

# **NMR Studies of Viral Envelope Proteins**

**BY**  
**JESSICA CELIGOY**  
**BS., Gonzaga University 2005**

## **THESIS**

**Submitted as partial fulfillment of the requirements  
for the degree of Doctor of Philosophy in Biochemistry and Molecular Genetics  
in the Graduate College of the  
University of Illinois at Chicago, 2011  
Chicago, IL**

### **Defense Committee:**

**Michael Caffrey, Chair and Advisor**

**Vadim Gaponenko**

**Jack Kaplan**

**Lijun Rong, Microbiology and Immunology**

**Michael Johnson, Medicinal Chemistry and Pharmacognosy**

*I would like to dedicate this thesis to my parents, Hank and Glenda.  
Thank you for all the love and encouragement, all the best things about  
me are because of you*

## ACKNOWLEDGMENTS

First and foremost, I would like to thank my advisor Dr. Michael Caffrey, for all the support and guidance over the years. It has been a true pleasure to work with you and I am a better scientist because of you. I have been a privilege to work for you.

To Ben Ramirez, thank you for all your help with the NMR experiments. Your patience and willingness to share your knowledge was greatly appreciated. You were a critical part of this project, thank you for all you did.

I would also like to thank my committee members Dr. Vadim Gaponenko, Dr. Jack Kaplan, Dr. Lijun Rong, Dr. Arnon Lavie, and Dr. Michael Johnson. Thank you for the suggestions and advice during this work.

Finally I would like to express my heartfelt gratitude to my friends and family who have loved and supported me throughout this process. Specific thanks to Natalie, Caroline, Kate and Ben; I was able to finish this work in large part because of you. I am incredibly blessed to have you in my life. Thank you.

## TABLE OF CONTENTS

<u>CHAPTER</u>	<u>PAGE</u>
1. INTRODUCTION AND BACKGROUND.....	1
1.1 Viruses.....	1
1.2 Entry of Enveloped Viruses.....	4
1.3 Class I Viral Fusion Proteins.....	8
1.4 Envelope Proteins as Therapeutic Targets.....	16
1.5 Thesis Research.....	19
2. MATERIALS AND METHODS.....	20
2.1 Characterization of SARS-CoV S2-HR2.....	20
2.1.1 Production and Purification of SARS S2-HR2.....	20
2.1.2 Analytical Ultracentrifugation of SARS S2-HR2.....	21
2.1.3 Circular Dichroism of SARS S2-HR2.....	21
2.1.4 Production and Purification of SARS Co-V S2-HR2 with a Paramagnetic Label.....	24
2.1.5 NMR Spectroscopy.....	25
2.2 Characterization of HIV Envelope.....	26
2.2.1 Peptide Synthesis.....	26
2.2.2 Preparation of Recombinant HIV Envelope Glycoproteins.....	26
2.2.3 NMR Experiments.....	29
2.2.4 Production of Virus Like Particles.....	29
3. SARS-CoV S2-HR2 IS A PARALLEL TRIMER.....	33
3.1 SARS Outbreak.....	33
3.2 Entry of SARS-CoV.....	34
3.3 SARS-CoV Spike Protein and Putative Homology to Viral Envelope Proteins.....	35
3.4 SARS-CoV S2-HR2.....	38
3.5 Structural Studies of SARS-CoV S.....	40
3.6 Overview of Work Presented in Chapter 3.....	41
3.7 Results and Discussion.....	43
3.7.1 SARS-CoV S2-HR2 Exists in a Monomer/Trimer Equilibrium Under Aqueous Conditions.....	43
3.7.2 Effect of Ionic Strength and pH on the Structure of S2-HR2.....	45
3.7.3 SARS-CoV S2-HR2 has a Parallel Orientation.....	53
3.8 Conclusions.....	57
4. PROBING THE HIV ENVELOPE PROTEIN STRUCTURE USING THE 12p1 PEPTIDE.....	58
4.1 HIV.....	58
4.2 Use of 12p1 to Probe Envelope Structure.....	59

## TABLE OF CONTENTS (CON'T)

<u>CHAPTER</u>	<u>PAGE</u>
4.2.1 Assignment of 12p1.....	61
4.2.2 STD Study of the 12p1 Interaction with Recombinant gp120.....	69
4.2.3 STD Study of the 12p1 Interaction with Recombinant gp140.....	71
4.2.4 Model for the Interaction Between 12p1 and gp120.....	75
4.2.5 WaterLOGSY of gp120 with 12p1 Peptide.....	77
4.2.6 STD Study of the 12p1 Interaction with Virus Like Particles (VLPs) Expressing HIV-1 Envelope Proteins.....	79
4.3 Summary.....	84
5. SUMMARY AND FUTURE STUDIES.....	85
5.1 Summary of Work.....	85
5.2 Future Studies.....	86
CITED LITERATURE.....	88
APPENDICES.....	96

## LIST OF FIGURES

<u>FIGURE</u>	<u>PAGE</u>
1.1 4 Basic Steps of the Viral Life Cycle.....	3
1.2 Modes of Viral Entry.....	7
1.3 Feature of the Class I Viral Fusion Proteins.....	10
1.4 General Mechanism of Class I Viral Entry.....	13
1.5 Examples of the Post-Fusion Structures Characteristic of Class I Fusion Proteins.....	14
1.6 Mechanism of Enfuvirtide Inhibition of HIV-1.....	18
2.1 Mutagenesis of SARS-CoV S2-HR2.....	23
2.2 Monoclonal Antibody Binding Analysis and Reactivity to CD4i and CD4- gp140Complex Specific mAbs of Purified R2gp140 Dimer.....	28
3.1 Schematic Diagram of SARS-CoV Spike and HIV Envelope.....	37
3.2 Available Structures of SARS-CoV S2-HR2.....	42
3.3 SV Studies of SARS-CoV S2-HR2.....	44
3.4 Circular Dichroism Studies of SARS-CoV HR2 at pH 7 and pH 5.....	46
3.5 Effect of Ionic Strength on Helical Content of S2-HR2 at pH 7 and pH 5.....	49
3.6 Helical Content as a Function of Ionic Strength.....	50
3.7 Electrostatic Profile of SARS-CoV S2-HR2.....	51
3.8 Model for SARS-CoV S2-HR2 Mediated Entry.....	52
3.9 a MTSL Labeling of SARS-CoV S2-HR2.....	55
3.9 b Effect of MTSL on <sup>15</sup> N Labeled S2-HR2.....	56
4.1 a Amide Region of 12p1 TOCSY.....	64
4.1 b Aliphatic Region of 12p1 <sup>13</sup> C-edited HSQC at Natural Abundance.....	65
4.1 c Aromatic Region of 12p1 <sup>13</sup> C-edited HSQC at Natural Abundance.....	66
4.1 d Secondary Chemical Shifts of 12p1.....	67
4.2 Saturation Transfer Difference NMR.....	68
4.3 STD NMR Spectra of 12p1 in the Presence of gp120 Constructs.....	72
4.4 a Schematic Diagram of gp120 and gp140 Constructs.....	73
4.4 b STD NMR of 12p1 in the Presence of gp140 Constructs .....	74
4.5 Model for 12p1 Binding to gp120.....	76
4.6 WaterLOGSY NMR of 12p1 in the Presence and Absence of R2gp120.....	78
4.7 STD NMR of VSVG-VLP and HIV-VLP.....	82
4.8 STD NMR of VLP Concentrated 10X.....	83

## LIST OF TABLES

<u>TABLE</u>	<u>PAGE</u>
1.1 Envelope Glycoproteins and Their Proteases.....	6
1.2 Key Differences Between Class I and Class II Viral Fusion Proteins.....	9
1.3 Surface and Transmembrane Units of Select Enveloped Viruses.....	11
1.4 Examples for the Three Types of Entry Inhibitors that Target Proteins from HIV-1..	17
2.1 Titers of VLP.....	32
2.2 Results from p24 Assay.....	32
3.1 HR1 and HR2 Derived Peptides.....	39
4.1 Chemical Shift Value Assignments for $^1\text{H}$ , $^{13}\text{C}$ , and $^{15}\text{N}$ of 12p1.....	62
4.2 Predicted and Observed Chemical Shift Values for P6 of 12p1.....	63

## ABSTRACT

Detailed structural information about viral envelope proteins is necessary for a better understanding of the viral entry mechanism, as well as for the development of new and improved anti-viral therapies, especially those targeting the pre-fusion and intermediate states of these conformations. The work presented here focuses on using various biochemical and NMR techniques to gather more information on the structure of the envelope proteins of SARS-CoV and HIV-1. Using these techniques I identified an intermediate state of S2-HR2 in the entry of SARS-CoV, which led to a more detailed proposed mechanism along with characterization of the prefusion state of S2-HR2. For the HIV-1 work I was able to use a peptide probe previously described to bind to HIV-1, 12p1, to probe the gp120/gp41 interface and gather details on the conformational changes that occur in this particular region during viral entry.

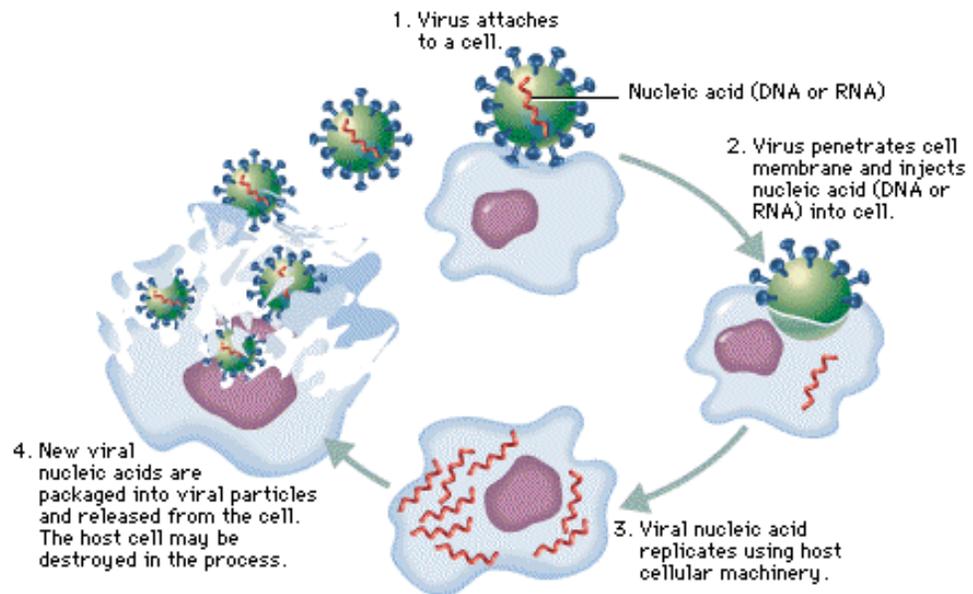
# 1. INTRODUCTION AND BACKGROUND

## 1.1 Viruses

Viruses are small infectious agents that depend on other living organisms for their survival. A virus must infect a host (viruses have been seen to attack all forms of life, i.e. animal, plant, and bacteria), transfer its genetic material into that host's cells, and then "hi-jack" the replicative machinery of that cell in order to reproduce and generate more viral particles. There are four basic steps in the viral life cycle: attachment, entry, replication, and packaging and release of new viral particles (Figure 1.1).

The first step of the viral life cycle is attachment of the virus to the target cell membrane. The viral attachment protein recognizes a specific receptor on the outside of the cell; this receptor may be a protein, carbohydrate, or lipid. This is the stage of the life cycle that leads to specificity of the virus; the target cell must contain the necessary receptor for the virus to initiate infection. After the virus has attached to the cell surface, the virus must release its genome into the cell. Entry of the virus can occur in one of two ways, direct fusion to the cellular membrane or being taken up into an endosome, after this entry step fusion occurs and the viral genetic material can be released into the cytoplasm of the target cell. Once the viral genome has entered the cell it takes over the host cell machinery, using host cell proteins to replicate the genetic material. The viral genome is transcribed and translated and the various viral proteins then produce everything needed to make new viral particles. Once the new viral particles are assembled, they make their way to the cell surface and are released from the host cell either by budding from the cellular membrane or, in some cases, the host cell is lysed and all viral particles are released *en mass* from the cell.

Viruses vary in shape (circular, rod shaped, head/tail), size, and the form the genetic material is transmitted (RNA vs DNA, double or single stranded nucleic acid), which allows many different ways to classify groups of viruses. One of the more broad classifications is based on the protein coat, or capsid, surrounding the virus. All viruses have a protein coat that surrounds the genetic material of the virus until the virus is ready to release the DNA/RNA into the host cell. For some viruses the capsid is the outer-most covering of the virus, these are called “non-enveloped viruses” or “naked viruses”; however, other viruses have a lipid membrane surrounding the capsid that is made up of host cell components and viral glycoproteins, these viruses are classified as “enveloped viruses”. This thesis will concentrate on the entry mechanisms of enveloped viruses.



**Figure 1.1.** 4 Basic steps of the viral life cycle.

## **1.2 Entry of Enveloped Viruses**

The entry of enveloped viruses occurs via membrane fusion either at the plasma membrane or at an intracellular location following endocytosis of the virus (Figure 1.2). Fusion at the plasma membrane is triggered by receptor binding induced conformational changes in the glycoproteins of the envelope. The second method of entry is mediated by a conformational change in the endosome. The envelope protein binds to the receptor on the cell surface of the target cell and is brought into the cell through endocytosis. Typically, viruses have evolved to use only one of the pathways for entry.

Receptor binding is how a virus maintains its specificity, the virus can only infect the cells that express the receptor that the virus has the ability to bind to. Identifying the cellular receptor is a critical step in knowing the entry mechanism of a particular virus and therefore the cellular receptors for the more well known enveloped viruses are known: HIV-1 binds to the CD4+ receptor of T cells, SARS-CoV binds to ACE2 (angiotensin converting enzyme 2), Ebola Virus binds to TIM-1 (T-cell immunoglobulin and mucin domain 1), and Influenza binds to a small sugar called sialic acid that is found attached to many proteins on the cell surface.

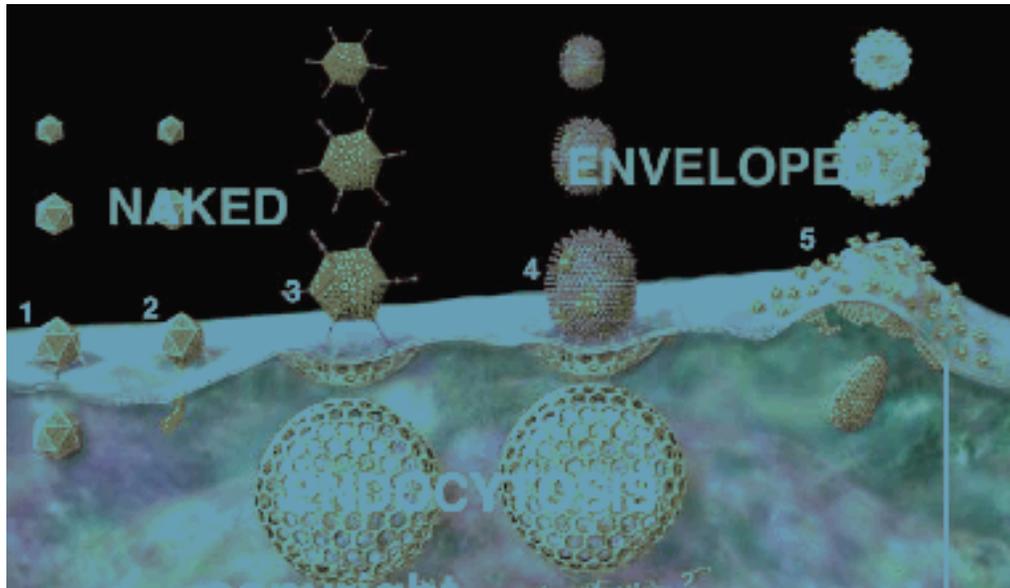
After receptor binding envelope proteins undergo a series of conformational changes. In the case of entry via the endosome, the viral envelope protein undergoes required conformational changes in response to the pH change or proteolytic cleavage in the endosomal environment. Proteolytic processing of the enveloped proteins results in a cleavage of the surface unit and the transmembrane unit of the envelope protein. The various viruses are susceptible to processing by a specific protease as shown in Table 1.1. The cleavage of the envelope proteins is a necessary step for viral entry, however, it does

not always occur at the same stage. For some viruses (i.e. SARS-CoV and Ebola) the proteolytic processing of the envelope proteins is a step during entry. For HIV the proteolytic processing occurs before the envelope proteins reach the surface of the virus and therefore they are in the cleaved form before the entry process is even initiated<sup>1</sup>. Similarly, Influenza HA is cleaved after newly synthesized virions are released from the cell<sup>2</sup>.

The conformational changes of envelope lead to fusion of the viral membrane with either the plasma or endosomal membranes, and the genetic material of the virus is released into the cell. The conformational changes both on the cell surface and within the endosome have been suggested to drive the fusion reaction by allowing the insertion of a hydrophobic domain, termed the fusion peptide, into the membrane of the target cell (either plasma membrane or endosomal membrane) leading to reorientation of the viral and cell membranes<sup>3,4,5</sup>. This reorientation of the membranes is possibly mediated by alignment of fusion peptide oligomers and transmembrane regions that leads to the formation of the fusion pore and enables entry of the viral genetic material<sup>6</sup>.

<b>Envelope Protein (SU/TM)</b>	<b>Protease</b>
HIV gp160 (gp120/gp41)	Furin
SARS-CoV Spike Protein (S1/S2))	Trypsin, cathepsin L, chymotrypsin, elastase
Ebola GP (GP1/GP2)	Cathepsin L
Influenza HA (HA1/HA2)	Bromelin

**Table 1.1.** Envelope glycoproteins and their proteases. It is necessary for envelope proteins to be cleaved into the surface unit (SU) and transmembrane subunits (TM) for the virus to be infectious. This table shows a list of some of the viral envelope proteins and their respective proteases responsible for this cleavage.



**Figure 1.2.** Modes of viral entry.

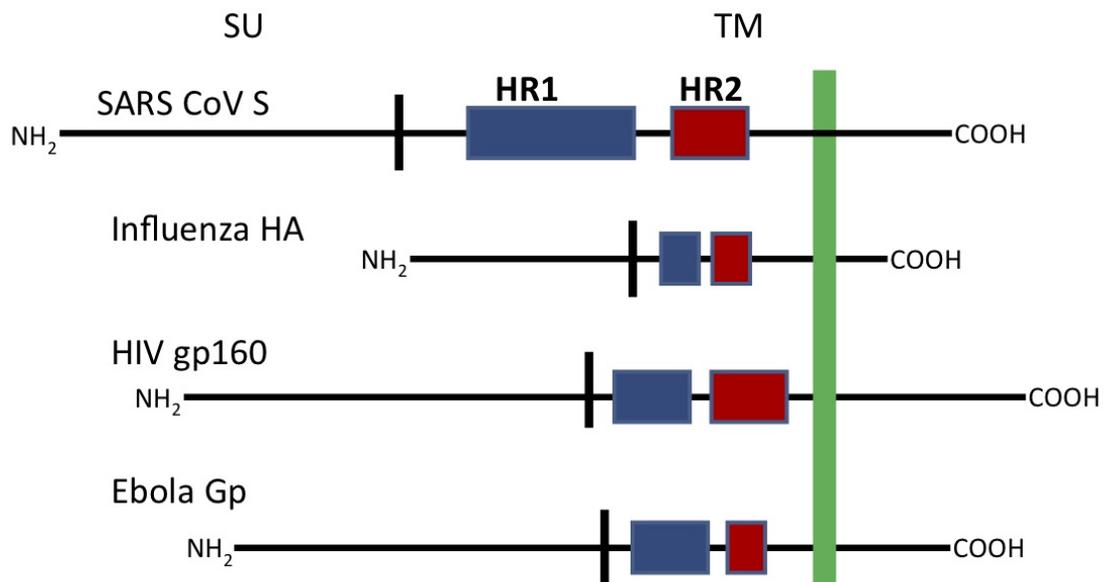
The enveloped viruses are distinguished by their envelope protein, which are expressed on the surface of the viral membrane; this protein is responsible for mediating viral entry into the host cell and has a general homology across the class of envelope proteins. Currently, there are two classes of enveloped viruses, Class I and Class II, which have been defined based on key structural differences. These key structural differences are summarized in Table 1.2. While there are currently only two classes of enveloped viruses there is a possibility that a virus doesn't fit any of these structural definitions of Class I or Class II and a third class of enveloped viruses will be necessary. Class I envelope, or fusion, proteins will be the focus of this thesis.

### **1.3 Class I viral fusion proteins**

Class I viral fusion proteins have the same structural features and domain organization: a surface unit (SU) that binds to receptor and a transmembrane unit (TM) that mediates fusion (Figure 1.3). Within the TM unit there are 5 subdomains: fusion peptide, heptad repeat 1 (HR1), heptad repeat 2 (HR2), transmembrane domain (TM), and cytoplasmic tail (CM). The surface units and transmembrane domains of Influenza, HIV, SARS-CoV and Ebola are listed in Table 1.3.

<b>Feature</b>	<b>Class I (e.g. Influenza HA)</b>	<b>Class II (e.g. SFV E1)</b>
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	Alpha helix	Beta Sheet
Post-Fusion structure	Trimer of hairpins with central alpha-helical coiled coil	Trimer of hairpins composed of Beta 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

**Table 1.2.** Key differences between Class I and Class II Viral Fusion Proteins

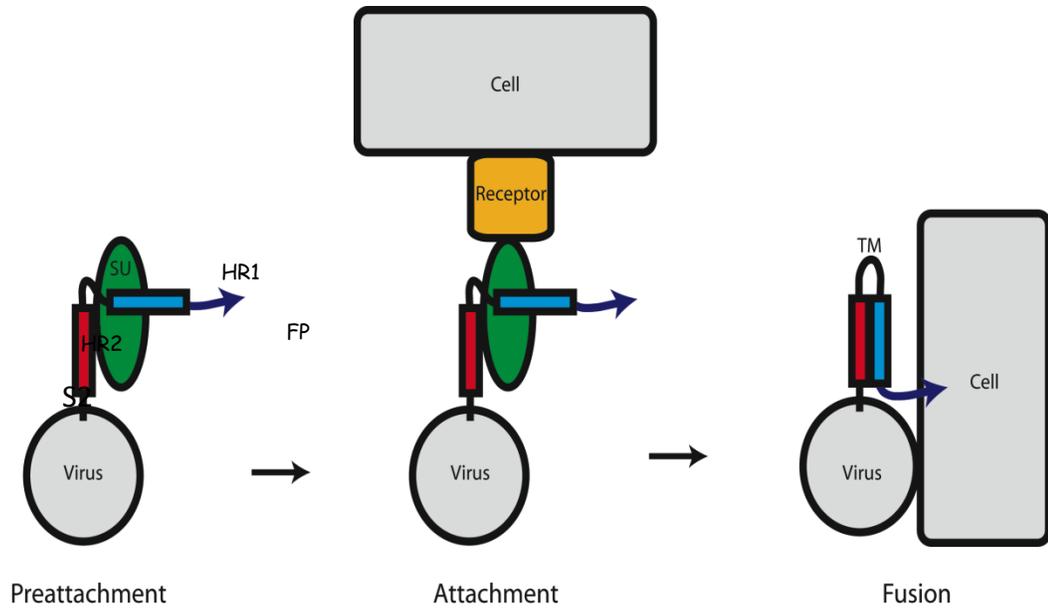


**Figure 1.3.** Features of the class I viral fusion proteins. Abbreviations: SARS, severe acute respiratory syndrome, CoV, coronavirus, HA, hemagglutinin, HIV, human immunodeficiency virus, Gp, glycoprotein, HR, heptad repeat. Not drawn to scale.

