S protein ectodomain with an HA affinity tag at the C terminal (S1188HA). The S protein that remains bound to peptides on the membrane was detected by rabbit anti-HA antibody followed by anti-rabbit HRP secondary antibody and finally chemiluminescence detection.

Figure 20. Immobilized peptide arrays with SARS CoV S protein ectodomain. Film was developed after exposure to the membrane that was treated with chemiluminescence reagents. The spots indicate sequence of self binding to S protein ectodomain (S1188HA). The peptide sequence can be retrieved based on their coordinates on the membrane.

3.2 Tabulation and analysis of results from screening

Based on their coordinates on the membrane, the peptide sequences of spots showing self-binding to S protein ectodomain (S1188HA) were deduced. Their amino acid sequence and positions on the S protein are tabulated in Table 8. These peptide sequences were further analyzed to predict their secondary structure, hydrophobicity and presence of potential glycosylation sites.

For prediction of secondary structure, the full length S protein sequence of SARS CoV was entered into the PSIPRED program to identify secondary structures. The PSIPRED results indicated α helices, β -strands or random coil based on the predictive program (Jones *et al.* 1999). The hydrophobicity of the peptides was calculated by adding the hydrophobicity values of individual amino acids (Black *et al.*, 1991) in the sequence, and obtaining the average hydrophobicity value for each sequence. Using a cutoff value of 0.5, sequences with hydrophobicity values of greater than 0.5 were predicted to be hydrophobic in the analysis.

Glycosylation sites within the S protein were predicted using the NetNGlyc 1.0 server and sequences containing glycosylation sites were indicated with the position of asparagine (N) where the glycosylation was predicted. The results of the analysis of the peptides were tabulated in Table 9.

Table 8. Tabulation of peptide sequences from screening of S peptide library

Spot Number	Amino acid position	Sequence	Relative Intensities
A6	21 – 40	TFDDVQAPNYTQHTSSMRGV	3.87
A8	29 – 48	NYTQHTSSMRGVYYPDEIFR	14.88
C6	213 – 232	FNTLKPIFKLPLGINITNFR	19.65
C7	217 – 236	KPIFKLPLGINITNFRAILT	2.99
C8	221 – 240	KLPLGINITNFRAILTAESP	1.27
D1	289 – 308	SVKSFEIDKGIYQTSNFRVV	0.59
D4	301 – 320	QTSNFRVPSGDVVRFPNIT	23.07
D14	341 – 360	ERKKISNCVADYSVLYNSTF	3.72
D23	377 – 396	LCFSNVYADSFVVKGDDVRQ	11.15
D24	381 – 400	NVYADSFVVKGDDVRQIAPG	1.05
E7	409 – 428	NYKLPDDEFMGCVLAWNTRNI	2.35
E11	425 – 444	TRNIDATSTGNYNKYRYLR	12.34
E13	433 – 452	TGNYNKYRYLRHGKLRPFE	2.82
E14	437 – 456	NYKYRYLRHGKLRPFERDIS	0.64
E24	477 – 496	PLNDYGFTTTTGIGYQPYRV	11.12
F12	525 – 544	VNFNFNGLTGTGVLTPSSKR	2.61
F13	529 – 548	FNGLTGTGVLTPSSKRFQPF	4.74
F14	533 – 552	TGTGVLTPSSKRFQPFQQFG	14.88
F15	537 – 556	VLTTPSSKRFQPFQQFGRDVS	4.20
G15	633 – 652	AGCLIGAEHVDTSYECDIPI	11.41
G19	649 – 668	DIPGAGICASYHTVSLRS	24.65
H13	721 – 740	NMYICGDSTECANLLLQYGS	3.43
H18	741 – 760	FCTQLNRALSGIAAEQDRNT	17.42
J2	869 – 888	TFGAGAALQIPFAMQMAYRF	5.76
J3	873 – 892	GAALQIPFAMQMAYRFNGIG	11.48
J5	881 – 900	AMQMAYRFNGIGVTQNVLYE	2.56
K2	965 – 984	RLDKVEAEVQIDRLITGRLQ	34.99
K3	969 – 988	VEAEVQIDRLITGRLQSLQT	29.35
K4	973 – 992	VQIDRLITGRLQSLQTYVTQ	5.83
K5	977 – 996	RLITGRLQSLQTYVTQQLIR	13.69
L5	1073 – 1092	REGVFVFNGTSWFITQRNFF	4.35
L6	1077 – 1096	FVFNGTSWFITQRNFFSPQI	2.34
L8	1085 – 1104	FITQRNFFSPQIITDNTFV	1.68
L12	1101 – 1120	NTEVSGNCDVVIGIINNTVY	2.19

Receptor Binding Domain

HR1 Domain

Fusion Peptide

Loop between HR1 and HR2

*Intensities of spots were measured by densitometry relative to Spot O20 (negative control for background)

Table 9. Analysis of peptide sequences from screening of S peptide library

Spot	Start	End	Secondary str prediction	Hydrophobicity value	Glycosylation sites	Region
A6	21	40	β /RC	0.45615	N29	
A8	29	48	β	0.48495	N29	
C6	213	232	β	0.63775	N227	
C7	217	236	β	0.6539	N227	
C8	221	240	β /RC	0.64135	N227	
D1	289	308	β /RC	0.51965		
D4	301	320	β	0.52675	N318	RBD
D14	341	360	β /RC	0.5114	N357	RBD
D23	377	396	β	0.5246		RBD
D24	381	400	β	0.51405		RBD
E7	409	428	β /RC	0.55775		RBD
E11	425	444	RC	0.46255		RBD
E13	433	452	RC	0.49075		RBD
E14	437	456	RC	0.4539		RBD
E24	477	496	β	0.60105		RBD
F12	525	544	β	0.54045		
F13	529	548	β	0.5737		
F14	533	552	β	0.53985		
F15	537	556	β / α	0.50535		
G15	633	652	β	0.5449		
G19	649	668	β	0.58825		
H13	721	740	α	0.5605		
H18	741	760	α	0.4581		
J2	869	888	α	0.64935		FP
J3	873	892	α	0.63005		FP
J5	881	900	α	0.57365		FP
K2	965	984	α	0.44295		HR1
K3	969	988	α	0.4804		HR1
K4	973	992	α	0.52435		HR1
K5	977	996	α	0.5355		HR1
L5	1073	1092	β	0.5749	N1080	Loop
L6	1077	1096	β	0.61965	N1080	Loop
L8	1085	1104	β /RC	0.5763		Loop
L12	1101	1120	β	0.5961	N1116	Loop

Positive peptide sequences binding to S protein were analyzed for their secondary structures, hydrophobicity and presence of glycosylation sites. α refers to α helix; β refers to β strands; RC refers to random coil; RBD is receptor binding domain, FP is fusion peptide; HR 1 is heptad repeat 1; Loop refers to the loop region between HR1 and HR2.

3.3 Validation by *S* protein pull down assay

To validate the results of the immobilized peptide library, sequences that were positive for strong binding to *S* protein were synthesized for use in pull down assay (Table 10). These peptides were synthesized by Fmoc chemistry as carboxamides and biotinylated at the C terminal through incorporation of Lys-biotin, an orientation similar to the peptide scan allowing their C terminals to be anchored by avidin beads. The peptides were analyzed by HPLC and confirmed by molecular weight determination using MALDI TOF mass spectrometry. The biotinylated peptides were used to pull down *S* protein ectodomain using avidin beads. First, the biotinylated peptides were incubated with the NeutrAvidin beads so as to couple the peptides to the beads due to the high affinity interaction (K_d of 10^{-15} M) between biotin and avidin. This was followed by blocking of the beads with a commercial blocking buffer to prevent non-specific interactions between the *S* proteins and the avidin beads. Subsequently, the peptide-coupled beads were incubated with *S* protein ectodomain under 2M urea binding condition. This is an optimized procedure because the absence of 2M urea in the binding condition produced non-specific binding of *S* protein to the beads, hindering correct interpretation of the pull down results. After incubation with *S* protein ectodomain, the beads were washed stringently before boiling with SDS sample buffer to dissociate the pulled down *S* protein for SDS PAGE analysis (Figure 21).

