8-2018

Contributions of the gp120 Variable Loops to Envelope Glycoprotein Trimer Stability in Primate Lentiviruses

Dane Bowder
University of Nebraska-Lincoln

CONTRIBUTIONS OF THE gp120 VARIABLE LOOPS TO ENVELOPE GLYCOPROTEIN TRIMER STABILITY IN PRIMATE LENTIVIRUSES

by

Dane M. Bowder

A DISSERTATION

Presented to the Faculty of
The Graduate College at the University of Nebraska
In Partial Fulfillment of the Requirements
For the Degree of Doctor of Philosophy

Major: Biological Sciences

(Genetics Cellular and Molecular Biology)

Under the Supervision of Professor Shi-Hua Xiang

Lincoln, Nebraska

August, 2018

CONTRIBUTIONS OF THE gp120 VARIABLE LOOPS TO ENVELOPE GLYCOPROTEIN TRIMER STABILITY IN PRIMATE LENTIVIRUSES

Dane Bowder, Ph.D.

University of Nebraska, 2018

Advisor: Shi-Hua Xiang

Human Immunodeficiency Virus (HIV) is the etiological agent of AIDS and is responsible for the AIDS pandemic worldwide. According to UNAIDS, as of 2016, there were approximately 36.7 million people living with HIV globally, and 1.8 million new infections that year. While antiretroviral therapies and education continue to reduce these numbers, a preventative vaccine is still required to curb this epidemic. The envelope glycoprotein trimer of HIV, which is the sole protein on the surface of the virus and facilitates entry of the virus into host cells, is of keen interest to the HIV vaccine and drug-development field. Recently, the structure of this trimer was solved which has opened the door for the design of new immunogens and drugs. These solved envelope protein trimer structures demonstrate that the major variable loops of the env protein are positioned at the apex of the trimer where they play a role in trimer stability. Data from our lab has shown previously that both the V2 and V3 loop contain elements that are critical for envelope trimer stability.

The work presented in this dissertation expands on these initial studies and confirms that the V2 and V3 loops contribute to envelope trimer stability and functionality in HIV-1, HIV-2 and in Simian Immunodeficiency Virus (SIV)

species. We have demonstrated that the hydrophobic patch near the tip of the V3 loop, which is present in all primate lentiviruses, is critical for trimer stability in subtype C HIV-1, HIV-2 and in SIV species. Additionally, we further explored the role of the twin-cysteine motif found in the V2 loop of HIV-2 and SIV species, but not in HIV-1. We have shown that this motif is critical for the stability of the envelope protein of these viruses. Both observations are supported by our experimental data and by molecular modeling. This work supports the released trimer structures and provides a further understanding of the forces that stabilize the envelope trimer of primate lentiviruses, and will aid in the development of stabilized envelope trimer based immunogens in the effort to develop an effective HIV vaccine.

For my family.

To Alison, Mom, Dad and Kyleah.

This wouldn't have been possible without you, and your support means everything.

From the bottom of my heart, I thank you.

ACKNOWLEDGEMENTS

I would like to begin by thanking my advisor, Dr. Shi-Hua Xiang. He has provided the environment for me to learn, conduct these studies and has been extremely supportive along the way. I am forever grateful for his wisdom and guidance in this process, and for taking a chance on me in my first year of graduate school. I would also like to thank my graduate committee members, Dr. Deb Brown, Dr. Qingsheng Li, Dr. Fernando Osorio and Dr. Charles Wood for their support, guidance and feedback throughout this process.

I also must acknowledge the role my family played in supporting me in this process, which was critical for the completion of this process. I thank my wife Alison for her incredible patience and undying support of me pursuing my dreams. I absolutely could not have done this without her by my side. I also must thank my parents, Lynn and Dale, and my sister Kyleah. They have told me for as long as I can remember that I can be anything that I wanted, and that I could do anything I set my mind to. They all have been integral to my success and have believed in me even when I did not have the capacity to believe in myself. I would also like to thank my parents in-law, Jane and Lowell for always checking and asking me about how things were going in the lab, and for their consummate support of my work. Finally, I would like to thank my extended family, particularly my grandmother Greta, for being one of my biggest cheerleaders and for weekly home-cooked lunches throughout graduate school; and my late grandmother Delores, who learned the word “virology” just so she could brag to her friends about me and my work. She is very dearly missed.

I would also like to thank my fellow graduate students and post-docs who played a role in my academic development. Specifically, I would like to thank Dr. Jesse Thompson, Dr. Chris Bohl and Dr. Wenzhong Wei for their technical advice and for teaching me many of the techniques that were crucial for this work. Additionally, I would like to thank my preparing future faculty mentor (and now colleague), Dr. Tessa Durham-Brooks. Her insights into teaching, research and professional development have been instrumental to my career success. I would also like to thank Dr. Alex Vogel, Dr. Landon Olp, Dr. Tyler Moore, Duoyi Hu, Josh Wiggins and Dr. Jessica Hargarten, and all the graduate students who worked in the NCV, for all the good times and conversations that we shared as graduate students together, it feels as though we have been to war together and share a collective experience that others will never understand. This work would also not have been possible without the technical staff in the lab who keep things functioning properly. Particularly, I would like to thank you to Michael Mellon and Danielle Shea who helped my research project move forward in ways that may have seemed trivial to them, but were very helpful to me.

Training undergraduate students in the laboratory has been one of the absolute most rewarding experiences I have had during graduate school, and many of the students who worked with me helped me to develop materials for my projects. I thank Kate Durst, Haley Hollingsead, Ngan Nguyen, Ryan Shogren, Jas Kaur, Jennifer Berger, Cindy Tan, Russell Milton and Michelle Quinn for your assistance and interest in learning about biology. These students have taught me as much, if not more, about being a good mentor and teacher, as I have taught

them about how to be a virologist. All these students are doing incredible things with their careers and I couldn't be prouder.

Finally, I need to thank my support system of friends who have been patient with my complicated schedule and limited time during graduate school. My friends have always been a source of light and levity during this program, which has a been a very challenging and at times, dispiriting. I am lucky to have had such an incredible group of supportive friends to lift me up when I needed it. I would specifically like to thank Derek, Micah, Caitlin, Kirsten, Lisa, Seth, Bob, Deepa, Colin, Ross, Andrea, Chris, Joe, Chris, Molly, Derrik and Tanner. All these people mean so much to me, and I look forward to many more years of fun and adventure with them.

The work presented in this dissertation was supported by funds from the National Institutes of Health, the Bill and Melinda Gates Foundation, The Nebraska Center for Virology, the University of Nebraska Agricultural Research Division and the School of Biological Sciences.

To conclude, this program has truly been the most difficult journey that I have ever embarked on. To everyone listed here, you have all played a role in my growth and development as a scientist and a person. None of this would have been possible without every single one of you, and I thank you all.

TABLE OF CONTENTS

List of Tables

List of Figures

Research Aims

The objective of this dissertation research was to further characterize the activity of the major variable loops (V2 and V3) of the primate lentiviral envelope protein in envelope trimer stability and functionality. The V2 and V3 loops have been shown previously to play an important role in the stability of the Human Immunodeficiency Virus (HIV) envelope trimer, and in turn, the functionality of this complex. The advent of near atomic structures of the trimer has revealed that V2 and V3 are packed tightly into the apex of the envelope trimer, where they exist as part of the trimer association domain [1-3]. Because of the potency and breadth of trimer specific broadly neutralizing antibodies (bNAbs), understanding the features of the envelope trimer structure and functionality can contribute to the design and development of a trimer based immunogen that could be used in the development of a successful vaccine and novel drugs that target the viral entry process [4].

Based on sequence homology, the V3 loop length is conserved in most primate lentiviral species and insertions or deletions to the length of the V3 loop have been shown to reduce envelope stability and functionality in subtype B HIV-1 viruses. Additionally, a patch of three hydrophobic residues at the tip of the V3 loop is conserved across primate lentiviruses. While the exact amino acid sequence of this motif is capable of some variability, the residues present are consistently hydrophobic. This hydrophobic patch motif has also been shown to contribute to envelope trimer stability and functionality in subtype B HIV-1 [5]. Because of the conservation of this length and hydrophobic patch in other HIV-1

subtypes and primate lentiviruses (PLVs), we sought to explore whether this function was conserved as well. Our hypothesis is that additions to the V3 loop length in subtype C HIV, HIV-2 and Simian Immunodeficiency Virus (SIV) species will result in a similar reduction in envelope stability and functionality. Additionally, substitution of the hydrophobic patch with hydrophilic and small neutral amino acids will also result in the destabilization of the envelope complex, regardless of PLV species.

Additionally, our lab recently identified and characterized a twin-cysteine motif (TCM) in the V2 loop of the SIVmac strain, SIVmac39. This motif is present in all SIV and HIV-2 species, except SIVcpz, where it is present in some strains, but not in others. This motif is notably absent from all HIV-1 strains. We showed that removal of either, or both, of these twin cysteines from SIVmac239 resulted in a dramatic decrease in envelope stability, indicated by an increase in gp120 shedding. This decrease in stability correlated with a decrease in the envelope's ability to initiate cell-cell fusion and viral entry [6]. We hypothesize that this twin cysteine motif is forming an important disulfide bond that stabilizes the envelope gp120 molecule, contributing to the overall stability of the trimeric complex. To further explore the role that these cysteines play, our hypothesis is that the removal of this motif from PLV envelopes where it is naturally present will result in a similar decrease in envelope stability and functionality, and that the addition of this motif to HIV-1 envelopes will further stabilize the envelope complex.

Understanding the structure and biological function of the primate lentiviral envelope is paramount to the development of innovative treatments and

preventative cures. In the development of an HIV vaccine, the envelope trimer is a strong candidate as an immunogen because of development of bNAbs. Many of these bNAbs target trimer specific epitopes and a trimer based, or trimer specific epitope based design will be required to generate these antibodies in a host. While the envelope modifications presented here may not singularly lead to the direct development of an envelope based HIV immunogen, these modifications could contribute to better design and development of a stabilized envelope trimer alongside other stabilizing modifications, which could be used as an HIV vaccine. The data in this dissertation contributes to the overall understanding of the stability and function of the viral envelope trimer and assists in validating the envelope trimer structures that have been released.

CHAPTER 1- LITERATURE REVIEW

Human Immunodeficiency Virus- Introduction

Human Immunodeficiency Virus Type-1 (HIV-1) is a human retrovirus in the Retroviridae family, and is part of the lentivirus genus. This virus is the causative agent of Acquired Immunodeficiency syndrome (AIDS) and has been the cause of a major global epidemic for the past 37 years. HIV-1 primarily infects CD4+ T-lymphocytes and causes a slow, long-term infection, eventually depletes the CD4+ T cell population from its host, causing severe immunodeficiency in advanced cases. This immunodeficiency puts the host at major risk for co-infection by opportunistic infections and cancers. Despite extensive research and many clinical trials, there is currently no cure for the disease, or a sterilizing preventative vaccine for this virus, though significant advancements have been made in antiretroviral and prophylactic drugs. The World Health Organization estimates that as of 2016, there were 36.7 million individuals living with HIV worldwide, and that there were 1.8 million new infections. In the same year, it was estimated that there were 1.0 million HIV related deaths [7]. Fortunately, the number of new infections each year and the number of deaths per year continue to decline, however, HIV continues to be a major health crisis, and effective treatment and prevention strategies are necessary to effectively halt the spread of this virus. While the treatment of the HIV epidemic has been improving globally, HIV continues to be a major global pathogen, with a majority of infections occurring on the African continent, specifically in Eastern and Southern Africa [8].

HIV was first discovered in 1983 as a retrovirus that was causing the AIDS epidemic occurring in gay men, intravenous drug users and hemophiliacs in the United States [9, 10]. Although HIV was first discovered in the US in the early 1980s, phylogographic studies examining viral sequence divergence and human populations have demonstrated the origin of the HIV epidemic occurred in the early part of the 20th century in central Africa. It is likely that since the original zoonotic transmission of a Simian Immunodeficiency virus (SIV) that infects chimpanzees (SIVcpz) in the early 1900s that HIV-1 spread through Africa undetected for the next fifty years until it was finally detected in the United States and Europe in the 1980s where it spread rapidly [11].

Virion Structure

HIV virions are typically between 120 and 150 nm in diameter, enveloped by a viral membrane derived from the host cell and have a cone-shaped core containing two copies of the viral RNA genome which is in complex with the nucleocapsid protein, p7 [12, 13]. The structural protein matrix (MA), p17 of Gag, is present immediately beneath the cellular membrane and binds to the cellular membrane forming the spherical outer shell of the core [14, 15]. The cone-shaped core is made up of the Gag gene product CA (p24) and encapsulates the viral RNA genome, and is characteristic of lentiviruses [13]. Within the virion, copies of the viral enzymes protease, integrase, reverse transcriptase and select accessory proteins are present and are transported to the host cell upon infection [16]. Envelope protein trimers are embedded in the membrane of the virion, with

10-15 envelope complexes per virion [17, 18] (Figure 1.1) The HIV-1 genome is approximately 9.7 kb single stranded RNA and contains nine open reading frames of three primary structural genes and enzymes, yielding a total of nine essential proteins for viral structure and replication (Gag proteins- Matrix, Capsid, p6, Nucleocapsid; Pol enzymes- Reverse Transcriptase, Integrase, Protease; and Envelope subunits- gp120, gp41), two essential regulatory elements (Rev, Tat) and four accessory proteins that enhance viral pathogenesis (Vif, Vpr, Vpu, Nef) [19, 20]. Flanking both the 5' and 3' ends of the viral genome is a sequence known as the long terminal repeat (LTR). This sequence contains the promoter region for transcription and is the binding site for cellular transcription factors and regulatory elements, as well as the poly-adenylation sequence (which is only functional on the 3' end of the genome) [20, 21]. The terminal ends of the genome are also important for the process of reverse transcription, as the 5' and 3' ends both contain repeated sequences, which allow for the first jump in the process of reverse transcription to occur, which results in the formation of the sequentially identical LTR in the integrated provirus [22]. Additionally, the LTRs are the targets of integration into the host genome [23, 24]. A schematic of the viral genome structure is shown in Figure 1.2 [19].

HIV Replication Cycle

The HIV life cycle begins with binding of the envelope protein on the surface of the virus to the primary receptor, CD4, at the outer membrane of a host cell. Prior to specific interaction between Env and CD4, non-specific

interactions with other cellular proteins including DC-SIGN, α4β7 integrins and heparans have been observed, but are not required for viral entry [25-27]. Following CD4 binding, a conformational change occurs within the envelope trimer on the virus, exposing the binding site for the coreceptor [28]. This allows envelope to engage with the coreceptor for the virus, either CCR5 or CXCR4. Next, fusion of the viral and cellular membranes occurs, which allows the release of the viral capsid into the cytoplasm of the cell. This is a pH independent process, and is thought to occur at the cellular membrane, but there have been reports that fusion may occur in endosomes [29-31]. Following fusion, the viral capsid is released into the cytoplasm, and partial uncoating occurs. At this point, reverse transcription begins to occur within the partial capsid. Reverse transcription is carried out by the viral reverse transcriptase enzyme present in the virion, and ultimately generates a double stranded DNA molecule from the viral RNA [32-34]. After reverse transcription, viral proteins, including Gag components, Vpr and cellular proteins come together to transport the double stranded viral DNA genome to the nucleus, this conglomeration is referred to as the pre-integration complex (PIC) [20, 35, 36]. Once in the nucleus, the viral DNA is integrated into the host genome via the activity of the viral integrase enzyme [37, 38]. Following integration, the viral genome serves as a template for transcription by host transcription machinery, producing full length viral RNA strands. Some of these strands are partially or completely spliced to produce viral proteins through translation [20]. Viral RNAs are exported from the nucleus by the activity of the Rev protein, which interacts with the rev response element

(RRE) located within an intron located the *env* gene sequence [20, 39, 40]. Translation of viral proteins occurs primarily in the cytoplasm, with Gag and Gag-Pol precursors being transported to the cellular membrane for post translational virus assembly [16, 41]. Simultaneously, the envelope protein is translated on the rough endoplasmic reticulum, and is glycosylated, cleaved by furin proteases and transported through the Golgi apparatus. Envelope trimers then interact with the MA protein at the cell surface and are incorporated by an unknown mechanism [15, 42]. Viral RNA dimers interact with the Gag precursor and are packaged into budding virions [43, 44]. Viral budding occurs from the cellular membrane and requires activity of the cellular endosomal sorting complexes required for transport (ESCRT) machinery, releasing immature viral particles from the cell [45]. Following viral particle release, maturation of the viral core occurs via the action of the HIV protease (PR) enzyme which is packaged into virions. PR proteolytically cleaves the Gag precursor into MA, CA, NC and p6 which results in the formation of the characteristic cone shaped core, and renders the viral particle infectious [16, 46, 47].

HIV Gene Products

Gag

The gag gene produces a precursor protein, pr55. This polyprotein is cleaved by the protease enzyme into the MA (p17), CA (p24), NC (p7) and p6 during viral maturation following budding from the host cell. Prior to maturation, the gag polyprotein makes contact with the cellular membrane via a

myristoylated glycine residue at the N-terminus of the Gag precursor [48]. This cleavage and maturation is absolutely required for viral infectivity [47, 49]. Interestingly, expression of the gag gene individually is sufficient to form viral particles [14, 50]. MA is the matrix protein and is located at the N terminus of the Pr55 precursor, it forms the outer shell of the virus and remains bound to cellular membrane. MA has also been shown to play a role in Env recruitment and interacts with envelope via the cytoplasmic tail of gp41 [15]. Directly to the C-terminal end of the MA domain is the capsid protein, CA (p24). Following cleavage, the capsid oligomerizes into primarily hexamers, and forms the inner cone-shaped core, a crystal structure of this the HIV cone-shaped capsid has been published which reveals the arrangement of this protein [46]. C-terminal to the CA protein is the nucleocapsid protein, NC (p7) which has a two zinc finger domains and is bound to the viral RNA [12]. NC plays a critical role in recruiting viral RNA to assembling virions by binding to the packaging signal (Psi), a series of four stem loops located in the 5' untranslated region (UTR), which is located 5' to the gag gene sequence [51-53]. The final gene product of the *gag* gene is p6, a protein domain found primarily in primate lentiviruses. This small protein appears to play an important role in viral budding and in the packaging of the accessory protein, Vpr into viral particles [41, 54-56].

Pol

The *pol* gene is synthesized along with the *gag* gene from a single RNA molecule that is the result of a ribosomal frameshift in the *gag* gene that occurs in only at a rate of about 5%. Because of this, a ratio of 20:1, Gag to Gag-Pol

proteins are maintained and are important for viral genome dimerization and function [57]. This produces a single Gag-Pol polyprotein from which some Gag products and the viral enzymes, protease (PR), reverse transcriptase (RT) and integrase (IN) are produced [16, 58]. The enzymes produced by *pol* are essential for infection and are packaged into virions [16, 47]. The viral protease enzyme is responsible for the cleavage of the Gag and Pol polyprotein precursors. This enzyme is a functional dimer and is essential for viral maturation and infectivity [47]. Because the action of PR is vital for viral function, it is a desirable target for anti-HIV drug development and several protease inhibitors are used with great effectiveness [59]. Despite this, because of the high mutation rate in HIV, resistance to protease inhibitors occurs [60].

RT is a multifunctional enzyme with RNA dependent DNA polymerase and DNA dependent DNA polymerase activities, which combined synthesize a double stranded DNA molecule from the viral RNA genome to be incorporated into the host genome [32, 34]. RT also has RNAseH activity, which is utilized to digest the viral RNA template in RNA-DNA intermediate molecules that are generated during reverse transcription [33, 61]. All these activities are critical for viral genome reverse transcription [34, 62]. This enzyme is a heterodimer of two proteins, p66 and p51 for which the crystal structure has been solved [63, 64]. Importantly, the RT lacks proofreading capabilities, so errors are made during the process of reverse transcription at a high rate, leading to the vast genetic diversity of HIV [65]. Again, because RT is essential for viral function, it is the

target of several FDA-approved antiretroviral drugs and is the source of continued drug development [66, 67].

Integrase (IN) incorporates the double stranded viral DNA product following reverse transcription. This occurs in two-steps: 3' end processing, in which two nucleotides are removed from the 3' end of both sides of the viral DNA, followed by DNA strand transfer, in which two strand breaks are made in the host genome, followed by attachment of the viral DNA to the host chromosome [37, 38, 68]. This activity is likely followed by host DNA repair mechanisms to remove the two base flap that is left by the action of IN, ultimately sealing the viral genome into the host genome permanently [69].

Env

The *env* gene produces the major glycoprotein complex that is present on the surface of the virion. This gene is initially translated into a 160 kD precursor protein, which is cleaved by furin proteases in the Golgi apparatus, into two viral protein products- gp41, the transmembrane protein (TM), and gp120 the surface glycoprotein (SU) [42, 70]. These two proteins are assembled into the envelope glycoprotein trimeric complex which is a trimer of heterodimers, in which three gp41 molecules are non-covalently attached to three gp120 molecules. About half of the mass of this protein is made up of N-linked glycans, which are important for immune evasion [71, 72]. Envelope trimers are then incorporated into budding virions via an unclear mechanism, though it has been shown that the activity of cytoplasmic tail of the gp41 subunit of envelope and the MA protein are both necessary for proper incorporation into virions [15]. The envelope trimer

is the primary immunogen on the virus, and is the primary target for antibody development [73, 74]. The primary function of the envelope protein is to bind to the cell surface receptors and facilitate fusion and entry into the cells. The specific binding of the envelope protein to cellular receptors determines cellular tropism [75]. The structure and biological function of the envelope protein will be further discussed at length later in chapter 1, on pages 20-26.

Tat

The tat gene encodes for the trans-activator of transcription, or Tat protein. This protein is 14kDa and is an essential regulatory element in the viral genome, and deletion of this protein results in replication-deficiency [76, 77]. The primary function of the Tat protein is to bind the Trans-Activation Response Region (TAR) at the 5' end of viral RNAs and increase the rate of transcription by interacting with cellular factors that increase the phosphorylation of the RNA pol II [20]. Viral transcription initially occurs without the action of Tat, ultimately resulting in the initial transcription and translation of Tat protein, which drastically increases the amount of transcription of viral gene products [20]. Interestingly, Tat protein is also secreted and contributes to HIV pathogenesis by encouraging the growth of Kaposi's sarcoma lesions and by initiating apoptosis in T-lymphocytes [78-80].

Rev

The second regulatory protein encoded by HIV is Rev, a 19 kDa protein responsible for nuclear export of viral RNAs. Rev stands for Regulator of

Expression of Viral Proteins, and this protein is essential to the viral life cycle [39, 40]. The Rev protein interacts primarily with a three-dimensional stem loop structure in RNA transcripts within the *env* gene sequence called the Rev Response Element (RRE) [20, 39]. Rev binds to the RRE and aids in the export of viral RNAs to the cytoplasm, where efficient translation of viral proteins can occur [81, 82].