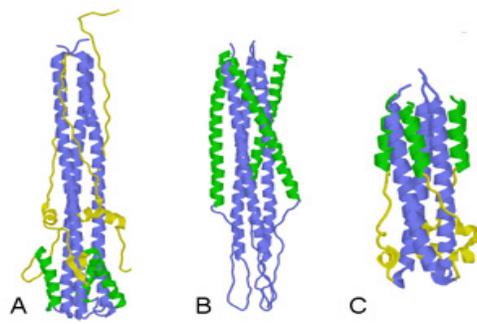
<b>Virus</b>	<b>Surface Unit (SU)</b>	<b>Transmembrane Unit (TM)</b>
Influenza	HA1	HA2
HIV	gp120	gp41
SARS-CoV	S1	S2
Ebola	GP1	GP2

**Table 1.3.** Surface and transmembrane units of select enveloped viruses.

Whether the virus is entering the cell at the plasma membrane or within an endosome after endocytosis, the envelope proteins behave in a manner that is characteristic of Class I fusion proteins (Figure 1.4). After the SU of the envelope protein binds to the receptor on the host cell the envelope protein goes through a series of conformational changes that expose the fusion peptide within the TM domain. These conformational changes may also result in the dissociation of the SU protein. The TM domain then undergoes a conformational change that leads to the highly stable post-fusion structure of the “six-helix bundle”. This “six-helix bundle” is formed by a trimer of parallel  $\alpha$ -helices (HR1) surrounded by antiparallel  $\alpha$ -helices (HR2) (Figure 1.5). This is a defining structure of the Class I fusion proteins and its presence has been used to justify including a virus within this particular class. This is a highly stable structure (e.g., the melting temperature of the HIV gp41 ectodomain is  $>100^{\circ}\text{C}^7$ ) and the energy released in the transition from the metastable pre-fusion state into this post-fusion state is believed to drive the fusion of the viral and host cell membranes.



**Figure 1.4.** General mechanism of class I viral entry. Upon binding the host cell receptor the SU domain and the TM domain undergo a series of conformational changes that bring the viral and host cell membranes in close proximity and promote fusion. In this diagram the envelope protein is depicted as a monomer for simplicity.



**Figure 1.5.** Examples of the post-fusion structures characteristic of Class I fusion proteins. A) Influenza HA2 B) SIV gp41 C) Ebola GP2. In this diagram the viral membrane is oriented at the top, the HR1 are colored blue and the HR2 are colored green.

Another defining characteristic of the Class I fusion proteins is the fusion peptide. These fusion peptides do not share sequence identity but generally consist of ~16-36 residues, tend to be rich in glycine and alanine residues, and contain several bulky hydrophobic residues<sup>8,9</sup>. The fusion peptide has different relative locations in different viruses. For example, the fusion peptide of HIV and Influenza is located at the N terminus of the TM subunit and the fusion peptide of Ebola is located internally. In some viruses, e.g. SARS-CoV, the fusion peptide has not yet been identified.

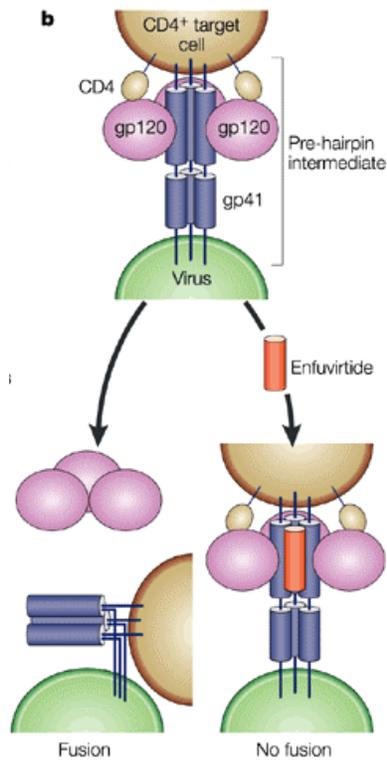
## **1.4 Envelope Proteins as Therapeutic Targets**

The physical and economic cost of the various enveloped viruses demands that there is a continued effort to find effective treatments against these viruses. For example, the Influenza outbreak of 1916 was responsible for over 50 million deaths, with some estimates as high as 100 million deaths<sup>10</sup>. As a consequence of the critical function of the envelope proteins, they are clearly attractive targets for antiviral treatments. As discussed by Caffrey<sup>11</sup> inhibitors targeting the envelope proteins can be divided into 3 groups based on the target of the drug: i) formation of the free state (preventing the envelope protein from maturing), ii) formation of the receptor-bound state, and iii) formation of the fusion state. Some examples of HIV-1 inhibitors are shown in Table 1.4.

One of the more successful entry inhibitors developed for HIV-1 is known as T-20/enfuvirtide, which would fall into the third group of Table 1.4. T20, which is a FDA-approved therapy that is widely used in the United States and Europe, exploits the transition from pre-fusion conformation to the post-fusion “six-helix bundle”. T-20 is a synthetic peptide that corresponds to the 36 amino acid sequence of the HR2 domain of gp41 (TM). It is thought to bind to the HR1 domain and prevents the transition from the pre-fusion state to the post-fusion state (Figure 1.6). It is believed that T-20 is effective during a kinetic window that begins at CD4<sup>+</sup> binding and ends at co-receptor attachment<sup>12</sup>.

<b>Target of Entry Inhibitor</b>	<b>Example from HIV-1 Therapies</b>
Formation of the free state (envelope maturation)	Furin Inhibitors: alpha-2 macroglobulin alpha-1-antitrypsin gp160 processing disruption: UK-201844
Formation of the receptor-bound state	Monoclonal antibodies: PRO 140 (CCR5) Ibalizumab (CD4) Peptides: 12p1 (gp120) Small Molecules : BMS-806 (gp120) Silver nanoparticles
Formation of tethered/fusion state	Peptide: T-20/enfuvirtide Small Molecule: ADS-J1

**Table 1.4.** Examples for the three types of entry inhibitors that target envelope proteins from HIV-1<sup>11</sup>.



**Figure 1.6.** Mechanism of enfuvirtide inhibition of HIV-1. After gp120 binds to its receptors, conformational changes occur that leave the viral envelope proteins in an intermediate, metastable state. It is at this intermediate state that enfuvirtide binds to the HR1 domain of gp41, preventing the formation of the “six-helix bundle” and therefore inhibiting entry<sup>13</sup>.

## **1.5 Thesis Research**

Detailed structural information about viral envelope proteins is necessary for a better understanding of the viral entry mechanism, as well as for the development of new and improved anti-viral therapies, especially those targeting the pre-fusion and intermediate states of these conformational states. My work focuses on using various biochemical and NMR techniques to gather more information of the structure of the envelope proteins of SARS-CoV and HIV.

In the first part of the thesis, I have used sedimentation velocity experiments to show that SARS-CoV S2-HR2 is in a dynamic equilibrium between monomer and trimeric states. I then used circular dichroism (CD) to show that the S2-HR2 helix is stabilized at physiological ionic strength and at endosomal pH. Finally, I used NMR studies of spin-labeled protein to establish that the S2-HR2 are in a parallel orientation. Together these studies give insight into the mechanism of SARS-CoV entry, as well as the inhibitory effectiveness of HR2-based peptides

In the second part of the thesis, I studied HIV-1 envelope conformation using NMR. For this study I used a 12-mer peptide, 12p1, which had previously been shown to inhibit HIV entry, to probe gp120 conformation in the presence and absence of gp41 domains. I first characterized 12p1 in the free state by NMR. Next, I used Saturation Transfer Difference (STD) and WaterLOGSY NMR to demonstrate binding of 12p1 to gp120 and identify the 12p1 <sup>1</sup>H that are in closest contact. Interestingly, I found that the 12p1 interaction with gp120 is different in the presence of gp41 domains, suggesting that the gp120 conformation is sensitive to long-range interactions with gp41. Finally, I used STD NMR to show that 12p1 binds to HIV envelope on virus like particles. Together

these studies, give new insights into the 12p1 inhibitor binding, the gp120 conformation, and demonstrate the power to NMR to characterize large complex systems.

## 2. MATERIALS AND METHODS

### 2.1 Characterization of SARS-CoV S2-HR2

#### 2.1.1 Production and Purification of SARS CoV S2-HR2

SARS-CoV HR2 was prepared as a HIS-PG-HR2 fusion protein as previously described<sup>14</sup>. The construct contains a 6XHIS tag for purification, followed by protein G (the IgG binding domain of streptococcus protein G)<sup>15</sup>, and a TEV (tobacco etch virus) cleavage site, sequence=ENLYFQGS<sup>16</sup> for removal of the expression tag. Briefly, protein expression was achieved by growing *E. coli* strain SG13009 (Qiagen, Valencia, CA) in the presence of the appropriate plasmid in 1L of LB media supplemented with 100 µg/mL ampicillin and 50 µg/mL of kanamycin at 37°C until they reached an OD<sub>600</sub> of 0.6. Expression was induced by addition of 0.8 mM IPTG and the culture was grown for an additional 4-5 hours at 37°C. The HIS-PG-HR2 fusion protein was purified from the soluble fraction using a Ni<sup>2+</sup> fast-flow Sepharose column (Qiagen, Valencia, CA). The protein was then cleaved using TEV protease and run once more over the Ni<sup>2+</sup> column to remove HIS-PG and TEV protease, which also contains a polyhistidine tag, as well as uncut HIS-PG-HR2. The flow-through fraction containing HR2 was then dialyzed extensively against 10 mM NaPO<sub>4</sub>, pH 7.0 and concentrated by ultrafiltration (YM3; Amicon, Billerica, MA). The purity and identity of HR2 was confirmed using SDS-PAGE on a 12% Bis-Tris Gel (Invitrogen, CA) and MALDI-TOF mass spectrometry.

### 2.1.2 Analytical Ultracentrifugation of SARS-CoV S2-HR2

Sedimentation velocity experiments were performed in a Beckman ProteomeLab XL-1 analytical ultracentrifuge using an An60Ti rotor at 20°C. To test for self-association, two S2-HR2 concentrations were assayed (110 and 220  $\mu\text{M}$ ). Epon double-sector centerpieces were filled with 400  $\mu\text{L}$  of sample and buffer (10mM NaPO<sub>4</sub>, pH 7.0) and centrifuged at 50,000 rpm. Absorbance data were acquired at 280 nm with a time interval of 5 min. Buffer viscosity, protein partial specific volumes, and frictional ratios were calculated with SEDNTERP (<http://www.rasmb.bbri.org/>) developed by Hayes, Laue, and Philo. The data were fit using a continuous  $c(s)$  Lamm equation distribution model with the program Sedfit<sup>17</sup>.

### 2.1.3 Circular Dichroism of SARS Co-V S2-HR2

Circular dichroism (CD) spectra were measured on a Jasco-710 spectropolarimeter. Wavelength spectra were recorded from 190 to 260 nm at peptide concentrations of 0.7 mg/ml (110  $\mu\text{M}$ ) in cells of 0.2 mm path length. For all experiments the spectra were corrected with the subtraction of a blank corresponding to the buffer of that experiment. Buffer conditions were 10 mM NaPO<sub>4</sub>, pH 7 or 5, and 0-300 mM NaCl at 20°C. The percentage of alpha-helical residues was calculated by the observed molar ellipticity at 222 nm divided by the theoretical molar ellipticity ( $\Theta_t$ ), where ( $\Theta_t$ ) = 40,000 x (1-4.6/N) and N=55 (the total number of residues).

A)  
| 1 10 20 30 40 50  
GSHTSPDV~~DL~~GDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIK

B)

HR2-S14C forward primer

5'-CTT GGC GAC ATT TGC GGC ATT AAC GCT-3'

HR2-S14C reverse primer

5'-AGC CTT AAT GCC GCA AAT GTC GCC AAC-3'

**Figure 2.1.** Mutagenesis of SARS-CoV S2-HR2. A) Amino acid sequence of S2-HR2<sup>14</sup>.

The construct consists of residues 1141-1193, however, it has been renumbered as 1-55 for simplicity. The mutated residue has been underlined. B) Forward and reverse primers ordered from IDT (Coralville, IA) for preparation of S2-HR2 S14C.

#### **2.1.4 Production and Purification of SARS Co-V S2-HR2 S14C with a Paramagnetic Label**

The SARS-CoV S2-HR2 S14C was prepared by introducing cysteine at residue 14 (Figure 2.1A) using the Quick Change II site directed mutagenesis kit from Stratagene. The annealing temperature for PCR was 55°C for 1 hour using primers described in Figure 2.1B. The mutation was confirmed by DNA sequencing at the UIC RRC. The SARS-CoV S2-HR2 domain containing the cysteine mutation was subcloned into the *BamHI/HindIII* restriction sites of a modified pQE30-expression vector (Qiagen, Valencia, CA).

Protein expression of <sup>15</sup>N-labeled S14C S2-HR2 was achieved by growing *Escherichia coli* strain SG13009 (Qiagen, Valencia, CA) in 1L of LB media supplemented with 100 µg/mL ampicillin and 50 µg/mL of kanamycin at 37°C until they reached an OD<sub>600</sub> of 0.8. The cells were then pelleted by centrifuging at 5,000 x g for 30 minutes, washed once with M9 salts, resuspended in 250 mL of M9 minimal media supplemented with 1 g/L <sup>15</sup>NH<sub>4</sub>Cl (Martek Biosciences, Columbia, MD), and set to recover in the absence of antibiotic selection. Protein expression was induced after 1 hour by addition of IPTG to a final concentration of 1 mM and grown for an additional 4 hours at 37°C.

The labeled S14C HIS-PG-HR2 fusion protein was purified from the soluble fraction using a Ni<sup>2+</sup> fast-flow Sepharose column (Qiagen, Valencia, CA). The protein was then cleaved using TEV protease, which was prepared in our lab, and run once more over the Ni<sup>2+</sup> column to remove HIS-PG and TEV protease, which also contains a polyhistidine tag, as well as uncut HIS-PG-HR2. The flow-through fraction containing

S14C S2-HR2 was then dialyzed extensively against 10 mM NaPO<sub>4</sub>, pH 7.0 and concentrated by ultrafiltration (YM3, Amicon, Billerica, MA). The purity of S14C S2-HR2 was confirmed using SDS-PAGE on a 12% Bis-Tris Gel (Invitrogen, CA) under reducing conditions.

Purified protein was quantified by UV-VIS absorbance; any possible disulfide bonds were reduced using 5X molar excess of DTT. Excess reductant was removed using a Micro-Bio Spin chromatography column (Bio-Rad, Hercules, CA), then the reduced protein was incubated with a 7X molar excess of MTSL (*S*-(2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl) methyl methanesulfonothioate) overnight at room temperature in 10 mM TRIS-HCl, pH 7.4, 0.02% sodium azide, to label the protein with the MTSL spin label at the cysteine residue.

### **2.1.5 NMR Spectroscopy**

The experimental conditions for the NMR experiments were 1 mM S2-HR2 in 10 mM NaPO<sub>4</sub>, pH 7.0, 30% TFE-*d*<sub>3</sub> (Cambridge Isotope Laboratories, Andover, MA) and 10% <sup>2</sup>H<sub>2</sub>O or 1 mM S2-HR2 in 10 mM NaPO<sub>4</sub>, pH 7.0, and 10% <sup>2</sup>H<sub>2</sub>O. Spectra were recorded at 25°C on a Bruker AVANCE 800 MHz spectrometer equipped with a cryogenic triple resonance probe. Processing of the spectra was done with NmrPipe and visualized with NmrDraw<sup>18</sup>.

## **2.2 Characterization of HIV Envelope**

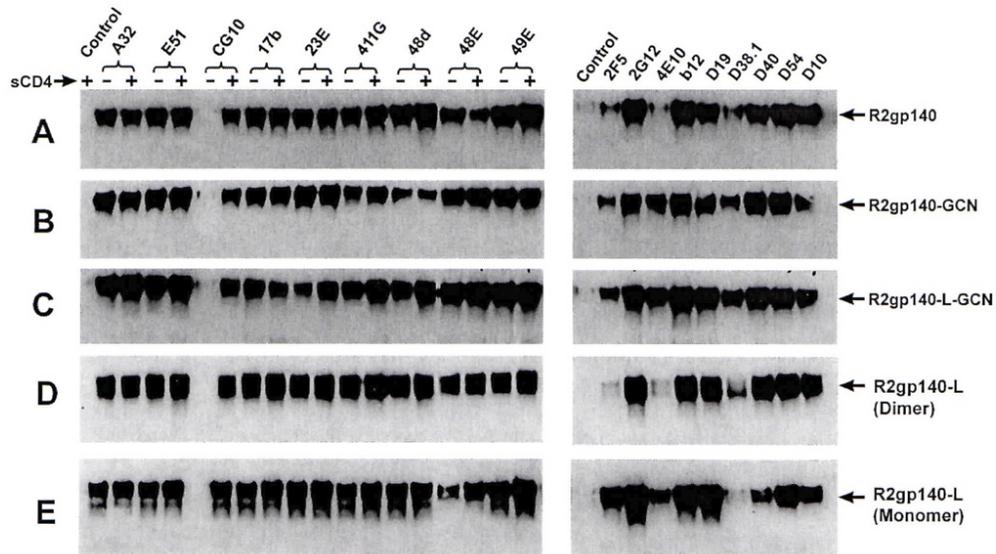
### **2.2.1 Peptide Synthesis**

The 12-mer peptide, 12p1 (sequence = RINNIPWSEAMM), was synthesized by solid phase peptide synthesis with unmodified N- and C- terminal groups by the UIC RRC. The peptide was purified by HPLC and the mass was verified by MALDI-TOF mass spectrometry. In our initial work we observed additional resonances within the NMR spectra that were from oxidation of the tryptophan indole, which is especially sensitive to oxidation, as previously observed for other peptides and proteins<sup>20</sup>. This oxidation was verified by MALDI-TOF mass spectrometry; and thus the peptide solid was stored under vacuum and all NMR solutions were stored in the presence of argon to avoid oxidation.

### **2.2.2 Preparation of recombinant HIV envelope glycoproteins**

The gp120 and gp140 from strain R2 were prepared in the laboratory of Christopher Broder (Uniformed Services University of the Health Sciences, Bethesda, MD). Briefly, a codon optimized R2 Env gene was cloned into the vector phCMVhygro, which was modified from the plasmid phCMV (Genlantis, Inc) by introduction of a hygromycin resistance gene (phCMV is a CMV promoter and enhancer driven cassette). For preparation of R2 gp140, the R2 Env gene was truncated at the transmembrane domain and the arginine residues at both gp120-gp41 cleavage sites were mutated to serine (R517S, R520S) and the GCN4 trimerization domain was fused to the carboxy terminus of R2 gp140 to form R2 gp140-GCN, as described<sup>20, 21</sup>. Monomeric R2 gp120

was produced as previously described<sup>22</sup>. The R2 gp140 envelopes were produced by stable cell line expression. All R2 envelope proteins were prepared by expression in serum-free conditions using 293 T cells, and purified using a sequential combination of DEAE adsorption, lentil lectin affinity, and size exclusion chromatography. The purity of the recombinant R2 envelope protein preparations was assessed by SDS-PAGE. The oligomerization state of the R2 gp140 envelopes were assessed by Blue Native PAGE; R2 gp140-GCN and R2gp140-linker-GCN migrate as trimers (molecular weight 720 kDa). The conformational integrity of the recombinant R2 envelope proteins was demonstrated by their binding to soluble CD4, CCR5, and a series of monoclonal antibodies (Figure 2.2). Gp120 from strain YU2 was obtained from Immune Technology Corporation. Gp120 from strain 96ZM65 was obtained from the NIH AIDS Research and Reference Reagent Program. Note that all gp120 and gp140 constructs produced in eukaryotic cells are expected to be fully glycosylated.



**Figure 2.2** Monoclonal antibody binding analysis (right panel) and reactivity to CD4i and CD4-gp140 complex specific mAbs (left panel) of purified R2 gp140 dimer (A), R2 gp140+GCN trimer (B), and R2 gp140+linker+GCN trimer (C). For the right panel, 1  $\mu$ g of different versions of purified R2 gp140 was incubated with or without excess (3  $\mu$ g) of sCD4 in 700  $\mu$ l of reaction buffer (PBS containing 0.5% Triton X-100) at 4°C for overnight, followed by 2  $\mu$ g of the indicated mAb for an additional 4 hours. For the left panel, 1  $\mu$ g of different versions of purified R2 gp140 was incubated with the indicated mAbs in 700  $\mu$ l of reaction buffer for 4 hours at 4°C. In both cases, the complexes were then precipitated with 50  $\mu$ l of Protein G Sepharose (20% solution) for an additional 2 hours at 4°C. The samples were washed three times with lysis buffer (0.1 M Tris-HCl, pH 8.0, 0.1M NaCl, 0.1% Triton X-100). The precipitated complex was resuspended in SDS-PAGE sample buffer, boiled for 5 min and resolved on 4-12% Bis-Tris SDS-PAGE followed by western blotting. The blots were then probed with a polyclonal rabbit anti-gp140 antiserum. Note that these experiments were performed in the Broder laboratory.

### 2.2.3 NMR Experiments

NMR experiments were performed on Bruker 600, 800, and 900 MHz spectrometers equipped with room temperature or cryogenic temperature triple resonance probes. Buffer conditions were either 50 mM PO<sub>4</sub>/pH 7.2 or PBS (20 mM PO<sub>4</sub>/pH 7.2, 150 mM NaCl) in 90% <sup>1</sup>H<sub>2</sub>O, 10% <sup>2</sup>H<sub>2</sub>O or 100% <sup>2</sup>H<sub>2</sub>O at 25°C. NMR assignments were obtained by a series of TOCSY, ROESY (mixing time=100 msec), NOESY (mixing time=200 msec), <sup>13</sup>C-edited HSQC, and <sup>15</sup>N-edited HSQC. STD NMR experiments were performed with a 50 msec Gaussian-shaped saturating pulse for 2.5 sec with “on” resonance saturation at -1.5 ppm and “off” resonance saturation at 30 ppm (the total relaxation delay was 3.5 sec). The WaterLOGSY experiments were performed as previously described<sup>23</sup>. In the WaterLOGSY experiments the data was collected with a sweep width of 12820 Hz, and acquisition time of 1.28 seconds and a relaxation delay of 2.5 seconds. Water was selectively saturated using a 2 msec square shaped pulse. All data were processed by NMRPipe with referencing to the water peak at 4.773 and using a 10 Hz line-broadening window.