Figure 22 shows the pull down assay using different peptides selected from screening of the peptide library. Consistent with the pepscan results, positive peptides were able to pull down *S* protein ectodomain (S1188HA) from insect cell lysate. However, it can

Table 10. Peptides from SPOT screening results for synthesis

Spot Number	Sequence Position	Amino Acid Sequence
A8	29 – 48	NYTQHTSSMRGVYYPDEIFR
C6	213 – 232	FNTLKPIFKLPLGINITNFR
D4*	302 – 321	TSNFRVVPBGDVVRFPNITN
D14	341 – 360	ERKKISNVCVADYSVLNSTF
D24	381 – 400	NVYADSFVVKGDDVRQIAPG
E11	425 – 444	TRNIDATSTGNYNKYRYLR
E24	477 – 496	PLNDYGFYTTTGIGYQPYRV
F14	533 – 552	TGTGVLTPSSKRFQPFQQFG
G15	633 – 652	AGCLIGAEHVDTSYECDIPI
H18*	743 – 762	TQLNRALSGIAAEQDRNTRE
J2	869 – 888	TFGAGAALQIPFAMQMAYRF
J5	881 – 900	AMQMAYRFNGIGVTQNVLYE
K2*	964 – 983	SRLDKVEAEVQIDRLITGRL
K5	977 – 996	RLITGRLQSLQTYVTQQLIR
L5	1073 – 1092	REGVFVFNGTSWFITQRNFF
L6	1077 – 1096	FVFNGTSWFITQRNFFSPQI

Asterisks (*) indicates that original sequences were shifted 1 or 2 amino acids upstream or downstream to minimize difficulties in synthesis and subsequent purification. All sequences were kept at 20 amino acids in length and synthesized as carboxamides with biotin at the N terminal.

Figure 21. Schematic of S protein pull down assay

Biotinylated peptides are first coupled to the avidin beads via biotin-avidin interactions. Subsequently, the peptide-coupled beads were incubated with S protein ectodomain (S1188HA) under 2M urea binding condition. The beads were washed stringently before boiling with SDS sample buffer to dissociate the pulled down S protein for SDS PAGE analysis.

Figure 22. Pull down of S ectodomain (S1188HA) using biotinylated peptides chosen from screening peptide array.

Biotinylated peptides were loaded onto NeutrAvidin beads and the peptide loaded beads were then used to pull down S1188HA from Sf9 insect cell lysate. Final pull down S1188HA were analyzed on SDS PAGE and detected on Western Blot using anti-HA antibodies. The estimated size of S1188HA is ~200kDa.

be observed that the strength of binding of the different peptides varies. The pull down ability of S1188HA by the soluble peptide-biotin construct and its corresponding counterpart used in the SPOT screen does not always correlate. A case in point is H18 versus J2. H18 and J2 showed weak intensity in the pull down assay in Figure 22 but H18 had stronger intensity in the screening library in Figure 20. Inversely, J2 showed a correspondingly weak spot (Figure 20). H18 is a soluble peptide when used for pull down, but H18 may exist as multimer on the SPOT membrane, thus having the ‘enhanced avidity’ effect, although synthesized as monomer. Therefore, this further illustrated the importance of incorporating the additional pull down step to validate the positive interactions as in this case.

3.4 Additive effects of peptides through noncovalent heterovalency

To improve the affinity of peptides to pull down S1188HA, we looked at the effects of mixing peptides that binds to different regions of S protein. Peptides that bind to different binding sites on S protein were mixed and loaded to the avidin beads before performing the S protein pull-down assay. The results of pull-down assay (Figure 23) show that E11 and E24 did not exert any effects of enhancing pull down of S protein, and dilution effect was observed. Additive effects of two different peptides on S protein pull down were observed when using G15 and H18. These peptides when used individually did not display a strong affinity to pull down S protein. However, when used together, the mixture was able to pull down a substantially higher amount of S protein (Figure 23). Loading of both peptides onto the beads were confirmed using MALDI TOF mass spectrometry.

To further demonstrate the additive effects of G15 and H18, varying molar ratio of G15 to H18 was performed on the pull down assay. It was shown that the highest amount of S protein being pulled down was by an equal ratio of G15 and H18 (Figure 24a). The investigation on additive effects of peptides was extended to the other positive peptides. Another pair of peptides that could enhance the pull down affinity was found to be: C6 and D4 (Figure 24b). In summary, there were two pairs of peptides found to act additively to enhance the binding affinity to S protein in pull down experiments. They were G15/ H18 and C6/ D4.

3.5 Homodimer increases pull down affinity to S protein

For the homodimer studies, a Lys-based linker was coupled to the resin before synthesizing the respective monomer sequence which will yield homodimers after final cleavage from resin. However, we could only obtain C6 homodimers successfully despite several attempts. MALDI-TOF mass spectrometry of other final homodimer products did not yield the expected mass, with most of them having truncations of amino acids. We proceeded to examine the ability of C6 homodimer in pulling down S protein compared to C6 monomer. Pull down assay was carried out as mentioned before, and it was observed that the C6 homodimer produced a much darker S protein band on Western blot than C6 monomer. The results of pull down assay (Figure 25) showed that C6 homodimer enhanced the pull down of S protein more than that of C6 monomer.

Figure 23. Pull down of S ectodomain (S1188HA) using combination of biotinylated peptides.

Biotinylated peptides were pre-mixed and loaded onto NeutrAvidin beads equally and the peptide loaded beads were then used to pull down S1188HA from Sf9 insect cell lysate. Final pull down S1188HA were analyzed on SDS PAGE and detected on Western Blot using anti-HA antibodies. The estimated size of S1188HA is ~200kDa.

Figure 24a. Pull down of S ectodomain (S1188HA) using varying concentrations of G15 and H18 peptides.

Biotinylated G15 and H18 peptides were pre-mixed and loaded onto NeutrAvidin beads equally and the peptide loaded beads were then used to pull down S1188HA from Sf9 insect cell lysate. Final pull down S1188HA were analyzed on SDS PAGE and detected on Western Blot using anti-HA antibodies.

Figure 24b. Pull down of S ectodomain (S1188HA) using a combination of biotinylated peptides showing additive effects of C6 and D4.

Biotinylated peptides were pre-mixed and loaded onto NeutrAvidin beads equally and the peptide loaded beads were then used to pull down S1188HA from Sf9 insect cell lysate. Final pull down S1188HA were analyzed on SDS PAGE and detected on Western Blot using anti-HA antibodies. The estimated size of S1188HA is ~200kDa.

Figure 25. Pull down of S1188HA using biotinylated C6 monomer and C6 dimer peptides. Biotinylated C6 monomer and C6 dimer peptides were loaded onto NeutrAvidin beads. The peptide loaded beads were then used to pull down S1188HA from Sf9 insect cell lysate. Final pull down S1188HA were analyzed on SDS PAGE and detected on Western Blot using anti-HA antibodies.

Figure 26. C6 dimers aggregate to form fibril-like structures. Aggregation from C6 dimers were mounted on glass slides and viewed under 10x (left panel), 40x (middle panel) and 100x (right panel) magnification.

3.6 Determination of peptide binding affinity using surface plasmon resonance

We made use of BIAcore to perform the surface plasmon resonance experiments. The molecular interactions were studied by immobilizing one of the reactants (ligand) on the sensor surface while the other is being passed over that surface in solution. When interaction occurs, the response is recorded in resonance units (RU) and is proportional to the molecular mass on the surface. In our study, we used BIAcore to study the binding kinetics of two peptides A8 and C6. We initially used S protein as ligands with the peptides as analyte and later the respective peptides as ligands with S protein as the analyte. Data obtained from different concentrations were used to obtain the dissociation constant for the binding between the peptides and S protein. Both A8 and C6 have an average dissociation constant of 10^{-6} (Figure 27-29). However, it was noted that at concentrations higher than 20 μ M, A8 seems to show signs of aggregation.

C6 dimer binding to S protein was also examined by surface plasmon resonance. Binding was observed (Figure 30a) but we were unable to obtain quantitative data on its dissociation constant as aggregation appeared to occur at concentrations greater than 2 μ M (Figure 30b). Comparison with C6 monomer showed that C6 dimer was able to bind with higher affinity as there was an approximately five fold increase in binding, in terms of RU, when 1 μ M of peptides were used to perform binding to S protein.

Figure 27. Surface plasmon resonance of binding between purified S protein and biotinylated A8 peptide. Graph of response against concentrations for determination of dissociation constant (K_D) for binding of S protein (analyte) to immobilized biotinylated A8 peptide. The data points are obtained from combined sensograms for binding of S protein to A8 peptide at S protein concentrations of 0.125μM, 0.25μM, 0.5μM, 1.0μM and 2.0μM (shown in bottom right panel).

Figure 28. Surface plasmon resonance of binding between purified S protein and biotinylated A8 peptide. Graph of response against concentrations for determination of dissociation constant (K_D) for binding of S protein (analyte) to immobilized biotinylated A8 peptide. The data points are obtained from combined sensograms for binding of S protein to A8 peptide at S protein concentrations of 0 μM , 1.25 μM , 2.5 μM , 5.0 μM and 10.0 μM (shown in bottom right panel).