Vif

The *vif* gene codes for a 23 kDa protein called Vif, or Viral Infectivity Factor. The primary role of Vif is to counteract the action of cellular APOBECs, a family of cytidine deaminases, which cause hyper mutations of C to U in the HIV viral genome. The APOBECs serve as potent restriction factors of HIV infection, but their action is reduced by the presence of Vif [83, 84]. Initially, the role of Vif in the viral life cycle was somewhat unclear, but it was known that Vif was essential in making some cells susceptible to HIV infection, and that the presence of Vif was important for the generation of infectious particles. It was later discovered that Vif was likely counteracting the action of a cellular factor, which prevented HIV infection, later revealed to be APOBEC [85-87]. Vif counters the action of APOBEC by targeting it for degradation, which prevents APOBEC from being packaged into budding virions, where it can initiate C to U mutations. Vif also appears to play a role in viral morphology and RNA packaging and viral assembly [88, 89].

Vpr

The *vpr* gene codes for a 14kDa protein called Viral Protein R, or Vpr [90-92]. Vpr is a protein with multifaceted functions during the viral life cycle. Vpr is packaged into budding virions via interaction with the p6 domain of the Gag polyprotein precursor [93]. One of the functions of Vpr is to play an important role in the transport of the pre-integration complex. Vpr has been shown to localize at the nucleus and is important for import of the PIC into the nucleus [94-96]. Vpr has also been shown to be involved in trans-activation of the LTR and to interact with various other transcription factors [92, 97, 98]. Additionally, Vpr has been implicated in the G2 cell cycle arrest which is important for early events in viral infection. This has been shown to be mediated via an interaction between Vpr and cell cycle proteins [99-101].

Vpu

Viral Protein U (Vpu) is encoded by the HIV accessory gene, *vpu*, and is a 16 kDa viral protein that resides in the cellular membrane [102]. Vpu is associated with two primary functions. The first is the degradation of cellular CD4 in infected cells by targeting CD4 to the E3 ubiquitin ligase complex [103-105]. Degradation of CD4 is thought to reduce interactions between CD4 and envelope proteins, which can inactivate the envelope protein and prevent its incorporation into forming virions [106]. The second function that is associated with Vpu is to counter the action of the viral restriction factor Tetherin (which is also known as BST-2), resulting in more efficient budding and release of new virions [107, 108]. Tetherin is a protein that inhibits viral infection by preventing the complete release of virions by tethering budding virions to the cell surface [107]. Like its

action on CD4, Vpu encourages the degradation of Tetherin by targeting it for ubiquitination and subsequent degradation in the proteasome, though there is evidence that this degradation also can occur in lysosomal compartments [108-110].

Nef

The *nef* gene overlaps by about 50% with the 3' LTR and encodes the Negative Factor protein (Nef), which is about 37kDa and is incorporated into HIV virions [20]. Nef appears to play an important role in viral pathogenesis, as Nef mutations have been associated with reduced pathogenesis in humans. Additionally, *nef* deleted SIVs are unable to maintain infection in the macaque experimental model, indicating an important role in viral pathogenesis [111]. Nef also plays a role in the down regulation of CD4 by functioning as a bridge molecule between CD4 and the clathrin adaptor protein, AP-2. This ultimately results in the lysosomal degradation of CD4 [112, 113]. In addition to CD4 downregulation, Nef alters the display of other cell surface molecules including MHC-I, DC-SIGN, CD28 and CTLA-4 to help the virus evade immune recognition and disrupt normal immune function [114-117]. Nef has been shown to interact with several cellular protein kinases, and may play a role in cell signaling, though the effect of these interactions is not well characterized [118-120]. Nef is clearly an important protein for viral pathogenesis, but it is still unclear exactly which of these activities are so critical for viral infection in vivo.

The Evolution of HIV

HIV-1

HIV-1 originally entered the human population via a zoonotic event, wherein a related Simian Immunodeficiency Virus that infected chimpanzees (SIVcpz) crossed the species barrier to infect humans [121, 122]. There are over 40 species of non-human primates that are infected by SIVs, and the disease range in natural hosts varies greatly [123]. For example, African green monkeys infected with their natural SIV counterpart, SIVagm, do not develop any notable disease despite having very high viral titres and having an estimated prevalence rate of over 50%, with females more commonly infected than males [124]. In contrast, chimpanzees, when infected with SIVcpz, often develop Simian AIDS, which has been observed both in captivity and in the wild [125-127]. Importantly, when SIV infects non-natural hosts, immunodeficiency is much more common [128-131]. SIVsmm, the species of SIV that infects sooty mangabeys, does not cause notable disease in its host, despite a high viral load and prevalence of over 50% [132, 133]. SIVsmm crossed into a macaque population due to co-habitation in a primate facility performing experiments on Kuru [134], resulting in SIVmac, an SIV species that causes simian AIDS in macaques which often serves as an experimental animal model for HIV-1 infection [135, 136]. Cross-species transmission and co-infection can also result in recombination of these viruses, resulting in new strains with altered pathogenicity. SIVcpz itself likely derived from a recombination of SIVrcm (the SIV that infects red capped mangabeys) and SIVgsn (the SIV that infects greater spot nosed monkeys) [137].

HIV-1 has entered the human population on at least four separate occasions giving rise to the four main HIV-1 groups M, N, O and P [11]. Group M is the major group and accounts for more than 90% of worldwide HIV-1 infections [138]. Group O, the outlier group, is responsible for tens of thousands of infections, but has not spread as notably as group M HIV-1. Group O infections are primarily restricted to central African countries, specifically Cameroon [139-142]. Groups N and P infections are far fewer and are restricted to a few individual infections [143-146]. Group N HIV-1, like Group M is derived from a strain of SIVcpz [146-148] but Group O and P HIV-1 are very closely related to an SIV found in gorillas, SIVgor [149-151]. It is important to note that SIVgor is very closely related to a specific species of SIVcpz, there is a postulation that rather than direct transmission of SIVgor to humans, groups O and P HIV are actually SIVcpz derived, and this SIVcpz was transmitted to both gorillas and humans [152]. Because group N and group P HIV-1 are found in so few individuals, these viruses are geographically restricted to Cameroon and the surrounding countries [146, 151]. Studying SIVs is important because of the number of different SIV species, and their potential to continue to be transferred into the human population. Also, understanding the interactions between SIV species and the host immune system could help to develop more effective HIV therapies and prevention strategies.

Phylogeographical and molecular clock studies have projected that Group M HIV-1 crossed the species barrier during the early part of the 20th century in Central Africa [153]. This virus then diversified rapidly, dividing into 9 subtypes

(A-H) that are characterized by a difference in genome sequence of 30% [154]. Of the subtypes (or clades), subtype C is the most prominent, accounting for almost 50% of HIV-1 infections globally. This subtype is the predominant strain in sub-Saharan Africa and southeast Asia, while subtype B predominates in Western countries [155]. Because subtype B is more prevalent in the United States and Western Europe, it is the subtype that has been used for most foundational studies on HIV-1, though more and more studies are considering a diverse array of subtypes. There is evidence that there are differences in the transmission and pathogenesis of different subtypes, specifically subtype C [156-162]. Because subtype C is responsible for most global infections, it is important that this subtype be included in drug development and in work to develop a broadly neutralizing vaccine.

The diversity of HIV-1 is further complicated by the ability of HIV to undergo recombination within a host. Circulating recombinant forms (CRFs) result when an individual is dually or coinfected with multiple HIV-1 subtypes at once, giving rise to an unlinked recombinant form (URF) of HIV-1. When this URF is found in more than three individuals, it is reclassified as a CRF [154]. The diversity of CRFs is vast and expanding, and CRFs currently account for about 20% of HIV-1 infections worldwide [155].

HIV-2

HIV-2 also entered the human population via zoonotic events, where SIV that infects sooty mangabeys (SIVsm) crossed into the human population [163]. This cross-species transmission has occurred at least 8 times, accounting for the

8 subtypes of HIV-2 (A-H), with evidence of a ninth group indicating that SIVsm continues to cross the species barrier [164, 165]. Molecular clock evidence suggests that the primary circulating types of HIV-2 entered the human population in the middle part of the 20th century, around 1940 [166]. At this point, HIV-2 is primarily restricted to West African countries, and subtypes A and B are the most prominent, as subtypes C-H have not spread beyond isolated infections [167, 168]. Subtypes C-H appear to be restricted to single infections, though there is evidence that subtype F has been found in two patients [169]. HIV-2 does not cause as severe of disease in humans and individuals who are HIV-2 positive often remain long-term non-progressors [170]. HIV-2 infection is also characterized by a higher level of neutralizing antibodies, lower proviral load and immune control [165, 171-175]. This is echoed in natural SIVsmm infections of sooty mangabeys, as infected animals carry a high viral load but don't have symptoms of immunodeficiency [132]. It has been shown that SIVsmm can cause AIDS like symptoms, but these reports usually consider virus that has been isolated in macaques prior to infection [176]. One report has shown that SIVsmm can cause disease, as researchers discovered an animal in captivity that developed symptoms of immunodeficiency after 18 years of SIVsmm infection. This provided evidence that the progression to an AIDS like state takes longer than the natural life span of the animal [177].

HIV Envelope Protein Biology and Function

The envelope protein of primate lentiviruses, including both HIV species and SIVs is of great interest in the HIV field because it is the only viral protein on the virus that is exposed to the immune system, and is highly immunogenic, as a result, the envelope trimer is an attractive prospect for a protein based HIV vaccine. Additionally, the process of viral entry, which is facilitated by the envelope protein, is an attractive target for anti-HIV drugs because its action does not require infection of the cell before it becomes effective [178]. Envelope is an HIV protein that is uniquely capable of handling a high level of mutation, and shows an incredible level of diversity, attributed to the error-prone RT enzyme [22, 65]. This characteristic contributes to pathogenesis by allowing the virus to evade immune recognition [179]. As such, the entry of HIV is targeted by several anti-retroviral drugs that are included in highly active antiretroviral therapy (HAART), including Enfuvirtide, which binds to gp41 complexes and acts as a membrane fusion inhibitor, and Maraviroc, a drug that binds to CCR5 on the host cell and prevents gp120 contact with the coreceptor [180, 181]. Recently, the native envelope trimeric structure has been solved, which is an important step in understanding the envelope trimer's function and in the development of effective novel anti-retroviral drugs and a broadly neutralizing vaccine [2, 3].

HIV Entry

The primary function of the envelope protein of HIV is to bind to receptors on the surface of the cell and to facilitate entry via viral and cellular membrane fusion, and this process occurs in several well-characterized steps [31, 182]. Prior to specific interaction between the envelope protein and cellular receptors,

the virus often contacts a susceptible cell through non-specific interactions with cell surface molecules, though this step is not required for viral entry. Molecules that have been shown to interact with the HIV envelope protein and play a role in this initial attachment include heparan sulfate, dendritic cell-specific intercellular adhesion molecular 3-grabbing non-integrin (DC-SIGN) or α4β7 integrin [25-27, 183]. This adsorption process brings the virus and cell into proximity, which allows specific interactions between envelope and cellular receptors to occur. Next, the envelope trimer binds to the primary receptor, CD4 on the surface of a susceptible cell. Following binding, the trimer undergoes a conformational change which exposes the coreceptor binding site [184-186]. In this conformational change, the monomeric gp120 subunits spread apart from each other and the variable loops (V1, V2 and V3) undergo rearrangement [28, 187]. A result of this rearrangement is the formation of the bridging sheet, a four β-strand domain which connects the inner and outer domains of gp120 which plays a role in maintaining the correct conformation for coreceptor binding [28]. Part of this conformational change is the exposure of the V3 loop, which forms the coreceptor binding site along with the bridging sheet [28, 188-190]. The newly formed coreceptor binding site contacts either CCR5 or CXCR4. Coreceptor usage is determined largely based on the sequence of the V3 loop, and viruses that utilize either CCR5 or CXCR4 are termed R5-tropic or X4-tropic, respectively, while viruses that can use both CCR5 and CXCR4 are termed R5X4 or dual-tropic [191]. Following binding to the coreceptor, the gp41 molecules transition to an extended form, which insert into the cellular membrane

and tether the viral and cell membranes to one another [192-194]. Next, the heptad repeat regions of gp41 (HR1 and HR2) fold onto themselves, forming the 6-helix bundle (6HB), which brings the two membranes into close contact with one another, resulting in fusion of the two membranes [192-194]. Following membrane fusion, the viral capsid and genetic material are released into the host cell where reverse transcription, replication, integration and viral assembly can occur. A schematic of the viral entry process is shown in Figure 1.3. The location of entry for HIV remains somewhat unclear. Initially, studies suggested that entry occurs at the cell membrane, however, additional research has shown that entry can occur in the endosomal compartment, but that acidification is not required for this process [29, 30].

Cell Tropism

The cellular tropism for HIV is primarily determined by the binding and fusion activity of the envelope protein, and the expression of viral receptors on the target cell membrane. HIV specifically infects cells that express CD4 and a suitable coreceptor, CCR5 or CXCR4. A small minority of viruses have been shown to interact with alternate coreceptors, though their relevance to HIV cellular tropism remains unclear [195]. Sexual transmission of HIV relies on the use of the coreceptor CCR5 as this initial transmission usually requires infection of tissue resident macrophages, but later in infection, some viruses develop the use of CXCR4 through viral evolution [196, 197]. Cells that express the receptors required for HIV infection are the primary targets of the virus. This includes macrophages, and most notably, CD4+ T-lymphocytes. Additionally, subsets of

dendritic cells express CD4 as well as CCR5 and CXCR4, have been shown to be susceptible, albeit at a low level, to HIV infection [198, 199]. Macrophages and dendritic cells, including Langerhans cells in the mucosa are early targets of viral infection in sexual transmission, and contribute to the spread of HIV within the host as infected cells travel to the lymph nodes where productive infection of macrophages and CD4+ T cells can occur [200-205]. Transmission of HIV-1 appears to be restricted primarily to R5 and R5X4 tropic viruses, with X4 viruses appearing within a patient late during infection [197, 206]. A shift to the use of X4 as a coreceptor also is usually accompanied by an increase in disease progression [207]. Macrophages often serve as reservoirs for HIV infection, and contribute to the spread of the virus even years into infection [208]. Following productive infection, CD4+ T cells are killed, and eventually the population of this subset of cells is depleted from the host, resulting in severe immunodeficiency and increased susceptibility to opportunistic infections [196].

Envelope Trimer Structure

Historically, a major effort in the field has been to generate an atomic level structural image of the native HIV envelope trimer. Solving the structure of the envelope trimer has been an attractive prospect because the information can be useful for the development of drugs and vaccines in addition to a more thorough understanding the viral entry process. In 2009, the HIV trimer was listed as one of Nature magazine's most important structures yet to be solved, since then, significant progress has been [209]. One of the first structures of the envelope protein solved was the ectodomain of gp41, which revealed a triple-coil structure

with similarities to the pH induced conformation HA2 of the influenza virus. Two research groups generated similar structural models in the same year, and both came to very similar conclusions about the role gp41 plays [192, 193]. In 1998, a crystal structure of the gp120 core (with the variable loops removed) bound to CD4 and a CD4 induced conformational antibody, 17b. This structural study revealed the inner and outer domains of the gp120 core, as well as the bridging sheet [28]. Generating a structure of the envelope protein trimer has been complicated, because it is membrane bound, heavily glycosylated, is conformationally flexible and is a non-covalently attached trimer, all of which have made X-ray crystallography challenging. However, significant advancements have been made in this area using cryo-electron microscopy and X-ray crystallography techniques. One study has generated a cryo-EM structure of an uncleaved envelope trimer that has a mutation preventing gp120/gp41 cleavage and a truncated cytoplasmic tail [210, 211], though this structure was the subject of extreme controversy among those in the field [212-215]. This structure of the envelope trimer revealed three primary points of contact between subunits at the gp41 transmembrane region, the gp41 ectodomain, and importantly, the gp120 trimer association domain, which consisted primarily of the V1, V2 and V3 domains, at the apex of the trimer. Additionally, this envelope structure revealed the presence of a large space at the center of the trimer, which was consistent with previous lower resolution structures (Figure 1.4) [210]. A cryo-EM and crystal structure of a cleaved trimer, from the subtype A virus BG505, that has been solubilized was described in the same year [1-3]. This

envelope trimer contains a disulfide bridge between the gp120 and gp41 subunits, and includes an isoleucine to proline mutation in gp41 to aid in trimer stabilization [216, 217]. Trimers with these two modifications are referred to as SOSIP trimers, and have been the subject of extensive study into immunogen design in recent years and have shown some promise [218-220]. This trimer structure differed substantially from the uncleaved envelope trimer previously described, primarily in the positioning of the gp120 core and gp41 helices. Notably, these trimeric structures did not include the open-space at the center of the complex (Figure 1.5). Many subsequent structural studies have been performed with the SOSIP trimer and have included envelopes from viruses of other clades, as well as additional modifications. These studies have further characterized the presence and conformation of the glycan shield and bNAb binding sites [1, 221, 222]. Additionally, in 2016, a cryo-EM structure of a fully glycosylated, uncleaved native trimer based on the JRFL envelope in contact with the bNAb PGT151 was released. This structure revealed striking similarity to the SOSIP structure, and included the membrane proximal external region (MPER), which was notably absent from the SOSIP structures released earlier [223].

Despite specific differences in the structures of the envelope trimer, an important consistency in all the structures is that the variable loops are packed into the apex of the trimer in a specific conformation, and contribute to trimer association, which is consistent with molecular stability studies (Figure 1.5, 1.6) [5, 6, 224]. In earlier structural studies, the variable loops were considered

dispensable components of the trimer, and were flexible, so they were often removed from the envelope, but the current structures demonstrate that they are indeed fixed in the apex prior to CD4 binding and fusion, and may play a role in the trimerization of the envelope protein, though the exact conformation of the loops is slightly different in each structure.

Envelope Processing

The env gene product is produced as a single gp160 precursor protein. This precursor protein is covered with N-linked glycans, which account for almost half of the molecular weight of the envelope trimer and cover the surface of the glycoprotein [71, 223]. This glycan shield is important for viral immune evasion, but several well characterized bNAbs rely on the presence of specific N-linked glycans, namely, those antibodies in the PGT family which include the glycan at N332 proximal to the V3 loop in their binding site [225]. The gp160 molecules are then processed and trafficked to the cell surface through the Golgi apparatus, where they are cleaved by furin proteases into gp120 and gp41 at a K/R-X-K/R-R motif by furin proteases [42, 70]. Trimerization then occurs and three gp120 molecules are joined and are non-covalently attached to three gp41 molecules. These trimers are then transported to the plasma membrane where they are incorporated into virions. Though the exact mechanism of envelope incorporation into virions is still somewhat unclear, it has been shown that this process relies on both the activity of the cytoplasmic tail of gp41 and the matrix protein [15, 42].

Envelope Stoichiometry

The envelope stoichiometry for entry into a cell is an area that is of interest. An HIV virion contains between 5 and 15 envelope trimers, but the number of trimers required to bind and enter the cell is yet to be determined in the literature as different assays have yielded differing results [17, 226, 227]. This work requires mathematical modeling to identify the stoichiometry required for viral entry, as visualization techniques are either unable to distinguish between individual trimers in real time, or require samples to be fixed to slides or frozen [227]. Some models have suggested that as few as a single envelope trimer is sufficient for entry, while others suggest as many as 19 trimers are required for successful entry [227-231]. Electron microscopy tomography has revealed the formation of the “entry-claw”, a group of five to seven complexes at the interface between virus and cell. It is unclear whether all the complexes in the claw are required for entry, or even represent functional trimers, but this work has revealed the interactions between virus and cell in higher resolution than ever before [232]. The stoichiometry of individual viral strains can differ, which drives the efficiency of entry for these viruses and may be related to the energy required to initiate membrane fusion. Identifying the number of trimers required for entry by HIV, and specifically by a given viral strain helps to enhance the knowledge of the viral entry process, and in turn contributes to the design of more effective envelope and entry inhibitors. This information will also help to elucidate the required stoichiometry of neutralizing antibodies. Additionally, an understanding of envelope stoichiometry can help elucidate to elucidate the amount of protein

required for the entry process within individual species, which could reveal the role of entry in the initial infection and spread of HIV.

Primate Lentiviral Envelope Variable Loops

The HIV envelope protein is divided into five conserved regions and five variable regions (Figure 1.6). As their name implies, the conserved regions are well conserved in length and sequence, and are named C1-C5. The variable loops, named V1-V5 have more variety in sequence, but several of the loops have fixed lengths and conserved motifs [5, 6, 233, 234].The variable loops were once thought to have flexible conformations because of their variability, and because in early structural studies, they were often found in differing conformations, and as such were intentionally truncated or removed in many early gp120 structural studies [235]. With the advent of atomic trimer structures, it is apparent that these loops, specifically V1/V2 and V3 have fixed, important structural roles within the trimeric complex, found in the apex of the trimer where they are clearly exposed on the virus, and appear to make inter-protomer contacts [1-3, 210, 211]. Importantly, research has shown that although the variable loops typically take a fixed structural role, point mutations to the envelope protein can cause the V1/V2 and V3 loops to move out of these structurally constraining positions, and that the V3 loop is able to move independently of the V2 loop, which is surprising, as the V3 loop is positioned directly underneath the V2 loop according to the SOSIP structures [236, 237]. The variable loops were often somewhat disregarded in terms of form and

function early in the study of envelope, due to their variability and perceived flexibility [28], but have since become an important focus as they are integral functional components of the envelope, and as targets for epitope based vaccines as well as entry inhibiting drugs and small molecules.

V1/V2 Loop

The V1 and V2 loops are often considered in combination because of the double loop structure that forms them. While there have been some characterized monoclonal antibodies to the V1 loop, its antigenic potential is somewhat limited because its sequence and length is widely variable among HIV-1 species. Because of this high level of sequence length and variability [238], it has been proposed that the structure of the V1 loop is more flexible than its other variable loop counterparts, though high resolution trimeric structures suggest that it is packaged into the apex of the envelope trimer with the V2 loop [2, 3]. The V1 loop is involved in forming the coreceptor binding site, but does not dictate coreceptor usage. The V2 loop is essential for trimer stability and entry by HIV-1 [6, 239, 240], though early studies suggested that viruses can overcome deletion of the entire V1/V2 loop during long term infections [241]. Other reports have shown however that deletion of the V1/V2 loop completely disables functionality of the envelope protein, as these mutants are unable to initiate viral entry or fusion, and causes Env proteins to be more susceptible to binding by CD4 binding site antibodies [239, 240]. There is evidence that V2 loop also contributes to the early dissemination and spread of HIV-1 by binding to the α4β7 integrin receptor, which is involved in the homing of HIV infection to gut tissue

and activation of LFA-1. The activation of LFA-1 promotes the formation of the virological synapse and cell-cell spread [26], although studies suggest that this feature may be limited to only a few HIV-1 species [242]. The V2 loop has a much higher antigenic potential than the V1 loop because it is the target of many broadly neutralizing, quaternary specific monoclonal antibodies including 2909, PG9 and PG16 [243, 244]. Based on high resolution structural analysis, the V2 loop is packaged into the apex of the envelope trimer, and plays an important role in trimerization of gp120 [6]. The length of the V2 loop is relatively well conserved between viral species and is typically around 40 residues. There are some viruses that have V2 loops longer than 40 residues, but very few that have V2 loops less than 40 residues in length. This suggests that structurally, the envelope complex can tolerate slightly longer V2 loops, but not shorter [238].