### 2.2.4 Production of Virus Like Particles

The virus like particles (VLPs) were prepared in the laboratory of Amy Jacobs (State University of New York, Buffalo), as previously described<sup>24</sup>. Briefly, the cell line used for all transfection experiments was 293T, which are human embryonic kidney cells (HEK). The cells were grown in DMEM containing 4.5 g/L glucose and L-Glutamine (Mediatech) supplemented with 10% FBS and 1% penicillin-streptomycin (Mediatech) and 0.5 mg/mL G418 (Mediatech).

The plasmids used to produce HIV-1 envelope proteins were pHXB2-env (NIH AIDS Reference and Reagent Program) and JRFL (NIH AIDS Reference and Reagent Program). Both plasmids are HIV-1 Rev dependent and therefore require co-transfection with a Rev expression construct (e.g. pCMV-rev or pNL4-3.Luc.R-E-) for envelope protein expression. The plasmid pNL4-3.Luc.R-E- (NIH AIDS Reference and Reagent Program) contains the HIV-1 cDNA with the Firefly luciferase gene inserted into the *nef* gene and two frame shifts resulting in blunted *env* and *vpr* genes. Due to these mutations three independent mutations would have to occur to produce a fully replicative virus, therefore this plasmid can be co-transfected with envelope expression plasmids to generate a non-replicative HIV-1 virus-like particle (VLP).

The 293T cells were plated at a density of  $2 \times 10^6$  per  $10 \text{ cm}^2$  dish the day before transfection. A total of 20  $\mu\text{g}$  DNA per plate was transfected using a standard calcium phosphate protocol (10  $\mu\text{g}$  each of pHXB2-env/JRFL and pNL4-3.Luc.R-E-). 48 hours post transfection the medium of 293T cells, which contains the non-replicative virus-like particles released from the cells, was harvested, filtered through sterile 0.45 micron filter (Nalgene) and stored at  $-70^\circ\text{C}$  for later use. Virus titer was determined by X-Gal staining based on the titration method adapted as previously described<sup>25</sup>. Briefly,  $8 \times 10^4$  TZM-bl cells per well were seeded in a 12 well plate with complete DMEM the day before infection. The following day, the medium was removed and replaced with diluted virus in a total volume of 200  $\mu\text{L}$  in the presence of 20  $\mu\text{g}/\text{mL}$  DEAE-dextran. The virus inoculated cells were incubated for two hours at  $37^\circ\text{C}$ , then 1 mL of complete DMEM was added, followed by a further incubation of 2 days. The medium was then removed and the cells were fixed with 2 mLs of fixation buffer (1% formaldehyde, 0.2%

gluteraldehyde, PBS) for 5-10 minutes at room temperature. The fixed cells were washed three times with PBS and incubated with 500  $\mu$ L of X-Gal staining buffer (0.4 mg/mL X-Gal in 4 mM potassium ferrocyanide, 4 mM potassium ferricyanide, 2 mM magnesium chloride, PBS) for 24 hours at 37<sup>0</sup>C. Colonies of blue cells were counted under the microscope at a magnification of 100X and infection viral titers were quantified as summarized in Table 2.1.

The VLPs were further characterized using a p24 antigen capture assay, performed by the Chicago Developmental Center for AIDS Research (CFAR) . Briefly, 20  $\mu$ L of 5% Triton X-100 was added to all uncoated microplate wells (except the blank) then 200  $\mu$ L of either sample, negative control or positive control was added to the designated wells. The plate was then sealed and incubated for two hours at 37°C. The plate was washed 6 times with 300  $\mu$ L of wash buffer per well. After the washing step, 100  $\mu$ L of Detector Antibody is added to all wells (except the blank), the plate was sealed and incubated for 60 minutes at 37°C. After this incubation the plate was washed as described above and 100  $\mu$ L of diluted Streptavidin-Horse Radish Peroxidase (SA-HRP) was added to all wells (except the blank), the plate was sealed and incubated for 30 minutes at room temperature. The plate was washed in the same manner as above then 100  $\mu$ L of ortho-phenylenediamine-HCl (OPD) Substrate solutions was added to all wells, the plate was sealed and incubated for 30 minutes at room temperature in the dark. The reaction was stopped by adding 100  $\mu$ L of Stop Solution to all wells and the plate was read at 490 nm using a plate reader. Final results of the p24 assay are shown in Table 2.2.

label	plasmid	volume (ul)	# vials	titer (infectious units/mL)	virus/vial	total virus
No env	-	25	35	-	-	-
VSVG	pVSVG	25	63	14,500,000	36,250	2,283,750
JRFL (J1)	pJRFL-160	25	27	550,000	13,750	371,250
HXB2	pHXB2-160	25	81	26,500	663	53,663

**Table 2.1** Titers of VLP.

Sample ID	1 uL (pg/mL)	24 uL (pg/mL)
No ENV	7,681	8,542
VSVG	7,856	7,397
J1	7,393	9,612
- ENV	3,255	7,118
H5	3,041	8,081
H1	294	2,995
Marburg	7,504	9,225

**Table 2.2** Results from P24 Assay. The samples were performed at 2 different concentrations (1  $\mu$ l and 24  $\mu$ l of stock vlp) to assess reproducibility.

### 3. SARS-CoV S2-HR2 IS A PARALLEL TRIMER

#### 3.1 SARS Outbreak

In November of 2002 an atypical pneumonia emerged in the Guangdong Province of China, quickly spreading to 20 countries in Asia, North America, South America and Europe. This disease, known as Severe Acute Respiratory Syndrome (SARS), affected more than 8,000 people and was the cause of over 700 deaths by the time the outbreak was contained with quarantine measures and travel restrictions in the summer of 2003. Strikingly, the death rate from SARS was ~10%, which is ~4X that of the influenza outbreak of 1918, a pandemic that resulted in greater than 40 million deaths<sup>26</sup>.

The causative agent of SARS was identified as a new coronavirus named SARS Coronavirus (SARS-CoV); previous to the identification of SARS-CoV, coronaviruses in humans had only been associated with mild diseases. Coronaviruses had previously been divided into three separate groups, group 1 and 2 containing mammalian viruses with group 3 containing only avian viruses, however, SARS-CoV has a low sequence homology to other members in the *Coronaviridae* family and therefore has been classified in a new group, group 4<sup>27-30</sup>.

The 2002-2003 outbreak of SARS was quickly contained and no human infections have been reported since July of 2003, however, there are still animal reservoirs of coronaviruses and therefore it is impossible to say that SARS-CoV or another coronavirus will not return. That distinct possibility, along with a lack of

effective anti-viral therapy for SARS-CoV, makes it important to learn as much as we can about how this virus works, especially the mechanism of entry, to aid in developing anti-viral therapies and vaccines. Moreover, SARS-CoV bears similarities to other enveloped viruses including Ebola, HIV and Influenza. As a consequence, study of SARS-CoV is expected to lend new insights into envelope-mediated viral entry mechanisms.

### **3.2 Entry of SARS-CoV**

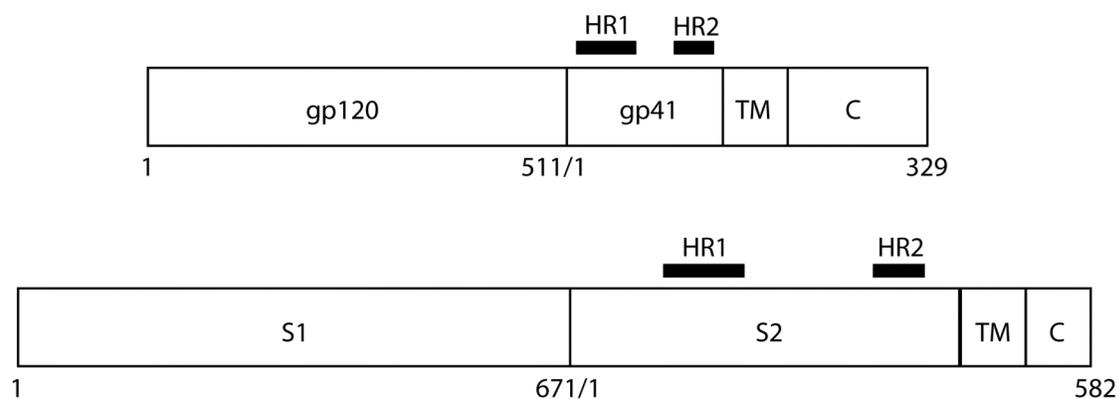
As discussed in Chapter 1 enveloped viruses enter via one of two pathways: extracellular or endosomal. Identification of the route of entry is a vital step in developing entry inhibitors and SARS appears to utilize both routes of entry. Early studies showed that SARS-CoV infection was blocked by lysosomotropic agents, which inhibit the acidification of the endosome, suggesting that SARS-CoV enters cells through the endosomal pathway and that the low pH environment of the endosome is a vital part of the entry process<sup>31, 32</sup>. However, it has also been reported that SARS-CoV is sensitive to proteolytic cleavage by the enzymes trypsin, chymotrypsin and elastase<sup>31, 33, 34</sup>. Treatment of virus attached to the target cells by these proteases resulted in viral infectivity, even when the endosomal entry pathway was blocked with endosome-tropic agents, which leads to the conclusion that extracellular entry via proteolytic cleavage may also be utilized by SARS-CoV. In fact, entry that is triggered by proteolytic cleavage was 100 times more efficient than infection via an endosomal pathway<sup>34</sup>. All this information taken together has led to a model of SARS-CoV entry that is tissue specific.

It appears that when in the lung tissue, a primary site of SARS-CoV entry and an environment that has a relatively high concentration of proteases, SARS-CoV enters via the more efficient extracellular pathway mediated by proteolytic cleavage. However, in other cell types where there are not high levels of proteases, the virus enters via the endosomal pathway and is then cleaved by cathepsin L within the endosome<sup>35</sup>. Regardless which pathway SARS-CoV enters target cells, the envelope protein (termed spike or S) is playing a critical role in that entry process. Moreover, S is the only protein present on the viral surface and thus the only target for neutralizing antibodies. Knowing as much information as possible about S can aid in the discovery of new therapies and possibly vaccines for SARS-CoV.

### **3.3 SARS-CoV Spike Protein and Putative Homology to Viral Envelope Proteins**

SARS-CoV S consists of 2 extracellular domains termed S1 and S2, which are followed by transmembrane and cytoplasmic domains. Based on studies of related CoV, the S1 domain binds to cellular receptors and the S2 domain is responsible for mediating fusion of the viral and target membranes<sup>36, 37</sup>. In Figure 3.1, a schematic diagram of SARS-CoV spike is shown. The SARS-CoV S1 possesses multiple cysteines, which are often involved in the formation of disulphide bonds<sup>38</sup>, and numerous putative glycosylation sites. With respect to the SARS-CoV S1, S2 possesses fewer cysteines (10 versus 19) and fewer putative glycosylation sites (4 versus 13). The SARS-CoV S2, as well as the S2 of other CoV, is predicted to contain 2 heptad repeat regions, termed HR1

and HR2, which are expected to form coiled coils and be involved in fusion of the viral and target membranes<sup>30, 39</sup>. Interestingly, CoV spike share many similarities to the well-characterized envelope proteins of retroviridae, herpesviridae, filoviridae and paramyxoviridae<sup>40-42</sup>. For example as depicted in Figure 3, the N-terminal region of the HIV envelope protein, termed gp120, is responsible for binding to receptor; the C-terminal domain of the HIV envelope protein, termed gp41, is responsible for the fusion process<sup>43</sup>. Moreover, HIV envelope exhibits a similar distribution of cysteines and glycosylation sites (i.e. gp41, the S2 homolog, possesses fewer cysteines and glycosylation sites than gp120, the S1 homolog). Thus, SARS-CoV spike is expected to be functionally and structurally, at least at the level of gene arrangement, similar to other well-characterized viral envelope proteins.



**Figure 3.1.** Schematic diagrams of SARS-CoV spike (bottom) and HIV envelope (top). Numbering corresponds to that of the cleaved forms. The locations of the heptad repeats are denoted by horizontal bars. TM and C represent the transmembrane and cytoplasmic domains, respectively.

### **3.4 SARS-CoV S2-HR2**

As stated in Section 1.3, peptides that mimic parts of the envelope proteins have been proven to be effective inhibitors for other enveloped viruses (i.e. T-20 for HIV-1). Consequently, the analogous peptide regions in SARS-CoV envelope, termed S2- HR1 and S2-HR2, have become attractive areas of study. For example, a number of peptides that are derivatives of the HR1 and HR2 regions of the S2 domain have exhibited inhibition of entry (Table 3.1). Of the HR1/HR2 derivatives tested the best inhibitor had an  $EC_{50}$  value of 17  $\mu$ M; compared to T-20 for HIV-1, which has an  $EC_{50}$  value in the nanomolar range, and thus the HR2 are relatively poor inhibitors and need to be improved. Interestingly, the inhibitory properties are very sensitive to the length and sequence of the HR2 region. A better understanding of the mechanism of entry, especially the conformational changes that the envelope proteins undergo, is a step towards the development of better inhibitory peptides.

Peptide	Amino Acid Sequence	EC <sub>50</sub> $\mu$ M
HR2		
sHR2-1	ELDSPKEELDLYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEAVAKNLNEDLIDLQELGKYE	43 +/- 6.4
sHR2-2	PKEELDLYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEAVAKNLNEDLIDLQELGKYE	24 +/- 2.8
sHR2-3	LDLYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEAVAKNLNEDLIDLQELGKYE	>50
sHR2-4	FKNHTSPDVDLGDISGINASVVNIQKEIDRLNEAVAKNLNEDLIDLQELGKYE	>50
sHR2-5	TSPDVDLGDISGINASVVNIQKEIDRLNEAVAKNLNEDLIDLQELGKYE	>50
sHR2-6	DISGINASVVNIQKEIDRLNEAVAKNLNEDLIDLQELGKYE	>50
sHR2-7	ELDSPKEELDLYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEAVAKNLNEDLIDLQELGKYEQYIK	>50
sHR2-8	ELDSPKEELDLYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEAVAKNLNEDLIDLQEL	17 +/- 3.0
sHR2-9	ELDSPKEELDLYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEAVAKNLNEDLID	34 +/- 4.0
sHR2-10	ELDSPKEELDLYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEAVAKNLNEDLIDLQELGKYE	>50
HR1		
sHR1	AYRFNGIOVTQNVLYE- NQKQIANQFNKAISQIQESLTTTSTALGKQLQDVVNQNAQALNTLVKQLSSNFGAISSVNLNDILSRDKVEAEVQIDRLIT	>50
sHR1a	NQKQIANQFNKAISQIQESLTTTSTALGKQLQDVVNQNAQALNTLVKQLSSNFGAISSVNLNDILSRDKVEAEVQIDRLIT	>50
sHR1b	NQKQIANQFNKAISQIQESLTTTSTALGKQLQDVVNQNAQALNTLVKQLSSNFGAISSVNLNDILSRDKVEAEVQIDRLIT	>50
sHR1c	NQNAQALNTLVKQLSSNFGAISSVNLNDILSRDKVEAEVQIDRLIT	>50

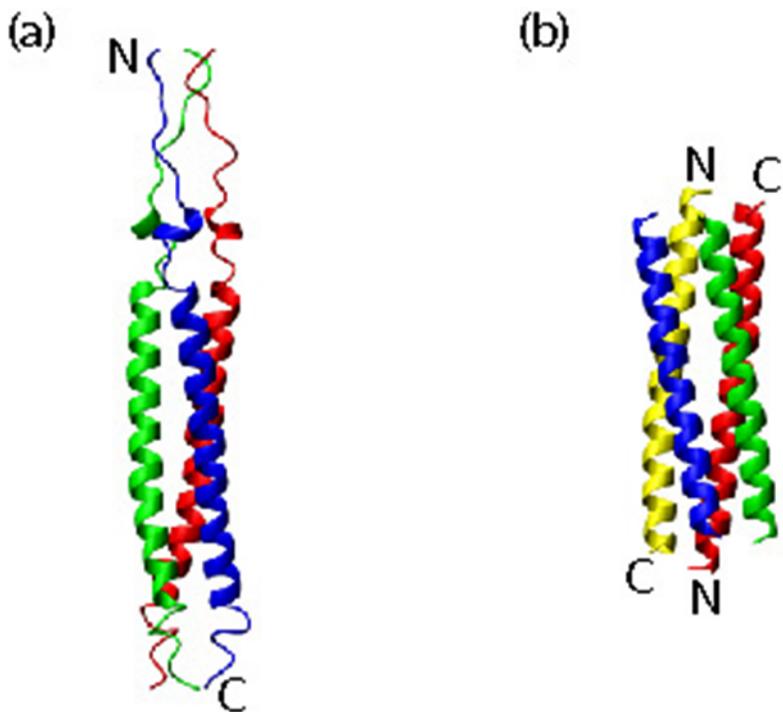
**Table 3.1.** HR1 and HR2 derived peptides were tested for as inhibitors for SARS Co-V entry. Amino acid sequence and EC<sub>50</sub> values are shown<sup>33</sup>.

### **3.5 Structural Studies of SARS-CoV S**

Structural studies of the HR1 and HR2 of SARS-CoV, as well as the HR1 and HR2 of mouse hepatitis virus (MHV), a related CoV, demonstrate that they form a "six helix bundle" that is analogous to that formed by other TM proteins such as HIV/SIV gp41<sup>33,44-51</sup>. More recently the S2-HR2 have been studied in isolation to lend insight into prefusion conformations of S2<sup>14, 51-54</sup>. Interestingly, the previous studies of SARS-CoV S2-HR2 have resulted in two opposing structures for the prefusion state<sup>14, 51</sup>. The TFE stabilized NMR structure indicated that the prefusion state is a parallel trimer of  $\alpha$ -helices, while the x-ray solution is that of an anti-parallel tetramer of  $\alpha$ -helices (Figure 3.2). In both cases, analytical ultracentrifugation confirmed the oligomeric state (the NMR construct was trimeric in the presence of 30% TFE and the x-ray construct was tetrameric under aqueous conditions). Nonetheless, there could be several explanations for the disparity. First, the NMR structure was determined in the presence of TFE, a co-solvent used to stabilize helical structures. Second, the NMR construct was 11 residues larger than the x-ray construct as shown in Figure 3.2. Third, the determination of a protein complex by NMR is relatively demanding because it is based on a small number of intermolecular NOEs that could possibly be misinterpreted. Therefore, it will be important to resolve this issue for a better understanding of S2-mediated viral entry and the mechanism of S2-HR2 inhibitory peptides.

### **3.6 Overview of Work Presented in Chapter 3**

My work consists of three parts. In the first part, I used analytical centrifugation to demonstrate that under aqueous conditions the SARS-CoV S2-HR2 is in equilibrium between monomeric and trimeric forms. In the second part, I used circular dichroism to show that the S2-HR2 helix is stabilized at physiological ionic strength and the pH of the endosome. Finally, I used paramagnetic relaxation enhancement (PRE) to confirm that the solution structure of S2-HR2 consists of parallel helices. Together these studies will be exploited to lend insight into the mechanism of SARS-CoV entry.



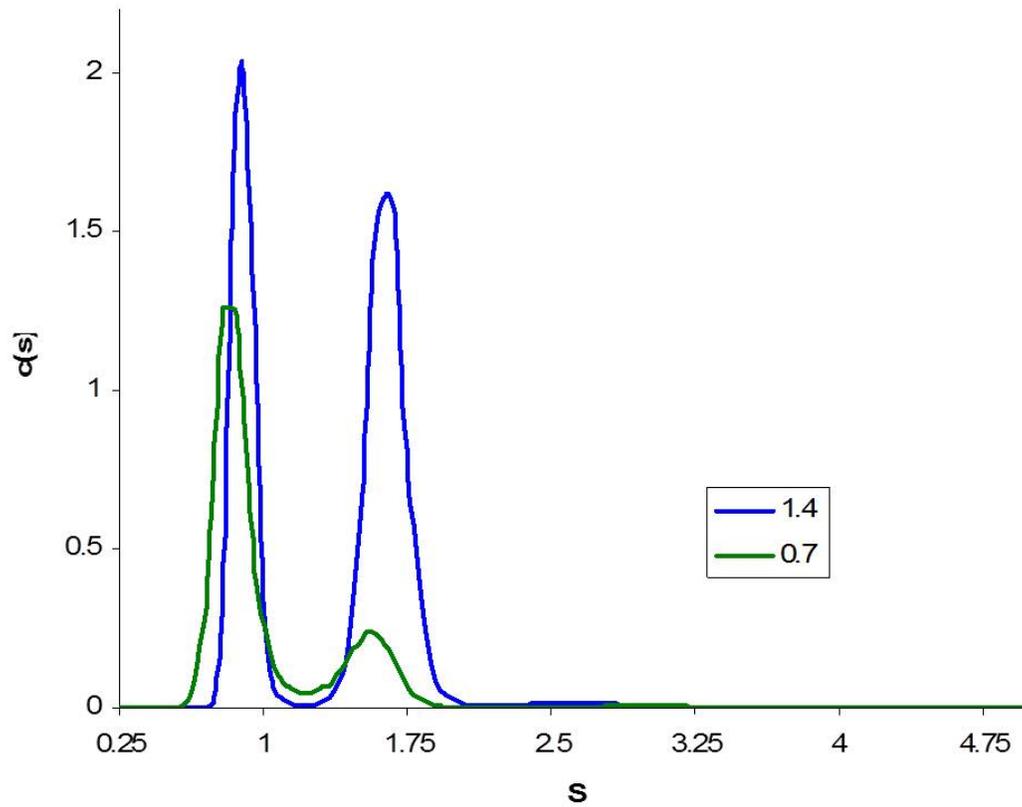
<b>Method</b>	NMR	X-Ray
<b>Organization</b>	Parallel trimer	Anti parallel tetramer
<b>Concentration</b>	1 mM	10 mg/mL (~1.5 mM)
<b>Buffer Conditions</b>	10 mM NaPO <sub>4</sub> /pH 7.0	0.1 M TrisHCl/pH 8.5, 10mM NiCl <sub>2</sub>
<b>Additives</b>	30% TFE	PEG MME 2000
<b>Residues</b>	1141-1193 (55 res total)	1150-1193 (44 res total)

**Figure 3.2.** Available structures of SARS-CoV S2-HR2. (A) The NMR Solution structure was solved as a parallel trimer of  $\alpha$ -helices<sup>14</sup>. (B) The x-ray structure was solved as a tetramer of anti-parallel  $\alpha$ -helices<sup>51</sup>. (C) Conditions for each structure determination.

## **3.7 Results and Discussion**

### **3.7.1 SARS-CoV S2-HR2 exists in a monomer/trimer equilibrium under aqueous conditions**

To determine the oligomeric state of the S2-HR2 construct under aqueous conditions (i.e. in the absence of TFE) I did a series of sedimentation velocity (SV) experiments using an analytical ultracentrifuge. The SV studies were able to show that the S2-HR2 construct exists in two distinct oligomeric states, monomer and trimer; and that higher concentrations of the peptide favored the trimeric state, suggesting that the monomer and trimer are in equilibrium (Figure 3.5). My SV result is consistent with previous studies that suggested a monomer-trimer equilibrium based on intermediate  $^{13}\text{C}\alpha$  secondary shifts,  $\text{H}_\text{N}$  exchange rates, and relaxation rates with respect to the TFE stabilized trimer<sup>14, 52, 53</sup>.