Figure 29. Surface plasmon resonance of binding between purified S protein and biotinylated C6 peptide. Graph of response against concentrations for determination of dissociation constant (K_D) for binding of S protein (analyte) to immobilized biotinylated C6 peptide. The data points are obtained from combined sensograms for binding of S protein to C6 peptide at S protein concentrations of 0μM, 1.25μM, 2.5μM, 5.0μM and 10.0μM (shown in bottom right panel).

Figure 30. Surface plasmon resonance of binding between purified S protein and biotinylated C6 dimer. (a) Sensogram of 1 μ M C6 dimer binding to S protein. (b) Combined sensograms for binding of C6 dimer to immobilized S protein.

3.7 C6 homodimers aggregates to form fibril-like structures

It was found that C6 homodimers formed visible fibre-like aggregates after storage at 4°C. The fibrous particles were removed from solution, mounted on slides to view under high power microscope, and the images revealed fibril-like structures (Figure 26). Such aggregates visible to the naked eye were not observed to be formed for the C6 monomer and the other monomeric peptides.

3.8 Functional application: Comparison of conventional Ab vs. synthetic Ab from their usage in immunofluorescence assay of expressed S protein

Self interaction between peptide and protein can be exploited such as pull-down and protein purification and other applications. These applications are similar to a certain extent to “synthetic antibody” for a lab reagent. To compare whether the “synthetic antibody” can indeed perform like a conventional antibody, we used it to perform immunofluorescent staining on expressed full-length S protein on 293T cells, alongside anti-SARS S antibodies. After transfected 293T cells were fixed with 4% paraformaldehyde and permeabilized with Triton X-100, the cells were stained with either anti-S antibodies or biotinylated peptide followed by their respective secondary antibodies accordingly. Confocal microscopy shows that C6 peptide was able to effectively stain the full-length S protein on 293T cells, with comparable intensities as S protein on cells stained by conventional anti-S antibodies (Figure 31). A8 peptide could also stain expressed S protein, albeit with lower fluorescence than that observed in conventional anti-S antibody stained cells (Figure 32). The results showed that C6

and A8 peptides could be use to stain expressed S protein by immunofluorescence similar to a conventional antibody.

Figure 31. Confocal imaging of immunofluorescent stained 293T cells expressing full length SARS S protein. 293T cells were transfected to express full length S protein using Lipofectamine 2000. After 24 hours, transfected 293T cells fixed with 4% paraformaldehyde and permeabilized with Triton X-100. The cells were stained with either antibodies or biotinylated C6 peptide followed by their respective secondary antibodies accordingly. The cells were then mounted on glass slides using ProLong Gold AntiFade with DAPI.

Figure 32. Confocal imaging of immunofluorescent stained 293T cells expressing full length SARS S protein. The cells were stained with either antibodies or biotinylated A8 peptide followed by their respective secondary antibodies accordingly. The cells were then mounted on glass slides using ProLong Gold AntiFade with DAPI. Cells labeled with NeutrAvidin Dylight 488 only (top left panel), cells labeled with D-biotin and NeutrAvidin Dylight 488 (top right), cells labeled with biotinylated A8 peptide and NeutrAvidin Dylight 488 (bottom left) and cells labeled with anti-SARS mAb and anti-mouse FITC antibodies (bottom right).

4. Discussion

To establish the technology of the proposed strategy, we use the SARS-CoV S protein as an example to discover rules that govern self interactions that could further our understanding of predictive rules for use in other systems. Using an insect cell expressed HA-tagged S protein ectodomain (S1188HA) to screen the S protein peptide library that consisted of 310 peptides of 20 amino acids long and a frameshift of 4 amino acids with each sequence, 34 peptides were found to interact with S protein (Table 8). These peptides mapped to RBD, fusion peptide, HR1 domain and loop region between HR1 and HR2. Some peptides were also found within previously uncharacterized regions. Their properties were further analyzed to determine rules governing self-binding peptide sequences in proteins (Table 9).

From the analysis in Table 9, all peptides with three exceptions were found to have distinct ordered secondary structure elements. Almost all peptides exhibit high propensity to α -helical, β -stranded or both secondary structural elements. Analysis of their hydrophobicity revealed that about 80% of the sequences were hydrophobic. Also, 10 of the peptides contain N-glycosylation sites, suggesting binding to carbohydrates. We propose that self-binding sequences can be categorized into four general groups: (i) complementary binding sequences, (ii) self-binding β -strand sequences, (iii) hydrophobic interacting sequences, and (iv) carbohydrate-binding sequences.

Group 1 - Complementary binding α -helical peptides

The complementary binding sequences, bind to a specific domain and is best

exemplified by HR1 domain binding to HR2 domain in S protein to form 6-helix bundle. They interact through a Leu/ Ile zipper type and is very well-characterized in Class I virus glycoprotein. These sequences tend to be predominantly α -helical in nature. A diagrammatic representation of HR1 and HR2 interaction is shown on Figure 33 (Liu *et.al.*, 2004).

Group 2 - Self-binding β -stranded peptides

For self-binding sequences, they refer to the given amino acid sequence binding back to the same amino acid sequence. This type of interaction is supported by Sawaya *et al.* (2007), who found that self-binding sequences cause aggregation. The mechanism is attributed to a pair of β -sheets, with the facing side chains of the two sheets interdigitated in a dry 'steric zipper'. This common structure is categorized into 8 classes (Figure 34), and can be found in amyloid- β , tau, PrP prion, insulin, islet amyloid polypeptide (IAPP), lysozyme, myoglobin, α -synuclein and $\beta(2)$ -microglobulin proteins (Sawaya *et al.* 2007). This type of self-binding sequences may also explain the peptides within the RBD that are able to bind to S protein, as the RBD is made up of 7 anti-parallel β -sheets. These first 2 groups of self-binding sequences possess ordered secondary structures, which agreed with the concept of molten globule of protein folding central to our proposed strategy.

Group 3 - Hydrophobic self-interacting peptides

The hydrophobic interacting sequences refer to sequences that are hydrophobic in nature and interact with other hydrophobic segments. This can explain why fusion peptide region can bind back S protein, as the fusion peptide composition is well known to be hydrophobic and rich in Ala and Gly.

Figure 33. Interaction between NP-1 (HR1) and CP-1 (HR2) of SARS CoV spike protein as predicted by molecular modeling

A homology model of NP-1 (HR1) and CP-1 (HR2) association was built on the basis of the X-ray crystal structure of HIV-1 gp41 core formed by gp41 N-peptides and C-peptides. NP-1 showed a typical coiled-coil structure. Most of CP-1 had an α -helical conformation with the exception of a kink in the middle of the structure. (Taken from Liu *et al.* 2004)

Figure 34. The eight classes of steric zippers.

Two identical sheets can be classified by: the orientation of their faces (either ‘face-to-face’ or ‘face-to-back’), the orientation of their strands (with both sheets having the same edge of the strand ‘up’, or one ‘up’ and the other ‘down’), and whether the strands within the sheets are parallel or antiparallel. Both side views (left) and top views (right) show which of the six residues of the segment point into the zipper and which point outward. Green arrows show two-fold screw axes, and yellow arrows show translational symmetry. Below each class are listed protein segments that belong to that class. (Taken from Sawaya *et al.* 2007)

Group 4 - Carbohydrate binding peptides

The carbohydrate-binding sequences refer to sequences that interact with carbohydrate residues. These sequences contain glycosylation sites and will normally be glycan shielded in S protein, which prevents self interaction by our synthetic peptide probes. Carbohydrate-binding peptides have literature precedents (Yamamoto *et al.* 1992, Konami *et al.* 1992, Roberts and Goldstein 1984). They have been isolated and characterized and there appears to be a close relationship between carbohydrate-binding specificity and the sequence of the carbohydrate-binding peptides (Konami *et al.* 1992, Roberts and Goldstein 1984). These two groups (Groups 3 and 4) of interacting peptides, differ from the first two groups (Groups 1 and 2), and tend to be non-specific in nature and are not solvent exposed. Thus, synthetic peptides from Groups 1 and 2 may prove to be more useful.

It is interesting to note that the peptides found to be positive in the screening may contain sites of trimerization as the S protein is a trimeric structure on the virus surface.

When the biotinylated peptides were synthesized and used for S protein pull down assay, varying affinity of the peptides to S protein were observed. In order to improve the avidity of the peptides to S protein, (i) we used different peptides that bind to two different regions on the S protein, and (ii) we synthesized the biotinylated peptides in dimeric form. Indeed, as shown in the results (Figures 23 and 24), synergistic effects could be seen when certain peptides were mixed. The two pairs of peptides that showed synergistic effects were analyzed, and each pair was about a hundred over amino acids apart from each other. If the sequences bind back to their own sequence,

these could suggest a distal effect allowing the mixed peptides to bind and enhance avidity, as peptide pair distal by primary amino acid structure may be close together if one considers the conformational structure of their native protein.