Because of the position of the V2 loop in the trimeric apex, its antigenic potential and its relationship to well-characterized broadly neutralizing antibodies, it is an attractive candidate for the development of vaccines targeting the envelope trimer [245]. The V2 epitope approach is supported by analysis of the RV144 Thai vaccine trial which showed that V1/V2 antibody responses correlated with a lower risk of infection [246]. Further analysis of patient samples from the RV144 vaccine trial showed the presence of antibodies to V2 loop epitopes early after vaccination at higher levels than naturally infected individuals [247]. Additional V2 loop based vaccine studies have characterized SOSIP based immunogens based on viruses that have a ‘glycan hole’ in the envelope trimer apex, which lack the N-linked glycosylation site at N130 in the V1 loop, as

well as in the V2 loop hypervariable region. These immunogens elicited neutralizing autologous antibodies by taking a V2 focused approach [219]. Another group has used an adenoviral capsid display method to generate a vaccine based on the sequence of the V1/V2 loops together as well as the V2 loop by itself. Follow up analysis of mice that were exposed to the variable loop displaying Ad5 vectors showed the presence of a strong anti-V2 antibody response, suggesting that V2 antibodies may be successfully elicited by relying on the V2 loop sequence individually [248]. Altogether, there is evidence that targeting the immune system to V2 specific epitopes may be an effective HIV vaccine strategy.

V3 Loop

The V3 loop, like the V2 loop, is a β-hairpin loop created by a disulfide bond positioned in the apex of the envelope trimer and is an important functional component of the envelope protein. This loop is divided into three primary regions, the base, the stem and the crown (or tip). The base and crown of the V3 loop are very well conserved, and most sequence variation occurs in the stem [249]. The V3 loop is the major determinant of coreceptor usage, which is largely based on the specific amino acid sequence of the hypervariable region of the loop. Although there is much research into this topic, and much data concerning the use of specific coreceptors, the selection of a coreceptor appears to be a multifaceted process dependent on structural elements, specific amino acid residues and electrostatic charge [250-252]. Deletion of the V3 loop abolishes infection and drastically reduces envelope function, indicating that this region is

an essential component of the envelope protein [239, 240, 253]. The V3 loop has also been shown to play a role in the stability of the envelope trimer in HIV-1 subtype B species [5], a fact which is supported by the near atomic structures of the HIV-1 envelope trimer [2, 3]. The length of the V3 loop is largely conserved in HIV-1 species at 34-35 amino acids, and alterations to this length result in massive destabilization of the envelope trimer [5].

The V3 loop is highly immunogenic, and has been considered an epitope-based vaccine candidate because of its exposed nature, important structural role and the fact that there are important conserved elements in its sequence. The V3 loop is an immunodominant regions, and is the target of a large proportion of antibodies, especially early in infection [238, 254, 255]. Some studies have shown that V3 loop antibodies have promise, with the V3 mAbs 2219, 2557 and 3074 neutralizing around 30% of pseudoviruses they were tested against from both tier 1 viruses, which are sensitive to antibody neutralization, and tier 2 viruses, which are more resistant to antibody neutralization, including representatives from clades A, B and C [256]. This indicates that V3 does may have the potential to serve as a neutralization target for a plethora of viruses, despite its characterization as a region with high variability. Primarily, antibodies generated against the V3 loop bind to residues in the crown of the loop, which contains a highly conserved GPGQ/R sequence that facilitates the turn of the loop. Many of the neutralizing antibodies that target the V3 loop bind to residues near this motif, suggesting that the crown is an important structure to consider when generating epitope based vaccines [234].

Recently, the SOSIP soluble trimers have been used as immunogens and have revealed more about the antigenicity and immunodominance of the V3 loop. Studies using the SOSIP.664 trimer as an immunogen have recently characterized the V3 loop as an antigenic distraction because its presence in trimer constructs results in the generation of antibodies that do not neutralize either autologous or heterologous tier 1 viruses, but do neutralize tier 2 heterologous viruses [220]. Some work has been done showing that the V3 loop can be stabilized through the addition of a tryptophan residue at amino acid position 316 [257], and by adding hydrophobic residues to the already hydrophobic crown [258]. These stabilization mutations lock the loop in the pre-fusion state, preventing its exposure, resulting in lower induction of V3 non-NAbs.

HIV Envelope Protein as a Vaccine Candidate

Although anti-retroviral drugs have significantly improved the lives of those infected with HIV, there were still 1.8 million new HIV infections in 2016, indicating clearly that a preventative vaccine for this virus is paramount [8]. Despite decades of work, an effective HIV vaccine remains elusive. A successful HIV vaccine will likely result in the development of bNAbs capable of neutralizing viruses of multiple subtypes. To date, six vaccine candidates have entered human trials, taking a wide array of approaches, none of which have been efficacious enough to move forward for clinical use. The first two of these trials were VAX003 and VAX004, which utilized the AIDSVAX B/E and B/B designs, respectively. AIDSVAX B/E was a combination of recombinant gp120 from a

subtype B strains (MN) and a subtype CRF01_AE (A244); AIDSVAX B/B was composed of recombinant gp120 from two subtype B strains (MN and GNE8). Overall, these were unsuccessful attempts, and were considered non-efficacious [259-261]. The STEP and Phambili vaccine trials relied on the use of a combination of replication deficient Adenovirus vectors expressing HIV gag, HIV nef or HIV pol. These trials were preliminary ended because there was evidence of increased HIV infection in some participants [262, 263]. The most successful HIV vaccine trial to date has been the RV144 trial conducted in Thailand. This vaccine trial utilized a canarypox vector expressing HIV Gag and Pro proteins as well as HIV envelope protein. This was followed by two boosts with the AIDSVAX B/E immunogen. This trial did show efficacy at 31%, and the presence of antibodies specific to the V1/V2 loop of HIV gp120 was inversely correlated with HIV infection [246, 247, 264]. The results from the RV144 trial elicited new interest in the V2 loop as a potential immunogen, and supported a significant role for the variable loops in the development on neutralizing antibodies [245, 265]. The most recent HIV vaccine trial was a DNA vaccine, HVTN 505. This trial included a DNA prime consisting of six plasmids coding for viral genes. Clade B Gag, Pol and Nef genes were included as well as *env* gene plasmids from clade A, B and C. This was followed up with a boost with Ad5 vectors expressing the Gag-Pol precursor from subtype B viruses and *env* from clades A, B and C. Again, although the vaccine initiated a specific immune response, including HIV specific T-cell responses and production of reactive IgG, this was not efficacious and the trial was ended prematurely [266]. In all, the trials that have entered

human trials have been largely disappointing, indicating the need for new and innovative immunogen design and immunization strategies.

In addition to these studies that have gone to clinical trials, countless immunogen studies using the envelope protein as the central immunogen have been performed with varying degrees of success. A current focus of vaccine and immunogen design is the use of trimeric envelope proteins [220, 267-272]. gp120 trimers are of specific interest as an immunogen because they have been shown to be more immunogenic than monomeric gp120, and because of the discovery and development of broadly neutralizing antibodies that specifically target trimer specific (quaternary) epitopes [271, 273-278]. Varying conformations of the SOSIP trimer have shown promise, but have not resulted in the production of antibodies that neutralize tier 2 viruses beyond the homologous strain used to create the immunogen [220]. In natural infection and immunization with intact envelope trimers, many antibodies are generated against the V3 loop because it is very immunogenic and can move somewhat independently to the outside of the trimer [237, 255]. However, these antibodies are not broadly neutralizing and are not a measure of a successful immunogen, so effort has been taken recently to stabilize the V3 loop to reduce the number of non-neutralizing V3 antibodies. These efforts have indeed been successful in stabilizing the V3 loop, though they have yet to yield an immunogen with the capability to initiate broadly neutralizing antibodies [257, 258, 270]. The HIV envelope trimer is clearly a solid candidate for the development of an effective HIV vaccine because of its surface exposure on the virus and its natural immunogenicity, but further antibody and structural

studies of this trimer are required for the development of an immunizations strategy that will successfully neutralize HIV of multiple origins.

FIGURE 1.1

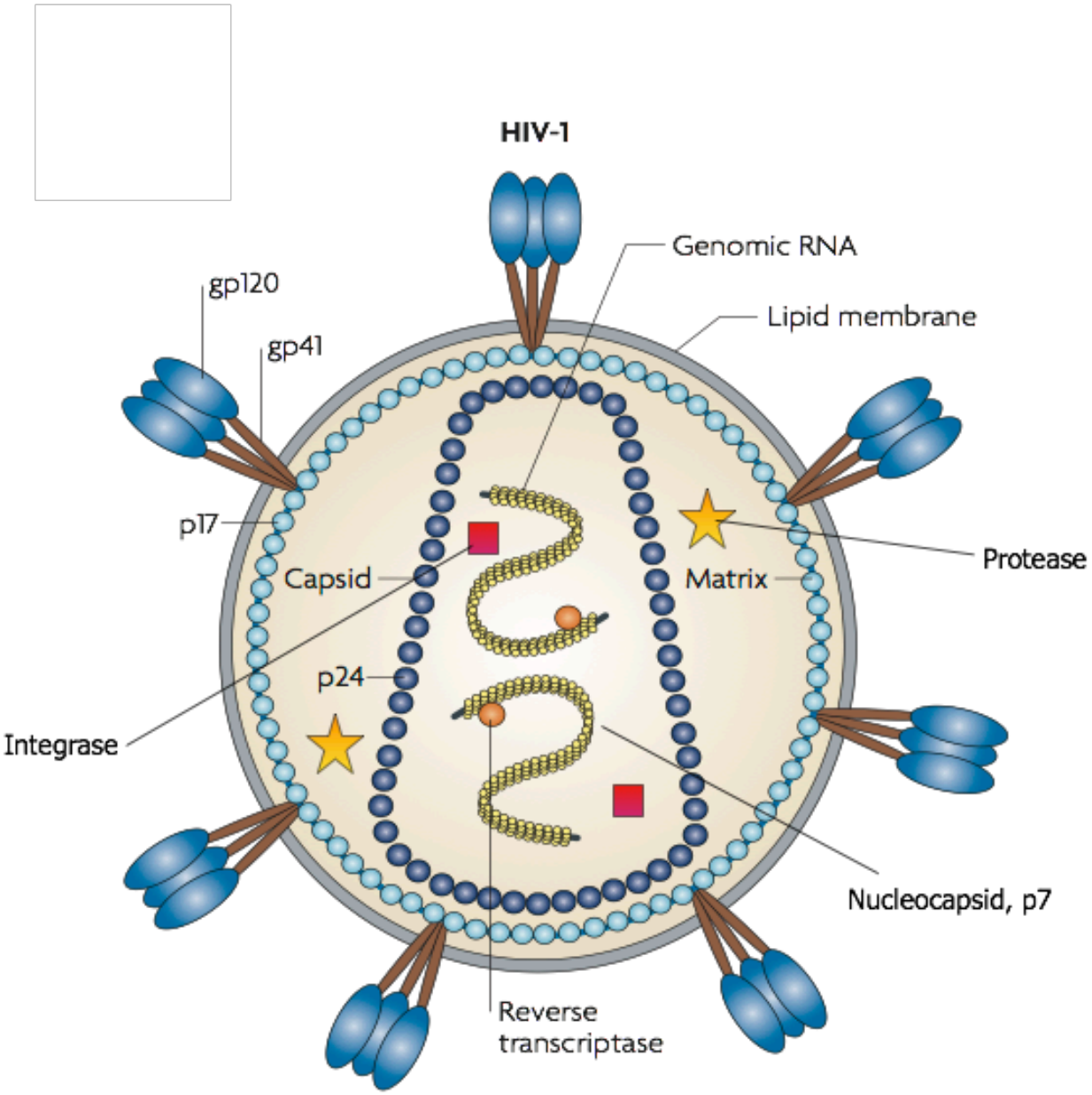
HIV-1
Genomic RNA
gp120
gp41
Lipid membrane
p17
Capsid
Matrix
Protease
p24
Integrase
Nucleocapsid, p7
Reverse
transcriptase

FIGURE 1.1. Schematic of an HIV virus. In an HIV virion, the outer membrane is host derived and typically 5-15 HIV envelope trimers are embedded. Directly interior to the membrane is the MA protein derived from the Gag precursor protein, which forms an outer shell for the virus. The characteristic cone-shaped viral capsid of lentiviruses is located within the virion and contains two viral RNA genomes in contact with the NC protein. Additionally, copies of the viral enzymes reverse transcriptase (RT), integrase (IN), and protease are packaged into virions. Figure adapted from Hedestam et. al., 2008 [279].

FIGURE 1.2

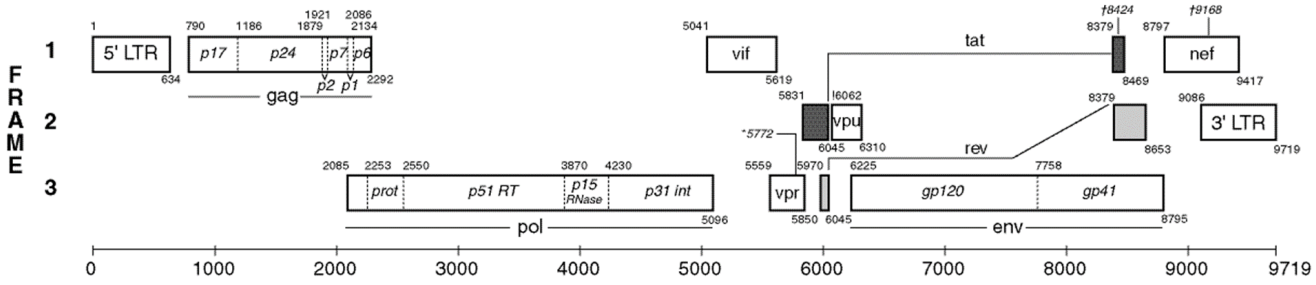
FRAME
1
2
3
5' LTR
p17
p24
p7
p6
p2
p1
gag
vif
tat
nef
3' LTR
vpu
rev
prot
p51 RT
p15
RNase
p31 int
pol
vpr
gp120
gp41
env
0
1000
2000
3000
4000
5000
6000
7000
8000
9000
9719

FIGURE 1.2. HIV genome structure and protein products. The HIV genome is approximately 9.7 kb and includes 9 open reading frames that code for 15 functional proteins. Subunits of the Gag, Pol and Env precursor proteins are indicated by the presence of a dotted line. Figure from Los Alamos National Laboratory HIV Sequence Database [19].

FIGURE 1.3

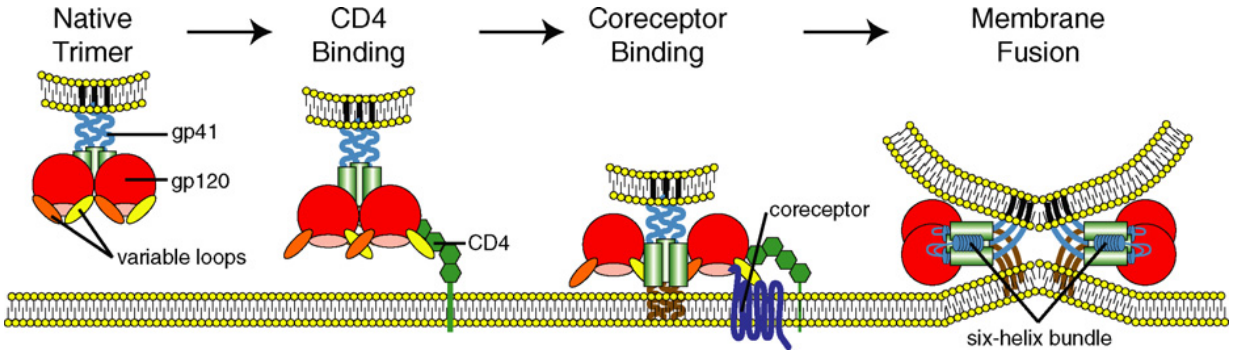
Native Trimer
CD4 Binding
Coreceptor Binding
Membrane Fusion
gp41
gp120
variable loops
CD4
coreceptor
six-helix bundle

FIGURE 1.3. Primary steps in the HIV entry process. HIV viral entry is initiated by binding of the native envelope trimer to the primary receptor CD4 (green). Following binding to CD4, a conformational change in the envelope protein occurs, resulting in the rearrangement of the variable loops, exposing the coreceptor binding site. The V3 loop contacts the coreceptor (CCR5 or CXCR4- shown in blue) and a second conformational change occurs, which allows for fusion of the viral and cellular membranes. Figure from Tilton and Doms, 2010 [182].

FIGURE 1.4

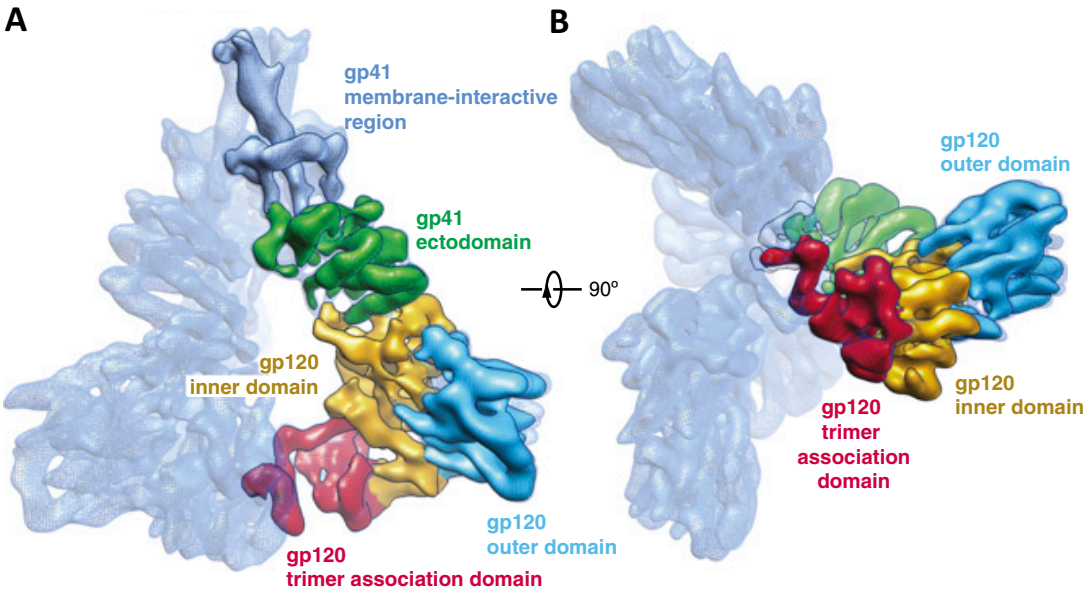
A
B
gp41
membrane-interactive
region
gp41
ectodomain
gp120
inner domain
gp120
outer domain
gp120
trimer association domain
90°
gp120
outer domain
gp120
inner domain
gp120
trimer
association
domain

FIGURE 1.4. Cryo-EM structure of the uncleaved HIV envelope trimer. This cryo-EM structure of the uncleaved HIV envelope trimer reveals a hollow center of the trimer and the positioning of the variable loops in the “trimer association domain” at the central apex of the envelope trimer (shown in red). A, side view of the trimer; B, top down view of the trimer. Trimer association domain, red; gp120 outerdomain, cyan; gp120 inner domain, gold; gp41 ectodomain, green; gp41 membrane interactive region, light blue. Figure adapted from Mao, et. al., 2013.

FIGURE 1.5

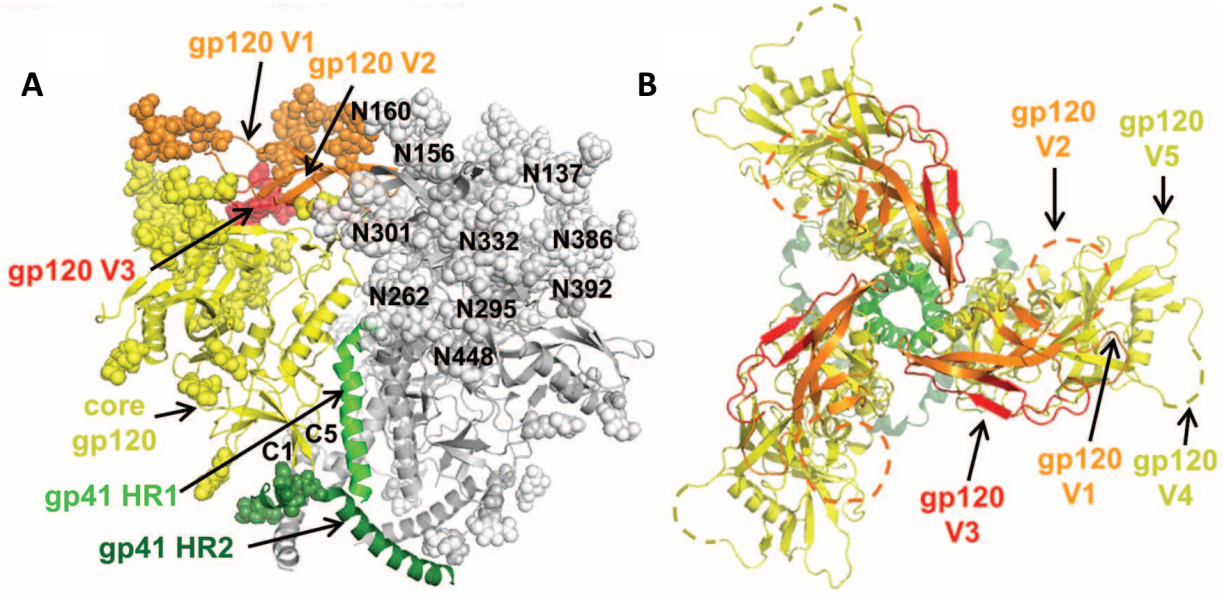
A
gp120 V1
gp120 V2
N160
N156
N137
N301
N332
N386
gp120 V3
N262
N295
N392
N448
core
gp120
C5
C1
gp41 HR1
gp41 HR2
B
gp120
V2
gp120
V5
gp120
V1
gp120
V4
gp120
V3

FIGURE 1.5. Crystal structure of the BG505 SOSIP.664 HIV envelope trimer. The trimer structure of the SOSIP trimer reveals no central space within the trimer and the V1/V2 and V3 loops are positioned at the apex of the trimer as shown in crystal structure. A, side view; B. top view. V1/V2 loop, orange; V3 loop, red; gp120 core including V4 and V5 loops, yellow; gp41 segments, green. Figure adapted from Julien, et. al., 2013.