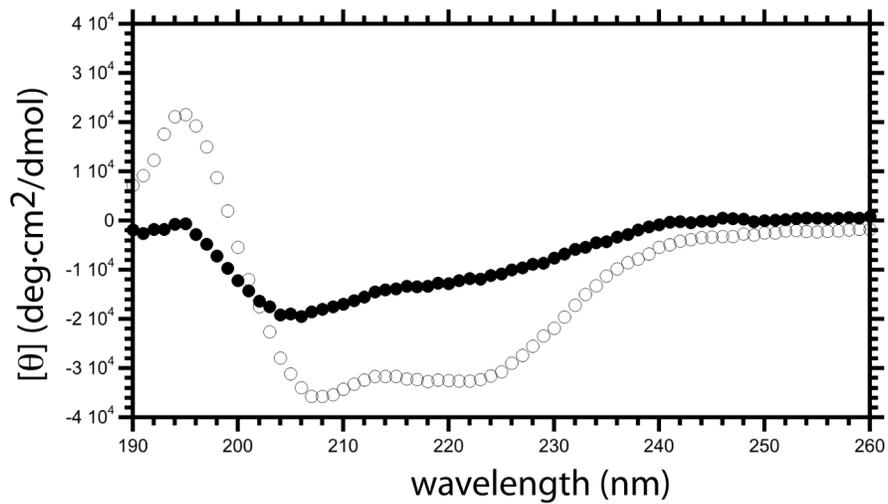


**Figure 3.3.** SV studies of SARS-CoV S2-HR2. The blue line represents a protein concentration of 1.4 mg/mL sample, the green line represents a protein concentration of 0.7 mg/mL. The peak with an S value of  $\sim 1$  corresponds to monomeric peptide, the peak at  $\sim 1.75$  corresponds to trimeric peptide.

### 3.7.2 Effect of ionic strength and pH on the structure of S2-HR2

In this part of the study I used circular dichroism (CD) experiments to determine the helical content of S2-HR2 under different solution conditions (ionic strength and pH). This is important in two respects. First, it has previously been shown that increased helical content is correlated with formation of the trimer<sup>14</sup>; therefore, the helical content under different solution conditions is expected to give insight into the oligomerization state of S2-HR2. Note that the S2-HR2 monomer would be expected to form a better interaction with HR1 and thus be a better inhibitor. Second, in most cell types SARS-CoV envelope-mediated membrane fusion occurs in the endosome at pH 5.

Accordingly, I first tested the effect of pH on SARS-CoV HR2 structure at pH 7 and 5, the pH of the extracellular and endosomal spaces, respectively. As shown in Figure 3.4, the circular dichroism spectrum at pH 7 reveals the presence of helical structure (as deduced by the minima at 208 and 222 nm), which is estimated to be ~32% (~18 of 55 residues). At pH 5, the helical content of HR2 increases substantially to ~90% (~50 of 55 residues).



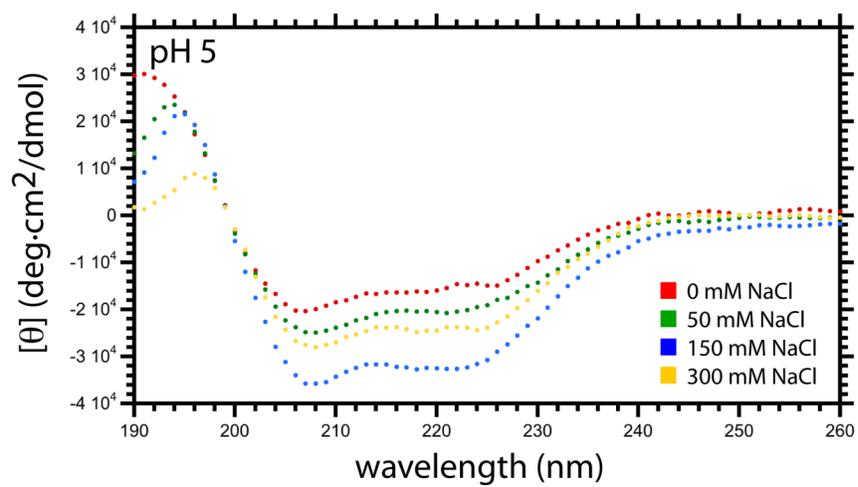
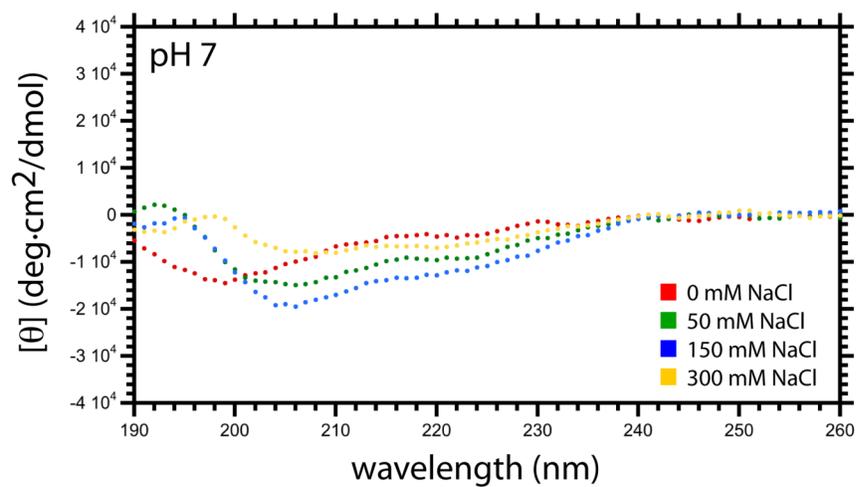
**Figure 3.4** Circular dichroism studies of SARS-CoV HR2 at pH 7 (filled circles) and pH 5 (open circles). The experimental conditions were 110  $\mu\text{M}$  HR2 in 10 mM  $\text{NaHPO}_4$ , 150 mM  $\text{NaCl}$  at 20°C.

I next tested the effects of ionic strength on SARS-CoV HR2 structure at pH 7. As shown in Figure 3.5a, at pH 7 the helical content of HR2 increases from ~13% at low ionic strength to a maximum of 32% at physiological ionic strength. In a similar pattern, at pH 5 the helical content of HR2 increases from ~40% at low ionic strength to a maximum of ~90% at physiological ionic strength (Figure 3.5b). At both pH, the helical content decreases at the highest ionic strength measured (Figure 3.6). The effects of ionic strength are in agreement with previous observations for other peptides, for which NaCl has helix stabilizing properties at low concentration and helix destabilizing properties at relatively high concentrations. This effect has been attributed to the interaction of Hofmeister ions<sup>55, 56</sup>.

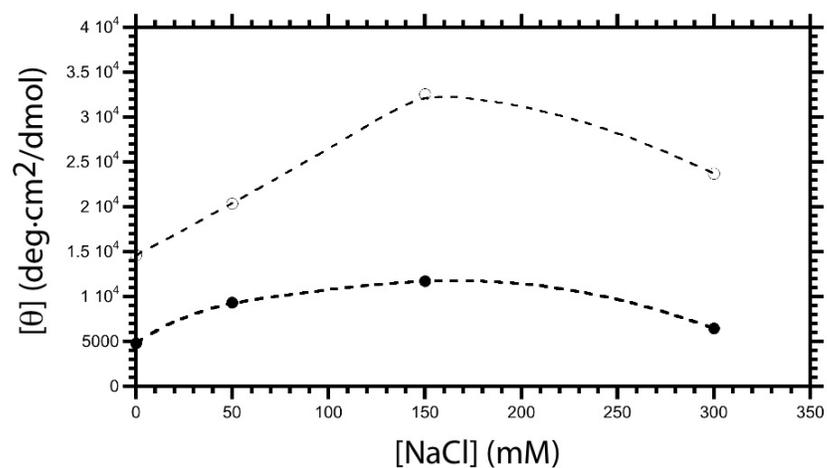
To interpret the effects of pH and ionic strength, it is of interest to consider the structure of the SARS-CoV HR2 trimer<sup>14</sup>. As shown in Figure 3.7, there are numerous patches of like-charged sidechains in the TFE-stabilized helical trimer. For example, there are 4 regions of potential charge repulsion: (1) the negatively charged sidechains of D7, D9 and D12; (2) the positively charged sidechains of K25 and R29; (3) the negatively charged sidechains of E26, D28 and E32; (4) and the negatively charged sidechains of E39, D43 and E46. Increasing ionic strength would be expected to reduce the charge repulsion present in the helical form and thus stabilize the helical form with respect to the extended coil form. Furthermore, at pH 5 the acidic groups are expected to be partially protonated, thereby further reducing charge repulsion. Interestingly, in Influenza HA2, a region that connects two heptad repeat domains also exhibits pH-induced helix formation, presumably due to partial protonation of acidic sidechains and

the reduction of electrostatic repulsion in the helical conformation<sup>57</sup>. In the case of Influenza, the formation of helix in this region triggers a large conformational change in HA2 structure that is intrinsic to the viral entry mechanism in the endosome<sup>57, 58</sup>. Thus, it is tempting to speculate that the SARS-CoV HR2 has evolved to favor the transition from coil to helix in the endosome. Moreover, based on the present results HR2-based peptides would form helical trimers in the endosome, possibly reducing their ability to interact with HR1 and thus reducing efficacy as an inhibitor.

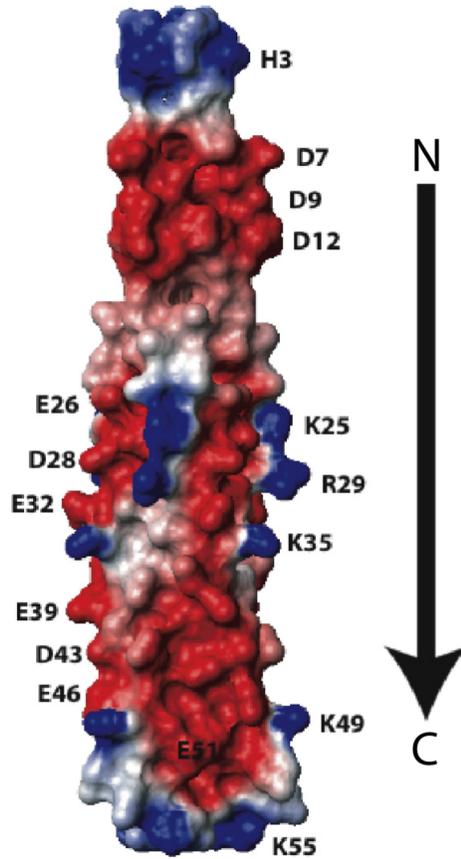
The Caffrey group has previously presented a model for SARS-CoV entry in which the coil-helix equilibrium of S2-HR2 played a role in allowing repositioning of the HR1 and HR2, based on biochemical and biophysical studies at pH 7<sup>52</sup>. In light of my results at pH 5, the model needs to be modified. In the modified model, at pH 7 S2-HR2 is in equilibrium between extended coil and helix with the extended coil being favored (Figure 3.8). Upon entering the endosome (i.e. a transition to pH 5) the helix conformation of S2-HR2 is greatly favored. Subsequently, the S2-HR2 helix interacts with the S2-HR1 helix to form the highly stable 6 helix bundle, which brings the viral and endosomal membranes into close proximity and allows fusion to occur. As a consequence of this model, isolated S2-HR2 helix is transiently present in the endosome, which therefore may present a novel target for therapeutic intervention. Finally, I note that it will be interesting to consider whether pH-induced structural changes occur for the HR2 of other viruses that enter via endocytosis such as Ebola and Influenza.



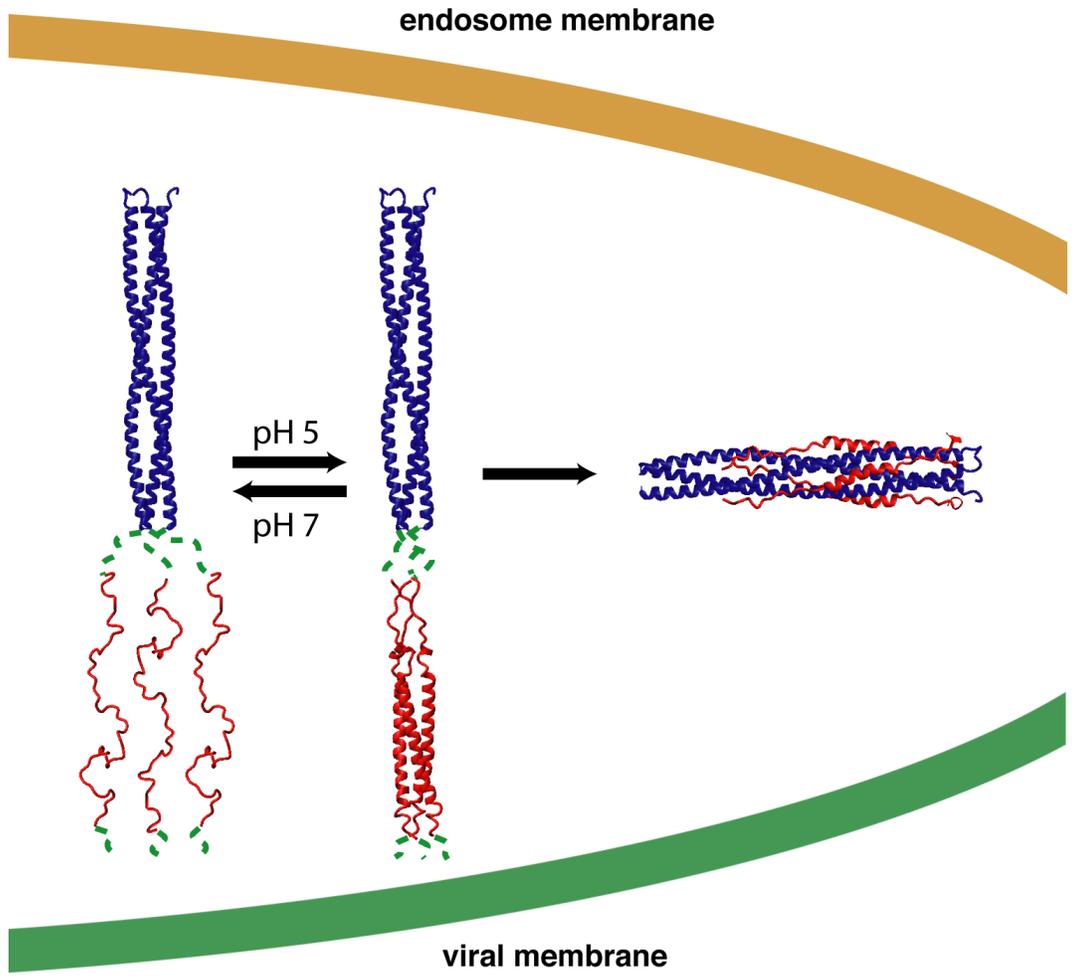
**Figure 3.5.** Effect of ionic strength on helical content of S2-HR2 at pH 7 and pH 5. The experimental conditions were 110  $\mu$ M HR2 in 10 mM NaHPO<sub>4</sub> and 0-300 mM NaCl at 20°C.



**Figure 3.6** Helical content as a function of ionic strength. SARS-CoV HR2 helix content as a function of ionic strength at pH 7 (filled circles) and pH 5 (empty circles). The experimental conditions were 110  $\mu$ M HR2 in 10 mM NaHPO<sub>4</sub> and 0-300 mM NaCl at 20°C.



**Figure 3.7** Electrostatic profile of SARS-CoV S2-HR2. The structure is taken from Hakansson-McReynolds et al., 2006<sup>14</sup>.



**Figure 3.8** Model for SARS Co-V S2-HR2 mediated entry. S2-HR1 and S2-HR2 are depicted in blue and red, respectively. Unknown structural domains are depicted as green dashes. Coordinates for isolated S2-HR2 prefusion state are taken from Hakansson-McReynlds et al., 2006<sup>14</sup>. Coordinates for the S2-HR1 and S2-HR1/S2-HR2 are taken from Duquerroy et al., 2005<sup>50</sup>.

### 3.7.3 SARS-CoV S2-HR2 has a parallel orientation

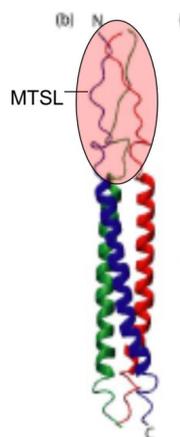
As discussed in the introduction, there are significant differences in the structure of S2-HR2 determined by NMR and x-ray crystallography (Figure 3.2). To determine if S2-HR2 is organized in a parallel or anti-parallel manner under aqueous conditions, I used an NMR technique that exploits the effect that a paramagnetic label has on the relaxation rate of nearby nuclei (termed paramagnetic relaxation enhancement or PRE). When the transverse relaxation rate of a particular nucleus is increased it results in a broadening of the associated peak in the spectrum. MTSL (S-(2,2,5,5,-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methansulfonothioate) is a compound that can be covalently attached via a disulfide bond to a cysteine residue. Introduction of MTSL will result in a paramagnetic effect on the relaxation rate of nuclei that are in proximity of the labeled residue, up to 25 angstroms. By attaching MTSL to one terminal region of S2-HR2, residue 14, I am able to determine if the peptide organizes in a parallel or anti-parallel manner by monitoring which residues experience the paramagnetic effect. (Figure 3.6) For example, in the case of a parallel orientation (i.e. the NMR structure), N-terminal residues would be expected to be affected; in the case of an anti-parallel orientation (i.e. the x-ray structure), N- and C-terminal residues would be expected to be affected.

In a first step, I measured the  $^{15}\text{N}$ -edited HSQC of S2-HR2 S14C-MTSL in the presence of 30% TFE to test the validity of the previously determined structure of Hakansson-McReynolds et al., 2006<sup>14</sup>. In Figure 3.9a, I show the intensities of the H- $^{15}\text{N}$  correlations as a function of residue number. The residues near position 14, the site of

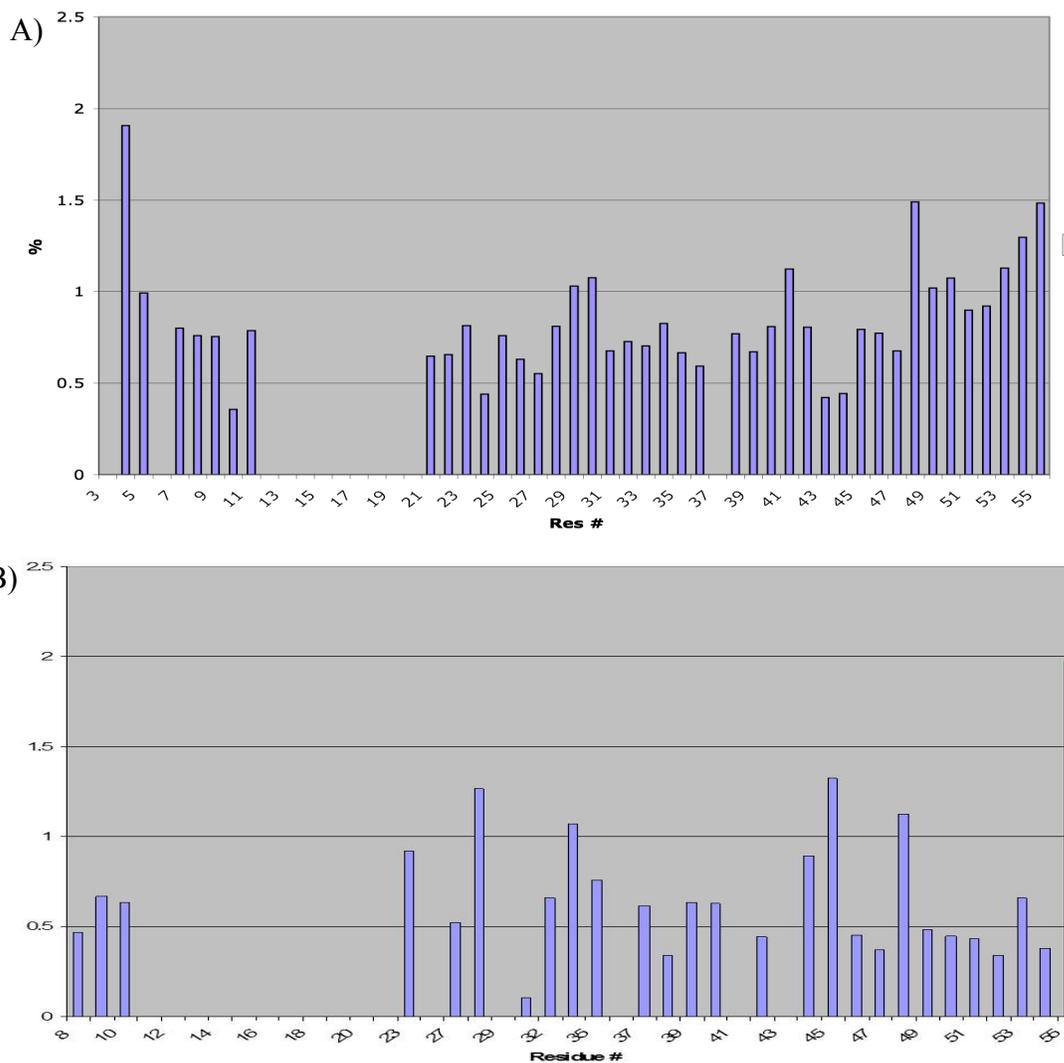
the MTSL, in sequence exhibit greatly reduced intensities with respect to S2-HR2 not containing the spin label. This observation implies that under these conditions, S2-HR2 is a parallel trimer, as previously suggested<sup>14</sup>.

In the next step, I measured the <sup>15</sup>N-edited HSQC of S2-HR2 S14C-MTSL under aqueous conditions (i.e. in the absence of TFE). In Figure 3.9b, I show the intensities of the H-<sup>15</sup>N correlations as a function of residue number. Again the residues near position 14, the site of the MTSL, in sequence exhibit greatly reduced intensities with respect to S2-HR2 not containing the spin label. This observation is consistent with the notion of a parallel trimer. Interestingly, the intensity of the other residues is also affected. I attribute this effect to be due to the equilibrium between unstructured monomer, in which the spin label may approach distant residues, and structured trimer.

Taken together, my studies strongly support the parallel orientation of S2-HR2 helices in the presence and absence of TFE. The differences between the NMR and x-ray structures may therefore be attributed to differences in the construct. For example the NMR construct contains 16 additional residues at the N-terminus (2 residues due to a cloning artifact and 14 native residues) with respect to the x-ray construct. The propensity of heptad repeat regions to self associate in different oligomerization states indicates that caution should be exercised in the interpretation of isolated domains. Nonetheless, it is important to note that trimers are generally observed for heptad repeat regions found in viral envelope proteins.