To improve the avidity of the peptides, dimers of the peptides used in this study were synthesized. The peptides were linked to a specially designed linker that has almost equal atomic distances on the both arms (Figure 35). C6 dimer was able to enhance pull down of more S protein compared to C6 monomer. Quantitative data on the affinity of peptide binding to S protein and their respective kinetics obtained via surface plasmon resonance places the affinity of the tested peptides and S protein at μM range. BIAcore data further confirms C6 dimer's enhanced binding affinity to S protein compared to C6 monomer by approximately five-fold. Further optimization of candidate peptides should be able to attain peptides with binding affinity in sub-micromolar to nanomolar range.

It is interesting to note the difficulties encountered when synthesizing dimeric versions of the candidate peptides. Only C6 dimer was successfully made after several attempts. C6 dimer was observed to aggregate to form fibril-like structures, and this could imply aggregation as the main problem during synthesis of dimer which prevented accurate assembly and formation of the final dimeric products. This is further supported by our observation for A8 where aggregation sets in during BIAcore experiments at high concentrations ($>20\mu\text{M}$) while C6 dimers also showed signs of aggregation in BIAcore at lower concentrations ($>2\mu\text{M}$). This is especially so during synthesis where elongating peptide chains exist at high concentration on the resins. Thus, the aggregation problem may be further aggravated resulting in ultimate failure

Figure 35. Structure of linker for synthesis of peptide homodimer

Peptides dimers will be synthesized on this biotinylated linker. β -alanine is linked to lysine to create a symmetrical linker with equivalent spacing on both arms to which peptides are attached.

in synthesis of the dimers. The fibril-like structures observed in C6 dimer solution resemble amyloid fibril formation. This could be due to the dimeric C6 priming and speeding up the aggregation process to form fibrils. As such, C6 peptide (not particularly hydrophobic in nature) has self-interacting ability, and this could imply that the C6 peptide binds to the same C6 sequence on S protein. For future work to experimentally demonstrate that C6 peptide binds to the same C6 sequence on the S protein, C6 sequence can be mutated or deleted on S protein and subsequently using it to perform pull down experiments.

In this study, we carried out preliminary trials using some candidate peptides for downstream application as lab reagents. We showed that it is possible to use them as antibody substitute for immunofluorescent staining of the target protein with equivalent efficiencies. They can also be used for other downstream applications as probes, substitutes for antibodies or even purification of proteins. However, it is important to note that the development of a “synthetic antibody” based on self-interacting peptides is novel and contrary to current thinking. Thus, it is crucial to extend the new technology to other systems to test the robustness of this concept. Candidate systems include other Class I virus fusion glycoprotein, Class II virus fusion glycoprotein and other soluble cytoplasmic proteins. It is also important to exclude the possibility that the peptides may bind to host proteins that may give rise to undesirable non-specific reactions.

Chapter 5

Conclusions and Perspectives

1. Studies on MPER of SARS CoV S protein

In recent years, we saw the emergence of novel infectious diseases such as SARS, as well as the re-emergence of other infectious diseases, such as H1N1 and H5N1 influenza viruses. Although the SARS epidemic was brought under control, the threat of SARS to the global community remains as laboratory-acquired SARS infections were reported (WHO update 7, Apr 2004) and its potential use in biological warfare is worrisome. The re-emergence of H1N1 influenza has developed into a global pandemic. Drawing experiences from HIV and influenza, there have been serious concerns on the effectiveness of therapeutic prophylactics such as blockers of proton channels and neuraminidase leading to the rapid emergence of mutant strains. Therefore, there is an urgent demand for effective prophylactic agents and inhibitors. This study was conceived during the SARS period and continued into the current H1N1 pandemic.

The importance of MPER in SARS S protein mediated infectivity and its feasibility as an entry inhibitor target was examined in my thesis. Characterizing the quaternary protein mimetics of MPER provides further understanding of the different intermediate states of fusion, allowing further insights into the SARS fusion mechanism and aiding fusion inhibitor design. By targeting such conserved functional regions on S protein critical for orchestrating highly conserved fusion events,

potentially mutation-resistant SARS entry inhibitor candidates that directly block virus fusion were developed. With the findings from this thesis, I hope to lay foundation and pioneer the development of a novel class of SARS CoV peptidic-based prophylactics that are low cost, orally available, have long shelf life and amenable to large-scale production.

Enveloped viruses enter host cells by membrane fusion. Classical examples are the Class I enveloped viruses such as influenza virus, which attaches to the host membrane before being taken up by receptor-mediated endocytosis. The influenza virus fusion protein, hemagglutinin (HA), then undergoes large conformational changes when endosomal acidification causes the pH in endosome to decrease to around 5, initiating the fusion of viral and host membranes. In some HA-mediated fusion models, it is suggested that interactions between transmembrane and fusion domains may facilitate late steps of membrane fusion (Tamm 2003). This fusion mechanism is highly conserved among the Class I viruses (eg. SARS CoV, HIV and influenza). Similar to other Class I enveloped viruses, SARS CoV entry into host cell proceeds in a similar manner, via the endocytic pathway, although it has been revealed recently that SARS CoV may involve a novel clathrin- and caveolae-independent endocytic pathway (Wang *et al.* 2008).

The importance of Trp residues of MPER in SARS CoV infectivity and virus fusion was confirmed in this thesis as shown by results described in Chapter 2. It was found that the importance of Trp in MPER function can be attributed to two aspects: (i) residual importance and (ii) positional importance. The residue identity of Trp in MPER is critical due to the indole ring effect that can be activated upon endosomal

acidification to proceed with fusion events, which cannot be replaced with other aromatic residues. Positional importance of Trp in MPER was also found to be crucial in MPER function. From the results obtained, a premature increase in helical structure of the W1194A mutant causes the MPER to overshoot the fusion stage prior to meeting with the host membrane, rendering it fusion ineffective. This suggests that increase in helicity may be a key conformational change in S protein-mediated fusion. The virus fusion protein must have the flexibility to take on different conformations during fusion through different intermediate states. In summary, we believe that the key to the importance of Trp in MPER is its contribution to the plasticity of the fusion protein structure during fusion events.

Oligomeric protein mimetics of MPER were also designed to mimic the various fusion states of MPER, so as to understand the different intermediate states of fusion and to allow further insights into the SARS fusion mechanism. We found that NT dimer mimics an intermediate fusion state where only partially exposed surfaces of MPER are available for interactions. On the other hand, CT dimer mimics a fusogenic intermediate state during fusion where all the available surfaces of MPER are exposed for interactions. A summary of the activities of the protein mimetics were listed in Table 7 of Chapter 3.

Our studies support a current model for the hypothetical role of MPER during fusion, which suggests that when the S protein undergoes conformational changes whereby HR2 folds back to HR1 to form the 6HB, it not only pulls both membranes close to each other in preparation for fusion but also brings the two hydrophobic regions of FP and MPER to close proximity, sandwiched between the virus and host membrane

(Figure 36). This may allow formation of a continuous track of hydrophobic surfaces that provide a low energy barrier path for disrupted lipids to flow, allowing the fusion pore to form. Transfer of viral genetic material to the host cell for replication and successful infection then occurs via this fusion pore. However, in this thesis, I have attempted to further dissect the fusion kinetics of this process by providing additional details of the fusion mechanism at a molecular level. We postulate that the pre-fusion state S protein (before receptor binding) is a trimer with its MPER held closely together via hydrophobic interactions, in the initial fusion-inactive ‘locked’ state. Its fusion potential can only be unleashed when fusion is activated via receptor binding, which triggers conformational changes that permits it to progress to the next intermediate state. As the endocytosed virus undergoes acidification in the endosome, pH-induced conformational changes may result in increased lipidic environment surrounding the protein, causing the indole N atom on Trp in MPER to become charged by protonation. As the MPERs become more positively charged (under lipidic environment), charge repulsion occurs between MPERs which dissociates the trimeric structure. The key is in the Trp residues of MPER that contributes flexibility to the structure to allow conformational changes. This fusion-active ‘opened’ intermediate state, having all surfaces exposed for interaction, then allows MPER to exert its membrane destabilizing potential on the host membrane before going into the post-fusion non-fusogenic ‘locked’ conformation that is more α -helical than pre-fusion, and ‘locked’ from reverting back to the fusion-active form again.

Previous findings by our laboratory (Lu *et al.* 2008) confirm the involvement of receptor-associated lipid rafts serving as an entry portal for SARS CoV entry into host cells. It was also found that clustering of ACE2 receptors on host membrane may be

Figure 36. Conformational changes of S protein during SARS CoV fusion events.