FIGURE 1.6

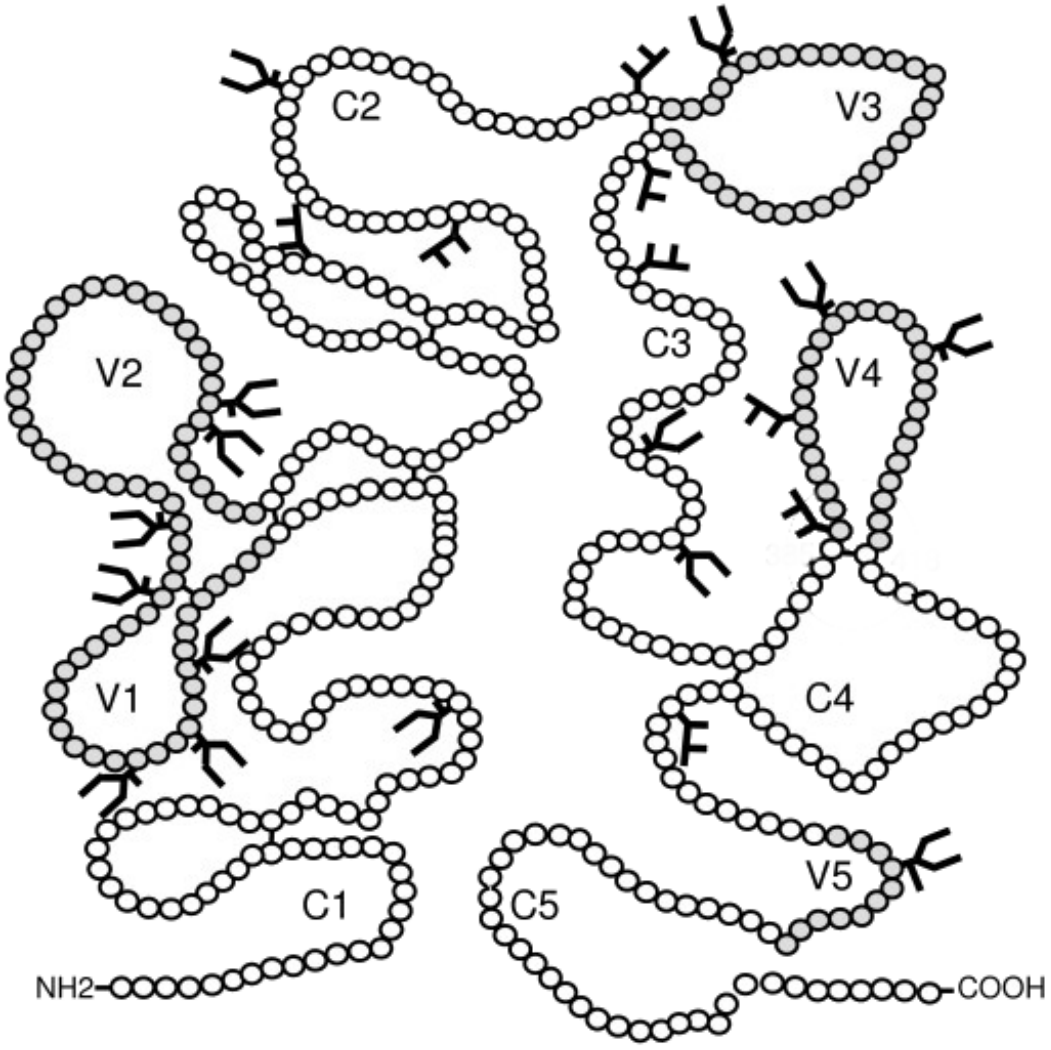

FIGURE 1.6. Schematic of the variable loops and conserved regions of the HIV gp120. The HIV gp120 protein is divided into five conserved regions (white circles), as indicated in this figure by C1-C5, and five variable loop regions (grey circles), indicated as V1-V5. Disulfide bonds that form the variable loops are indicated by bold lines connecting two amino acids, and conserved glycosylation sites are indicated. Figure adapted from Sanders, et. al., 2008.

CHAPTER II- MATERIALS AND METHODS

Cell Lines

293T cells (ATCC) were used for envelope expression experiments, as effector cells for the fusion assay and to produce the single-round virus for entry assays [280]. 293T were grown in complete Dulbecco's modified Eagle Medium (DMEM) containing 10% fetal bovine serum, 8mM L-Glutamine and 100 µg/mL penicillin/streptomycin solution.

TZM-bl cells (NIH AIDS Reagent Program) were also grown in complete DMEM and were used as the target cells for the fusion assay as well as entry assays using single-round, replication deficient virus. This virus is capable of infecting a cell, but due to a deficiency in the envelope gene, is not able to reproduce in cells. TZM-bl cells are engineered HeLa cells that stably express human CD4 and CCR5 [281-283]. They also stably express firefly luciferase and β-galactosidase genes under a Tat-inducible promoter, so that when Tat protein is present in the cells (by transient means or via a virus), both the luciferase and β-galactosidase enzymes are expressed as reporters [284].

Cf2Th-CD4/CCR5 and Cf2Th-CCR5 (Dr. J. Sodroski Laboratory) are canine thymus cells that stably express CD4 and CCR5 and CCR5 under antibiotic selection, respectively [5, 285]. Cf2Th-CD4/CCR5 cells were also grown in complete DMEM supplemented with 50 µg/µL g418 and 150 µg/mL. Cf2Th- CCR5 cells were also grown in complete DMEM supplemented with 50 µg/µL.

Envelope Expression Plasmids

Plasmids for envelope expression were generated by subcloning. The ZM249MB_10 (ZM249), HIV-2 ST and SIVagmTan1 envelope expression plasmids were generated by PCR amplification of the envelope gene from existing envelope expression plasmids or proviral constructs and ligated into pCDNA3.1+ for expression in mammalian cells. The *env* gene of ZM249 was amplified from the plasmid pZM249MB_10 (NIH AIDS Reagent Program) by PCR using primers that added a 5'-HindIII site, and 3'-EcoRI site. The HIV-2 ST *env* gene was amplified from the pHIV2/ST molecular clone adding a 5'-AflII site and a 3'-BamHI site and subcloned into pCDN3.1+ [286, 287]. The SIVagmTan-1 *env* gene was amplified from the pSIVagmTan-1 molecular clone (NIH AIDS Reagent Program), adding a KpnI site to the 5' end of the gene and an EcoRI site to the 3' end of the gene, by including these sequences on the 5' and 3' ends of the amplification primers respectively (Appendix I), followed by subcloning into pCDNA3.1+ [288]. 1084i (Dr. C. Wood Lab) was subcloned into the pSVIII HXBc2 env construct via the conserved KpnI and BamHI sites. A BamHI site corresponding to the existing BamHI site in HXBc2 was introduced to 1084i by site directed mutagenesis. Following subcloning, all envelope constructs were verified by sequencing the *env* gene to identify any mutations in the sequence. V3 loop hydrophobic patch mutations were also made directly to pZM249M_B10 for pseudovirus production. HIV-1 ADA, JR-FL and YU2 envelope expression plasmids (pSVIII) were previously available (Dr. J. Sodroski Lab), and any mutations made to these were made directly by site directed mutagenesis. A

complete list of plasmids and mutations used in these studies is available in table 2.1. All ligations were performed using the T4 DNA Ligase enzyme and buffers from NEB. Sequences of primers used for gene amplification are shown in Appendix I.

Envelope Gene Cloning by Reverse Transcriptase PCR

The SIVagm155-4 envelope gene was cloned into the pCDNA3.1+ expression vector by RT-PCR using live virus (NIH AIDS Reagent Program) [289]. First, viral RNA was extracted and purified. from virus containing supernatants using the QIAmp Viral RNA Mini Kit (Qiagen). A cDNA library was generated from this viral RNA using the Superscript III Reverse Transcriptase kit (Invitrogen), substituting oligo(dT)20 to serve as the primer for the reaction. This cDNA library was then subjected to gene amplification PCR using env gene specific primers, containing a 5'-XhoI site and a 3'-XbaI site for subcloning, using the TaKaRa ExTaq kit (Takara). The resulting DNA band of the correct size (2307 bp) was gel extracted using a gel extraction kit (Qiagen), digested with XhoI and XbaI, and subcloned into the pCDNA3.1+ vector using T4 DNA Ligase (NEB). The resulting clones were verified by restriction digest with XhoI and XbaI to verify the appropriate insert size, and the *env* gene was fully sequenced to verify the sequence identify any potential mutations gained in the cloning process.

Site-Directed Mutagenesis

V3-loop hydrophobic patch and twin cysteine mutations were made to the primate lentiviral envelope expression constructs by site-directed mutagenesis, using PfuUltra (Agilent) and the QuikChange XL protocol. Primers for mutagenesis were synthesized by Integrated DNA Technologies (IDT). A complete table of envelope constructs and their mutations is shown in Table 2.1. A list of the primers used to introduce mutations to the V3 loop hydrophobic patch of ZM249 can be found in Appendix A. A complete list of primers used to introduce V3 mutations to 1084i is found in Appendix B. Mutagenesis primers for the V3 loop hydrophobic patch in HIV-2 ST, SIVagmTan-1 and SIVmac239 are found in Appendix C, D and E respectively. The primer sequences used to introduce the twin cysteine motif (TCM), a pair of cysteines at positions 183 and 191 in the V2 loop) into HIV-1 are listed in Appendix F; the primers used to initiate the V2 Swap mutations to the HIV-1 species in Appendix G. The primer sequences used to remove the twin cysteine motif from HIV-2 ST, SIVagmTan-1 and SIVagm155 are all listed in Appendix H.

Overlapping PCR

An overlap extension PCR based method was used to generate the V2 swap mutations discussed in chapter 3. Initially, two segments were generated in two PCR reactions. Fragment 1 was generated using primer A, a forward primer extending from the 5’ end of the HIV-1 (ADA or JRFL) envelope gene sequence,

containing the 5'-KpnI site for cloning, and primer B, a reverse primer extending from the V2 loop twin cysteine region. Primer B contained the SIVmac239 or HIV-2 V2 twin cysteine sequence in place of the HIV-1 V2 twin cysteine region in frame. Fragment 2 was generated using primer C, a forward primer that overlapped with primer B at the V2 twin cysteine region from SIVmac239 or HIV-2 in frame with the HIV-1 envelope sequence, and primer D, a reverse primer that amplified the 3' end of the HIV-1 envelope gene, containing the BamHI site for subcloning. Both fragments were generated using the E7/ADA or E7/JRFL WT envelope plasmids. Each fragment was gel purified and included as the template in a second round of PCR, which used primers A and D to amplify the entire envelope gene with the V2 twin cysteine region replaced. The resulting envelope gene was digested with BamHI and KpnI and subcloned into the original pSVIII vector. Positive colonies were screened for the swap mutation by sequencing, and were then amplified for use in transfections. This process is diagrammed and explained in detail in Figure 2.1, and the primers sequences used are listed in Appendix G. V2 swap constructs are designated by listing the HIV-1 species first, followed by an abbreviation of the species the V2 region is derived from and designated as "V2Swap". For example, the ADA envelope containing the SIVmac239 V2 twin cysteine region is termed "ADA macV2Swap". Agarose gels showing the resulting PCR bands and sequencing results are shown in Figure 2.2.

Cell-cell Fusion Assay

On Day 0, 293T cells were plated at a density of $3.5x10^5$/well in a 6-well plate. The following day, the cells were transfected with 2 µg of WT or mutant envelope expression constructs and 1 µg Tat-expressing plasmid using polyethylynamine (PEI). 24 hours post transfection, the media was changed on transfected cells and TZM-bl cells were set in an opaque, black 96-well plate at a density of $1.0x10^4$ per well. On Day 3, the transfected 293T cells were detached from the plate using gentle pipetting and counted in a hemocytometer. The transfected effector cells were set on top of the previously plated TZM-bl cells at a density of $1.0x10^4$ and were co-incubated for 16 hours. Following co-incubation, the cells were washed once with PBS and lysed using 1X Passive Lysis Buffer (Promega) and frozen at -80°C. The plates were read for luciferase activity following thaw using the beetle luciferin substrate (Promega) and data was collected using a Veritas luminometer and the Veritas software. Values for WT were set to 100 percent and mutant values are expressed as a percentage of the WT value. Each fusion data set is representative of three separate experiments performed in quadruplicate.

gp120 Shedding Assay

293T cells were set at a density of $3.5x10^5$/well in a 6-well plate. The following day, the cells were transfected to express a WT or mutant envelope protein. One day post transfection, the cells were metabolically labeled overnight using the EasyTag EXPRESS^{35}S Protein Labeling Mix [^{35}S] (PerkinElmer). Following labeling, the supernatants were harvested and clarified of cell debris by

gentle centrifugation. Separately, the cells were lysed using an NP-40 based lysis buffer. The cell lysates were precleared using appropriate uninfected serum for 8 hours as previously described [6]. Following preclear, the cell lysates and supernatant fractions were separately subjected to immunoprecipitation with PLV infected patient serum from the corresponding host [290-292] and pull down was completed using protein-A sepharose beads (GE Health Sciences). The samples were run on SDS-PAGE and the gels were dried and exposed to phosphor screen for 12 hours at room temperature. Phosphor screens were developed using a personal molecular imager (Biorad) and analyzed using the QuantityOne software (Biorad). Values for the supernatant and cell lysate band were determined and the association index was calculated as previously described [293].

Pulse-Chase Assay

On day 0, 293T cells were set at a density of $5.0x10^6$/well in a 6-well plate. 24 hours later, four wells were transfected to express WT ADA envelope, or ADA envelope with the twin cysteine substitution mutation. The following day, the cells were serum starved for thirty minutes by adding serum free media. After serum starvation, the expressed proteins were labeled by adding 1 mL cys- met- media containing EasyTag EXPRESS^{35}S Protein Labeling Mix [^{35}S] (PerkinElmer) for a 1 hour pulse. The pulse media was removed and replaced with non-labeled complete media, and the cell lysates and supernatants were harvested separately at 0 h, 2 h, 4 h and 8 h post pulse. Once all harvests were complete,

immunoprecipitation was performed on the cell lysates and supernatants separately using HIV infected patient serum, and analyzed by SDS-PAGE.

Generation of Pseudoviruses

293T cells were plated at a density of $3.5x10^5$ cells per well in a 6-well plate. 24 hours later, the cells were transfected with a WT or mutant envelope expression plasmid and an envelope deficient viral backbone ($pSG3^{\Delta Env}$ [72, 281] in the case of ZM249 and SIVmac239 constructs or pHIV-Luc and pCMV Gag-Pol packaging plasmid [5, 294, 295] for 1084i constructs) using polyethylynamine (PEI) at a stock concentration of 1 μg/μL. One day following transfection, additional media was added to the wells and 3 days post transfection, the supernatants were harvested, briefly spun down to remove cell debris and stored at -80° C. The pseudovirus containing supernatants were then quantified using the RT assay.

Reverse Transcriptase Activity Assay

To quantify the amount of pseudovirus in a sample, the reverse transcriptase assay was performed. In duplicate, 500μL of pseudovirus containing supernatant was spun at 14,000xg for 2 hours at 4° C. Following the spin, the supernatant was removed and the viral pellet was resuspended in a Triton X-100 based suspension buffer and vortexed followed by a rapid freeze-thaw cycle a total of three times to properly lyse the virus. 50μL reaction mix was

added, which contained Oligo-dT Poly-A (Roche), and ^{3}H-dTTP (PerkinElmer) and the samples were incubated at 37° C for 1 hour in a heat block. Following incubation, the samples were pipetted onto cut squares of DEAE Filtermat paper (PerkinElmer), followed by three ten-minute washes in 2X SSC buffer, and one ten second wash in absolute ethanol. The filters were dried at room temp and then analyzed using a scintillation counter to quantify the CPM of ^{3}H-dTTP that was incorporated by the RT enzyme. The average CPM values from each duplicate were then used to normalize the amount of viral supernatant that was used in subsequent single-round viral entry assays.

Single Round Viral Infection Assay

Cf2Th-CD4/CCR5 or TZM-bl cells, which express both CD4 and CCR5 were set at a density of $8.0x10^{3}$ per well (Cf2Th that stably express CD4 and CCR5 in the case of 1084i pseudoviruses or TZM-bl in the case of ZM249, ADA, JR-FL and SIVmac239 pseudoviruses). The following day, a normalized volume of virus was added to each well based on the level of RT activity in the supernatant in a final volume of 100µL, diluted in complete DMEM. One day post infection, the viral supernatants were removed, the cells were gently washed one time with PBS and fresh complete media, containing 10% FBS, penicillin/streptomycin and L-glutamine, was added to each well. On day 3 post infection, the supernatants were removed, the cells were washed once with PBS and the cells were lysed in 1X Passive Lysis Buffer and frozen at -80° C. The

plates were then thawed and luciferase activity was measured using the Veritas luminometer and Beetle luciferin substrate (Promega).

CD4 Binding Assay

^{35}S labeled soluble gp120 proteins were assessed for CD4 binding by performing an immunoprecipitation assay using CD4-Ig. Input quantities were determined by immunoprecipitation using the same volume of supernatant and appropriate patient serum. For the CD4 binding of HIV-2ST, three times volume of supernatant was used in comparison to the WT because of a reduced level of binding to CD4-Ig compared to HIV-1 species [296-298]. Following IP with CD4-Ig, the bands were quantified using the Quantity One software (Biorad) and the volumes were compared to the volumes of the input gel. The ratios from WT was adjusted to 1, and the amounts of mutant protein species were compared to the WT.

CCR5 Binding Assay

Supernatants (~1.8 mL) containing ^{35}S labeled gp120 from WT or mutant envelope proteins were co-incubated for 1 hour at room temperature with 8 µg sCD4 and 5.0×10^6 CCR5 expressing Cf2-Th cells in a 2 mL reaction volume. Expression of human CCR5 on these cells was confirmed by flow cytometry. The cells were then spun down, lysed with a non-denaturing lysis buffer and cell debris was removed. An immunoprecipitation with the appropriate serum from

infected patients was then performed for one hour at room temperature and the results were analyzed by SDS-PAGE followed by exposure to a Personal Molecular Imager (Biorad), and the bands were analyzed using the QuantityOne software (Biorad).

Statistical Analysis

For all statistical analyses of cell-cell fusion and single round viral infection assays, experimental groups were compared using the Student's T-test. Each experiment was performed three separate times, with quadruplicate values. Each mutant was converted to percentage of the WT value, which was set to 100%. Significant differences between the average value of the WT and the mutants were determined using a p-value of ≤0.05. All statistical analysis was performed using the Prism 6 software (GraphPad).

Reagent Acknowledgements

The following reagents were obtained through the NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH: SIV_{agm}155-4 from Dr. Vanessa Hirsch and Dr. Philip Johnson; Anti-SIVmac251 Polyclonal; SIVagm tan-1 infectious molecular clone from Drs. Marcelo Soares and Beatrice Hahn; Antiserum to HIV-2ST gp120 from Dr. Raymond Sweet, SmithKline Beecham Pharmaceuticals; HIV-2 ST infectious molecular clone from Dr. Beatrice Hahn and Dr. George Shaw; TZMbl cells and $pSG3^{\Delta env}$ from Drs. John C. Kappes and Xiaoyun Wu.

The 1084i env gene, and clade C patient serum was a generous gift from Dr. Charles Wood. Cf2-Th-CD4/CCR5 and Cf2-Th-CCR5 cells, and pSVIII ADA, JRFL and YU2 envelope constructs were originally obtained from Dr. Joseph Sodroski.

FIGURE 2.1

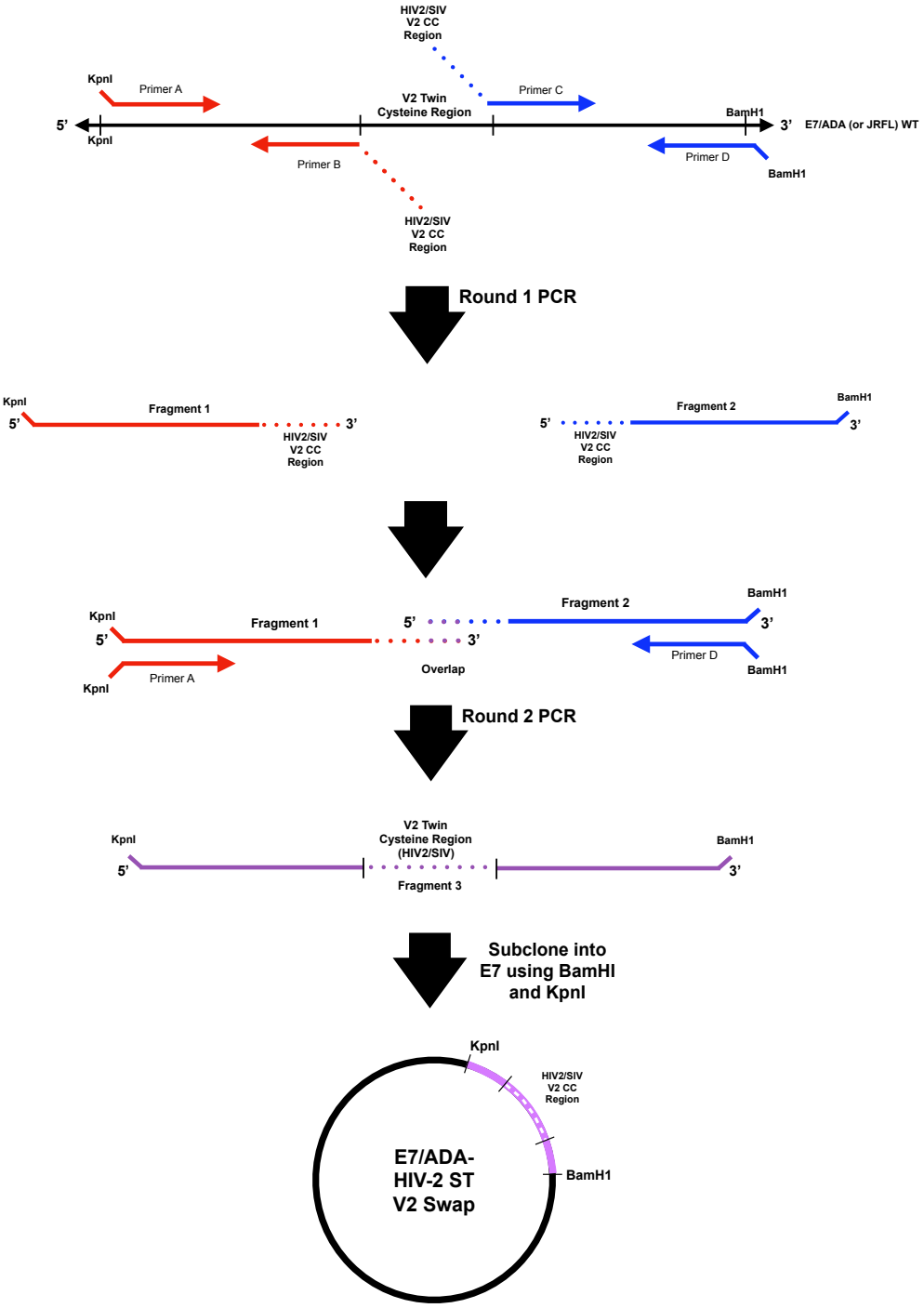
HIV2/SIV
V2 CC
Region
KpnI
Primer A
V2 Twin
Cysteine Region
Primer C
BamH1
5'
3'
E7/ADA (or JRFL) WT
KpnI
Primer B
Primer D
BamH1
HIV2/SIV
V2 CC
Region
Round 1 PCR
KpnI
Fragment 1
5'
3'
HIV2/SIV
V2 CC
Region
Fragment 2
BamH1
5'
HIV2/SIV
V2 CC
Region
3'
KpnI
Fragment 1
5'
3'
Fragment 2
5'
BamH1
3'
Primer D
BamH1
KpnI
Primer A
Overlap
Round 2 PCR
KpnI
V2 Twin
Cysteine Region
(HIV2/SIV)
BamH1
5'
3'
Fragment 3
Subclone into
E7 using BamHI
and KpnI
KpnI
HIV2/SIV
V2 CC
Region
E7/ADA-
HIV-2 ST
V2 Swap
BamH1

FIGURE 2.1. Overlapping PCR method to generate V2 swap mutants. Overlapping PCR was used to generate HIV-1 V2 swap mutants where the V2 twin cysteine regions from SIVmac239 and HIV-2 ST were swapped into both ADA and JRFL. The primers used for overlapping PCR are included in Appendix G. This figure only shows the swap process into HIV-1 ADA, but these steps were also performed using JRFL as the template. Dotted lines indicate non-HIV-1 V2 twin cysteine sequence.