**Figure 3.9a.** MTSL labeling of SARS-CoV S2-HR2. A) The MTSL compound used as the paramagnetic label. B) The MTSL was attached to the cysteine at position 14 (introduced by site directed mutagenesis of S14). The shaded red area encompasses nuclei expected to show an increase in relaxation rates.



**Figure 3.9b.** Effect of MTSL on  $^{15}\text{N}$  labeled S2-HR2. Values are relative peak heights as compared to spectra without the spin label ( $(\text{labeled value}/\text{unlabeled value}) \times 100$ ). Experimental conditions: A) 1 mM S2-HR2, 10 mM  $\text{NaPO}_4$ , pH 7.0, 30% TFE-d3 B) 1 mM S2-HR2, 10 mM  $\text{NaPO}_4$ , pH 7.0. Residues 1, 2, 3, 6 and 37 of (A) and residues 1-7, 24-26, 29, 30, 34, 36, 41, and 43 of (B) were not assigned due to spectral overlap.

### **3.8 Conclusions**

By using a combination of biophysical techniques, my studies on SARS-CoV S2-HR2 have given new insights into SARS-CoV S2 structure and function. First, my SV studies have demonstrated that S2-HR2 is in a concentration-dependent monomer-trimer equilibrium, which is in agreement with previous NMR and CD studies of the Caffrey laboratory<sup>52</sup>. Second, my CD studies suggest that the S2-HR2 helix is stabilized by the pH of the endosome and physiological ionic strength. This observation necessitated a modification of the SARS-CoV entry model. Exploitation of this newly defined transition state may be a new path for development of entry inhibitors. Finally, my NMR studies on spin-labeled S2-HR2 validate the parallel orientation of the S2-HR2 helices.

## **4. PROBING THE HIV ENVELOPE PROTEIN STRUCTURE USING THE 12P1 PEPTIDE**

### **4.1 HIV**

Since its identification in the early 1980s, AIDS (acquired immunodeficiency syndrome) has become a global threat affecting essentially every region of the world. Although AIDS is a preventable disease it continues to be a growing epidemic with an estimated 5 million new infections each year<sup>59, 60</sup>. The UNAIDS report for 2009 estimated that 33 million people were living with AIDS (about 2/3 of those people are living in Sub-Saharan Africa), with an estimated 1.8 million AIDS-related deaths. After HIV (human immunodeficiency virus) was identified as the causative agent of AIDS, there has been a massive effort to learn as much about the virus as possible in an effort to find therapeutic agents that are effective.

This effort has led to the development of therapies directed towards the HIV proteins reverse transcriptase (RT) and protease (PR), and more recently the envelope proteins that mediate membrane fusion. Due to the high rate of mutation leading to drug resistance<sup>61</sup>, the anti-virals directed against RT and PR are currently taken in combination with each other in a treatment known as HAART (highly active anti-retroviral treatment)<sup>62</sup>. Although HAART has the ability to suppress the virus and therefore reduce the number of AIDS related deaths, it does not have the ability to eliminate the virus from the body. Also, various HIV strains have shown that over time they have the ability to become resistant to the HAART making it ineffective against those strains. Due to these

complications with current HIV treatments, it is vital that new treatments are developed such as those directed against the envelope proteins<sup>11</sup>.

Development of anti-viral therapy can be significantly improved with structural knowledge of the virus. While there has been progress in obtaining a full picture of HIV entry, there are still significant pieces missing that have impeded progress in rational drug design. One area where we are lacking a full picture is the interaction between gp120 and gp41. The envelope proteins gp120 and gp41 form a non-covalent complex on the surface of HIV and mediate viral entry<sup>63</sup>. Gp120 initiates viral entry by binding to receptors found on the target cell surface, CD4 and a chemokine co-receptor (CR), and gp41 subsequently brings the target and viral membranes in close proximity, thereby allowing membrane fusion and entry. Conformational changes of gp120 and gp41 are thought to be essential to the entry process<sup>64, 65</sup>, and because of their critical role in HIV entry; both have been intensively studied by x-ray crystallography and NMR spectroscopy<sup>11</sup>. For example there is structural information for the gp120 core in the free state and in complex with CD receptor domains and various monoclonal antibodies<sup>66-68</sup> and there are numerous structures of the gp41 extracellular domain in the fusion state<sup>69-71</sup>. However, all gp120 structures to date are missing significant regions of gp120, and perhaps more importantly, domains of gp41<sup>11</sup>. In fact, the journal Nature has recently listed the gp120/gp41 complex as one of the most desirable structures in biology<sup>72</sup>.

#### **4.2 Use of 12p1 to probe envelope structure**

The critical role of gp120 has led to efforts to find molecules that bind to gp120 and inhibit its interaction with receptors<sup>11</sup>. For example, the peptide 12p1

(sequence=RINNIPWSEAMM), which was discovered by phage display against recombinant HIV-1 HXB2 gp120, binds to gp120 and blocks its interaction with CD4 with IC50 ranging from 0.3 to 25  $\mu$ M for different strains<sup>73</sup>. Subsequent studies established that 12p1 binds to gp120 with a stoichiometry of 1:1<sup>74</sup>. It is also of note that the assumed binding site is one that is adjacent to but not overlapping the CD4 or CR binding sites, and that 12p1 inhibits binding of CD4 in an allosteric manner<sup>74</sup>. These results suggest that 12p1 and its derivatives have the potential to be developed into HIV entry inhibitors with a unique binding site that has not previously been exploited<sup>75, 76</sup>. In this chapter, I will be presenting work in which I exploited the 12p1 to probe gp120 structure in the presence and absence of gp41 domains by NMR; specifically I will present the assignment of 12p1 <sup>1</sup>H, <sup>13</sup>C, and <sup>15</sup>N, resonances and use Saturation Transfer Difference (STD) NMR<sup>77</sup> to assess gp120 conformation and to identify 12p1 <sup>1</sup>H in close contact with the gp120 binding surface, which is expected to aid in the improvement of 12p1 as an anti-viral therapeutic.

### 4.2.1 Assignment of 12p1

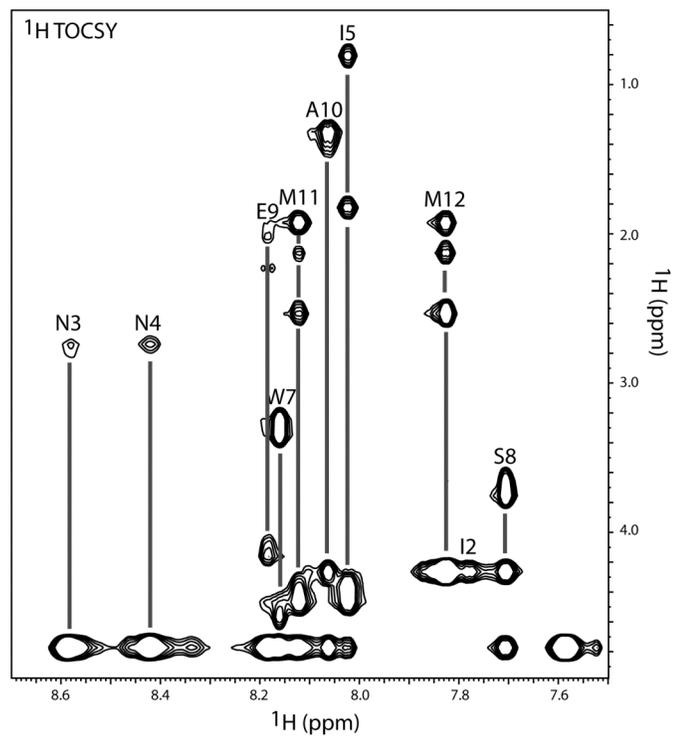
As a first step, I performed a series of NMR experiments to assign the  $^1\text{H}$ ,  $^{13}\text{C}$ , and  $^{15}\text{N}$  resonances of 12p1 as a free peptide. Examples of the TOCSY amide region and  $^{13}\text{C}$ -edited HSQC spectra with assignments are shown in Figure 4.1; the  $^{13}\text{C}$ -edited HSQC was especially useful in the assignment process due to the significant overlap of 12p1  $^1\text{H}$ . In total, assignments were obtained for over 95% of the  $^1\text{H}$ ,  $^{13}\text{C}$ , and  $^{15}\text{N}$  (Table 4.1). It was next of interest to assess the secondary structure of the peptide from secondary chemical shift differences, which are defined as  $\delta_{\text{observed}} - \delta_{\text{random coil}}$ <sup>79</sup>. As shown in Figure 4.1d, the secondary chemical shifts of the  $^1\text{H}\alpha$  and  $^{13}\text{C}\alpha$  indicate that 12p1 exhibits the absence of regular secondary structure, which is not surprising for a peptide. Interestingly, P6 exhibits chemical shifts indicative of a *trans* conformation. For example, the observed  $^1\text{H}\alpha$ ,  $^{13}\text{C}\alpha$ ,  $^1\text{H}\beta$ , and  $^{13}\text{C}\beta$  of P6 are 4.36, 63.4, 2.21, 1.80 and 31.9 ppm, respectively, which are closer to the values observed for *trans* proline than for *cis* proline (Table 4.2). The central seven residues of 12p1, N3-S8, which include P6, have previously been shown by single site alanine substitution to be critical for the interaction with gp120<sup>73</sup>, thus the conformational state of P6 is expected to be of importance to binding, however at this point we have no evidence of preferential binding to either the *cis* or the *trans* conformations of proline.

Residue	$H_N/^{15}N$	$H\alpha/^{13}Ca$	$H\beta/^{13}C\beta$	Others
R1		3.84, 55.51	1.83, 30.01	$H^{13}C\gamma$ : 1.55, 26.28; $H^{13}C\delta$ : 3.14, 43.18
I2	7.78, 119.55	4.27, 61.08	1.77, 38.61	$H^{13}C\gamma1$ : 1.44, 1.15, 27.25; $H^{13}C\gamma2$ : 0.87, 17.40; $H^{13}C\delta$ : 0.86, 12.73
N3	8.59, 123.18	4.77, 52.95	2.78, 38.94	$H^{15}N\delta$ : 7.58, 6.88, 112.92
N4	8.42, 119.49	4.76, 53.18	2.76, 38.77	$H^{15}N\delta$ : 7.57, 6.84, 112.75
I5	8.02, 122.31	4.37, 58.69	1.81, 38.61	$H^{13}C\gamma1$ : 1.44, 1.09, 26.83; $H^{13}C\gamma2$ : 0.81, 17.01; $H^{13}C\delta$ : 0.83, 12.60
P6		4.36, 63.37	2.21, 1.80, 31.90	$H^{13}C\delta$ : 3.80, 3.57, 50.92; $H^{13}C\gamma$ : 1.94, 27.35
W7	8.16, 122.43	4.61, 57.88	3.30, 3.22, 29.37	$H^{13}C\delta1$ : 7.24, 127.08; $H^{13}C\epsilon3$ : 7.55, 120.78; $H^{13}C\eta2$ : 7.20, 124.62; $H^{13}C\zeta2$ : 7.45, 114.59;
				$H^{13}C\zeta3$ : 7.11, 122.02
S8	7.71, 117.47	4.24, 57.90	3.63, 3.76, 64.05	
E9	8.18, 121.68	4.13, 56.74	2.00, 30.01	$H^{13}C\gamma$ : 2.27, 36.29
A10	8.06, 123.75	4.27, 52.46	1.36, 19.02	
M11	8.12, 119.61	4.42, 55.28	2.11, 1.95, 32.67	$H^{13}C\gamma$ : 2.53, 2.45, 32.02; $H^{13}C\delta$ : 2.04, 16.87
M12	7.82, 126.46	4.27, 57.03	2.11, 1.97, 33.75	$H^{13}C\gamma$ : 2.53, 2.46, 32.34; $H^{13}C\delta$ : 2.04, 16.87

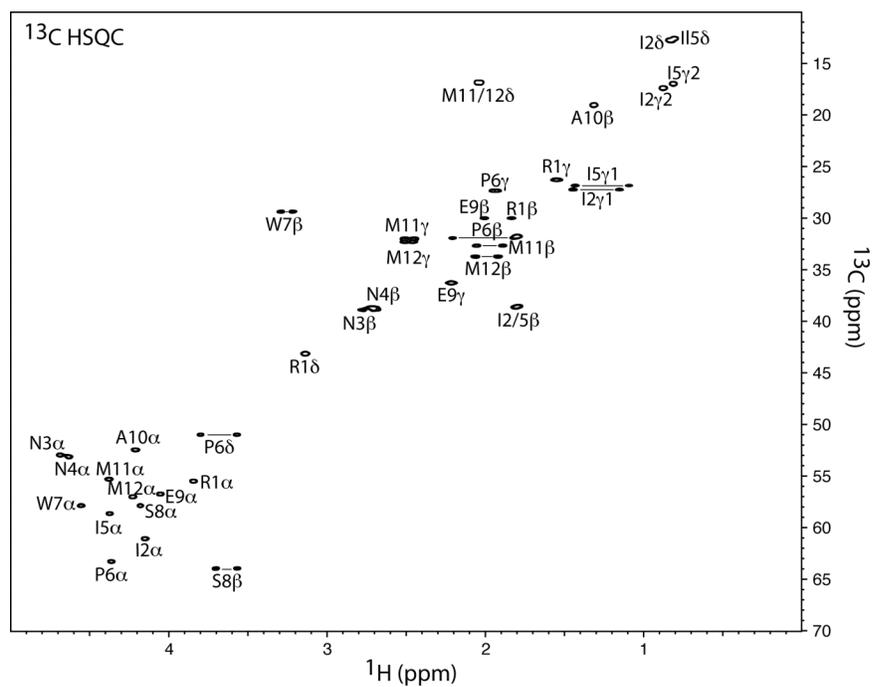
**Table 4.1.** Chemical shift value assignments for  $^1H$ ,  $^{13}C$ , and  $^{15}N$  of 12p1.

	<b>Predicted <i>cis</i> value</b>	<b>Predicted <i>trans</i> value</b>	<b>Observed value</b>
<sup>1</sup> H $\alpha$	4.60	4.45	4.36
<sup>13</sup> C $\alpha$	63.0	63.7	63.4
<sup>1</sup> H $\beta$	2.39/1.8	2.29/1.99	2.21/1.80
<sup>13</sup> C $\beta$	34.8	32.2	31.9

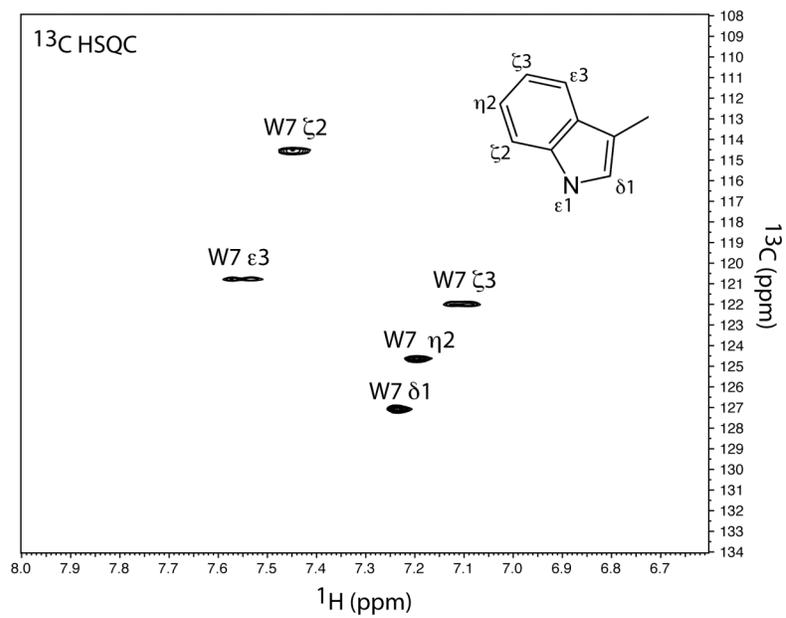
**Table 4.2.** Predicted and observed chemical shift values for P6 of 12p1. The proline at position 6 of 12p1 is predicted to be in the *trans* conformation based on the chemical shift values presented here.



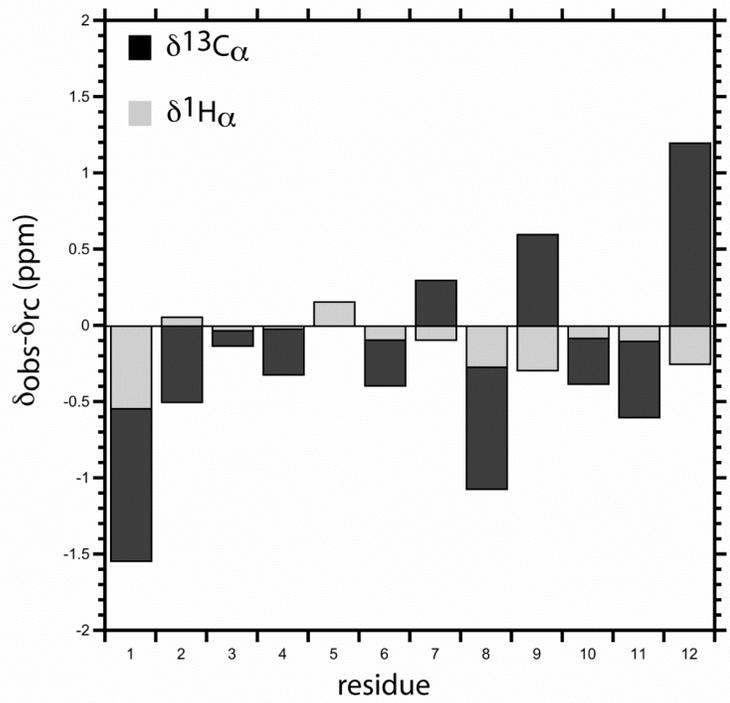
**Figure 4.1a.** Amide region of 12p1 TOCSY. The experimental conditions for the TOCSY were 2 mM 12p1 in 20 mM  $\text{PO}_4/\text{pH}$  7.2, 90%  $^1\text{H}_2\text{O}/10\%$   $^2\text{H}_2\text{O}$ .



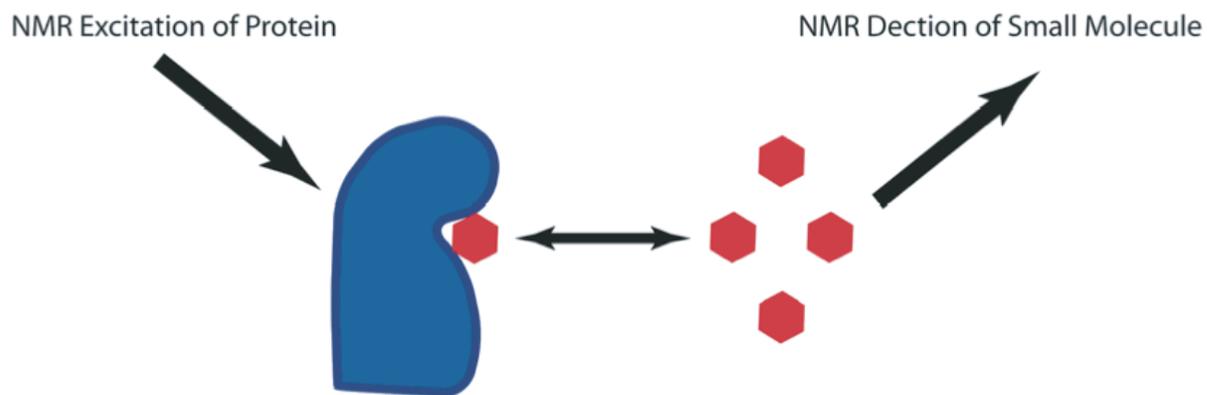
**Figure 4.1b.** Aliphatic region of 12p1  $^{13}\text{C}$ -edited HSQC at natural abundance. The experimental conditions for the  $^{13}\text{C}$ -edited HSQC were 2 mM 12p1 in 50 mM  $\text{PO}_4/\text{p}$  7.2, 100%  $^2\text{H}_2\text{O}$ .



**Figure 4.1c.** Aromatic region of 12p1 <sup>13</sup>C-edited HSQC at natural abundance. The experimental conditions for the <sup>13</sup>C-edited HSQC were 2 mM 12p1 in 50 mM PO<sub>4</sub>/p 7.2, 100% <sup>2</sup>H<sub>2</sub>O.



**Figure 4.1d.** Secondary chemical shifts of 12p1. Random coil values were taken from Schwarzinger et al., 2000<sup>78</sup>.



**Figure 4.2.** Saturation Transfer Difference NMR. The large molecule (blue) is selectively irradiated, magnetization is transferred to the small molecule (red) as it binds to the larger one, then detected. The difference with respect to a reference spectrum in which the larger molecule is not irradiated identifies  $^1\text{H}$  of the small molecule in close contact in the bound state.