(a) Before fusion is initiated, the virus fusion protein exists in a metastable conformation as a trimeric complex on the virus membrane, also known as the prefusion state. (b) Upon receptor binding on the host membrane and exposure to low pH found in endosomes after the virus is endocytosed, fusion is activated allowing conformational changes that expose the hydrophobic FP to insert into host membrane. (c) The virus fusion protein continues to undergo changes in its structure, pulling the host and virus membrane closer in the process, in preparation for membrane fusion to occur. (d) As HR1 and HR2 moves and folds back to each other to form a hairpin structure known as six helix bundle (6HB), this brings FP and MPER to close proximity, that may serve as a continuous hydrophobic track for mixing of lipids, initiating formation of fusion pore.

complemented by clustering of S proteins on viral surfaces as observed by cryoEM in HIV and SIV. From the notion of this information, together with findings from this thesis, we further postulate that MPER and the TM region may serve to anchor and cluster the S proteins on viral surface, within lipid rafts enriched regions, facilitating binding to clustered ACE2 receptors on host membranes for triggering the virus fusion process.

The 6HB-mediated membrane fusion is common in all Class I envelope virus including SARS CoV, HIV, influenza, Ebola, etc. Therefore, it was our plan initially to develop mutation resistant SARS CoV antivirals by exploiting the conserved virus fusion mechanism. This conserved fusion mechanism involves the interplay of several functional elements such as the FP, 6HB formation by HR1 and HR2, MPER, etc, to synchronize and perform the virus fusion event. The importance and essential role of MPER has been demonstrated in this thesis, and serves as a potential target to inhibit this critical fusion event. Any mutations that escape this inhibitory mechanism would then be lethal to the virus as mutations will directly prevent the fusion event from happening, resulting in unsuccessful infection.

The key design of this novel protein mimetic approach for antivirals against MPER is the multivalent parallel peptide chains which mimic the multivalent interactions of MPER. Such quaternary protein mimetics has protein-like properties and are druggable. Studies from our laboratory shows that such quaternary protein mimetics are resistant to proteolytic degradations, have prolonged stability in serum and able to overcome aggregation. The ability to overcome the disadvantages reported so far by monomeric peptide inhibitors makes our compound an attractive candidate for further

antiviral development. Future work should focus on improving the lead compound, more *in vivo* and clinical studies for development as an antiviral against SARS CoV.

2. Quaternary protein mimetics as ‘Synthetic Antibodies’

Antibodies, also known as immunoglobulins (Litman *et al.* 1993), are gamma globulin proteins found in blood and bodily fluids of vertebrates, produced by plasma cells for recognition and neutralization of antigens foreign to the host (eg. viruses, bacteria, etc) by the immune system. Generally, a single unit of antibody is made up of two large heavy chains and two small light chains, which can assemble to form the various isotypes. The antigen binding fragment (Fab) contains six hypervariable complementarity-determining regions (CDRs) (Kabat *et al.* 1987) that present a large contiguous surface for antigen recognition (Amit *et al.* 1986, Jones *et al.* 1986).

With the advancement of molecular biology and recombinant technology, came the possibility for producing monoclonal antibodies by the hybridoma method (Schwaber *et al.* 1973, Kohler *et al.* 1975), which was awarded the Nobel Prize in Physiology or Medicine in 1984. Monoclonal antibodies have proven to be invaluable tools for many areas including biological research and as therapeutic agents, predominantly in the field of oncology (Hudson 1999, Schaedel *et al.* 2006, Trikha *et al.* 2002). Conventional hybridoma technology, which involves the process of immunizing animals, is still the prevailing source for monoclonal antibodies. However, production of monoclonal antibodies is expensive, time consuming and not easily amenable to large scale production. Antibodies, being protein in nature, presents stringent storage

conditions and not easily druggable orally, making it difficult for high throughput approaches.

Previous efforts over the last 15 years have allowed for the development of *in vitro* immune repertoires and selection methodologies that can be used to derive antibodies without the need for direct immunization of a living host. More recently, producing antibodies with specifically tailored characteristics using synthetic antibody libraries, in which antigen-binding sites are completely artificial, can perform as well or even surpass that of natural immune repertoires (Sidhu and Fellouse 2006). However, this involves the use of the antibody phage display libraries, and antibodies are still proteinaceous in nature.

In this thesis, we put forward a novel approach to obtain high affinity peptide probes for any target protein sequences. Briefly, this novel platform technology aims to produce ‘synthetic antibodies’ for a target protein, which involves fishing self-binding sequences from a given protein by a peptide library of overlapping peptides derived from the given protein sequence, and probing for interactions using the same protein. Positive candidate peptides are validated for interactions with the target protein, and different formats of the peptides with enhanced binding affinity to the protein can then be deployed for downstream applications. As our ‘synthetic antibody’ is a small peptide, it harbors advantages that can overcome those of conventional antibodies. Production is comparatively more rapid and economical than conventional antibodies, and can be easily scaled up. Peptides are druggable and can be stored as lyophilized form without compromised activity. This technique will also eliminate the need for immunization in living animals.

In the context of this thesis, we found self interacting peptides in functional and previously uncharacterized regions, consisting of ordered domains and can be categorized into four general groups, namely: (i) complementary binding sequences, (ii) self binding β -strand sequences, (iii) hydrophobic sequences and (iv) carbohydrate binding sequences. These four groups of peptides have literature precedence, and may represent the general rules governing this system. Taking inspiration from the Y-shaped antibody structure, improved avidity and affinity of binding via multivalency and heterovalency was attained. The dimeric structure is reminiscent of the antibody structure, and the heterodimer may surpass the antibody structure, allowing two antigen binding sites on the same molecule, which does not exist in the natural repertoire. The self interacting ability of the peptides, exemplified by fibril forming properties of C6 dimer, implies that the mechanism of action for such peptides may be by binding back to itself. Concluding the findings in this thesis, trials of 'synthetic antibody' prototype, used as research analytical tool against SARS S protein, showed that it was able to function as well or even better than conventional antibodies as lab reagents.

Using this novel peptide-based technology, it is hoped that reliable and fast diagnosis of infected samples can be developed. Current standards of RT-PCR are too lengthy, costly and inefficient for rapid diagnosis of large clusters of community outbreaks, which is critical for effective healthcare administration and situational control. This is especially applicable in rapidly evolving situations, well exemplified by the H1N1 pandemic in 2009 and any possible future outbreaks of infectious diseases.

The possibility to create and produce ‘synthetic antibodies’ with functionalities equivalent or superior to current antibodies, brings new excitement and hope in biological research, diagnostics and therapeutics. In Chapter 4, we showed that it is possible to use them as antibody substitute for immunofluorescence staining and small scale purification of protein using the pull down assay. They can also be used for other downstream applications substituting the use of antibodies, protein purification without the need for attaching affinity tags and as high affinity probes that can be linked to a multitude of reporters. We hope that making use of this technology, peptides that bind different viruses can be immobilized on arrays or chips for diagnostic use. Sera samples can then be incubated with such diagnostic chips, giving characteristic signatures that identify the causative virus and even strain. This can efficiently aid control of outbreaks rapidly and accurately, limiting viral spread to epidemic or pandemic level.

The development of ‘synthetic antibodies’ can only further accelerate advances in this relatively young field, extending the utility of antibodies into areas of research and therapy that are beyond the scope of present technologies. For instance, it has been suggested that gene therapy may make intracellular delivery of intrabodies possible, and the use of specialized synthetic libraries will be crucial to the success of such approaches (Sanz *et al.* 2004). With the progress and standardization of this technology, it is possible to couple automation for high throughput antibody generation in a proteomics environment (Bradbury *et al.* 2003, Chambers *et al.* 2005). Recent advancements in recombinant antibody technology have facilitated a transformative effect on therapeutic biology (Sidhu and Fellouse 2006). With the possibility to synthetically create antibody-like molecules, this may be the first step to

gaining control over antibody design, and perhaps even antibody function and specificity.

Concluding this thesis, I have made use of the quaternary protein mimetic strategy to understand MPER during fusion, to derive potential entry inhibitor candidates, and to describe a novel platform technology for diagnostic uses. However, more information is required to fill the gaps in the field of SARS CoV over time. Crystal structures of S protein and MPER at various intermediate states of fusion would be invaluable to obtaining a clearer understanding of the fusion mechanism. It is also important to note that the development of a “synthetic antibody” based on self-interacting peptides is unconventional. Thus, it is crucial to extend the new technology to other systems to test the robustness of this concept. Candidate systems includes other Class I virus fusion glycoproteins, Class II virus fusion glycoproteins and other soluble cytoplasmic proteins. To use the novel platform technology for diagnostics, it still needs to be validated in clinical trials using a statistically adequate number of infected specimens versus matched control samples from patients.