FIGURE 2.2

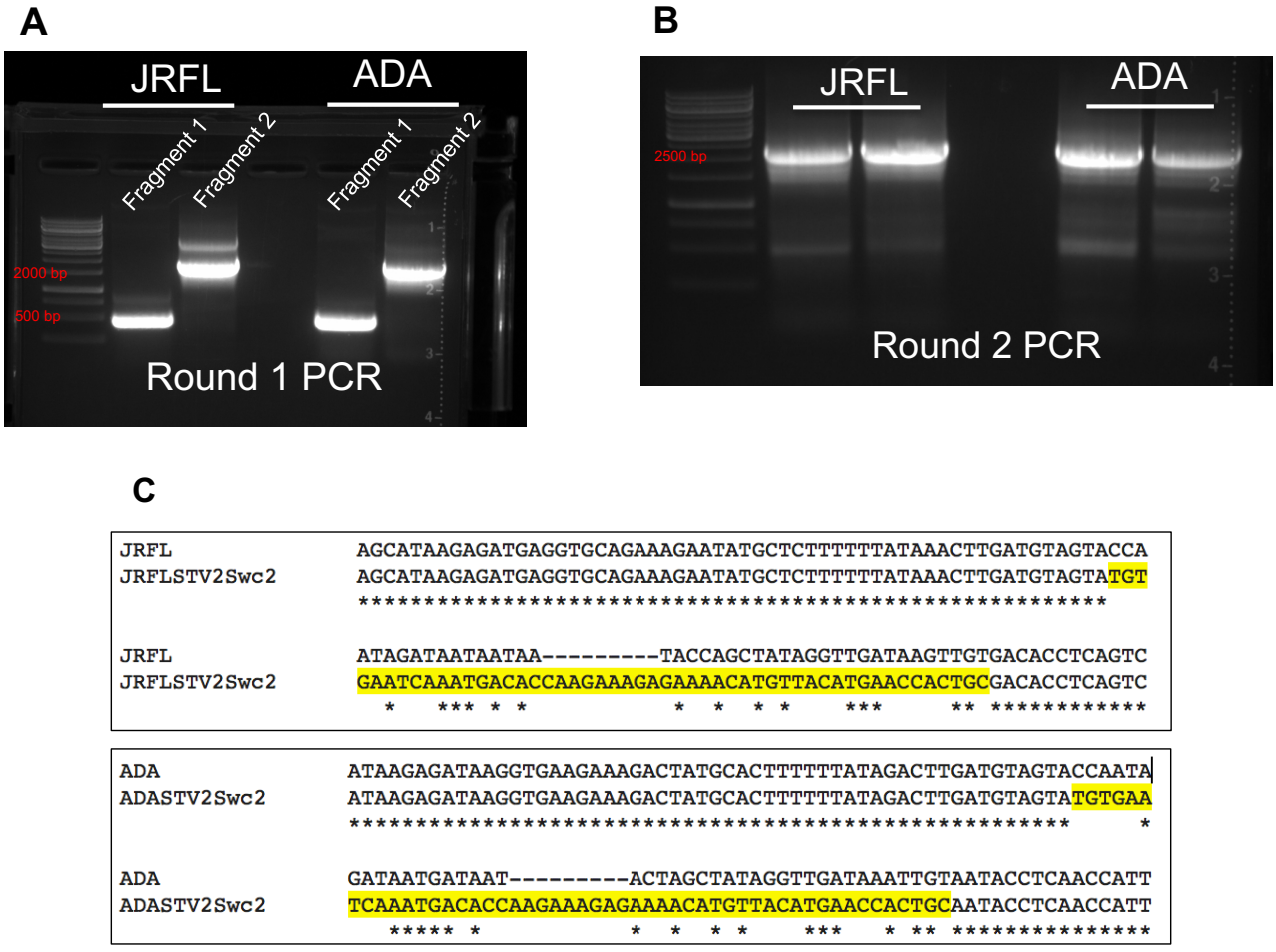
A
JRFL
ADA
Fragment 1
Fragment 2
Fragment 1
Fragment 2
2000 bp
500 bp
Round 1 PCR
B
JRFL
ADA
2500 bp
Round 2 PCR
C
JRFL AGCATAAGAGATGAGGTGCAGAAAGAATATGCTCTTTTTTATAAACTTGATGTAGTACCA
JRFLSTV2Swc2 AGCATAAGAGATGAGGTGCAGAAAGAATATGCTCTTTTTTATAAACTTGATGTAGTATGT
JRFL ATAGATAATAATAA----------TACCAGCTATAGGTTGATAAGTTGTGACACCTCAGTC
JRFLSTV2Swc2 GAATCAAATGACACCAAGAAAGAGAAAACATGTTACATGAACCACTGCGACACCTCAGTC
ADA ATAAGAGATAAGGTGAAGAAAGACTATGCACTTTTTTATAGACTTGATGTAGTACCAATA
ADASTV2Swc2 ATAAGAGATAAGGTGAAGAAAGACTATGCACTTTTTTATAGACTTGATGTAGTATGTGAA
ADA GATAATGATAAT----------ACTAGCTATAGGTTGATAAATTGTAATACCTCAACCATT
ADASTV2Swc2 TCAAATGACACCAAGAAAGAGAAAACATGTTACATGAACCACTGCAATACCTCAACCATT

FIGURE 2.2. Overlapping PCR to create HIV-2 ST V2 Swap Mutants. (A) Round 1 PCR produced fragments 1 and 2 from ADA and JRFL with the V2 region from HIV-2 ST. Fragment 1 was generated by PCR using primer A and Primer B. Expected size for both ADA and JRFL is 573 bp. Fragment 2 was generated by PCR using primer C and D. Expected band for JRFL was 2049 bp and ADA was 2067 bp. **(B)** Second round PCR included both fragments from round 1 and produced the full length JRFL and ADA env gene containing the V2 region from HIV-2 ST. Expected size for JRFL was 2598 bp, and ADA was 2616 bp. **(C)** Sequencing confirmed the presence of the HIV-2 ST V2 swap for ADA and JRFL.

TABLE 2.1

Plasmid Name	Vector Name	Envelope Gene Species	Plasmid/Gene Source	Mutations Generated	Notes
Envelope Expression Plasmids					
pSVIII/1084ienv	pSVIII	HIV-1 1084i	pSVIII- J. Sodroski Lab; 1084i env- C. Wood Lab	303Gins, 322Gins, M307A, M307E, M307L, I309A, I309E, I309L, F317A, F317E, F317L	-
pCDNA3.1+/ZM249env	pCDNA 3.1+	HIV-1 ZM249MB_10	pCDNA3.1+ - Invitrogen; pZM249MB_10- B. Hahn (NIH/AIDS Reagent Program)	303Gins, M307A, M307E, M307L, I309A, I309E, I309L, F317A, F317E, F317L	-
pCDNA3.1+/HIV2STenv	pCDNA 3.1+	HIV-2 ST	pCDNA3.1+ - Invitrogen; pHIV-2/ST- B.Hahn and G. Shaw (NIH/AIDS Reagent Program)	303Gins, I307A, I309A, F317A, C183A, C191A, C183A/C191A	-
pCDNA3.1+/SIVmac239env	pCDNA 3.1+	SIVmac239	J. Sodroski Lab	303Gins, V307A, I309A, F317A	-
pCDNA3.1+/SIVagmTan1env	pCDNA 3.1+	SIVagmTan-1	pCDNA3.1+ - Invitrogen; pSIVagmTan-1- B. Hahn and M. Soares (NIH/AIDS Reagent Program)	303Gins, V307A, I309A, F317A, C183A, C191A, C183A/C191A	-
pSVIII/ADAenv	pSVIII	HIV-1 ADA	J. Sodroski Lab	S183C, P191C, 183CinsL/191Cins, 183CinsR/191Cins, SIVmacV2Swap, HIV2STV2Sw	-
pSVIII/JRFLenv	pSVIII	HIV-1 JRFL	J. Sodroski Lab	S183C, P191C, 183CinsL/191Cins, 183CinsR/191Cins, SIVmacV2Swap, HIV2STV2Sw	-
Pseudoviral Production Plasmids					
pHIV-1Luc	-	-	J. Sodroski Lab	-	Transfected with Env expression construct and CMVPack to produce luciferase tagged, replication deficient viral particles. Expresses firefly luciferase gene.
pCMVPack	-	-	J. Sodroski Lab	-	CMV HIV packaging plasmid, encodes for Gag-Pol genes
pSG3Δenv			J. Kappes, X. Wu (NIH/AIDS Reagent Program)	-	Env deficient provirus based on pSG3.1. Contains a defective Env gene due to a 4 bp insertion, and an early stop codon. Also has a defective vpu gene.
Other Plasmids					
pTat	-	-	pTat- J. Sodroski Lab	-	Expresses the HIV-1 Tat protein under a CMV promoter. Cotransfected in fusion assays to act upon the Tat-dependent promoter for luciferase in TZM-bl cells.

TABLE 2.1. Complete listing of env genes, point mutations and plasmids generated for the studies in this dissertation.

CHAPTER 3- ASSOCIATION OF THE gp120 V3 LOOP WITH ENVELOPE TRIMER STABILIZATION IN PRIMATE LENTIVIRUSES

Data from this chapter were published in *Virology*, 2018. Volume 521, August 2018, Pages 158-168.

Contributing authors to this publication include:

Dane Bowder, Haley Hollingsead, Kate Durst, Duoyi Hu, Wenzhong Wei, Joshua Wiggins, Halima Medjahed, Andrés Finzi, Joseph Sodroski, Shi-Hua Xiang

INTRODUCTION

Entry of HIV-1, HIV-2 or SIV is facilitated by the envelope protein trimer complex which is displayed on the surface of the virus. This complex is composed of three molecules of gp120, the surface glycoprotein (SU), non-covalently bound to three molecules of gp41 which are transmembrane proteins (TM). The gp120 first binds to CD4 which initiates a conformational change in the gp120 molecule to expose the coreceptor (CCR5 or CXCR4) binding site [299-301]. Following coreceptor binding, gp41 undergoes an additional major conformational change, which brings the host and viral membrane into proximity, to initiate membrane fusion.

The V3 loop is one of five variable regions in the HIV envelope protein which plays a major role in contacting the coreceptor for viral entry. Thus, the V3 loop is essential for the function of the envelope protein as deletion of the V3 loop results a complete reduction of envelope initiated fusion and viral entry [240, 241]. The length of this loop is well conserved across viral strains at 34-35 amino acids [234, 238, 249, 302, 303]. The V3 loop can be divided into three parts: the base, stem and tip. The base and tip are usually more conserved in sequence, while the stem region is more variable and flexible [234, 249, 304]. The sequence of the V3 loop determines the coreceptor selection of CCR5 or CXCR4 [302, 305-309]. In addition to its role in coreceptor binding, the V3 loop has also been found to be associated with envelope trimer stabilization [5, 236]. This work revealed that insertions or deletions in the V3 sequence cause a major destabilization of the envelope complex as indicated by increased envelope

gp120 shedding. Additionally, conserved hydrophobic residues in a region known as the “hydrophobic patch” in the tip of the V3 loop have also been shown to contribute to the stability of the envelope trimer, as substitution of these residues with neutral amino acids, or hydrophilic amino acids results in a similar decrease in envelope trimer stability [5]. This was the first study to show that the variable loops may play a significant role in envelope stability, as the variable loops were originally thought to have a flexible conformation and were thus removed from the protein in early structural studies of this protein [28]. It is important to note that the effect of the hydrophobic patch has been shown exclusively in subtype B HIV-1 species. Subtype B is the subtype predominantly responsible for HIV-1 infections in the United States and Western Europe, and has been studied most extensively. Subtype C however, accounts for approximately 50% of HIV-1 infections globally, as it is the subtype that accounts for most of the HIV-1 infections in sub-Saharan Africa and India [155, 310]. Additionally, there is evidence that HIV subtype C differs from subtype B in terms of transmission, pathogenesis and disease progression, which stresses the importance of understanding the biology of subtype C viruses [156, 158, 311, 312].

Several near atomic structures of the HIV-1 envelope trimer have been released and have helped to make important progress in the field of HIV entry research and vaccine development. These structures show that the V3 loop is packaged into the apex of the trimer as part of the trimer association domain, which supports the finding that the V3 loop plays an important structural role in envelope trimer stabilization [2, 3, 210, 275]. The V3 loop is also highly

immunogenic, and is an important target of antibodies both in natural infection and in vaccination studies [255, 306, 313-315]. Because of this, recent efforts to use the SOSIP trimer as a vaccine immunogen have shown that non-cross neutralizing V3 antibodies are generated in high numbers [220, 257, 272, 277]. To reduce this effect, recent efforts have sought to further stabilize the V3 loop to reduce its exposure and antigenicity, while still generating heterologous neutralizing antibodies [257, 270, 272, 277]. Strategies have included from the targeted addition of glycans and additional hydrophobic residues flanking the hydrophobic patch described previously, indicating the importance of this hydrophobic region in the positioning of the V3 loop in the envelope protein complex [258, 270, 277].

Although the sequences and lengths of the V3 loops are variable between primate lentiviral species (including HIV-2 and SIV), the hydrophobic patch residues remain conserved [5]. There are some variations in the exact amino acid sequence at these positions, but these residues remain hydrophobic regardless of lentiviral species. This indicates that the hydrophobic patch is an evolutionarily relevant motif, and plays an important structural role. It is important to note, as HIV-2 represents an entirely separate transmission of an SIV into the human population, and because SIVs from other primate species often behave exceptionally different from HIV-1 in terms of pathogenesis, immune evasion, and antibody development [11, 123, 165]. The relationship between HIV-1 and SIV species is an important area to study, as fully understanding the way SIV

species function in the hosts will help provide understanding of the pathogenesis and transmission of HIV-1.

In this work, the role of the V3 loop hydrophobic patch in both HIV-1 subtype C, HIV-2 species, and SIV nonhuman primate lentiviruses were examined. The hypothesis that the V3 loop plays an important role in the stability of the unliganded envelope trimers of HIV-1 subtype C, HIV-2 and SIV nonhuman primate lentiviruses is tested through mutational studies of the envelope on its stability and functionality, as evidenced by the ability to initiate envelope fusion and infection into cells.

RESULTS

Generation of V3 Hydrophobic Patch Mutants in Primate Lentiviral Species

Based on the first report on the V3 Hydrophobic Patch in HIV-1 subtype B strains published in 2010 [5], an alignment of the V3 loops of multiple species of primate lentiviruses, the hydrophobic patch region is conserved across HIV-1 subtypes as well as in HIV-2 and SIV. Although there is some variability in the residues present at amino acid positions 307, 309 and 317, those residues are consistently hydrophobic (Figure 3.1.A). The hydrophobic patch residues are positioned in the tip of the V3 loop, with 307 and 309 on one side of the turn sequence, and 317 directly across these residues on the other side of the turn (Figure 3.1.B-C). Based on this sequence alignment, the recent structural models and our previously published data, it is our hypothesis that the hydrophobic nature of these residues contributes to the stability of all primate lentiviral

envelope trimers. To test this hypothesis, a panel of hydrophobic patch mutants was generated in two HIV-1 subtype C species, ZM249MB_10 (ZM249) and 1084i. In each of these subtype C strains, a glycine (G) insertion mutant was first generated to observe the effects on envelope shedding and decreased functionality between amino acids T303 and R304. V3 insertion mutations are known to decrease envelope complex stability, as they change the overall length of the V3 loop. Additionally, an alanine (A) substitution was generated at amino acids 307, 309 and 317 individually to represent a change to a relatively neutral amino acid. A glutamic acid (E) substitution was generated at each position to represent a change to a hydrophilic amino acid and a leucine (L) substitution was generated to represent a hydrophobic to hydrophobic change. All mutations were generated by site directed mutagenesis. Based on the natural hydrophobicity of the WT residues, our hypothesis is that the alanine and glutamic acid substitutions will cause a decrease in envelope stability and in turn cause a decrease in cell-cell fusion and infectivity, and that the leucine substitution will not cause as significant of an overall change, though small decreases or increases in stability and functionality may occur.

To broaden the scope of this study, a wider panel of primate lentiviruses, including HIV-2 and SIVs, was also included (Figure 3.1). Glycine insertion mutations between corresponding residues 303 and 304 and alanine substitution mutations to the hydrophobic patch were made to HIV-2 ST, SIVagmTan-1 and SIVmac239. A complete table of strains used and mutations introduced can be found in Table 2.1 in Chapter 2.

Neutral and Hydrophilic Substitutions to the V3 Loop Reduce Stability and Functionality of HIV-1 Subtype C Envelope Complex.

To observe the effect of the hydrophobic patch mutations on the stability and functionality of the envelope proteins, the gp120 shedding assay was used to assess the stability of the envelope trimer and is shown for ZM249 in Figures 3.2.A-D. In this assay, an immunoprecipitation of ^{35}S labeled proteins is performed on both the cell lysate and supernatant fraction of transfected cells. The amount of gp120 in the cell lysate fraction is compared to the amount of gp120 in the supernatant, and the association index is an established method to quantify the difference in gp120 between WT and mutant PLV envelopes [293]. Insertion of a glycine residue in the V3 loop resulted in a massive increase in gp120 shedding. Based on these results, it is evident that the length of the V3 loop is important for envelope trimer stability in subtype-C HIV-1, as was seen in subtype-B viruses [5]. Neutral and hydrophilic amino acid substitutions caused a significant decrease in envelope stability at all the hydrophobic patch locations. The hydrophobic residue (isoleucine, I or phenylalanine, F) to hydrophobic residue (leucine, L) mutations did not cause an increase in gp120 shedding. This suggests that hydrophobicity at this position is required, but not a specific residue in the hydrophobic patch. According to the SOSIP structure, the V3 loop hydrophobic patch appears to interact with other hydrophobic residues in the gp120 core which stabilizes the envelope trimer structure by reinforcing the overall monomeric gp120 structure.

To further understand the effect of these mutations, we assessed the ability of these envelopes to initiate cell-cell fusion. Correlated with the results of the shedding assay, the level of cell-cell fusion was largely reduced by all the mutants except the leucine mutations. The isoleucine or phenylalanine to Leucine mutations showed a much lower reduction in the fusion, just as they showed no major change in gp120 shedding (Figure 3.2.E-H). This reduction in fusion is the result of the decrease in envelope stability. The glycine insertion mutation caused a significant decrease in envelope fusion as expected (Figure 3.2.E).

Finally, to evaluate the effects of these mutations on viral entry, the WT and mutant envelopes were used to generate pseudotyped viruses and normalized infections based on an RT activity assay were carried out on TZM-bl cells. These pseudoviruses are only capable of entering a cell, not replicating, so this assay provided a clear picture of the entry process facilitated specifically by the envelope trimer. The results followed a very clear trend wherein the glycine insertion, alanine and glutamic acid substitution mutations caused a decrease in viral entry to the level of background, and an intermediate level of entry was achieved in the envelopes with leucine mutations (Figures 3.2.I-L). The leucine mutations at positions 307 and 309 did significantly reduce the entry of these viruses (Figure 3.2.J, K), although to a lesser degree than the alanine and glutamic acid mutations. The leucine mutation at position 317 did not cause a significant change in the entry of these single-round viruses (Figure 3.2.L).

Importantly, the mutations to the ZM249 V3 loop did not appear to affect the ability of this envelope protein to bind CD4. CD4-Ig was used to immunoprecipitate labeled gp120 from transfected supernatants, and no significant difference in binding compared to an input quantity was observed between the WT sample and the mutants (Table 3.1). In this experiment, alanine mutations were used to represent the larger panel of hydrophobic patch mutations. This indicates that the mutations made did not drastically affect the overall shape of the protein regarding the CD4 binding site, and that the reduced functionality is not due to impaired CD4 binding.