#### 4.2.2 STD study of the 12p1 interaction with recombinant gp120

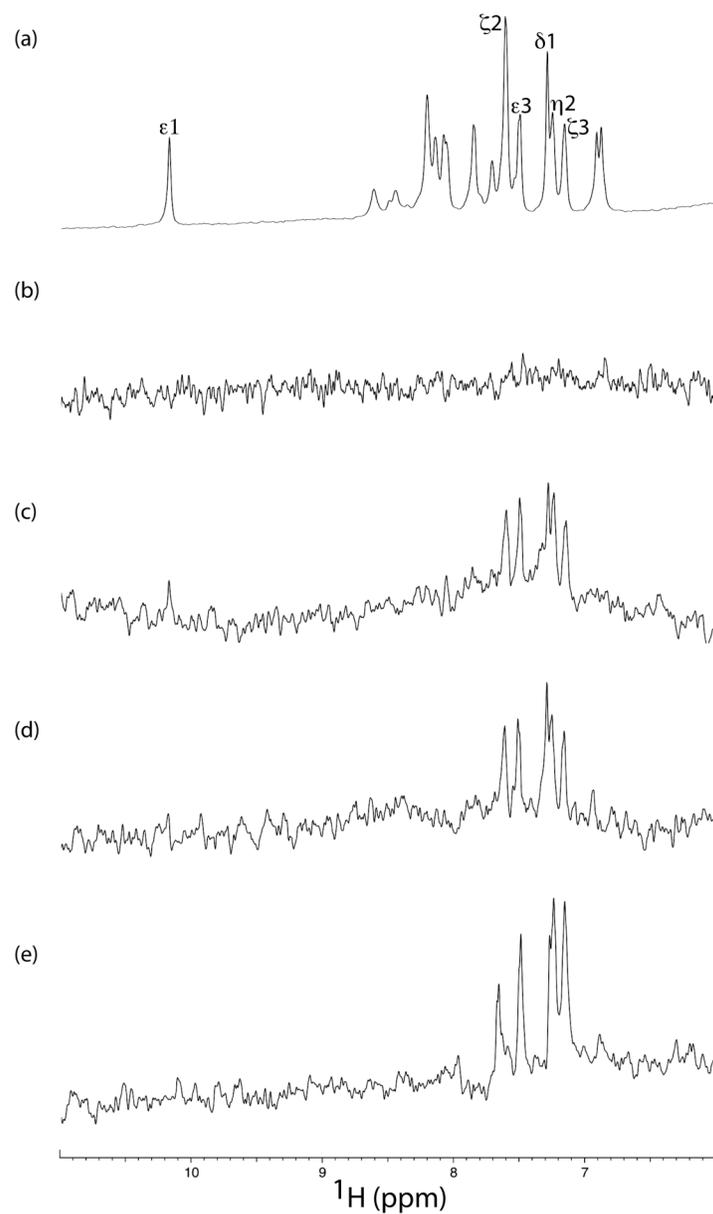
STD NMR presents a powerful method to characterize interactions between small molecules and large molecules<sup>77</sup>. In the STD NMR experiment, the resonances of the large molecule  $^1\text{H}$  are selectively irradiated, subsequently magnetization is transferred to the  $^1\text{H}$  of the small molecules that exchange between bound and free states during the irradiation period; the difference, with respect to a reference spectrum in which the large molecule  $^1\text{H}$  are not irradiated and hence no magnetization transfer occurs, identifies  $^1\text{H}$  in closest contact in the bound state (Figure 4.2). The 1D  $^1\text{H}$  NMR spectrum of the 12p1 downfield region is shown in Figure 4.3a as a reference with the W7 sidechain  $^1\text{H}$  identified. In Figure 4.3b, the STD NMR spectrum of 12p1 in the absence of gp120 exhibits the absence of signal as expected. In contrast, in Figure 4.3c, the STD NMR spectrum of 12p1 in the presence of gp120 from strain YU2, which is a 120 kDa protein in the monomeric state, suggests that in the bound state the W7 sidechain atoms  $\text{H}_{\delta 1}$ ,  $\text{H}_{\zeta 2}$ ,  $\text{H}_{\eta 2}$ ,  $\text{H}_{\zeta 3}$  and  $\text{H}_{\epsilon 3}$  of 12p1 are in close contact with gp120. Interestingly, the absence of other resonances in this spectral region (the backbone and sidechain  $^1\text{H}$  present in Figure 4.3a) indicates that these  $^1\text{H}$  are not in close contact to gp120.

The importance of the W7 sidechain to the 12p1-gp120 interaction is supported by previous studies of W7A and W7F versions of 12p1 that exhibited greatly reduced binding to gp120<sup>73, 74</sup>. In Figures 4.3d and 4.3e, similar STD NMR spectra are observed for 12p1 in the presence of gp120 from HIV-1 clade B strain R2 and clade C strain 96ZM65, respectively, suggesting similar interactions between 12p1 and gp120 in a

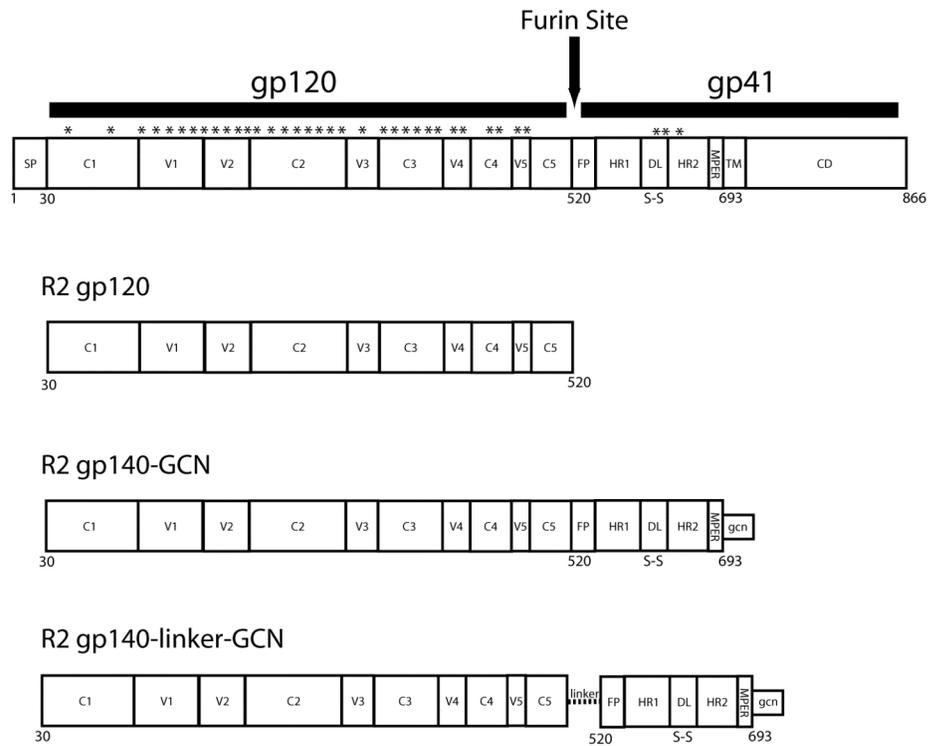
strain and clade independent manner. Previous studies have suggested that 12p1 binds to and inhibits viral entry of various clade B strains<sup>73, 74</sup>; however, the potential therapeutic potential to clade C, which is the most rapidly spreading subtype and is prevalent in Southern and Eastern Africa, as well as India<sup>79</sup>, makes 12p1 even more attractive for development as a therapeutic. Finally, we note that STD NMR signals are also observed in the aliphatic region of 12p1, which is indicative of binding; however, due to spectral overlap, specific interactions are difficult to identify (eg. I2 versus I5 <sup>1</sup>H $\delta$  and N3 versus N4 <sup>1</sup>H $\beta$ ).

### 4.2.3 STD study of the 12p1 interaction with recombinant gp140

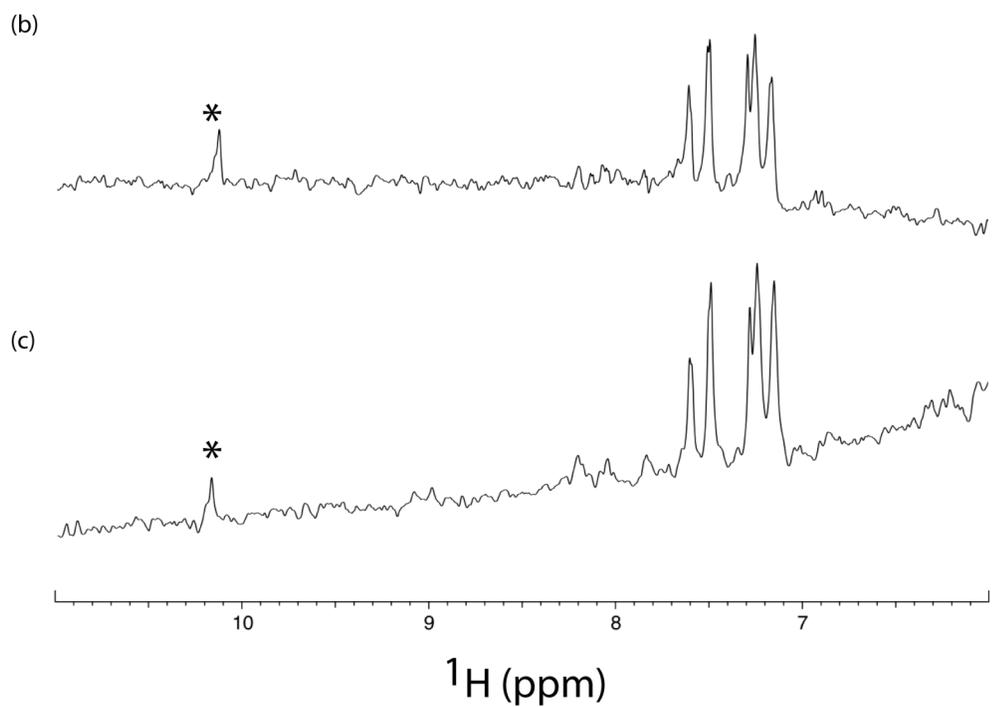
It is next of interest to contrast the 12p1 interaction with gp120 in the presence and absence of gp41 domains. For this part of the study, two more constructs were tested (Figure 4.4a). The first construct, denoted R2gp140, consisted of full length strain gp120 followed by gp41 in which the C-terminal transmembrane and cytoplasmic domains have been deleted. In this construct, the furin cleavage site has been mutated by arginine to serine substitutions and thus gp120 is covalently attached to gp41 in order to stabilize the gp120-gp41 interaction<sup>20</sup>. In addition, a GCN trimerization domain has been appended to the C-terminus of gp41 to stabilize the trimeric state<sup>20</sup>. The second construct, denoted gp140-linker, is similar to the gp140 except that a 15 residue flexible linker domain has been inserted between the gp120-gp41 interface with the rationale that the linker region would remove possible conformational strain present in the gp140 construct. In both cases, the purified gp140 constructs were trimers of 520 kDa. As shown in Figure 4.4b, the STD NMR spectrum of 12p1 in the presence of gp140 suggests that the sidechain atoms of W7 are in close contact to trimeric gp120; however, the additional signal observed for the W7 indole group (H $\epsilon$ 1) clearly indicates an additional interaction site with respect to monomeric gp120. This notion is further supported by the STD NMR spectrum of 12p1 in the presence of gp140-linker, as shown in Figure 4.4c. Interestingly, the similar spectra of the gp140 and gp140-linker also suggest that the presence of the linker does not measurably affect the interaction between 12p1 and trimeric gp120.



**Figure 4.3.** STD NMR spectra of 12p1 in the presence of gp120 constructs. (a) Reference 1D NMR spectrum of 12p1. (b) STD NMR spectrum of 12p1 in the absence of gp120. (c) STD NMR spectrum of 12p1 in the presence of gp120 from strain YU2. (d) STD NMR spectrum of 12p1 in the presence of gp120 from strain R2. (e) STD NMR spectrum of 12p1 in the presence of gp120 from strain 96ZM65. The experimental conditions were 1 mM 12p1 +/- 4  $\mu$ M gp120 in 20 mM  $\text{PO}_4/\text{pH}$  7.2, 90%  $^1\text{H}_2\text{O}/10\%$   $^2\text{H}_2\text{O}$ .



**Figure 4.4a.** Schematic diagram of gp120 and gp140 constructs. The glycosylation sites are denoted by asterisks.

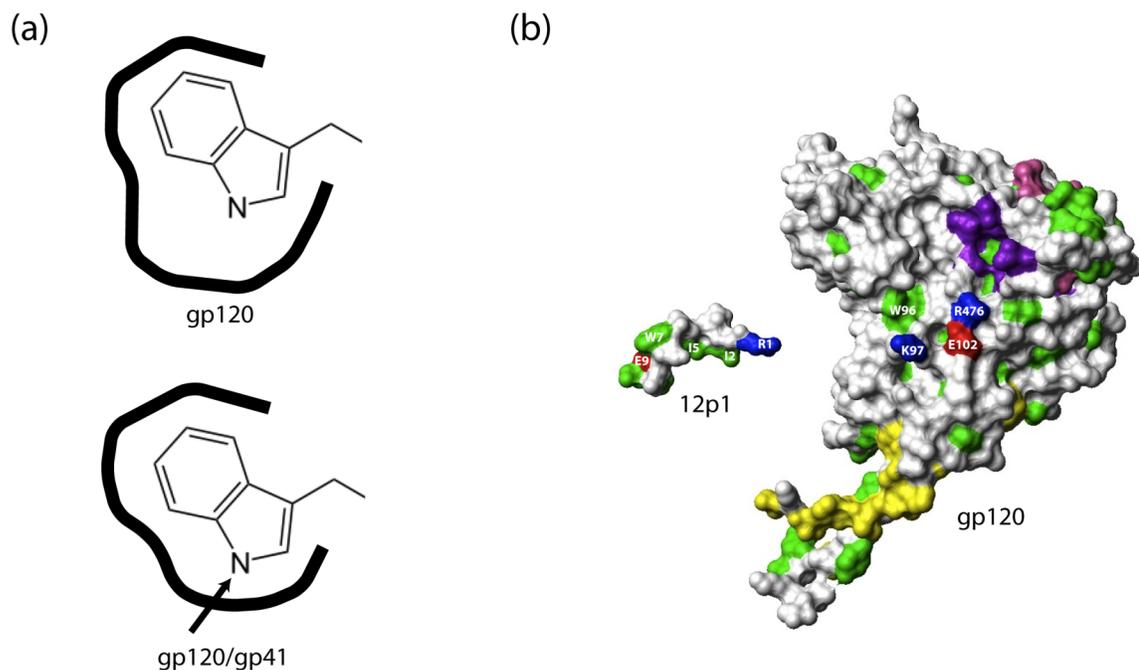


**Figure 4.4b.** STD NMR of 12p1 in the presence of gp140 constructs. (b) STD NMR spectra of 12p1 in the presence of gp140. (c) STD NMR spectrum of 12p1 in the presence of gp140-linker. Experimental conditions were 1 mM 12p1 +/- 4  $\mu$ M gp140 in 20 mM  $\text{PO}_4/\text{pH}$  7.2, 90%  $^1\text{H}_2\text{O}/10\%$   $^2\text{H}_2\text{O}$ .

#### 4.2.4 Model for the interaction between 12p1 and gp120

In Figure 4.5a, a model for the differences between the 12p1 interaction with the gp120 in the presence and absence of gp41 domains is presented to highlight the closer contact of the 12p1 Hε1 to gp120 in the presence of gp41 domains due to alteration of the binding pocket. These subtle differences in the interaction of 12p1 suggest subtle differences in the structure of gp120 in the presence of gp41 domains. The sensitivity of the interaction is further supported by previous work demonstrating that the addition of non-natural hydrophobic moieties to P6 significantly enhances the 12p1 affinity for gp120<sup>75, 76</sup>.

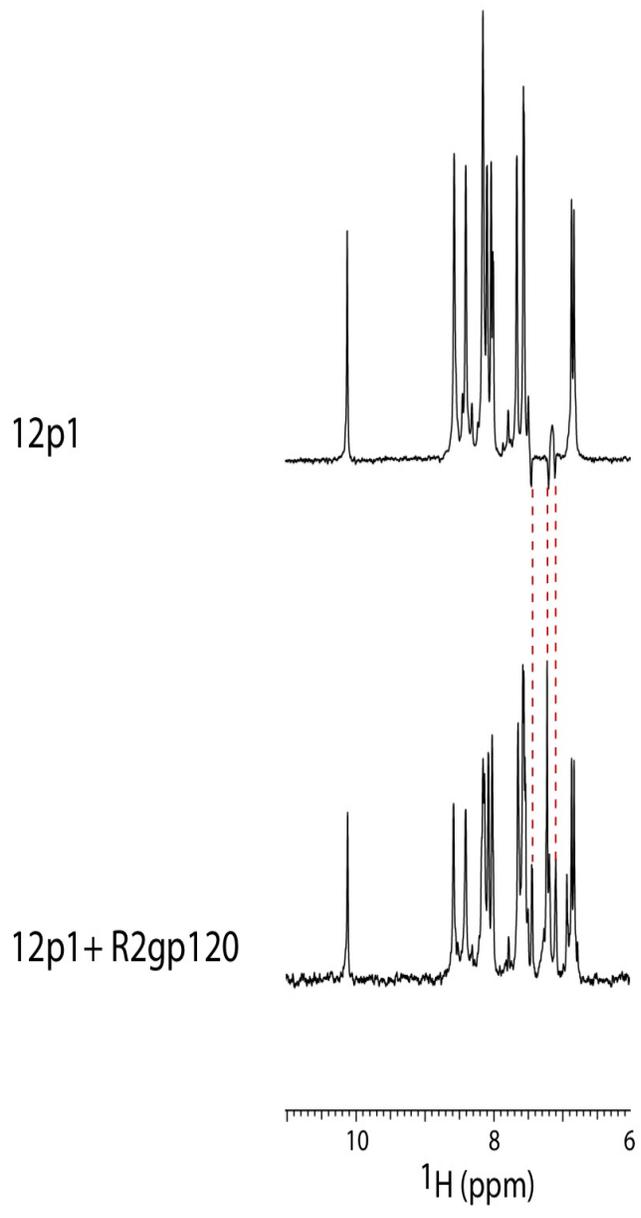
Figure 4.5b is a model for the hydrophobic interaction between 12p1 and gp120. A previous study of gp120 escape mutants suggested that the 12p1 interaction occurs near residues K97, E102, and R476 of gp120<sup>74</sup>. As noted in the previous study, the CD4 and CR binding sites are proximal but not overlapping with this region as is evident in the figure. Based on mutagenesis studies<sup>11, 80, 81</sup>, the interaction site with gp41 is also proximal but not overlapping with the putative 12p1 binding site, an observation that is in agreement with the subtle differences observed in the STD NMR studies presented here. As noted above, hydrophobicity at residues 6 and 7 of 12p1 appear to be important for its interaction with gp120. Accordingly, the exposed hydrophobic regions on the gp120 surface have been highlighted in Figure 4.5b. Interestingly, W96 of gp120 is relatively exposed and in close proximity to the site of the escape mutants; thus, W96 may present a potential interaction partner stabilizing electrostatic interactions between K97, E102 and R476 of gp120.



**Figure 4.5.** Models for 12p1 binding to gp120. (a) Schematic model for the interaction between the W7 sidechain of 12p1 and the gp120 binding surface in the monomeric and trimeric states. (b) Space filling representations for the interaction between 12p1 and gp120. The structure of 12p1 represents a random coil model determined by CNS<sup>81</sup> and the gp120 structure is taken from Pancera et al., 2010<sup>68</sup>. The exposed hydrophobic sidechains are colored green and select acidic and basic sidechains are colored, red and blue, respectively. For reference, the bindings sites of CD4, CR and gp41 binding are colored purple, pink and yellow, respectively.

#### 4.2.5 WaterLOGSY of gp120 with 12p1 peptide

WaterLOGSY (Water Ligand Observed via Gradient Spectroscopy) is another NMR technique available to identify binding atoms of compounds, which has the potential to be more sensitive than STD-NMR. This technique takes advantage of the water molecules that surround the peptide of interest and the displacement of those water molecules upon binding. In the WaterLOGSY experiment the bulk water is selectively magnetized and then magnetization is partially transferred via the protein-ligand complex to the free ligand in a selective manner<sup>23</sup>. The atoms that are binding show a change in the sign of their resonance compared to that of a reference spectrum. In Figure 4.6 we can see that the resonances assigned to the W7 sidechain atoms have an opposite sign in the presence of gp120, this indicates that 12p1 is interacting with gp120. With respect to the STD experiments, the WaterLOGSY was acquired with a lower concentration of gp120 (1  $\mu\text{M}$  versus 4  $\mu\text{M}$ ) and lower acquisition time (1 hour versus 12 hours). Consequently, the WaterLOGSY experiment presents great potential to be used for titrations and the characterizations of other ligand-protein interactions.



**Figure 4.6.** WaterLOGSY NMR of 12p1 in the presence and absence of R2gp120. Dotted redlines correspond to W7 sidechain resonances. Experimental conditions were 200  $\mu$ M 12p1 +/- 1  $\mu$ M R2gp120 in 20 mM PO<sub>4</sub>/pH 7.2, 90% <sup>1</sup>H<sub>2</sub>O/10% <sup>2</sup>H<sub>2</sub>O.

#### **4.2.6 STD study of the 12p1 interaction with Virus Like Particles (VLPs) expressing HIV-1 envelope proteins**

The final series of STD experiments I did consisted of using VLPs as the macromolecule 12p1 was interacting with (Figure 4.7a). As described in Section 2.2.4 HIV-1 VLPs were produced using the JR-FL plasmid (HIV-VLP) along with the control VLPs expressing VSVG envelope proteins (VSVG-VLP) or no envelope protein (-env-VLP). Both VSVG-VLP and -env-VLP act as negative controls; the -env-VLP could show that the ligand is not binding simply to the surface of the VLP and VSVG-VLP could show that the ligand is binding specifically to HIV envelope proteins, not any other proteins present. That being said, I chose to use VSVG-VLP as my control because it gives more information about the binding specificity than -env-VLP.

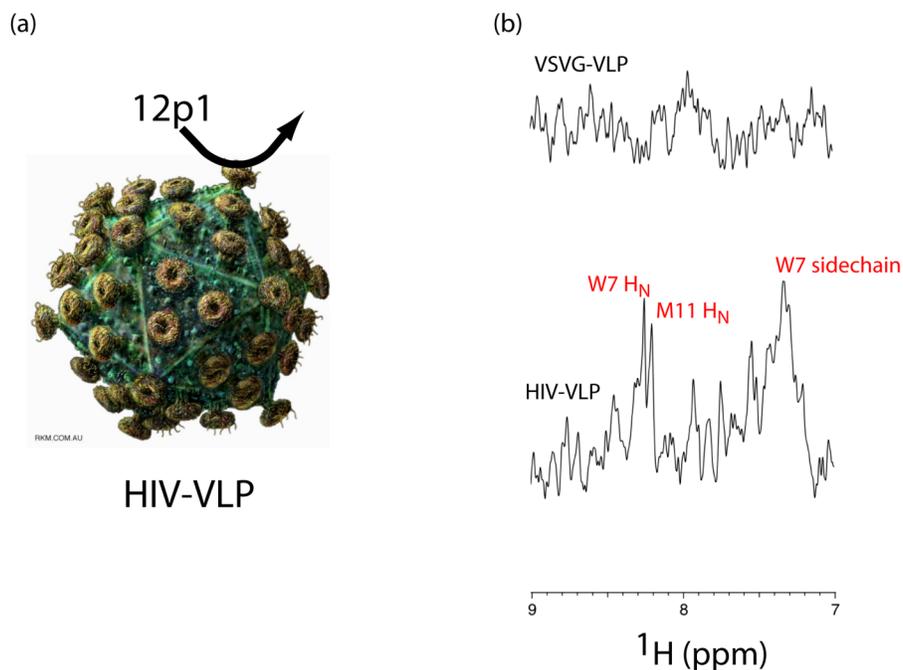
For comparing VLP experiments, it is important to consider how to normalize different preparations to each other. There are a number of ways to quantify VLPs: 1) count the number of particles; 2) determine infectious units (based on the ability of 1 VLP to infect 1 cell); 3) determine the p24 content. For the present study, the VLPs were quantified using methods 2 and 3, which produced somewhat different results in the viral titre (see Chapter 2). The MAGI assay (method 2) is based on the activation of a LTR- $\beta$ -galactosidase gene in a CD4<sup>+</sup> cell; the Tat protein from HIV transactivates the gene by the HIV-1 long terminal repeat (LTR) promoter in the CD4<sup>+</sup> cells. This assay can detect replication-deficient virus and has the ability to detect 1 infected cell in a background of  $10^6$  cells<sup>83</sup>. The p24 antigen capture assay (method 3) detects antigens to the HIV protein p24, a protein that is a core structural component of HIV-1. In this assay the reagents

provided disrupt the p24 antigen/antibody complexes allowing the antigen to be detected and measured in the samples. Determining the number of infectious units has the advantage that one is measuring the number of functional viruses (i.e. with correctly folded envelopes). The disadvantage is that different viral envelope proteins exhibit very different efficiency of infectivity (VSVG is much more infectious than HIV) and that this infectivity is sensitive to cell type/receptor used in the assay. Moreover, the –env VLP cannot be quantified in this manner because they are non-infective. Determining p24 content has the advantage that HIV-VLP have an established number of p24 molecules to HIV envelope (~2600:100). The disadvantage is that not all of the HIV envelope proteins on virus are functional (correctly folded) and different HIV strains exhibit different densities of envelope (but not p24). Furthermore, there are concerns that the VLP of VSVG versus HIV may in fact be different sizes. Taken together, it has not yet been established which method is best for NMR studies of VLP thus we have considered both infectious units and p24 content in the interpretation of our data. Our collaborators are currently studying average particle size by dynamic light scattering to compare HIV VLP to VSVG VLP. Ultimately, the best control may have to be a mutant of the same strain, but that was not available for me at this time.