Publications

1. **TL Neo**, Y. Lu, ZY Ong, KD Eom, SG Shochat and JP Tam. Self Interacting Sequences in SARS CoV Spike Protein Using a Novel Method – Implications in Mimetics of Antibodies. (In preparation)
2. **TL Neo**, JP Tam. Quaternary Protein Mimetics of SARS-CoV Spike Protein MPER of Intermediate States During Fusion. (In preparation)
3. Lu Y, **Neo TL**, Liu DX, Tam JP. (2008) Importance of SARS-CoV spike protein Trp-rich region in viral infectivity. *Biochem Biophys Res Commun*.
4. **TL Neo**, Y. Lu, DX Liu and JP Tam. (2006) The Importance of Trp-rich region in SARS-CoV Infectivity. 2006 *NIAID Research Conference Poster*

References

Akerstrom, S., A. Mirazimi, et al. (2007). "Inhibition of SARS-CoV replication cycle by small interference RNAs silencing specific SARS proteins, 7a/7b, 3a/3b and S." *Antiviral Res* 73(3): 219-27.

Amit, A. G., R. A. Mariuzza, et al. (1986). "Three-dimensional structure of an antigen-antibody complex at 2.8 Å resolution." *Science* 233(4765): 747-53.

Anderson, J. H., Jr., R. L. Brunelle, et al. (1997). "Reduction of postprandial hyperglycemia and frequency of hypoglycemia in IDDM patients on insulin-analog treatment. Multicenter Insulin Lispro Study Group." *Diabetes* 46(2): 265-70.

Anderson, J. H., Jr., R. L. Brunelle, et al. (1997). "Improved mealtime treatment of diabetes mellitus using an insulin analogue. Multicenter Insulin Lispro Study Group." *Clin Ther* 19(1): 62-72.

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.

Bisht, H., A. Roberts, et al. (2004). "Severe acute respiratory syndrome coronavirus spike protein expressed by attenuated vaccinia virus protectively immunizes mice." *Proc Natl Acad Sci U S A* 101(17): 6641-6.

Black, S. D. and D. R. Mould (1991). "Development of hydrophobicity parameters to analyze proteins which bear post- or cotranslational modifications." *Anal Biochem* 193(1): 72-82.

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.

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.

Bradbury, A., N. Velappan, et al. (2003). "Antibodies in proteomics II: screening, high-throughput characterization and downstream applications." *Trends Biotechnol* 21(7): 312-7.

Buchholz, U. J., A. Bukreyev, et al. (2004). "Contributions of the structural proteins of severe acute respiratory syndrome coronavirus to protective immunity." *Proc Natl Acad Sci U S A* 101(26): 9804-9.

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

Carr, C. M., C. Chaudhry, et al. (1997). "Influenza hemagglutinin is spring-loaded by

- a metastable native conformation." *Proc Natl Acad Sci U S A* 94(26): 14306-13.
- Carr, C. M. and P. S. Kim (1993). "A spring-loaded mechanism for the conformational change of influenza hemagglutinin." *Cell* 73(4): 823-32.
- Chambers, R. S. (2005). "High-throughput antibody production." *Curr Opin Chem Biol* 9(1): 46-50.
- 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.
- Chen, Z., L. Zhang, et al. (2005). "Recombinant modified vaccinia virus Ankara expressing the spike glycoprotein of severe acute respiratory syndrome coronavirus induces protective neutralizing antibodies primarily targeting the receptor binding region." *J Virol* 79(5): 2678-88.
- Chu, L. H., S. H. Chan, et al. (2008). "Fusion core structure of the severe acute respiratory syndrome coronavirus (SARS-CoV): in search of potent SARS-CoV entry inhibitors." *J Cell Biochem* 104(6): 2335-47.
- Coughlin, M., G. Lou, et al. (2007). "Generation and characterization of human monoclonal neutralizing antibodies with distinct binding and sequence features against SARS coronavirus using XenoMouse." *Virology* 361(1): 93-102.
- Czub, M., H. Weingartl, et al. (2005). "Evaluation of modified vaccinia virus Ankara based recombinant SARS vaccine in ferrets." *Vaccine* 23(17-18): 2273-9.
- Derewenda, U., Z. Derewenda, et al. (1989). "Molecular structure of insulin: the insulin monomer and its assembly." *Br Med Bull* 45(1): 4-18.
- DiMarchi, R. D., R. E. Chance, et al. (1994). "Preparation of an insulin with improved pharmacokinetics relative to human insulin through consideration of structural homology with insulin-like growth factor I." *Horm Res* 41 Suppl 2: 93-6.
- Du, L., Y. He, et al. (2009). "The spike protein of SARS-CoV--a target for vaccine and therapeutic development." *Nat Rev Microbiol* 7(3): 226-36.
- 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.
- Du, L., G. Zhao, et al. (2007). "Receptor-binding domain of SARS-CoV spike protein induces long-term protective immunity in an animal model." *Vaccine* 25(15): 2832-8.
- Dutta, N. K., K. Mazumdar, et al. (2008). "Search for potential target site of nucleocapsid gene for the design of an epitope-based SARS DNA vaccine." *Immunol Lett* 118(1): 65-71.
- Earp, L. J., S. E. Delos, et al. (2005). "The many mechanisms of viral membrane

fusion proteins." *Curr Top Microbiol Immunol* 285: 25-66.

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

Fouchier, R. A., T. Kuiken, et al. (2003). "Aetiology: Koch's postulates fulfilled for SARS virus." *Nature* 423(6937): 240.

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

Frank, R. and H. Overwin (1996). "SPOT synthesis. Epitope analysis with arrays of synthetic peptides prepared on cellulose membranes." *Methods Mol Biol* 66: 149-69.

Garoff, H., R. Hewson, et al. (1998). "Virus maturation by budding." *Microbiol Mol Biol Rev* 62(4): 1171-90.

Haim, H., Z. Si, et al. (2009). "Soluble CD4 and CD4-mimetic compounds inhibit HIV-1 infection by induction of a short-lived activated state." *PLoS Pathog* 5(4): e1000360.

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.

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

He, M. L., B. J. Zheng, et al. (2006). "Kinetics and synergistic effects of siRNAs targeting structural and replicase genes of SARS-associated coronavirus." *FEBS Lett* 580(10): 2414-20.

He, Y. (2006). "Immunogenicity of SARS-CoV: the receptor-binding domain of S protein is a major target of neutralizing antibodies." *Adv Exp Med Biol* 581: 539-42.

He, Y., J. Li, et al. (2006). "Identification and characterization of novel neutralizing epitopes in the receptor-binding domain of SARS-CoV spike protein: revealing the critical antigenic determinants in inactivated SARS-CoV vaccine." *Vaccine* 24(26): 5498-508.

He, Y., J. Li, et al. (2006). "Antigenic and immunogenic characterization of recombinant baculovirus-expressed severe acute respiratory syndrome coronavirus spike protein: implication for vaccine design." *J Virol* 80(12): 5757-67.

He, Y., J. Li, et al. (2006). "A single amino acid substitution (R441A) in the receptor-binding domain of SARS coronavirus spike protein disrupts the antigenic structure and binding activity." *Biochem Biophys Res Commun* 344(1): 106-13.

He, Y., J. Li, et al. (2006). "Cross-neutralization of human and palm civet severe acute respiratory syndrome coronaviruses by antibodies targeting the receptor-binding

domain of spike protein." J Immunol 176(10): 6085-92.

He, Y., H. Lu, et al. (2005). "Receptor-binding domain of severe acute respiratory syndrome coronavirus spike protein contains multiple conformation-dependent epitopes that induce highly potent neutralizing antibodies." J Immunol 174(8): 4908-15.

He, Y., Y. Zhou, et al. (2004). "Receptor-binding domain of SARS-CoV spike protein induces highly potent neutralizing antibodies: implication for developing subunit vaccine." Biochem Biophys Res Commun 324(2): 773-81.

He, Y., Q. Zhu, et al. (2005). "Identification of a critical neutralization determinant of severe acute respiratory syndrome (SARS)-associated coronavirus: importance for designing SARS vaccines." Virology 334(1): 74-82.

Hoekstra, D., T. de Boer, et al. (1984). "Fluorescence method for measuring the kinetics of fusion between biological membranes." Biochemistry 23(24): 5675-81.

Hofmann, H., K. Pyrc, et al. (2005). "Human coronavirus NL63 employs the severe acute respiratory syndrome coronavirus receptor for cellular entry." Proc Natl Acad Sci U S A 102(22): 7988-93.

Holmes, K. V. (2003). "SARS-associated coronavirus." N Engl J Med 348(20): 1948-51.

Hu, H., L. Li, et al. (2005). "Screening and identification of linear B-cell epitopes and entry-blocking peptide of severe acute respiratory syndrome (SARS)-associated coronavirus using synthetic overlapping peptide library." J Comb Chem 7(5): 648-56.

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.

Hudson, P. J. (1999). "Recombinant antibody constructs in cancer therapy." Curr Opin Immunol 11(5): 548-57.