Although the binding of CD4 was not drastically reduced, the binding of the alanine mutants to CCR5 was undetectable. This is not surprising, as significant reduction in CCR5 binding was also observed in subtype B viruses with the same mutations [5]. One of the biological roles of the V3 loop is to determine coreceptor usage and form the coreceptor binding site, and mutations to this region likely alter the binding to CCR5. A CCR5 expressing cell based assay was used in this experiment to bind shed gp120 from the supernatants of ^{35}S labeled envelope expressing cells. It should be noted that because ZM249 envelope naturally sheds gp120 at a low level. Because of this, it was difficult to detect binding of WT gp120 to CCR5 from the supernatant of WT envelope in this assay, so WT HIV-1 YU2 envelope (which does shed rapidly, similar in scale to the hydrophobic patch mutants themselves) was used as a positive control (an example of these results is shown in Figure 3.3). It is hard to conclude that the level of CCR5 binding was reduced in the hydrophobic patch mutations

compared to WT, however, the binding between the mutants and CCR5 was clearly undetectable, which is likely because the hydrophobic patch mutations are a contributing factor to the reduced level of cell-cell fusion and entry.

Consistent results were obtained when the 1084i envelope strain was used to perform the same experiments. Shedding (Figure 3.4.A-D), fusion (Figure 3.4.E, G-H) and viral entry (Figure 3.4.I, K-L) of 1084i with the panel of mutations at position I309 and F317 follow the same trend as ZM249. Alanine and glutamic acid mutations significantly reduced envelope stability, fusion and viral entry, which were affected to a lesser degree by leucine mutations. However, there were a few notable exceptions at position 307 in 1084i strain. The M307L mutation resulted in a significant increase in viral entry which was much higher than the WT, though its fusion ability was comparable to the WT. Another discordance is that the mutations of M307A and M307E at position 307 did not cause a significant change in the ability of this envelope to initiate cell-cell fusion (Figure 3.4.F, J). For the time being, we do not have an explanation for why the M307L mutant has an extremely high level entry or why the hydrophobicity reduction mutants at this position of 307, M307A and even M307E (a hydrophilic residue change) did not reduce the fusion ability, though increased affinity to either CD4 or more likely CCR5 is likely the cause of this change. This is certainly an interesting question but further investigation is needed to answer it.

Alanine Substitution Mutations to the V3 Loop Hydrophobic Patch of HIV-2 ST Reduce Envelope Stability and Functionality.

We next assessed the importance of the V3 loop hydrophobic patch in the HIV-2 strain, ST. Alanine mutations were made to the amino acid residues at positions 307, 309 and 317, and a glycine insertion was made between amino acids T303 and R304. As expected, all mutations caused a significant increase in gp120 shedding (Figure 3.5.A). As with the HIV-1 subtype C samples, this decrease in envelope stability correlated with a decrease in cell-cell fusion capability, the mutations significantly reduced the cell-cell fusion to the level of background (Figure 3.5.B). It is surprising that the Alanine mutations at positions 309 and 317 and the 303G insertion showed a significant decrease in the ability to bind to CD4-Ig, however, the overall binding of the HIV-2 ST envelope to CD4 is much lower than HIV-1 [296, 298], so even a minor reduction in binding appears to be significant (Figure 3.4.C). Overall, this data suggests that the structural role and function of the V3 hydrophobic patch is preserved in HIV-2 species, supporting our hypothesis.

The role of the V3 loop hydrophobic patch is conserved in nonhuman primate lentiviruses.

To determine whether the structural role of the V3 loop hydrophobic patch is also conserved in non-human primate lentiviruses, we tested the effects of these mutations in the SIVmac species. We generated a V3 loop glycine

insertion mutant, as well as a panel of alanine mutations to each of the residues in the hydrophobic patch in the SIVmac239 envelope (Figure 3.1.A), and carried out the same assays that were previously described (Figure 3.6.A-D). Interestingly, in SIVmac239, the I309A mutations resulted in less gp120 shedding than the 303G, V307A and the F317A mutations, as indicated by a much higher association index (Figure 3.6.A). This trend correlated with the fusion and entry data from these mutations, as fusion of I309A was not significantly different from WT envelope (Figure 3.6.B), and there was a much reduced but detectable level of entry by the I309A viruses, whereas the other mutations did not facilitate entry above background level (Figure 3.6.C). The binding of these envelopes to CD4-Ig was examined, and the level of binding by the mutants was not significantly different, indicating that our mutations did not initiate a major conformational change in the protein structure that affected the binding of this envelope to CD4. As in ZM249, the V3 mutants did not bind to CCR5 notably. WT SIVmac239 envelope was undetectable in this assay due to its low level of gp120 shedding, so WT YU2 gp120 was used as a positive binding control. The fact that I309 was not affected as majorly as the other residues in this strain of SIV is further evidence that there is a degree of strain specificity associated with the V3 loop hydrophobic patch. Although the role of the patch appears to be well conserved, some strains are more capable of carrying less hydrophobic amino acids at these positions. Clearly, SIVmac239 is more capable of carrying an alanine residue at position 309 than at positons 307 and 317, and is certainly more capable of

carrying this mutation than the other PLVs studied in this work, but the role of the hydrophobic patch in trimer stability is consistent.

The stability of the SIVagmTan-1 envelope is decreased by alanine substitution mutations to the V3 loop hydrophobic patch.

We also tested another SIV strain, SIVagmTan1, a typical SIV strain that naturally infects the African green monkey. A glycine insertion mutation was made between residues T303 and V304, and alanine mutations were made to each of the hydrophobic patch residues at amino acids 307, 309 and 317. The data from the insertion mutant 303G and the three hydrophobic patch mutants (V307A, I309A and F317A) supported our hypothesis and showed the same effect as in HIV-1. These mutations showed a very clear decrease in envelope stability as assessed by gp120 shedding (Figure 3.7.A), and this correlated to a significant decrease in cell-cell fusion initiated by these mutant envelopes (Figure 3.7.B). It is important to note that the overall levels of fusion initiated by the SIVagmTan1 envelope was quite low compared to the other strains described previously, though the mutations fused at levels very close to background. This may be a result of the TZM-bl target cells expressing human, not African green monkey, CD4 and CCR5.

Molecular modeling supports the role of the V3-loop hydrophobic patch in envelope trimer stabilization.

Our observations are supported by structural modeling of the envelope trimer complex. According the SOSIP envelope trimer structure in the native state (PDB 4ZMJ, [1]), it is clearly seen that the V3 loop (shown in magenta) sits in the apex of the trimer, where it plays a role in the association of the trimer along with the V2 loop (shown in cyan) (Figure 3.9.A, top-view). The interaction of the V2 and V3 loops in the CD4 unbound (native) state is depicted in Figure 3.9.B-C. The hydrophobic patch residues of the V3 loop sit perfectly in a hydrophobic cavity formed by the residues including F159, M161 and V172 in the V2 loop, forming a tight hydrophobic contact. Mutation of these residues to neutral, or hydrophilic residues disrupts these hydrophobic interactions, resulting in destabilization. Additionally, additions to the length of the V3 loop clearly will result in a crash of the V3 loop with the V2 loop, disrupting the fit conformation of these loops for trimer stability.

DISCUSSION

Previously, it has been shown that the V3 loop hydrophobic patch influences HIV-1 envelope stability using prototypic HIV-1 subtype B strains [5]. This work demonstrates that this conserved sequence region also has a conserved function in subtype C HIV-1, HIV-2, and SIV species. In subtype C species, we demonstrated that mutation of the hydrophobic residues at amino acid positions 307, 309 and 317 to hydrophilic or neutral amino acid residues resulted in a significant increase in gp120 shedding into the media, indicating a lack of envelope trimer stability. Our data supports envelope trimer structural

models, as the V3 loop is consistently packaged into the apex of the envelope trimer in these structures [1-3, 210, 211]. This work supports the fact that the V3 loop plays a structural role, likely stabilizing the envelope monomers to help them achieve trimeric conformation.

The changes seen in gp120 envelope stability correlate very closely to functional data. When envelope stability is disrupted by these mutations, similar detrimental effects are seen in cell-cell fusion and viral entry. We posit that this is an effect of the decreased stability of the envelope complex and a decrease in CCR5 binding. Because the mutant envelope trimers are less stable, and shedding increases drastically, significantly less envelope protein is present on the surface of the cell or the viral particle to facilitate these processes. Additionally, molecular modeling predicted that the effects of addition to the V3 loop and hydrophobic patch will be drastic in both the CD4 unbound and bound states. In the unbound state, the hydrophobic patch fits perfectly into the hydrophobic cavity of the V2 loop, due to hydrophobic interactions (Figure 3.8). Additions to the V3 loop destabilize the envelope trimer because of a crash with the V2 loop. This moves the hydrophobic patch out of the hydrophobic cavity, and the trimer dissociates due to a lack of stability. CCR5 binding is dependent on two contacts with the V3 loop, which are made possible by the conformational change initiated by CD4 binding [5, 306]. In the bound state, the V3 loop is exposed outside of the envelope trimer. In this bound state, the hydrophobic patch contacts the extracellular loop 2 (ECL2) of CCR5, participating in CCR5 binding. A sulfated tyrosine 14 (sTyr14) at the N-terminus of CCR5 makes

contact with N302 which is positioned in the V2 cavity to accommodate this sulfated tyrosine in the CD4-bound state [316, 317], and the sulfated tyrosine 173 (sTyr173) of the V2 loop of gp120 in the unliganded state, for which the electrostatic interactions are characterized [318]. The second contact is with the ECL2 for the hydrophobic patch by which the non-specific hydrophobic interactions are assumed [319-321]. Based on molecular modeling, when insertions are made to the loop, the hydrophobic patch is moved into a position where it is unable to contact CCR5, and fusion does not occur. The exact nature of the interaction between the V3 loop with the ECL2 of CCR5 is unknown, and this interaction will be explored in future studies.

An interesting result of these studies was found in HIV-1 1084i. A leucine substitution to the methionine residue at position 307 resulted in a very significant increase in viral entry, and the hydrophilic and neutral substitutions had no effect on cell-cell fusion. This is an interesting phenomenon and is likely due to strain specificity. 1084i as a pseudovirus does not enter cells at a high level when compared to other viruses, so this mutation may be useful when a highly infectious subtype C virus is required in the laboratory. Future experiments to study this phenomenon would be to substitute other hydrophobic amino acids to position 307 to determine whether other amino acids would initiate the same effect, though preliminary results do not show the same increased infection phenotype with an isoleucine substitution to this residue.

There are several directions to take the results of this work to better improve structural understanding of the envelope proteins of SIV and HIV. The

first is to continue to pursue hydrophobic residues that are not part of the V3 loop, but are close in physical proximity. Importantly, the V3 loop hydrophobic patch appears to be part of a much larger hydrophobic region we have coined the “hydrophobic core”, which contains elements of both the V1/V2 and V3 loop as well as other conserved regions in gp120. It appears that this region plays a similar role in HIV envelope structural stability, as mutations to these residues cause similar decreases in envelope trimer association. Another area that will be interesting to pursue in this research is the idea of species specificity to address the reasons why the hydrophobic patch of some primate lentiviral species is more capable of carrying mutations than others. It would be very interesting to consider the long-term fitness of SIVmac239 carrying an alanine at position 309 compared to the naturally occurring isoleucine residue. Does a single residue change like this affect the pathogenesis of the virus, or is this simply an interesting anomaly compared to other species studied in this work? Finally, this data will hopefully inform the development of new epitope and envelope trimer based vaccines. While not the direct intention of this study, identifying residues that could naturally stabilize the envelope in a trimeric state could be exploited in the development of a better vaccine against HIV-1.

In conclusion, our data clearly demonstrates that the V3 loop plays an important role in the stabilization of the envelope trimer in primate lentiviruses. The envelope trimer stability is controlled, in part, by the length of the V3 loop and the positioning of the hydrophobic patch in both the CD4 bound and unbound conformations. Additionally, the length of the V3 loop and the

hydrophobic patch are essential for proper function of the envelope trimer, alterations to either of these results in a very clear reduction of the functionality of the envelope.

FIGURE 3.1

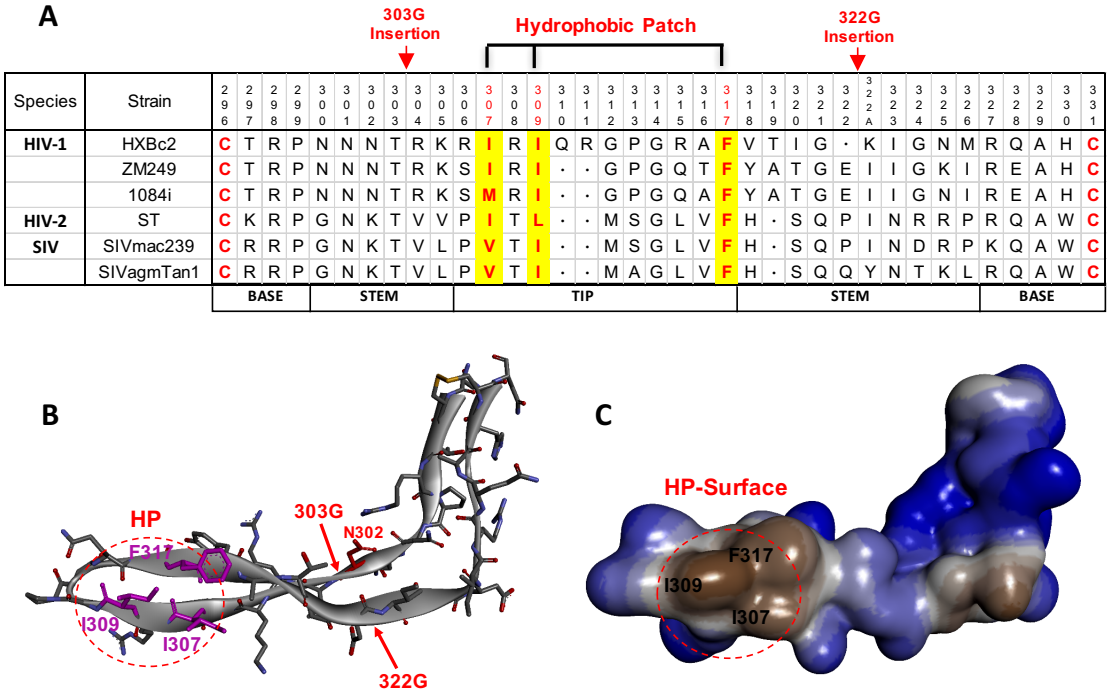
A
303G Insertion
Hydrophobic Patch
322G Insertion
Species
Strain
HIV-1
HXBc2
ZM249
1084i
HIV-2
ST
SIV
SIVmac239
SIVagmTan1
BASE
STEM
TIP
STEM
BASE
B
HP
F317
I309
I307
303G
N302
322G
C
HP-Surface
F317
I309
I307

FIGURE 3.1. Alignment of the V3 loop sequences of primate lentiviral species. A. Numbering was assigned according to standard HXBc2 numbering, which is included as a reference point [322]. The hydrophobic patch residues that were altered in the study are shown in yellow, and the location of glycine insertions are between the residues indicated by an arrow. **B.** Structural ribbon model of the V3 loop, highlighting the hydrophobic patch. The positions involved in the hydrophobic patch are circled in red, and the locations of insertion mutations are indicated with an arrow. **C.** Surface hydrophobicity model of the V3 loop and hydrophobic patch. The hydrophobic patch residues are indicated by a red circle. More hydrophobic areas are indicated in gray, and more hydrophilic in blue.

FIGURE 3.2

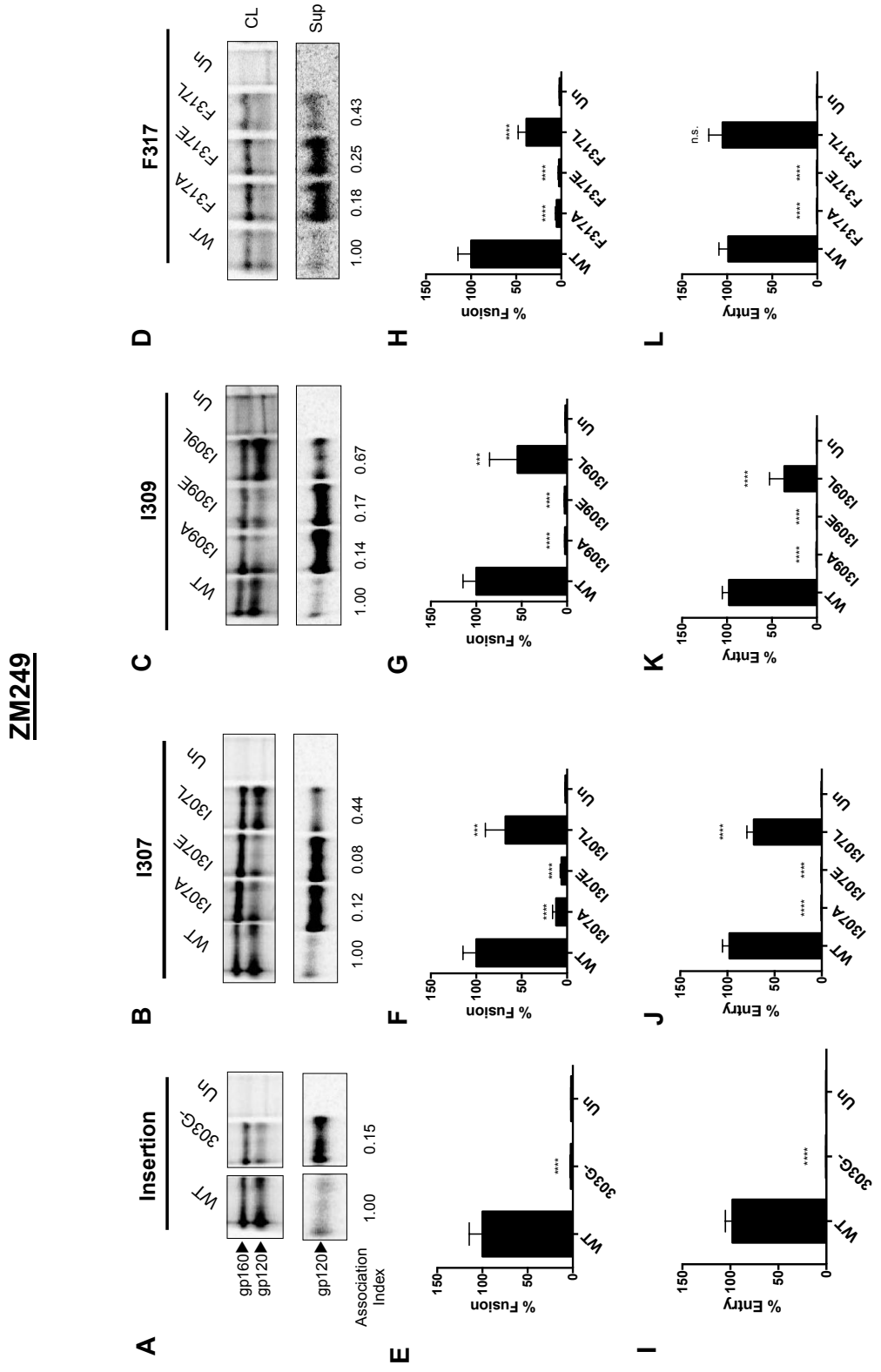
ZM249
A
Insertion
WT
303G-
Un
gp160
gp120
gp120
Association Index
1.00
0.15
B
I307
WT
I307A
I307E
I307L
Un
1.00 0.12 0.08 0.44
C
I309
WT
I309A
I309E
I309L
Un
1.00 0.14 0.17 0.67
D
F317
WT
F317A
F317E
F317L
Un
CL
Sup
1.00 0.18 0.25 0.43
E
% Fusion
F
% Fusion
G
% Fusion
H
% Fusion
I
% Entry
J
% Entry
K
% Entry
L
% Entry
n.s.

FIGURE 3.2. Envelope stability, cell-cell fusion, and viral entry in HIV-1 ZM249 with hydrophobic patch mutations. (A-D). gp120 shedding assay. SDS-PAGE analysis of immunoprecipitated cell lysate and supernatant fractions from cells expressing WT or mutant envelope ZM249 proteins. The gp160 and gp120 bands are indicated. The association index value was calculated as previously reported. Note that the gel shown in Figure 2A has been spliced and reformatted for presentation purposes. **(E-H).** Cell-cell fusion capabilities of the WT and mutant envelopes are expressed as percentages based on luciferase values as described in the materials and methods section. All figures are represented of at least three separate experiments performed with quadruplicate values. **(I-L).** Entry of WT or mutant envelope expressing pseudotyped virus is shown as a percentage based on luciferase values as indicated in the materials and methods section. All figures are representative of at least three separate experiments performed with quadruplicate values. Significance was determined by the student's T-test (*$p<0.05$).

FIGURE 3.3

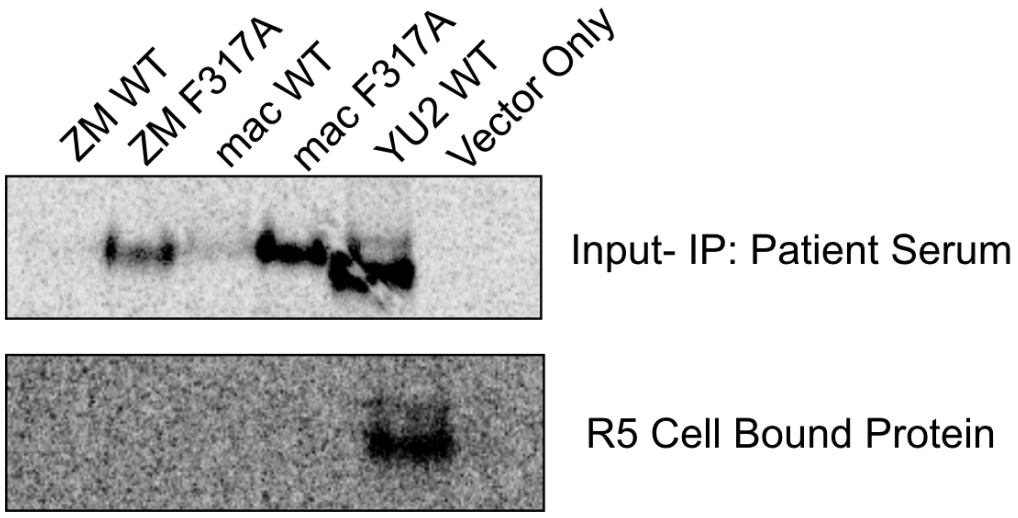
ZM WT
ZM F317A
mac WT
mac F317A
YU2 WT
Vector Only
Input- IP: Patient Serum
R5 Cell Bound Protein

FIGURE 3.3. Example of CCR5 Binding capabilities of hydrophobic patch mutations. Soluble monomeric gp120 was taken from the supernatant of WT and mutant PLVs and immunoprecipitated with appropriate patient serum to determine the input volume of gp120 (top panel). The same volume of supernatant was co-incubated with sCD4 and CCR5 expressing Cf2-Th cells for one hour. The R5 cells were then isolated, lysed and immunoprecipitation with patient serum of this lysate was carried out to determine the amount of CCR5 bound gp120 (bottom panel).