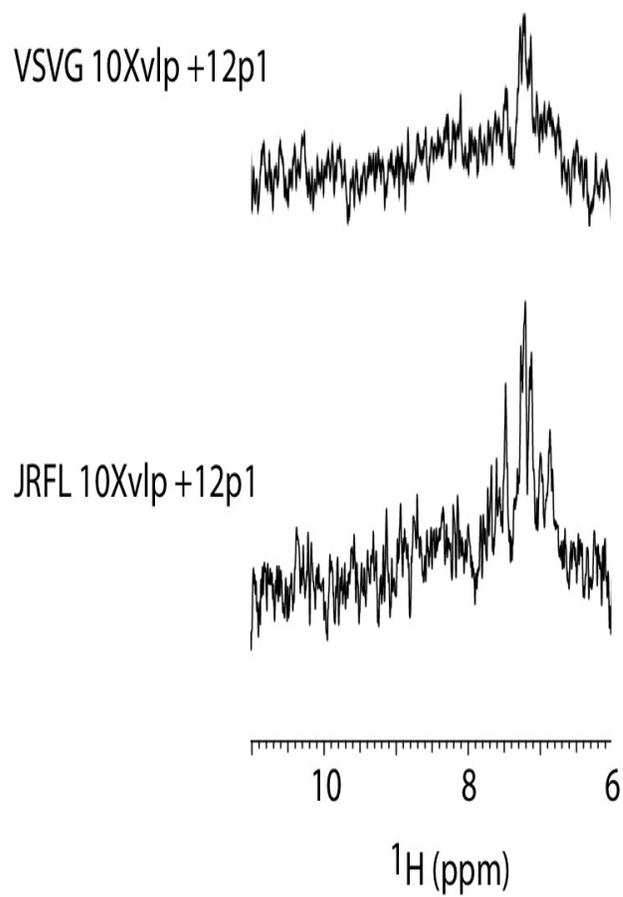
I first did a titration of VLP amount in the presence of 100  $\mu$ M 12p1 to determine the limit of STD-NMR sensitivity. I found that 50  $\mu$ L of the VLP sample (p24 values are comparable for each VSVG, -ENV, and HIV-VLPs so same amount was used for each; Table 2.1) is necessary for a detectable signal and used that amount for my preliminary studies. In the case of HIV envelope, this amount would correspond to a concentration of

~2  $\mu\text{M}$  envelope monomer. The STD-NMR spectra for VSVG-VLP shows no detectable interaction of the 12p1 protons with the VSVG envelope proteins; however, the STD-NMR of the HIV-VLP with 12p1 was able to detect interaction of the 12p1  $^1\text{H}$  from W7 sidechain (Figure 4.7b). This interaction of 12p1 is consistent with what I saw when measuring the STD effect with recombinant gp120 in the absence and presence of gp41 domains, although the VLP spectra exhibit lower signal to noise. This may be due to a lower concentration of envelope (the 2  $\mu\text{M}$  estimate for envelope is 50% of that of the 4  $\mu\text{M}$  used in the studies of recombinant gp120; however, the estimate is based on a number of assumptions and the actual concentration could be significantly lower). Moreover, misfolded or degraded envelope may also be present. Interestingly, 2 additional signals are observed in the STD spectrum of HIV-VLP with respect to recombinant gp120: the  $\text{H}_\text{N}$  of W7 and M11. The observation of additional signals may suggest that the binding mode of 12p1 to envelope in the virus is somewhat different.

To test the possible lower concentration, I concentrated the VLP samples to a 10X concentration and ran the STD experiments again (Figure 4.8). The 10X concentrated HIV-VLP sample showed better signal to noise in the STD-NMR, however I also started to see some non-specific binding to VSVG-VLP. It is possible that the non-specific binding to VSVG-VLP is due to the higher concentration of 12p1 in these samples (4X higher than the 1X VLP experiments). In addition, there is the possibility that the VLP are not very stable to the concentration process (several hours of ultrafiltration). Nonetheless, the STD experiments on the concentrated samples need to be repeated with 100  $\mu\text{M}$  12p1 to make a true comparison.



**Figure 4.7.** STD-NMR of VSVG-VLP and HIV-VLP. (a) Schematic diagram of 12p1 interaction with envelope present on virus. (b) The STD of HIV-VLP shows that the W7 sidechain atoms along with the H<sub>N</sub> of W7 M11 H<sub>N</sub> of 12p1 are directly interacting with the HIV VLP. 12p1 shows no interaction with VSVG-VLP. Experimental conditions were 50  $\mu$ L of VSVG-VLP or JRFL-VLP ([p24]  $\sim$  10,000 pg/mL) and 100  $\mu$ M 12p1 in PBS (20 mM PO<sub>4</sub>/pH 7.2, 150 mM NaCl) in 90% <sup>1</sup>H<sub>2</sub>O, 10% <sup>2</sup>H<sub>2</sub>O.



**Figure 4.8.** STD-NMR of VLP concentrated 10X. Experimental conditions were VSVG-VLP or JRFL-VLP ([p24] ~ 100,000 pg/mL) and 400  $\mu\text{M}$  12p1 in PBS (20 mM  $\text{PO}_4/\text{pH}$  7.2, 150 mM NaCl) in 90%  $^1\text{H}_2\text{O}$ , 10%  $^2\text{H}_2\text{O}$ .

### **4.3 Summary**

In this chapter, I assigned the NMR resonances of the peptide inhibitor 12p1. Using STD-NMR I was able to establish that 12p1 binds to HIV-1 clades B and C in a similar manner, this makes 12p1 an even more attractive potential therapeutic due to the rapid spread of HIV infections characterized as clade C. Not only does 12p1 bind in a clade independent manner to gp120, but the STD-NMR results also showed that there are subtle differences in the binding mode of 12p1 to gp120 in the presence of gp41 domains. These subtle differences highlight the importance the gp120-gp41 interface in the search for anti-virals that inhibit entry. Along with identifying binding atoms of 12p1 I was able to show that STD-NMR is a valid method of studying the binding of ligands to VLPs, information that could be used in future studies of envelope conformation on virus and the development of entry inhibitors.

## 5. SUMMARY AND FUTURE STUDIES

### 5.1 Summary of Work

As stated previously, the envelope proteins play critical roles in viral entry and there are many details about their structure and function that are missing. Moreover, the envelope proteins are attractive targets for new anti-viral therapies. In this thesis I have presented new information that is a step forward in gaining a full understanding of HIV and SARS-CoV entry. It is hoped that this information may eventually be exploited in the development of therapies, or vaccines, against HIV and SARS.

Using a combination of NMR, analytical ultracentrifugation, and circular dichroism techniques to study SARS S2-HR2, I obtained results that have led to a new model for the behavior of the S2-HR2 protein during entry. It appears that S2-HR2 is in an equilibrium between an unstructured monomer state and a helical trimer state. This equilibrium is important because the unstructured monomeric state of the HR2 peptide is expected to be a more potent inhibitor. My work showed that this equilibrium favors the helical state at physiological ionic strength and at pH 5; which are the conditions of the endosome, one site of SARS-CoV entry. The tendency for S2-HR2 to form a helical trimer in the endosome gives insight into why this HR2 peptide is a relatively weak inhibitor as compared to HIV HR2 peptide inhibitors.

In my work with HIV I showed that the use of a peptide probe, in combination with STD or WaterLOGSY NMR, is a useful method to gather more information on the envelope proteins, as well as a way to study envelope proteins in the context of the virus. Importantly, I identified the 12p1  $^1\text{H}$  that are in closest contact with gp120, an observation that may be exploited in the improvement of the 12p1 anti-viral activity. In addition, I demonstrated that 12p1 shows additional interactions with gp120 in the presence of gp41, suggesting that the structure of gp120 may be different when gp41 domains are present. This insight is just one more step closer to a complete picture of the gp120-gp41 interaction, which has been deemed one of the most desirable structures in biology.

## **5.2 Future Studies**

The work I have presented in this thesis has given new insights into the envelope proteins, however, there is still much work to be done. Some future experiments have been indicated based on the work that I have done. With respect to my SARS-CoV studies, it would be interesting to make mutant versions of the S2-HR2 that are designed to disrupt the helix formation at low pH and to test the effect of these mutations on SARS-CoV entry and the inhibitory properties of S2-HR2. It would also be of interest to examine whether other viral HR2 exhibit a monomer-trimer equilibrium and whether the helical trimer is stabilized under the conditions found in the endosome. This could be particularly important for viruses, such as Influenza and Ebola, which utilize the endosomal pathway of entry. With respect to my HIV studies, the NMR experiments can

be extended to the use of other probes of envelope conformation. For example there are numerous other small molecules that are known to bind to envelope proteins (e.g. CD4-, CDR3- or T20-based peptides). Alternatively, one could identify novel probes by NMR-based screens (e.g. the WaterLOGSY experiment could assay binding of 20 small molecules/hour and thus screen a library of 10,000 compounds in less than a month). Importantly, different conformational states of envelope proteins may be probed by the addition of receptor analogs. It would also be of interest to characterize 12p1, or other probe, structure in the bound state by trNOESY experiments to for improvement of anti-viral properties. With respect to my hypothesis concerning the binding site of 12p1 on gp120, I would suggest NMR studies of mutations to gp120 residues implicated in direct interactions (e.g. W96 or the escape mutants). Finally, my VLP studies are very promising but obviously need to be further developed. Importantly, these studies may eventually be extended to many other enveloped viruses.

## CITED LITERATURE

1. Katanen, M.L., Leinikki, P., Kiusmanen, E., *Endoproteolytic cleavage of HIV-1 gp160 envelope precursor occurs after exit from the trans-Golgo network (TGN)*. Arch. Virol, 1995. **140**(8): p. 1441-9.
2. Chaipan, C., Kobasa, D., Bertram, S., Glowacka, I., Stffen, I., Solomon tsegaye, T., Takeda, M., Bugge, T., Kim, S., Park, Y., Marzi, A., Pohlmann, S., *Proteolytic Activation of the 1918 Influenza Virus Hemagglutinin*. Virology, 2009. **83**(7): p 3200-11.
3. Bullough, P.A., Hughson, F.M., Skehel, J.J., Wiley, D.C., *Structure of influenza haemagglutinin at the pH of membrane fusion*. Nature, 1994. **371**(6492): p. 37-43.
4. Bron, R., Kendal, A.P., Klenk, H.D., Wilschut, J., *Role of the M2 protein in influenza virus membrane fusion: effects of amantadine and monensin on fusion kinetics*. Virology, 1993. **195**(2): p. 808-11.
5. Doms, R.W., Lamb, R.A., Rose, J.K., Helenius, A. *Folding and assembly of viral membrane proteins*. Virology, 1993. **193**(2): p. 545-35.
6. Martin, I. Ruyschaert, J.M., *Common properties of fusion peptides from diverse systems*. Biosci Rep. 2000. **20**(6): P 483-500.
7. Jacobs, A., Simon, C., Caffrey, M., *Thermostability of the HIV gp41 wil-type and loop mutations*. Protein Pept Lett, 2006. **13**(5): 477-80.
8. White, J.M., *Viral and cellular membrane fusion proteins*. Annu Rev Physiol, 1990. **52**: p 675-97.
9. Earp, L.J., Delos, S.E., Park, H.E., White, J.M., *The many mechanisms of viral membrane fusion proteins*. Curr Top Microbil Immunol, 2005. **285**: p 25-66.
10. Taubenberger, J.K., Morens, D.M., *The pathology of influenza virus infections*. Annu Rev Pathol, 2008. **3**: p. 499-522.
11. Caffrey, M., *HIV envelope: challenges and opportunities for development of entry inhibitors*, Trends Microbiol, 2011. **19**(4): p. 191-7.
12. Cervia, J.S., Miriam, A.S., *Enfuvirtide (T-20): A novel human immunodeficiency virus type 1 fusion inhibitor*. Rev Anti-Infective Agents, 2003. **37**: p. 1102-6.
13. LaBonte, J., Lebbos, J., Kirkpatrick, P., *Enfuvirtide*. Nat Rev Drug Discovery, 2003. **2**(5): p. 345-6.

14. Hakansson-McReynolds, S., Jiang, S., Rong, L., Caffrey, M., *Solution structure of the acute respiratory syndrom-coronavirus heptad repeat 2 domain in the prefusion state*. J Biol Chem, 2006. **281**(17): p. 1965-71.
15. Huth, J.R., Bewley, C.A., Jackson, B.M., Hinnebusch, A.G., Clore, G.M. and Gronenborn, A.M., *Design of an expression system for detecting folded protein domains and mapping macromolecular interactions by NMR*. Protein Sci, 1997. **6**(11): p. 2359-64.
16. Dougherty, W.G., Cary, S.M., Parks, T.D. *Molecular genetic analysys of a plant virus polyprotein cleavage site: a model*. Virology, 1989 **171**(2): p. 356-64.
17. Schuck , P., Radu, C.G., Ward, E.S., *Sedimentation equilibriu analysis of recombinant mouse FcRn with murin IgG1*. Mol Immunol, 1999. **36**(15-16): p. 1117-25.
18. Delaglio, F., Grzesiek, S., Vulster, G.W., Zhu, G., Pfeifer, J., Bax, A., *NMRPipe: a multidimensional spectral processing system based on UNIX pipes*. J BioMol NMR, 1995. **6**(3): p. 277-93.
19. Taylor, S.W., Fahy, E., Murray, J., Capaldi, R.A., Ghosh, S.S., *Oxidative post-translational modification of tryptophan residues in cardiac mitochondrial protein*. J Biol Chem, 2003. **278**(22): p. 19587-90.
20. Yang, X., Florin, L., Farzan, M., Kolchinsky, P., Kwong, P.D., Sodroski, J., Wyatt, R., *Modifications that stabilize human immunodeficiency virus envelope glycoprotein trimers in solution*. J Virol, 2000. **74**(10): p. 4746-54.
21. Yang, X., Lee, J., Mahoney, E.M., Kwong, P.D., Wyatt, R., Sodroski, J., *Highly stable trimers formed by human immunodeficiency virus type 1 envelope glycoproteins fused with the trimeric motif of T4 bacteriophage fibritin*. JVirol, 2002. **76**(0): p. 4634-42.
22. Zhang, P.F., Cham, F., Dong, M., Choudhary, A., Bouma, P., Zhang, Z., Shao, Y., Feng, Y.R., Wang, L., Mathy, N., Voss, G., Broder, C.C., Quinnan G.V. Jr., *Extensively cross-reactive anti-HIV-1 neutralizing antibodies induced by gp140 immunization*. Proc Natl Acad Sci USA, 2007. **104**(24): p. 10193-8.
23. Dalvit, C., Fogliatto, G., Stewart, A., Veronesi, M., Stockman., B., *WaterLOGSY as a method for primary NMR screening: practical aspects and range of applicability*. J Biomol NMR, 2001. **21**(4): p. 349-59.
24. Jacobs, A., Sen, J., Rong, L., Caffrey, M., *Alanine scanning mutants of the HIV gp41 loop*. J Biol Chem, 2005. **280**(29): p. 27284-8.

25. Yi, H.A., Diaz-Aguilar, B., Bridon, D., Quaraishi, O., Jacobs, A., *Permanent inhibition of viral entry by covalent entrapment of HIV gp41 of the virus surface*. Biochemistry, 2011. July 15 Epub ahead of print
26. Stadler, K., Masignani, V., Eickmann, M., Becher, S., Abrignani, S., Klenk, H.D., Rappuoli, R., *SARS-beginning to understand a new virus*. Nat Rev Microbiol, 2003 **1**(3): p. 209-18.
27. Drosten, C., Preiser, W., Gunther, S., Schmitz, H., and Doerr, H.W., *Severe acute respiratory syndrome: identification of the etiological agent*. Trends Mol Med, 2003. **9**(8): p.325-7.
28. Ksiazek, T.G., Erdman, D., Goldsmith, C.S., Zaki, S.R., Peret, T., Emery, S., Tong, S., Urbani, C., Comer, J.A., Lim, W., Rollin, P.E., Dowell, S.F., Ling, A.E., Humphry, C.D., Shieh, W.J., Guarner, J., Paddock, C.D., Rota, P., Fields, B., DeRisi, J., Yang, J.Y., Cox, N., Hughes, J.M., LeDuc, J.W., Bellini, W.J., and Anderson, L.J., *A novel coronavirus associated with severe acute respiratory syndrome*. N Engl J Med, 2003. **348**(20): p.1953-66.
29. Marra, M.A., Jones, S.J., Astell, C.R., Holt, R.A., Brooks-Wilson, A., Buttefield, Y.S., Khattra, J., Asano, J.K., Barber, S.A., Chan, S.Y., Cloutier, A., Coughlin, S.M., Freeman, D., Girn, N., Griffith, O.L., Leach, S.R., Mayo, M., McDonald, H., Montgomery, S.B., Pandoh, P.K., Ptrescu, A.S., Robertson, A.G., Schein, J.E., Siddiqui, A., Smailus, D.E., Stott, J.M., Yang, G.S., Plummer, F., Andonov, A., Artsob, H., Bastien, N., Bernard, K., Booth, T.F., Bowness, D., Czub, M., Drebot, M., Fernando, L., Flick, R., Garbutt, M., Gray, M., Grolla, A., Jones, S., Feldmann, H., Meyers, A., Kabani, A., Li, Y., Norman, S., Stroher, U., Tipples, G.A., Tyler, S., Bogrig, R., Ward, D., Watson, B., Brunham, R.C., Kraiden, M., Petric, M., Skowronski, D.M., Upton, C., and Roper, R.L., *The Genome sequence of the SARS-associated coronavirus*. Science, 2003. **300**(5624): p. 1399-404.
30. Rota, P.A., Oberste, M.S., Monroe, S.S., Nix, W.A., Campagnoli, R., Icenogle, J.P., Penaranda, S., Bankamp, B., Maher, K., Chen, M.H., Tong, S., Tamin, A., Lowe, L., Frace, M., DeRisi, J.L., Chen, Q., Want, D., Erdman, D.D., Peret, T.C., Burns, C., Ksiazek, T.G., Rollin, P.E., Sanchez, A., Liffick, S., Holloway, B., Limor, J., McCaustland, K., Olsen-Rasmussen, M., Fouchier, R., Gunther, S., Osterhaus, A.D., Drosten, C., Pallansch, M.A., Anderson L.J. and Belini, W.J., *Characterization of a novel coronavirus associated with severe acute respiratory syndrome*. Science, 2003. **300**(5624): p. 1394-9.
31. Simmons, G., Reeves, J.D., Rennekamp, A.J., Amber, S.M., Piefer, A.J., Bates, P., *Characterization of severe acute respiratory syndrom-associated coronavirus (SARS-CoV) spike glycoprotein-mediated viral entry*. Proc Natl Acad Sci USA, 2004 **101**(12): p. 4240-5.

32. Simmons, G., Gosalia, D.N., Rennekamp, A.J., Reeves, J.D., Diamond, S.L., Bates, P., *Inhibitors of cathepsin L prevent severe acute respiratory syndrome coronavirus entry*. Proc Natl Acad Sci USA, 2005 **102**(33): p. 11876-81.
33. Bosch, B.J., Martina, B.E., Van Der Zee, R., Lepault, J., Haijema, B.J., Versluis, C., Heck, A.J., De Groot, R., Osterhaus, A.D., Rottier, P.J., *Severe acute respiratory syndrome coronavirus (SARS-CoV) infection inhibition using spike protein heptad repeat-derived peptides*. Proc Natl Acad Sci USA, 2004. **101**(22): p. 8455-60.
34. Matsuyama, S., Ujike, M., Marikawa, S., Tashiro, M., Taguchi, F., *Protease-mediated enhancement of SARS coronavirus infection*. Proc Natl Acad Sci USA, 2005 **102**(35): p. 12543-7.
35. Nagata, N., Iwata-Yoshikawa, N., Taguchi, F., *Studies of severe acute respiratory syndrome coronavirus pathology in human cases and animal models*. Vet Pathol, 2010. **47**(5): p. 881-92.
36. Gallagher, T.M., Buchmeier, M.J., *Coronavirus spike proteins in viral entry and pathogenesis*, Virology. **279**(2): p. 371-4.
37. Song, H.C., Seo, M.T., Stadler, K., Yoo, B.J., Choo, Q.L., Coates, S.R., Uematsu, Y., Harada, T., Greer, C.E., Polo, J.M., Pileri, P., Eickmann, M., Rappuoli, R., Abrignani, S., Houghton, M., Han, J.H., *Synthesis and characterization of a native, oligomeric form of recombinant severe acute respiratory syndrome coronavirus spike glycoprotein*. J Virol, 2004. **78**(19): p. 10328-35.
38. Lavillette, D., Barbouche, R., Yao, Y., Besson, B., Cosset, F.L., Jones, I.M., Fenouillet, E., *Significant redox insensitivity of the functions of the SARS-CoV spike glycoprotein: comparison with HIV envelope*. J Biol Chem, 2006. **281**(14): p. 9200-4.
39. Chan, W.E., Chuang, C.K., Yeh, S.H., Chang, M.S., Chen, S.S., *Functional characterization of heptad repeat 1 and 2 mutants of the spike protein of severe acute respiratory syndrome coronavirus*. J Virol, 2006. **80**(7): p. 3225-37.
40. Hernandez, L.D., Hoffman, L.R., Wolfsberg, t.G., White, J.M., *Virus-cell and cell-cell fusion*. Annu Rev Cell Dev Biol, 1996. **12**: p. 627-61.
41. White, J.M., *Membrane fusion*. Science, 1992. **258**(5084): p. 917-24.
42. Dimitrov, D.S., *How do viruses enter cells? The HIV coreceptors teach us a lesson of complexity*. Cell, 1997. **91**(6): p. 721-30.