Isenberg, G., S. Doerhoefer, et al. (2002). "Membrane fusion induced by the major lipid-binding domain of the cytoskeletal protein talin." Biochem Biophys Res Commun 295(3): 636-43.

Ikeguchi, M., S. Kato, et al. (1997). "Molten globule state of equine beta-lactoglobulin." Proteins 27(4): 567-75.

Jardetzky, T. S. and R. A. Lamb (2004). "Virology: a class act." Nature 427(6972): 307-8.

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.

- Jiang, S., K. Lin, et al. (1993). "HIV-1 inhibition by a peptide." *Nature* 365(6442): 113.
- Jin, H., C. Xiao, et al. (2005). "Induction of Th1 type response by DNA vaccinations with N, M, and E genes against SARS-CoV in mice." *Biochem Biophys Res Commun* 328(4): 979-86.
- Jones, D. T. (1999). "Protein secondary structure prediction based on position-specific scoring matrices." *J Mol Biol* 292(2): 195-202.
- Jones, P. T., P. H. Dear, et al. (1986). "Replacing the complementarity-determining regions in a human antibody with those from a mouse." *Nature* 321(6069): 522-5.
- Kam, Y. W., F. Kien, et al. (2007). "Antibodies against trimeric S glycoprotein protect hamsters against SARS-CoV challenge despite their capacity to mediate FcγRII-dependent entry into B cells in vitro." *Vaccine* 25(4): 729-40.
- Kao, R. Y., A. P. To, et al. (2004). "Characterization of SARS-CoV main protease and identification of biologically active small molecule inhibitors using a continuous fluorescence-based assay." *FEBS Lett* 576(3): 325-30.
- Kao, R. Y., W. H. Tsui, et al. (2004). "Identification of novel small-molecule inhibitors of severe acute respiratory syndrome-associated coronavirus by chemical genetics." *Chem Biol* 11(9): 1293-9.
- Kielian, M. and F. A. Rey (2006). "Virus membrane-fusion proteins: more than one way to make a hairpin." *Nat Rev Microbiol* 4(1): 67-76.
- Kohler, G. and C. Milstein (1975). "Continuous cultures of fused cells secreting antibody of predefined specificity." *Nature* 256(5517): 495-7.
- Konami, Y., K. Yamamoto, et al. (1992). "Purification and characterization of carbohydrate-binding peptides from *Lotus tetragonolobus* and *Ulex europeus* seed lectins using affinity chromatography." *J Chromatogr* 597(1-2): 213-9.
- Konami, Y., K. Yamamoto, et al. (1992). "Correlation between carbohydrate-binding specificity and amino acid sequence of carbohydrate-binding regions of Cytisus-type anti-H(O) lectins." *FEBS Lett* 304(2-3): 129-35.
- 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.
- Lai, A. L., H. Park, et al. (2006). "Fusion peptide of influenza hemagglutinin requires a fixed angle boomerang structure for activity." *J Biol Chem* 281(9): 5760-70.
- Lai, A. L. and L. K. Tamm (2007). "Locking the kink in the influenza hemagglutinin fusion domain structure." *J Biol Chem* 282(33): 23946-56.
- Lai, L., X. Han, et al. (2006). "Quaternary structure, substrate selectivity and inhibitor design for SARS 3C-like proteinase." *Curr Pharm Des* 12(35): 4555-64.

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

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., R. G. Paterson, et al. (2006). "Paramyxovirus membrane fusion: lessons from the F and HN atomic structures." *Virology* 344(1): 30-7.

Leparc-Goffart, I., S. T. Hingley, et al. (1998). "Targeted recombination within the spike gene of murine coronavirus mouse hepatitis virus-A59: Q159 is a determinant of hepatotropism." *J Virol* 72(12): 9628-36.

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, B. J., Q. Tang, et al. (2005). "Using siRNA in prophylactic and therapeutic regimens against SARS coronavirus in Rhesus macaque." *Nat Med* 11(9): 944-51.

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

Li, T., J. Xie, et al. (2006). "Long-term persistence of robust antibody and cytotoxic T cell responses in recovered patients infected with SARS coronavirus." *PLoS One* 1: e24.

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.

Liang, W. N., T. Zhao, et al. (2006). "Severe acute respiratory syndrome--retrospect and lessons of 2004 outbreak in China." *Biomed Environ Sci* 19(6): 445-51.

Litman, G. W., J. P. Rast, et al. (1993). "Phylogenetic diversification of immunoglobulin genes and the antibody repertoire." *Mol Biol Evol* 10(1): 60-72.

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., S. Wu, et al. (2007). "HIV entry inhibitors targeting gp41: from polypeptides to small-molecule compounds." *Curr Pharm Des* 13(2): 143-62.

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.

- Liu, Y., J. Zhu, et al. (2004). "Six-helix bundle assembly and analysis of the central core of mumps virus fusion protein." *Arch Biochem Biophys* 421(1): 143-8.
- Lu, Y., T. L. Neo, et al. (2008). "Importance of SARS-CoV spike protein Trp-rich region in viral infectivity." *Biochem Biophys Res Commun* 371(3): 356-60.
- MacDonald, R. I. (1990). "Characteristics of self-quenching of the fluorescence of lipid-conjugated rhodamine in membranes." *J Biol Chem* 265(23): 13533-9.
- Martino, M. C., R. A. Stabler, et al. (2001). "Helicobacter pylori pore-forming cytotoxin orthologue TlyA possesses in vitro hemolytic activity and has a role in colonization of the gastric mucosa." *Infect Immun* 69(3): 1697-703.
- McRoy, W. C. and R. S. Baric (2008). "Amino acid substitutions in the S2 subunit of mouse hepatitis virus variant V51 encode determinants of host range expansion." *J Virol* 82(3): 1414-24.
- Mizuguchi, M., K. Masaki, et al. (1999). "The molten globule state of a chimera of human alpha-lactalbumin and equine lysozyme." *J Mol Biol* 292(5): 1137-48.
- Mo, H. Y., J. Xu, et al. (2005). "Evaluation by indirect immunofluorescent assay and enzyme linked immunosorbent assay of the dynamic changes of serum antibody responses against severe acute respiratory syndrome coronavirus." *Chin Med J (Engl)* 118(6): 446-50.
- Nayar, R., Cullis, P. R., et al. (1989). "Generation of large unilamellar vesicles from long-chain saturated phosphatidylcholines by extrusion technique." *Biochim Biophys Acta* 986(2): 200-06.
- Nie, Y., G. Wang, et al. (2004). "Neutralizing antibodies in patients with severe acute respiratory syndrome-associated coronavirus infection." *J Infect Dis* 190(6): 1119-26.
- Normile, D. (2004). "Infectious diseases. Mounting lab accidents raise SARS fears." *Science* 304(5671): 659-61.
- Normile, D. (2004). "Infectious diseases. Second lab accident fuels fears about SARS." *Science* 303(5654): 26.
- Orellana, C. (2004). "Laboratory-acquired SARS raises worries on biosafety." *Lancet Infect Dis* 4(2): 64.
- Park, H. E., J. A. Gruenke, et al. (2003). "Leash in the groove mechanism of membrane fusion." *Nat Struct Biol* 10(12): 1048-53.
- Peiris, J. S., S. T. Lai, et al. (2003). "Coronavirus as a possible cause of severe acute respiratory syndrome." *Lancet* 361(9366): 1319-25.
- Prabakaran, P., J. Gan, et al. (2006). "Structure of severe acute respiratory syndrome coronavirus receptor-binding domain complexed with neutralizing antibody." *J Biol Chem* 281(23): 15829-36.

Qin, Z. L., P. Zhao, et al. (2004). "Silencing of SARS-CoV spike gene by small interfering RNA in HEK 293T cells." *Biochem Biophys Res Commun* 324(4): 1186-93.

Rex, S. (1996). "Pore formation induced by the peptide melittin in different lipid vesicle membranes." *Biophys Chem* 58(1-2): 75-85.

Roberts, D. D. and I. J. Goldstein (1984). "Isolation from lima bean lectin of a peptide containing a cysteine residue essential for carbohydrate binding activity." *J Biol Chem* 259(2): 909-14.

Rockx, B., D. Corti, et al. (2008). "Structural basis for potent cross-neutralizing human monoclonal antibody protection against lethal human and zoonotic severe acute respiratory syndrome coronavirus challenge." *J Virol* 82(7): 3220-35.

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.

Ruigrok, R. W., S. R. Martin, et al. (1986). "Conformational changes in the hemagglutinin of influenza virus which accompany heat-induced fusion of virus with liposomes." *Virology* 155(2): 484-97.