FIGURE 3.4

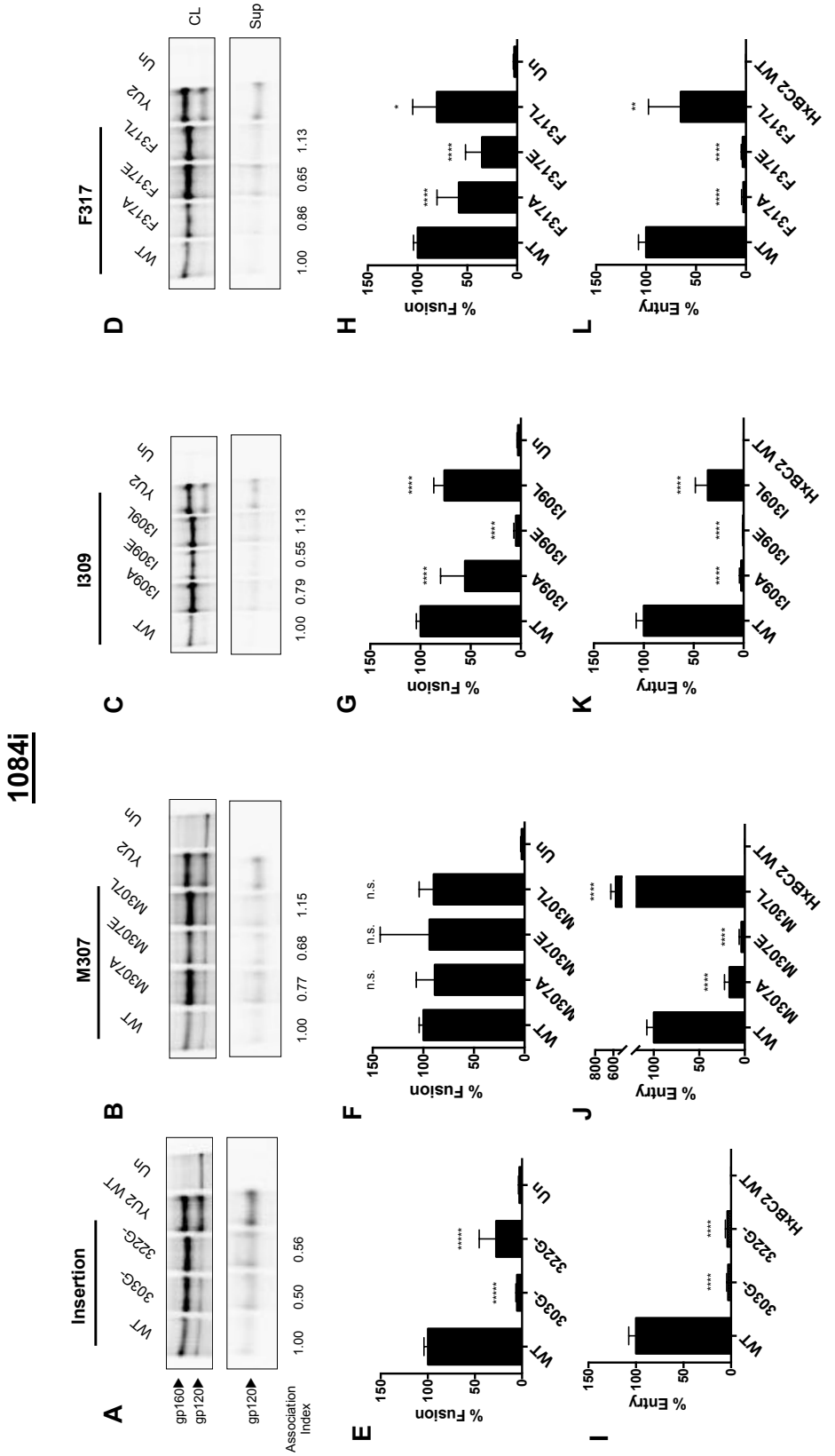
1084i
A
Insertion
WT
303G-
322G-
YU2 WT
Un
gp160
gp120
gp120
Association Index
1.00 0.50 0.56
B
M307
WT
M307A
M307E
M307L
YU2
Un
1.00 0.77 0.68 1.15
C
I309
WT
I309A
I309E
I309L
YU2
Un
1.00 0.79 0.55 1.13
D
F317
WT
F317A
F317E
F317L
YU2
Un
CL
Sup
1.00 0.86 0.65 1.13
E
% Fusion
F
% Fusion
G
% Fusion
H
% Fusion
I
% Entry
J
% Entry
K
% Entry
L
% Entry
HxBC2 WT
n.s.
*
**
****
*****

FIGURE 3.4. Envelope shedding, fusion and viral entry of HIV-1 subtype C strain 1084i with hydrophobic patch mutations. (A-D). Envelope gp120 shedding assay. The supernatant and cell lysate fractions of cells transfected with WT or mutant 1084i and ^{35}S labeled were harvested and immunoprecipitation was performed on each separately using infected patient serum, samples were analyzed by SDS-PAGE. The association index was calculated as previously described [293]. A low association index is an indicator of a less stable envelope complex and an increased level of gp120 shedding. **(E-H).** Cell-cell fusion assay to detect WT and mutant envelope functionality. Each fusion sample was performed in quadruplicate wells, and the graph is representative of three separate experiments. Significance was determined by using the Student's T-test (*$p<0.05$). **(I-L)** The WT and mutant envelope proteins were used to generate single-round luciferase tagged viruses and titer was calculated using the RT activity assay. A normalized amount of virus was used to infect CD4/CCR5 expressing CF2-Th cells and luciferase activity was measured as an indicator of viral entry. This was performed with quadruplicate values and the means are representative of three separate experiments. Significance was determined using the Student's T-test (*$p<0.05$).

FIGURE 3.5

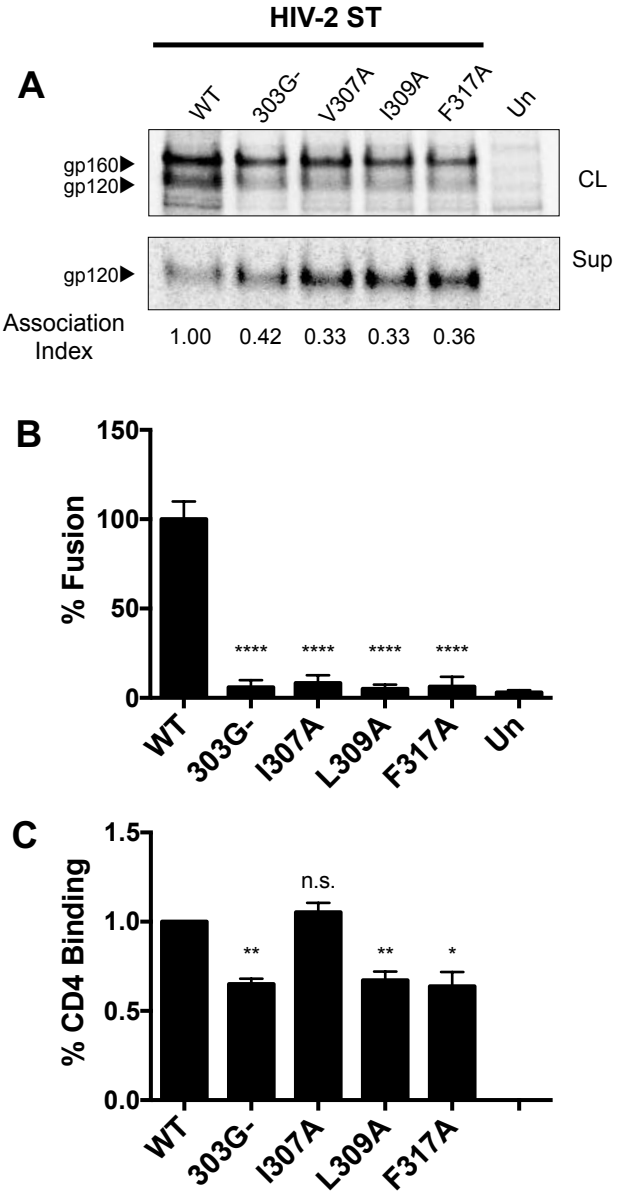
HIV-2 ST
A
WT
303G-
V307A
I309A
F317A
Un
gp160
gp120
CL
gp120
Sup
Association Index
1.00
0.42
0.33
0.33
0.36
B
% Fusion
150
100
50
0
****
****
****
****
WT
303G-
I307A
L309A
F317A
Un
C
% CD4 Binding
1.5
1.0
0.5
0.0
n.s.
**
**
*
WT
303G-
I307A
L309A
F317A

FIGURE 3.5. Envelope stability and cell-cell fusion of alanine substitution mutations to the hydrophobic patch of HIV-2 ST. (A). HIV-2 ST WT or mutant envelope constructs were expressed in 293T cells, labeled and harvested as described in the materials in methods. Immunoprecipitations of the cell lysates and supernatants were analyzed by SDS-PAGE. The gel shown is representative of three separate experiments, and the association index is the average value of those experiments. **(B).** The cell-cell fusion assay was performed using the WT or mutant envelopes and the results are included as a percentage of the WT luciferase values. This experiment was performed three separate times in quadruplicate. **(C).** Labeled WT or mutant gp120 protein was precipitated using CD4-Ig and the percent binding was calculated by comparing the quantified bands to an input immunoprecipitation with HIV-2 ST antiserum. The binding ratio of the WT was then compared to the binding ratio of the mutants. This figure is representative of three separate binding experiments. Significance was determined by the student's T-test (*$p<0.05$).

FIGURE 3.6

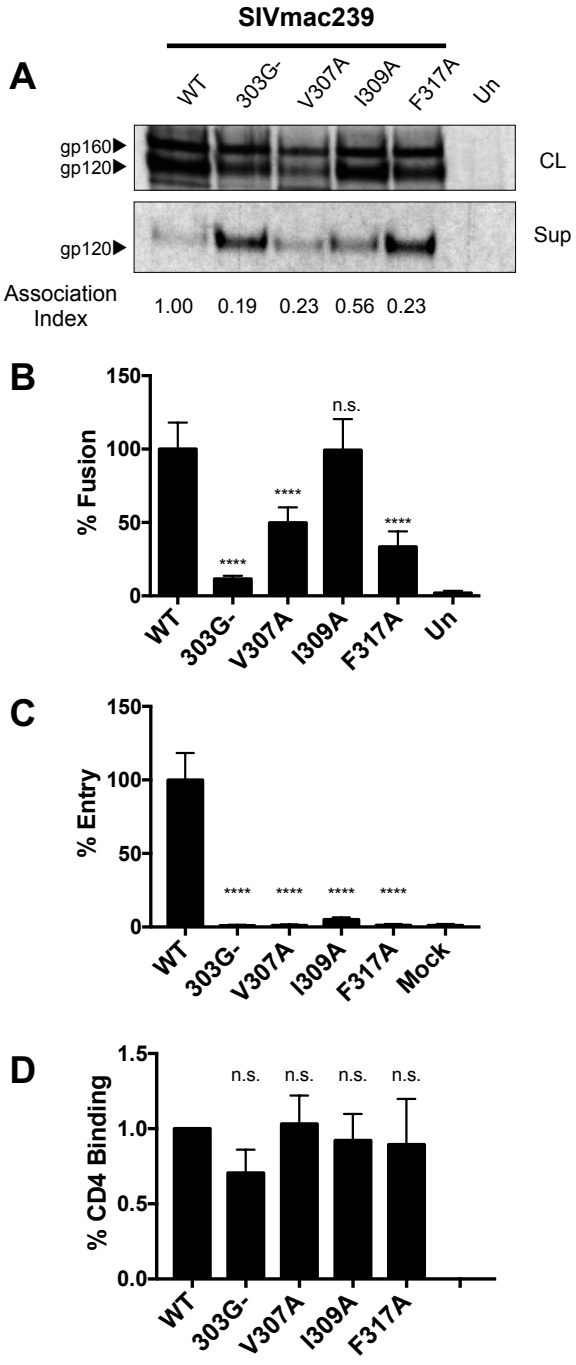
SIVmac239
A
WT
303G-
V307A
I309A
F317A
Un
gp160▶
gp120▶
CL
gp120▶
Sup
Association Index
1.00 0.19 0.23 0.56 0.23
B
% Fusion
150
100
50
0
n.s.
****
****
****
WT
303G-
V307A
I309A
F317A
Un
C
% Entry
150
100
50
0
****
****
****
****
WT
303G-
V307A
I309A
F317A
Mock
D
% CD4 Binding
1.5
1.0
0.5
0.0
n.s.
n.s.
n.s.
n.s.
WT
303G-
V307A
I309A
F317A

FIGURE 3.6. Envelope stability, viral entry and cell-cell fusion of SIVmac239 with alanine substitutions to the hydrophobic patch. (A) Envelope protein shedding assay was performed by ^{35}S labeling WT or mutant envelopes and harvesting the cell lysate and supernatant separately followed by immunoprecipitation with SIV patient serum. Association index values presented are the average of three separate experiments. **(B)** Fusion assay is representative of Tat-induced luciferase expression following fusion of envelope and tat protein expressing 293T cells with TZMbl. Mean cell-cell fusion values are representative of three separate experiments performed with quadruplicate replicate wells. **(C)** SIVmac239 pseudoviral particles were generated, and RT values were normalized prior to infection. Percent entry was determined by luciferase expression. Viral entry experiments were performed three times with quadruplicate values. **(D)** CD4 binding was determined by the amount of envelope protein precipitated by CD4-Ig compared to the amount of envelope protein bound by SIV infected monkey serum. Significance was determined by the student's T-test (*$p<0.05$).

FIGURE 3.7

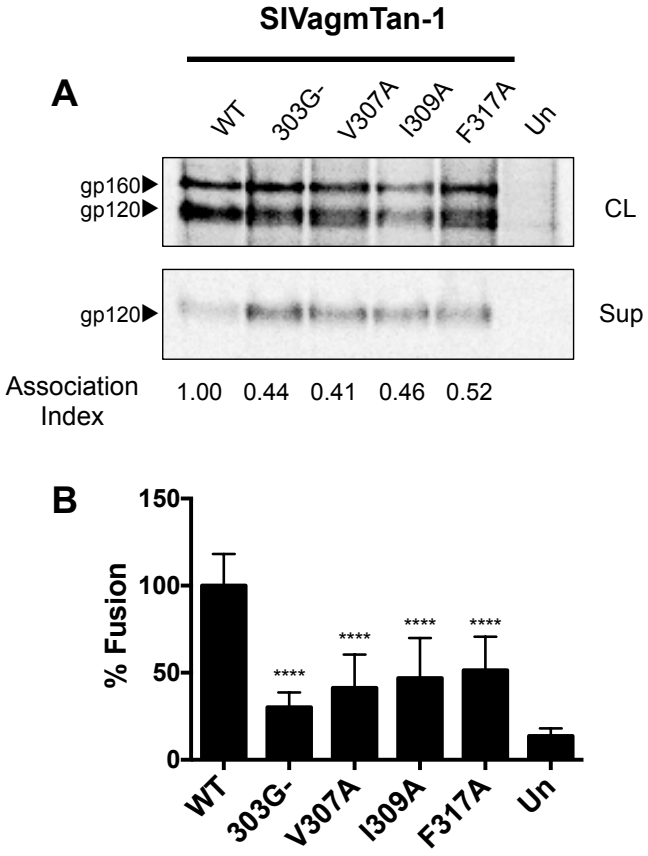
SIVagmTan-1
A
WT
303G-
V307A
I309A
F317A
Un
gp160
gp120
CL
gp120
Sup
Association Index
1.00 0.44 0.41 0.46 0.52
B
% Fusion
150
100
50
0
****
****
****
****
WT
303G-
V307A
I309A
F317A
Un

FIGURE 3.7. V3 loop hydrophobic patch mutations in the stability and functionality of the envelope complex of SIVagmTan-1. (A) ^{35}S labeled gp120 from SIVagmTan-1 precipitated in the supernatant fraction compared to the cell lysate fraction of the V3 mutant envelopes. Association index was calculated as previously described and is represented as the mean of three separate experiments. The immunoprecipitation was performed using a combination of SIV infected serum and HIV infected serum to increase pull-down. **(B)** Fusion assay with SIVagmTan-1 envelope was performed as previously described and is representative of three separate experiments performed, each with quadruplicate values. Significance was determined by the student's T-test (*$p<0.05$).

FIGURE 3.8

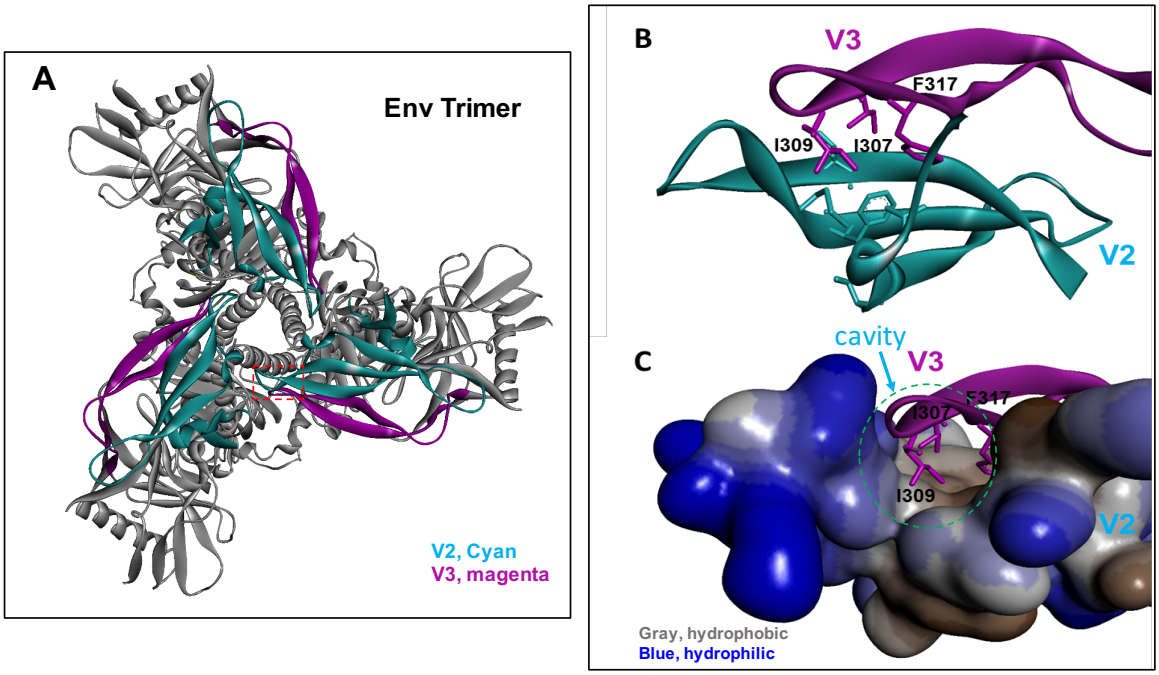
A
Env Trimer
V2, Cyan
V3, magenta
B
V3
F317
I309
I307
V2
C
cavity
V3
F317
I307
I309
V2
Gray, hydrophobic
Blue, hydrophilic

FIGURE 3.8. Molecular modeling of the HIV envelope trimer and the interaction between V3 and V2 loops. (A) HIV-1 trimer structure in the unbound (native) state (PDB: 4ZMJ). V3 loop is shown in magenta; V2 loop is shown in cyan. **(B)** Ribbon structure of the V2 and V3 loops showing their interaction with emphasis on the V3 hydrophobic patch residues at positions 307, 309 and 317. **(C)** V2 surface model showing the hydrophobic cavity (circled in cyan) and its interaction with the V3 loop hydrophobic patch. Gray, hydrophobic; Blue, hydrophilic.