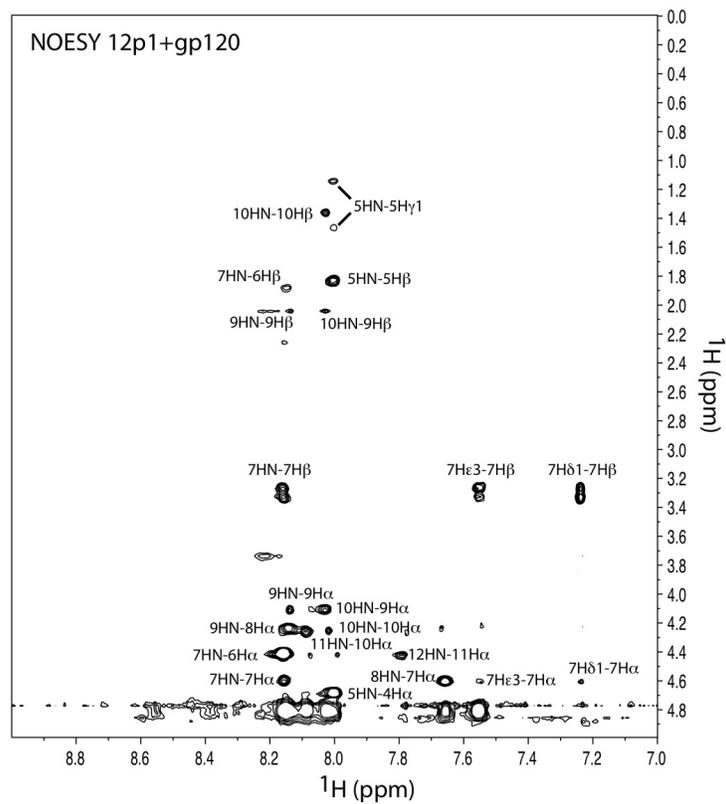
43. Freed, E.O., Martin, M.A., *the role of human immunodeficiency virus type 1 envelope glycoproteins in virus infection*. J Biol Chem, 1995. **270**(41): p. 23883-6.
44. Bosch, B.J., van der Zee, R., de Haan, C.A., Rottier, P.J., *The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex*. J Virol, 2003. **77**(16): p. 8801-11.
45. Liu, S., Xiao, G., Chen, Y., He, Y., Niu, J., Escalante, C.R., Xiong, H., Farmer, J., Debnath, A.K., Tien, P., Jiang, S., *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, 2004. **363**(9413): p. 938-47.
46. Tripet, B., Howard, M.W., Jobling, M., Holmes, R.K., Holmes, K.V., Hodges, R.S., *Structural characterization of the SARS-coronavirus spike S fusion protein core*. J Biol Chem, 2004. **279**(20): p. 20836-49.
47. Supekar, V.M., Bruckmann, C., Ingallinella, P., Bianchi, E., Pessi, A., Carfi, A., *Structure of a proteolytically resistant core from the severe acute respiratory syndrome coronavirus S2 fusion protein*, Proc. Natl. Acad. Sci. USA, 2004. **101**(52): p. 17958-63.
48. Xu, Y., Liu, Y., Lou, Z., Qin, L., Li, X., Bai, Z., Pang, H., Tien, P., Gao, G.F., Rao, Z., *Structural basis for coronavirus-mediated membrane fusion. Crystal structure of mouse hepatitis virus spike protein fusion core*. J Biol Chem, 2004. **279**(29): p. 30514-22.
49. Xu, Y., Lou, Z., Liu, Y., Pang, H., Tien, P., Gao, G.F., Rao, Z., *Crystal structure of severe acute respiratory syndrome coronavirus spike protein fusion core*. J Biol Chem, 2004. **279**(47): p. 49414-9.
50. Duquerroy, S., Vigouroux, A., Rottier, P.J., Rey, F.A., Bosch, B.J., *Central ions and lateral asparagine/glutamine zippers stabilize the post-fusion hairpin conformation of the SARS coronavirus spike glycoprotein*, Virology, 2005 **335**(2): p. 76-85.
51. Deng, Y., Liu, J., Zheng, Q., Yong, W., Lu, M., *Structures and polymorphic interactions of two heptad-repeat regions of the SARS virus S2 protein*. Structure, 2006. **14**(5): p. 889-99.
52. McReynolds, S., Jiang, S., Guo, Y., Celigoy, J., Schar, C., Rong, L., Caffrey, M., *Characterization of the prefusion and transition states of severe acute respiratory syndrome coronavirus S2-HR2*. Biochemistry, 2008. **47**(26): p. 6802-8

53. McReynolds, S., Jiang, S., Rong, L., Caffrey, M., *Dynamics of SARS-coronavirus HR2 domain in the prefusion and transition states*. J Magn Reson, 2009. **201**(2): p. 218-21.
54. Guo, Y., Tisoncik, J., McReynolds, S., Farzan, M., Prabhakar, B.S., Gallagher, T., Rong, L., Caffrey, M., *Identification of a new region of SARS-CoV S protein critical for viral entry*. J Mol Biol, 2009. **394**(4): p. 600-5.
55. Scholtz, J.M., York, E.J., Stewart, J.M., Baldwin, R.L., A neutral, water-soluble,  $\alpha$ -helical peptide: the effect of ionic strength on the helix-coil equilibrium, J. Am. Chem. Soc., 1991. **113**: p. 5102-4.
56. Baldwin, R.L., *How Hofmeister ion interactions affect protein stability*, Biophys. J., 1996. **71**(4): p. 2056-63.
57. Carr, C.M., Kim, P.S., *A spring-loaded mechanism for the conformational change of influenza hemagglutinin*. Cell, 1993. **73**(4): p. 823-32.
58. Skehel, J.J., Wiley, D.C., *Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin*. Annu Rev Biochem, 2000. **69**: p. 531-69.
59. Fauci, A.S. *Twenty-five years of HIV/AIDS*. Science, 2006. **313**(5786): p. 409
60. Simon, V., Ho, D.D., Abdool Karim, Q., *HIV/AIDS epidemiology, pathogenesis, prevention, and treatment*. Lancet, 2006. **368**(9534): p. 489-504.
61. Gaschen, B., Taylor, J., Yusim, K., Foley, B., Gao, F., Lang, D., Novitsky, V., Haynes, B., Hahn, B.H., Bhattacharya, T., Korber, B., *Diversity considerations in HIV-1 vaccine selection*. Science, 2002. **296**(5577): p. 2354-60.
62. Roux, K., and Taylor, K., *AIDS virus envelope spike structure*. Curr Opin Struct Biol, 2007. **17**: p. 1-9.
63. Delaney, M. *The development of combination therapies for HIV infection*. AIDS Res Hum Retroviruses. **26**(5): p. 501-9.
64. Eckert, D.M., and Kim, P.S. *Mechanisms of viral membrane fusion and its inhibition*. Annu Rev Biochem, 2001. **70**: p. 777-810.
65. Gallo, S., Finnegan, C., Viard, M., Taviv, Y., Dimitrov, A., Raway, S., Puri, A., Durrell, S., Blumenthal, R., *The HIV Env-mediated fusion reaction*. Biochimica Biophys Acta, 2003. **1614**(1): p. 36-50.
66. Kwong, P.D., Wyatt, R., Robinson, J., Sweet, R.W., Sodroski, J., Hendrickson, W.A., *Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody*. Nature, 1998. **393**(6686): p. 648-59.

67. Chen, B., Vogan, E., Gong, H., Skehel, J., Wiley, D., Harrison, S., *Structure of an unliganded simian immunodeficiency virus gp120 core*. Nature, 2005. **433**: p. 834-41.
68. Pancera, M., Majeed, S., Ban, Y.E., Chen, L., Huang, C.C., Kong, L., Kwon, Y.D., Stuckey, J., Zhou, T., Robinson, J.E., Schief, W.R., Sodroski, J., Wyatt, R., Kwong, P.D., *Structure of HIV-1 gp120 with gp41-interactive region reveals layered envelope architecture and basis of conformational mobility*. Proc Natl Acad Sci USA, 2010. **107**(3): p. 1166-71.
69. Chan, D., Fass, D., Berger, J., Kim, P., *Core structure of gp41 from the HIV envelope glycoprotein*. Cell, 1997. **89**: p. 263-73.
70. Weissenhorn, W., Dessen, A., Harrison, S., Skehel, J., Wiley, D., *Atomic structure of the ectodomain from HIV-1 gp41*. Nature, 1997, **387**: p. 426-30.
71. Caffrey, M., Cai, M., Kaufman, J., Stahl, S., Wingfield, P., Gronenborn, A., Clore, G., *Determination of the secondary structure and global topology of the 44 kDa ectodomain of gp41 of the simian immunodeficiency virus by multidimensional nuclear magnetic resonance spectroscopy*. J Mol Biol, 1997. **271**: p. 819-26.
72. Bhattacharya, A., *Protein structures: structures of desire*. Nature, 2009. **459**(7243): p. 24-7.
73. Ferrer, M., and Harrison, S.C., *Peptide ligands to human immunodeficiency virus type 1 gp120 identified from phage display libraries*. J Virol, 1999. **73**(7): p. 5795-802.
74. Biorn, A.C., Cocklin, S., Madani, N., Si, Z., Ivanovic, T., Samanen, J., Van Ryk, D.I., Pantophlet, R., Burton, D.R., Freire, E., Sodroski, J., Chaiken, I.M., *Mode of action for linear peptide inhibitors of HIV-1 gp120 interactions*. Biochemistry, 2004. **43**(7): p. 1928-38.
75. Gopi, H.N., Tirupula, K.C., Baxter, S., Ajith, S., Chaiken, I.M. *Click chemistry on azidoproline: high-affinity dual antagonist for HIV-1 envelope glycoprotein gp120*. ChemMedChem, 2006. **1**(1): p. 54-7.
76. Gopi, H.N., Umashankara, M., Pirrone, V., LaLonde, J., Madani, N., Tuzer, F., Baxter, S., Zentner, I., Cocklin, S., Jawanda, N., Miller, S.R., Schon, A., Klein, J.C., Freire, E., Krebs, F.C., Smith, A.B., Sodroski, J., Chaiken, I., *Structural determinants for affinity enhancement of dual antagonist peptide entry inhibitor of human immunodeficiency virus type-1*. J Med Chem, 2008. **51**(9): p. 2638-47.

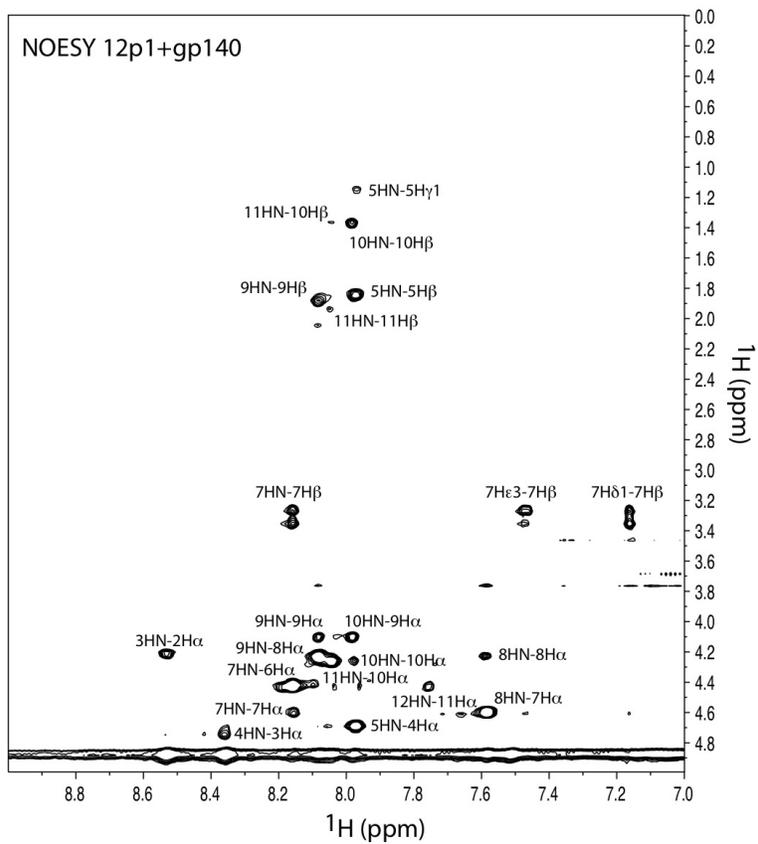
77. Meyer, B., Klein, J., Mayer, M., Meinecke, R., Moller, H., Neffe, A., Schuster, O., Wulfken, J., Ding, Y., Knaie, O., Labe, J., Palcic, M.M., Hindsgaul, O., Wagner, B., Ernst, B., *Saturation transfer difference NMR spectroscopy for identifying ligand epitopes and binding specificities*. Ernst Schering Res Found Workshop, 2004. **44**: p. 149-67.
78. Schwarzing, S., Kroon, G.J., Foss, T.R., Wright, P.E., Dyson, H.J., *Random coil chemical shifts in acidic 8 M urea: implementation of random coil shift data in NMRView*. J Biomol NMR, 2000. **18**(1): p. 43-8.
79. Archary, D., Gordon, M.L., Green, T.N., Coovadia, H.M., Goulder, P.J., Ndung'u, T., *HIV-1 subtype C envelope characteristics associated with divergent rates of chronic disease progression*. Retrovirology, 2010. **7**: p. 92.
80. Binley, J.M., Sanders, R.W., Clas, B., Schuelke, N., Master, A., Guo, Y., Kajumo, F., Anselma, D.J., Maddon, P.J., Olson, W.C., Moore, J.P., *A recombinant human immunodeficiency virus type 1 envelope glycoprotein complex stabilized by an intermolecular disulfide bond between the gp120 and gp41 subunits is an antigenic mimic of the trimeric virion-associated structure*. J Virol, 2000. **74**(2): p. 627-43.
81. Sen, J., Jacobs, A., Caffrey, M., *Role of the HIV gp120 conserved domain 5 in processing and viral entry*. Biochemistry, 2008. **47**(30): p. 7788-95.
82. Caffrey, M., *Model for the structure of the HIV gp41 ectodomain: insight into the intermolecular interactions of the gp41 loop*. Biochim Biophys Acta, 2001. **1536**(2-3): p. 116-22.
83. Kimpton, J., Emerman, M., *Detection of replication-competent and pseudotyped human immunodeficiency virus with a sensitive cell line on the basis of activation of an integrated beta-galactosidase gene*. J Virol, 1992. **66**(4): p. 2232-9.

## APPENDIX A



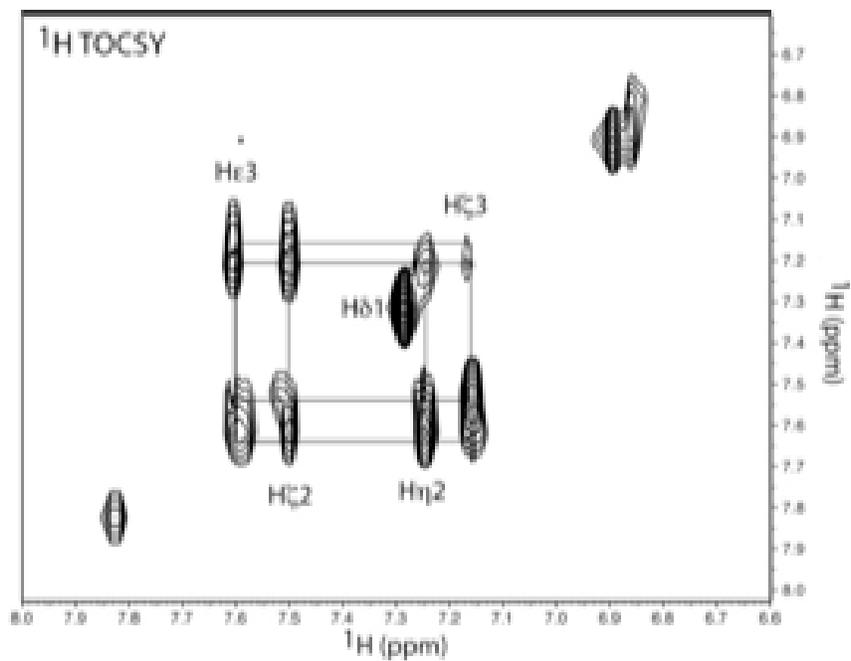
trNOESY of 12p1 in the presence of gp120. Experimental conditions were 400  $\mu\text{M}$  12p1 and 2  $\mu\text{M}$  gp120 in PBS at 25°C with a mixing time of 200 msec.

## APPENDIX B



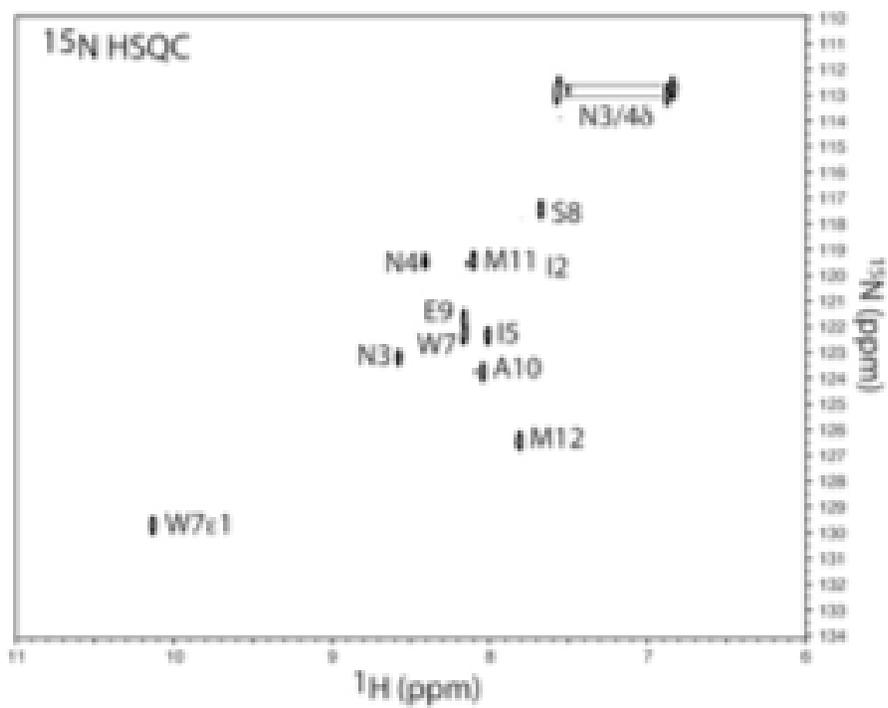
trNOESY of 12p1 in the presence of gp140. Experimental conditions were 400  $\mu\text{M}$  12p1 and 2  $\mu\text{M}$  gp140 in PBS at 25°C with a mixing time of 200 msec.

## APPENDIX C



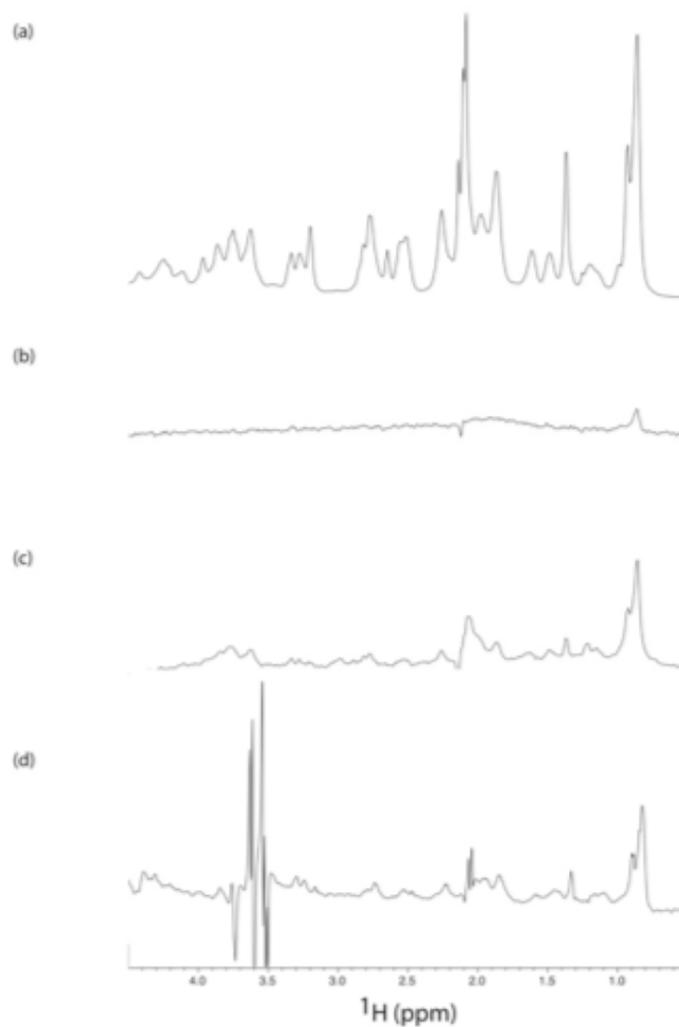
Aromatic region of 12p1 TOCSY. Experimental conditions were 2 mM 12p1 in 20 mM PO<sub>4</sub>/pH 7.2, 90% <sup>1</sup>H<sub>2</sub>O/10% <sup>2</sup>H<sub>2</sub>O.

## APPENDIX D



<sup>15</sup>N-edited HSQC at natural abundance. Experimental conditions were 2 mM 12p1 in 20 mM PO<sub>4</sub>/pH 7.2, 90% <sup>1</sup>H<sub>2</sub>O/10% <sup>2</sup>H<sub>2</sub>O.

## APPENDIX E



STD NMR spectra of the 12p1 aliphatic region in the presence of R2gp120 and gp140 constructs. A) Reference 1D NMR spectrum of 12p1. B) STD MR spectrum of 12p1 in the absence of gp120 or gp140. C) STD NMR spectrum of 12p1 in the presence of gp120 from strain R2. D) STD NMR spectrum of 12p1 in the presence of gp140 from strain R2. The experimental conditions were 1 mM 12p1 +/- 4  $\mu\text{M}$  gp120 in 20 mM  $\text{PO}_4/\text{pH}$  7.2, 90%  $^1\text{H}_2\text{O}/10\%$   $^2\text{H}_2\text{O}$ .

## VITA

- NAME: Jessica Celigoy
- EDUCATION: B.S. Chemistry, Gonzaga University, Spokane, WA 2005  
Ph.D., Biochemistry and Molecular Genetics, University of Illinois at Chicago, Chicago, IL, 2011
- PUBLICATIONS: Celigoy, J., Ramirez, B., Caffrey, M.: SARS-CoV heptad repeat 2 is a trimer of parallel helices. Protein Sci. Sep 15 Epub ahead of print
- Celigoy, J., McReynolds, S., Caffrey, M.: The SARS-CoV heptad repeat 2 exhibits pH-induced helix formation. Biochem Biophys Res Commun. 412(3):483-6, 2011
- Celigoy, J., Ramirez, B., Tao, L., Rong, L., Yan, L., Feng, Y.R., Quinnan, G.V., Broder, C.C., Caffrey, M.: Probing the HIV gp120 envelope glycoprotein conformation by NMR. J Biol Chem 286(27): 23975-81, 2011
- McReynolds, S., Jiang, S., Guo, Y., Celigoy, J., Schar, C., Rong, L., Caffrey, M.: Characterization of the prefusion and transition states of severe acute respiratory syndrome coronavirus S2-HR2. Biochemistry. 47(26): 6802-8, 2008