Saez-Cirion, A., J. L. Arrondo, et al. (2003). "Structural and functional roles of HIV-1 gp41 pretransmembrane sequence segmentation." *Biophys J* 85(6): 3769-80.

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

Sainz, B., Jr., J. M. Rausch, et al. (2005). "The aromatic domain of the coronavirus class I viral fusion protein induces membrane permeabilization: putative role during viral entry." *Biochemistry* 44(3): 947-58.

Salzwedel, K., J. T. West, Jr., et al. (1998). "Retention of the human immunodeficiency virus type 1 envelope glycoprotein in the endoplasmic reticulum does not redirect virus assembly from the plasma membrane." *J Virol* 72(9): 7523-31.

Sanchez, C. M., A. Izeta, et al. (1999). "Targeted recombination demonstrates that the spike gene of transmissible gastroenteritis coronavirus is a determinant of its enteric tropism and virulence." *J Virol* 73(9): 7607-18.

Sanz, L., B. Blanco, et al. (2004). "Antibodies and gene therapy: teaching old 'magic bullets' new tricks." *Trends Immunol* 25(2): 85-91.

Sawaya, M. R., S. Sambashivan, et al. (2007). "Atomic structures of amyloid cross-beta spines reveal varied steric zippers." *Nature* 447(7143): 453-7.

Schaedel, O. and Y. Reiter (2006). "Antibodies and their fragments as anti-cancer agents." *Curr Pharm Des* 12(3): 363-78.

Schibli, D. J., R. C. Montelaro, et al. (2001). "The membrane-proximal tryptophan-rich region of the HIV glycoprotein, gp41, forms a well-defined helix in dodecylphosphocholine micelles." *Biochemistry* 40(32): 9570-8.

Schwaber, J. and E. P. Cohen (1973). "Human x mouse somatic cell hybrid clone secreting immunoglobulins of both parental types." *Nature* 244(5416): 444-7.

See, R. H., M. Petric, et al. (2008). "Severe acute respiratory syndrome vaccine efficacy in ferrets: whole killed virus and adenovirus-vectored vaccines." *J Gen Virol* 89(Pt 9): 2136-46.

Shi, Z. and Z. Hu (2008). "A review of studies on animal reservoirs of the SARS coronavirus." *Virus Res* 133(1): 74-87.

Sidhu, S. S. and F. A. Fellouse (2006). "Synthetic therapeutic antibodies." *Nat Chem Biol* 2(12): 682-8.

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., A. J. Rennekamp, et al. (2006). "Proteolysis of SARS-associated coronavirus spike glycoprotein." *Adv Exp Med Biol* 581: 235-40.

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

Song, H. D., C. C. Tu, et al. (2005). "Cross-host evolution of severe acute respiratory syndrome coronavirus in palm civet and human." *Proc Natl Acad Sci U S A* 102(7): 2430-5.

Stadler, K., H. R. Ha, et al. (2008). "Amiodarone alters late endosomes and inhibits SARS coronavirus infection at a post-endosomal level." *Am J Respir Cell Mol Biol* 39(2): 142-9.

Stadler, K., V. Masignani, et al. (2003). "SARS--beginning to understand a new virus." *Nat Rev Microbiol* 1(3): 209-18.

Stiasny, K. and F. X. Heinz (2006). "Flavivirus membrane fusion." *J Gen Virol* 87(Pt 10): 2755-66.

Stockman, L. J., R. Bellamy, et al. (2006). "SARS: systematic review of treatment effects." *PLoS Med* 3(9): e343.

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.

Suarez, T., S. Nir, et al. (2000). "The pre-transmembrane region of the human

immunodeficiency virus type-1 glycoprotein: a novel fusogenic sequence." *FEBS Lett* 477(1-2): 145-9.

Sui, J., W. Li, et al. (2004). "Potent neutralization of severe acute respiratory syndrome (SARS) coronavirus by a human mAb to S1 protein that blocks receptor association." *Proc Natl Acad Sci U S A* 101(8): 2536-41.

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.

Takeuchi, H., A. Okada, et al. (2003). "Roles of the histidine and tryptophan side chains in the M2 proton channel from influenza A virus." *FEBS Lett* 552(1): 35-8.

Tamm, L. K. (2003). "Hypothesis: spring-loaded boomerang mechanism of influenza hemagglutinin-mediated membrane fusion." *Biochim Biophys Acta* 1614(1): 14-23.

Towler, P., B. Staker, et al. (2004). "ACE2 X-ray structures reveal a large hinge-bending motion important for inhibitor binding and catalysis." *J Biol Chem* 279(17): 17996-8007.

Traggiai, E., S. Becker, et al. (2004). "An efficient method to make human monoclonal antibodies from memory B cells: potent neutralization of SARS coronavirus." *Nat Med* 10(8): 871-5.

Trikha, M., L. Yan, et al. (2002). "Monoclonal antibodies as therapeutics in oncology." *Curr Opin Biotechnol* 13(6): 609-14.

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.

van den Brink, E. N., J. Ter Meulen, et al. (2005). "Molecular and biological characterization of human monoclonal antibodies binding to the spike and nucleocapsid proteins of severe acute respiratory syndrome coronavirus." *J Virol* 79(3): 1635-44.

Vishwanathan, S. A. and E. Hunter (2008). "Importance of the membrane-perturbing properties of the membrane-proximal external region of human immunodeficiency virus type 1 gp41 to viral fusion." *J Virol* 82(11): 5118-26.

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.

Weingartl, H., M. Czub, et al. (2004). "Immunization with modified vaccinia virus Ankara-based recombinant vaccine against severe acute respiratory syndrome is associated with enhanced hepatitis in ferrets." *J Virol* 78(22): 12672-6.

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

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, L., C. McKinlay, et al. (2004). "SARS coronavirus E protein forms cation-selective ion channels." *Virology* 330(1): 322-31.

Wimley, W. C. and S. H. White (1996). "Experimentally determined hydrophobicity scale for proteins at membrane interfaces." *Nat Struct Biol* 3(10): 842-8.

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.

Wu, C. J., H. W. Huang, et al. (2005). "Inhibition of SARS-CoV replication by siRNA." *Antiviral Res* 65(1): 45-8.

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, X. and X. Gao (2004). "Immunological responses against SARS-coronavirus infection in humans." *Cell Mol Immunol* 1(2): 119-22.

Yamamoto, K., Y. Konami, et al. (1992). "Carbohydrate-binding peptides from several anti-H(O) lectins." *J Biochem (Tokyo)* 111(4): 436-9.

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.

Yau, W. M., W. C. Wimley, et al. (1998). "The preference of tryptophan for membrane interfaces." *Biochemistry* 37(42): 14713-8.

Yi, L., Z. Li, et al. (2004). "Small molecules blocking the entry of severe acute respiratory syndrome coronavirus into host cells." *J Virol* 78(20): 11334-9.

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.

Zakhartchouk, A. N., C. Sharon, et al. (2007). "Immunogenicity of a receptor-binding domain of SARS coronavirus spike protein in mice: implications for a subunit vaccine." *Vaccine* 25(1): 136-43.

Zhang, Y., T. Li, et al. (2004). "Silencing SARS-CoV Spike protein expression in cultured cells by RNA interference." *FEBS Lett* 560(1-3): 141-6.

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.

Zheng, B. J., Y. Guan, et al. (2004). "Prophylactic and therapeutic effects of small interfering RNA targeting SARS-coronavirus." *Antivir Ther* 9(3): 365-74.

Zhi, Y., G. P. Kobinger, et al. (2005). "Identification of murine CD8 T cell epitopes in codon-optimized SARS-associated coronavirus spike protein." *Virology* 335(1): 34-45.

Zhi, Y., J. M. Wilson, et al. (2005). "SARS vaccine: progress and challenge." *Cell Mol Immunol* 2(2): 101-5.

Zhong, X., H. Yang, et al. (2005). "B-cell responses in patients who have recovered from severe acute respiratory syndrome target a dominant site in the S2 domain of the surface spike glycoprotein." *J Virol* 79(6): 3401-8.

Zhou, T., H. Wang, et al. (2004). "An exposed domain in the severe acute respiratory syndrome coronavirus spike protein induces neutralizing antibodies." *J Virol* 78(13): 7217-26.

Zhu, P., J. Liu, et al. (2006). "Distribution and three-dimensional structure of AIDS virus envelope spikes." *Nature* 441(7095): 847-52.

Zhu, Z., S. Chakraborti, et al. (2007). "Potent cross-reactive neutralization of SARS coronavirus isolates by human monoclonal antibodies." *Proc Natl Acad Sci U S A* 104(29): 12123-8.

Zwick, M. B., A. F. Labrijn, et al. (2001). "Broadly neutralizing antibodies targeted to the membrane-proximal external region of human immunodeficiency virus type 1 glycoprotein gp41." *J Virol* 75(22): 10892-905.