FIGURE 3.9

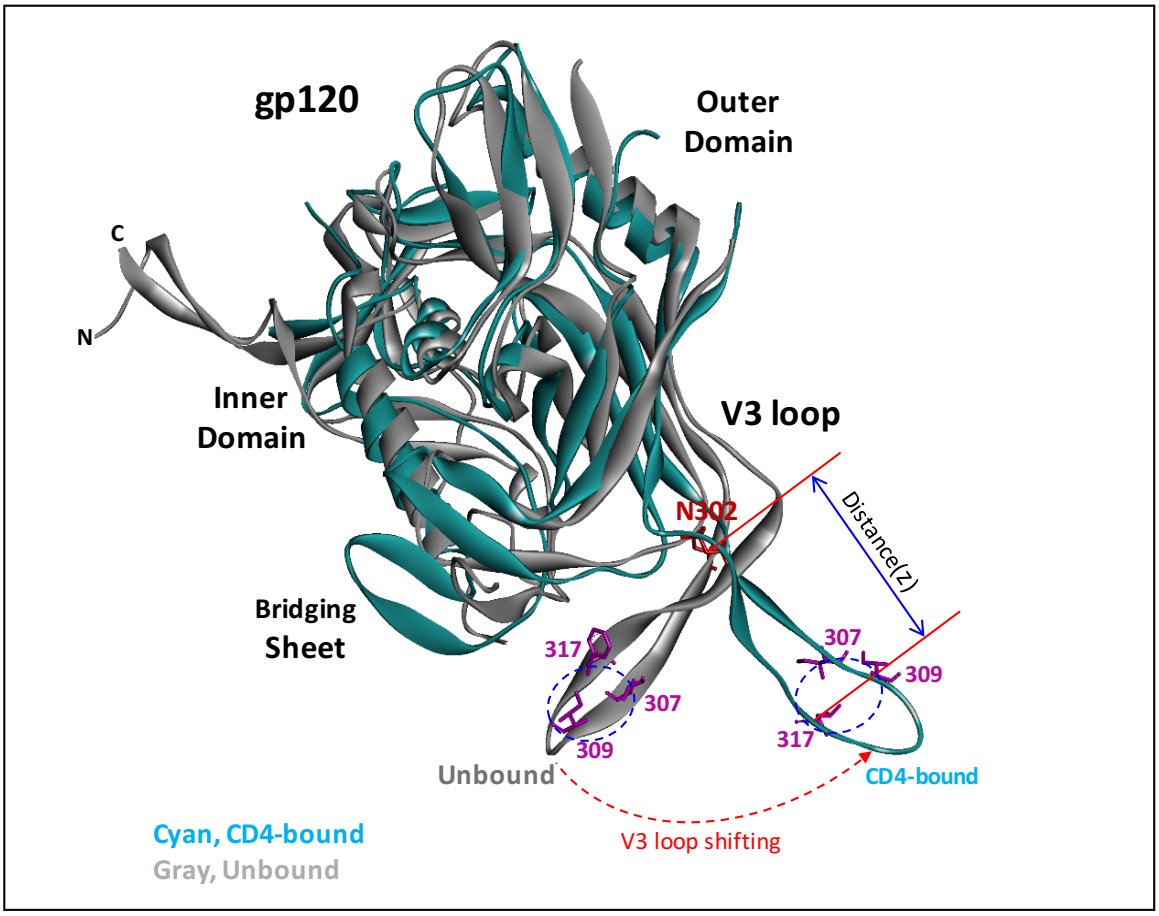
gp120
Outer Domain
C
N
Inner Domain
V3 loop
N302
Distance(z)
Bridging Sheet
317
307
309
307
309
317
Unbound
CD4-bound
Cyan, CD4-bound
Gray, Unbound
V3 loop shifting

FIGURE 3.9. Molecular modeling of the V3 loop shifting in the CD4 bound and unbound state of gp120. Model of superimposed gp120 structures in the CD4-unbound (gray) and CD-bound (cyan) states, showing the V3 loop shifting from CD4 unbound to the bound state. The hydrophobic patch residues of 307, 309 and 317 are labeled in red. The V3 base residue N302 is also indicated in red. Part of V2 loop in the unbound state is omitted for better presentation of the bridging sheet and V3 loop. The X-ray structures used for superimposition are PDB: 4ZMJ [1] for the unbound (native) state and PDB: 2QAD [317] for the CD4-bound state

TABLE 3.1

Env and Mutation	Association Index *a*	Cell-cell Fusion (Percent) *b*	Viral Entry (Percent) *c*	CD4 Ig Binding *d*
1084i				
WT	1.00	100.0	100.0	n.d.
303G-	0.50	5.4 *	3.2 *	n.d.
322G-	0.56	27.6 *	3.9 *	n.d.
M307A	0.77	88.7	16.9 *	n.d.
M307E	0.68	93.9	3.6 *	n.d.
M307L	1.15	89.8 *	573.5 *	n.d.
I309A	0.79	55.7 *	2.5 *	n.d.
I309E	0.55	5.0 *	0.7 *	n.d.
I309L	1.13	76.1 *	35.9 *	n.d.
F317A	0.86	58.6 *	2.7 *	n.d.
F317E	0.65	35.2 *	3.4 *	n.d.
F317L	1.13	80.7 *	65.2 *	n.d.
ZM249M-B10				
WT	1.00	100.0	100.0	1.00
303G-	0.15	2.5 *	0.6 *	0.98
I307A	0.08	12.2 *	0.6 *	0.97
I307E	0.05	6.2 *	0.6 *	n.d.
I307L	0.29	67.9 *	72.1 *	n.d.
I309A	0.08	2.4 *	0.7 *	1.03
I309E	0.13	2.8 *	0.5 *	n.d.
I309L	0.82	57.4 *	36.3 *	n.d.
F317A	0.13	5.3 *	0.6 *	0.90
F317E	0.21	2.7 *	0.5 *	n.d.
F317L	0.40	38.8 *	105.2	n.d.
SIVmac239				
WT	1.00	100.0	100.0	1.00
303G-	0.14	11.5 *	1.0 *	0.71
V307A	0.19	49.8 *	1.1 *	1.03
I309A	0.48	99.3	5.0 *	0.92
F317A	0.14	33.5 *	1.2 *	0.89
SIVagmTan1				
WT	1.00	100.0	n.d	n.d.
303G-	0.45	30.2 *	n.d	n.d.
V307A	0.41	41.3 *	n.d	n.d.
I309A	0.46	46.9 *	n.d	n.d.
F317A	0.53	51.4 *	n.d	n.d.
HIV-2 ST				
WT	1.00	100.0	n.d	1.00
303G-	0.43	6.0 *	n.d	0.65
I307A	0.33	8.3 *	n.d	1.03
L309A	0.34	5.1 *	n.d	0.67
F317A	0.37	6.3 *	n.d	0.63

TABLE 3.1. Phenotypes of primate lentivirus V3 loop mutants.

a. The association index is a measure of the envelope stability as indicated by the amount of radiolabeled WT or mutant gp120 shed into the media. The association index was calculated as follows: ([mutant gp120]$_{cell}$ x [wild-type gp120]$_{supernatant}$)/([mutant gp120]$_{supernatant}$ x [wild-type gp120]$_{cell)}$. The association index of the mutants is shown as a proportion of the WT envelope, which is set to 1.0.

b. Cell-cell fusion was assessed by coincubation of 293T cells expressing WT or mutant envelope and tat protein with TZMbl target cells in a ratio of 1:1, and luciferase was measured. Cell-cell fusion values of the respective WT strains were set to 100%, and mutants were expressed as a percentage of the WT value. *, indicates significant difference compared to WT Env, as determined by student's T-test ($p<0.05$).

c. Viral entry was measured as by infecting CD4+CCR5+ cells with replication deficient viruses expressing WT or mutant envelope proteins. Luciferase was measured as an indication of viral entry. Viral entry values of WT viruses were set to 100% and the values for the mutants are shown as a percentage of the WT value. *, indicates significant difference compared to WT, as determined by student's T-test ($p<0.05$).

d. CD4-Ig binding of WT gp120 was established by comparing the quantity of cell supernatant bound by CD4-Ig to the total gp120 immunoprecipitated by infected host serum; values for the mutants were compared to this value. n.d. not determined.

CHAPTER 4- CHARACTERIZATION OF THE TWIN CYSTEINE MOTIF IN THE V2-LOOP REGION OF gp120 IN PRIMATE LENTIVIRUSES

Data from this chapter were published in *Virology*, 2018. Volume 519, June 2018, Pages 180-189.

Contributing authors to this publication include:

Dane Bowder, Jesse Thompson, Kate Durst, Haley Hollingsead, Duoyi Hu, Wenzhong Wei, Shi-Hua Xiang

INTRODUCTION

Based on phylogeographic studies which consider sequence divergence and the rate of mutation in HIV, the major group of HIV-1 entered the human population via zoonotic transmission into humans by SIVcpz, the SIV species that is associated with chimpanzees [122]. Statistical analyses have indicated that this transmission event likely occurred in southern Cameroon in the early part of the twentieth century [153]. Group M HIV-1 accounts for the vast majority of infections with HIV-1, but there have been 4 transmissions of SIVs from chimpanzees or gorillas to humans accounting for group O, N and P infections [11], though they account for a very small minority of infections with HIV-1 globally [141, 142, 146, 148, 323]. HIV-2, a typically less pathogenic human lentivirus that is restricted primarily to West Africa represents a transmission of SIVsmm, the SIV species that primarily infects sooty mangabeys, into humans [163, 324]. This transmission of SIV from sooty mangabeys has occurred at least eight times, giving rise to eight clades of HIV-2 [164, 165]. Importantly, infection with HIV-2 typically does not result in AIDS, though immune suppression is seen, albeit to a much lower extent [170]. Patients with HIV-2 also show a marked increase in neutralizing antibodies compared to patients infected with HIV-1 [325-327].

It has been well-known for years that the envelope complex of HIV and SIVs exist as a trimer of heterodimers, but the structure of this complex was only recently solved in 2013. The SOSIP envelope trimer structure, based on the HIV-1 subtype A strain BG505, is a near native trimer that has an engineered

disulfide bridge added to the interface between gp120 and gp41 as well as an isoleucine to proline mutation in the heptad repeat 1 (HR1) region of gp41. The SOSIP trimer structure was solved through both X-ray crystallography as well as cryo-electron microscopy, resulting in a model of the HIV-1 envelope complex at an atomic level [1-3]. The development of these high-resolution structures has aided in the development of trimer based immunogens that have shown some levels of success and specificity by producing higher proportions of autologous trimer specific antibodies, though most of these antibodies do not neutralize heterologous tier 2 viruses [220]. Generating an immunogen that initiates neutralization of a broad spectrum of viruses, including tier 2 viruses is a continued challenge, and will be an important feature of a successful vaccine. Additionally, these structures have revealed specific epitopes for broadly neutralizing antibodies that were previously unknown [328-332]. Further understanding of the higher order structure of the envelope glycoprotein complex, and its interaction with broadly neutralizing antibodies will be crucial for the design of more effective envelope-based vaccine candidates to elicit broadly neutralizing antibodies against multiple strains.

The HIV envelope region is divided into five conserved regions and five variable regions [233, 333]. As their nomenclatures imply, the sequence of the conserved regions remains relatively well conserved in different HIV-1 strains, and the variable regions have more sequence variability. However, it is now known that the variable regions are more conserved than initially suggested, and consist of a relatively set number of residues and include conserved features and

residues [5, 6, 238, 318, 334]. The length of the variable loops is important for viral infection, as insertions or deletions to the loops can result in loss of envelope functional activity [238], and deletion of the loops altogether has a drastic effect on envelope function [239]. The advent of the envelope trimer structure has revealed that the variable loops are packaged into the apex of the HIV envelope trimer [2, 3, 275], which is consistent with reports that these loops contribute to the stability of the trimer complex [5, 6] and with similar near-atomic structure of the uncleaved, membrane-bound envelope trimer [211]. The V2 loop is of interest because of its role in envelope stability, and because it is an important immunogen in HIV infection. In the modestly successful RV144 Thai HIV vaccine trial, protection against infection with HIV correlated with the presence of V2 antibodies [247], which has stimulated significant interest in the potential of the V2 loop in the design of envelope based vaccines and anti-HIV drugs [219, 245, 248, 335].

Recently, our group identified a pair of cysteine residues present in the V2 loop of SIV species and in HIV-2 that are notably absent from the V2 loop of SIVcpz and HIV-1 species. We showed that the removal of this twin-cysteine motif (TCM) from SIVmac239 resulted in a major decrease in envelope stability, supporting our hypothesis that these cysteine residues form a disulfide bond that stabilizes the envelope trimer complex [6]. This study seeks to explore the twin-cysteine motif in more depth and to characterize its role in more SIV strains and in HIV-2. Additionally, this work seeks to determine the impact of adding the TCM back to HIV-1 strains.

RESULTS

The removal of the twin cysteine motif from HIV-2/ST results in a decrease in envelope stability.

HIV-2, like SIVmac, is originally derived from SIVsmm [11, 134, 336, 337], but unlike HIV-1, the twin-cysteine motif is preserved in HIV-2 through this cross-species transmission (Figure 4.1). Therefore, it is believed that the twin-cysteine motif in HIV-2 strains should play the same role as in SIVmac strains, such as SIVmac239. To further characterize the role of the twin cysteine motif in Primate Lentivirus (PLV) species where it is naturally present, the cysteines were removed by substitution with alanine residues both singly and doubly from the HIV-2 ST envelope glycoprotein gp120. A shedding assay was then performed on each of these removal mutants, where a clear increase in gp120 shedding (as indicated by less gp120 in the cell lysate fraction and more gp120 in the supernatant) and a decrease in the association index was seen (Figure 4.2.A). This decrease in envelope trimer stability correlated well with a decrease in envelope functionality, as the cell-cell fusion capability of these mutations dropped to the level of background (Figure 4.2.B). This indicates that the twin cysteines play a role in the stability of the envelope trimer in both SIVmac239 and HIV-2 ST and that this phenomenon is conserved.

Alteration to the twin cysteine motif changes envelope processing in SIVagm.

The SIVagm lineage is a separate lineage of PLV which is closely related to SIVsyk [338]. In this lineage, the twin-cysteine motif is also present in all SIVagm strains. For this experiment, SIVagmTan-1 was selected and the TCM was removed by substituting the cysteine residues with alanine residues, both individually and in tandem. A shedding assay was then performed with these mutations, and an unexpected phenotype arose. A clear accumulation of gp160 and reduction in gp120 in the cell lysate fraction occurred, along with a complete absence of gp120 found in the supernatant. It is important to note that SIVagmTan-1 already sheds gp120 to the media at a very low level, but in the mutations, gp120 was completely undetectable in the supernatant fraction, even after extended exposure (Figure 4.3.A). This result seems to indicate a change in envelope processing, as it appears that no gp160 is being cleaved into gp120. This phenotype also correlated to a major reduction in the processing index, further supporting the observation that processing was severely impaired in these mutants (Table 4.1). When the same mutants were used in a fusion assay, there was a significant reduction in cell-cell fusion, as was the case with both SIVmac239 and HIV-2 ST. This is not surprising, as cleavage of gp160 by furin proteases in the Golgi apparatus is required for its functional activity [42, 70, 339, 340]. The removal of the twin cysteines may directly affect the cleavage site or the accessibility for the protease. To further characterize this phenomenon, the envelope gene of an additional SIVagm strain, SIVagm155-4, was cloned into pCDNA3.1+ by RT-PCR for expression and mutagenesis. This strain also carries the twin cysteine motif. These residues were removed by site-directed

mutagenesis both individually and doubly. These mutations were then used to perform a cell-cell fusion assay, and the same phenomenon was observed as in SIVmac239, SIVagmTan-1 and HIV-2 ST. The single and double mutations all significantly reduced the fusion capability of these envelopes (Figure 4.3.C). The shedding experiment was not performed using agm155, because pull-down of this envelope glycoprotein was not achieved using the infected serums available, so a more specific SIVagm infected serum would be needed to complete this experiment.

Addition of the twin-cysteine motif to the V2 loop of HIV-1 species alters processing and reduces envelope functionality.

To explore the hypothesis that the twin cysteine motif was evolutionarily lost during the transition from SIV to HIV-1, the twin cysteine motif was added back to HIV-1 species. Because of a well-conserved proline residue in HIV-1 species, three types of addition mutants were generated. The first was a direct substitution of the twin cysteine motif into HIV-1 based on sequence alignment (Figure 4.4). This addition mutant substituted proline 183 and serine 191 of HIV-1 for cysteine residues, coined “SubCC”. Because proline is an amino acid that has important structural role in making turns in proteins, insertion mutations to both the left and right of proline 183 alongside an insertion mutation between serine 191 and tyrosine 192 were generated, coined “insL” and “insR” respectively (Figure 4.4). The shedding assay using the SubCC and insL mutants in both ADA and JRFL revealed another altered processing phenotype. As is the case in

SIVagmTan-1, these mutations retain very little gp120 in the cellular fraction, and uncleaved gp160 is found at higher levels than in WT. However, in the HIV-1 SubCC and insL mutations, shedding of gp120 appears to be increased over WT, suggesting that any gp120 that is cleaved is released into the supernatant immediately. The insR mutants did not exhibit this phenomenon to the same extent, so in the case of insR constructs, any changes in envelope shedding do not appear to be due to as drastic of an alteration in envelope processing. The processing index was calculated from the shedding gels for all of these mutations, and the SubCC and insL mutations resulted in a major reduction in this value, while the insR mutation showed a value closer to 1.0, indicating a lesser effect on envelope processing (Table 4.1). This was consistent in both ADA and JRFL HIV-1 strains (Figures 4.5 A-B). Interestingly, this change in envelope processing correlated well with changes in envelope functionality. In both ADA and JRFL, the SubCC and insL mutations significantly reduced the ability of these envelope proteins to initiate cell-cell fusion, but the insR did not, causing only a small, though significant, decrease in JRFL and no significant change in ADA (Figure 4.5 C-D). This fusion phenotype correlated with the ability of the mutant pseudoviruses to enter cells, as the JRFL insR mutation did not deviate significantly from WT, while the SubCC and insL mutations completely abolished viral entry. Unlike JRFL, ADA insR mutation did significantly reduce viral entry, though the decrease was not as severe as in the case of SubCC and insL (Figure 4.5 E-F). This reduction is likely because the unprocessed gp160 is

not being expressed on the cell surface, and any gp120 being transported to the cell surface appears to be shed into the media.

To attempt to characterize this altered processing in more detail, a pulse-chase assay was performed to track the envelope processing of these mutations over a period of 8 hours. This assay revealed that at no point after expression of the protein is gp120 ever retained in the cell lysate fraction, it is immediately shed. This assay did not show the gp160 retention in the cell lysate to the extreme level seen in the shedding assay, likely due to the limited amount of labeled envelope protein from a one hour pulse label, but did show the same gp120 shedding phenotype seen previously (Figure 4.6).

HIV-1 envelopes containing the V2 region from HIV-2 ST and SIVmac239 do not exhibit the altered processing phenotype.

Because SIV species and HIV-2 contain the twin cysteine residues, but do not exhibit the altered processing phenotype seen when the twin cysteines were added to HIV-1, we added the entire V2 twin cysteine region from SIVmac239 and HIV-2 ST to both HIV-1 ADA and JRFL species by overlapping PCR as described in the methods section and shown in figures 2.1 and 2.2. The HIV-1 envelopes with the mac239 V2 twin cysteine region did not show any notable differences in gp120 shedding, and only HIV-1 ADA with the V2 twin cysteine region from HIV-2 ST showed an increase in envelope shedding, JRFL ST V2 Swap showed no difference in envelope shedding. Importantly, in addition to no

change in envelope shedding levels, the processing of these envelopes was normal, and gp120 was observed in the cell lysate fraction. When these envelopes were used to test cell-cell fusion capabilities, largely, there were no significant changes in envelope fusion except for ADA SIVmac239 V2 Swap which demonstrated a minor significant increase in envelope fusion (Figure 4.7.B-C). To further test the phenotype of the V2 swap mutations, the envelopes were used to make single round luciferase virus and TZM-bl cells were infected. The SIVmac239 V2 Swap constructs did not alter the entry level of these viruses when compared to WT, but interestingly, the HIV-2 ST V2 Swap constructs caused a significant reduction in viral entry (Figure 4.7.D-E). It is unclear what causes this decrease in infectivity, but the slight increase in envelope shedding seen in ADA may be the cause.

Molecular modeling shows the formation of a disulfide bond by the TCM in SIVmac251

The twin-cysteine motif is an important feature in envelope trimer stabilization in primate lentiviruses, but we cannot know the exact mechanism of how this stabilization works since there is no structural information available for the twin-cysteine region in the V2 loop from any solved monomer or trimer based structures. For this reason, we have conducted structural modelling to predict the twin-cysteine structure in the model of SIVmac251 envelope trimer generated from the HIV-1 SOSIP structure (PDB: 4ZMJ) [1] (Figure 4.8). According to this

structural model, the twin-cysteines clearly formed a disulfide bond and generated a twin-cysteine loop (TCL), which interacts with the V1 loop. In SIVs, such as SIVmac251, the gp120 usually has a large V1 loop which can circle each subunit and covers the trimer apex along with the V2 loop (Figure 4.8, 4.9.A-B). The results from our gp120 experiments also imply that the twin-cysteines may act through the loop motif because the removal of single or double cysteines has the same effect. In this structural model, it appears that an intra-subunit disulfide bond can be formed, but not an inter-disulfide bond that connects subunits. It is evident that the intra-subunit disulfide-bond can help in stabilizing the V2 loop as well as the V1 loop (Figure 4.9.C and D). In addition, based on the interactions between the TCL and the V1 loop (Figure 4.9.E), it is likely that the TCL may have the potential role of regulating the V1 loop to affect trimer stability, which is noteworthy for further investigation.

DISCUSSION

In this study, we have shown that although the twin cysteine motif is conserved in primate lentiviruses, and that the effects of removing the twin cysteine motif from PLVs where the motif is present are species specific. Removal of the twin cysteine motif from HIV-2 ST resulted in a significant loss of envelope trimer stability which caused in turn, a reduction in envelope functionality. This result agreed with our previously published data, where a similar effect was observed in SIVmac239 [6]. Interestingly, removal of the twin

cysteine motif from SIVagmTan1 resulted in an altered processing state wherein gp160 is not appropriately cleaved into gp120 and presented on the cell surface. Because the envelope protein of this virus sheds at a very low level naturally, our evidence suggests that this does not cause a decrease in envelope stability, but instead a change in the processing of gp160 into gp120. However, the removal of this motif did result in a significant decrease in envelope functionality, as envelope fusion was reduced to background level in both single and double twin cysteine mutations. This indicates that the removal of these residues does affect the structure of the envelope protein in a way that renders it non-functional.

We also examined the effects of adding this twin cysteine motif back into the HIV-1 envelope, where it was lost along the path of evolution from SIVcpz to HIV-1. A simple substitution or insertion of these residues into the sequence did not stabilize the envelope trimer as we hypothesized. Instead, we saw a processing phenotype similar to SIVagmTan-1. In both the ADA and JRFL HIV-1 strains, the SubCC and insL mutations resulted in an accumulation of gp160 in the cell lysate fraction along with a complete lack of cellular gp120 and no major change in shed gp120. This indicated that gp120 is produced to a lesser extent than in WT HIV-1, and all the protein that was produced was shed, leaving none bound to the cell surface. Interestingly, the insR mutation resulted in a processing phenotype like WT, though there was a slight increase in gp160 retention in the cell lysate fraction. This WT processing phenotype correlated to a partial or complete rescue of envelope functionality. It is likely that the insR sequence of amino acids is closer in functionality to the TCM, when present in

SIV, though simply adding this motif to the HIV-1 sequence was not capable of stabilizing the envelope trimer to the level of SIV on its own. It is very likely that there are compensatory amino acid differences in other parts of the envelope sequence that work in cooperation with the TCM to stabilize the envelope trimers of SIV and HIV-1. The advent of high resolution structures of HIV-1 will help to dissect this interaction, though comparisons with an equally high resolution structure of an SIV or HIV-2 species where the motif is present will be essential to truly understanding how these structures differ. Our modeling does suggest that the TCM indeed forms an intra-disulfide bond that stabilizes the gp120 monomer, and can be used to identify additional trimer stabilizing features of the SIV and HIV-2 envelope proteins.

Processing changes were found consistently in these experiments, thought the exact effects that these mutations have on envelope processing remain unknown. Our pulse chase data reveal that at no point is soluble gp120 found in the cellular fraction, and that all gp120 that is produced is released into the media. Future experiments would be needed to elucidate the role that this region of the V2 loop is playing in HIV envelope processing, but it is possible that an alteration to the overall structure of the gp120 molecule is reducing the furin recognition of the envelope protein. It seems unlikely that these mutations directly alter the furin recognition site however, as the furin protease cleaves the gp160 at the junction between gp120 and gp41, which is distant from the V2 loop. It is more likely that an overall change in the flexibility or structure of the envelope protein could cause this site to be masked in some way, or cause a

level of allosteric interference. Antibody binding and structural experiments would be necessary to fully understand this process.

Identifying the role of features within the envelope sequence that contribute to the stability of the trimer is an important step in developing an effective trimer based envelope vaccine. Many important broadly neutralizing HIV-1 antibodies, like PG9, PG16, and PGT-145, target trimer specific quaternary epitopes so targeting the immune response to these epitopes only found when the envelope is in the unbound trimeric state is crucial [225, 243, 274, 328]. Previous vaccines that have attempted to do this have largely failed, though the advent of the SOSIP structure is likely to enhance the generation of whole trimer-based and trimer-specific epitope based vaccines.

In the future, it will be interesting to investigate the twin cysteine motif in more depth in the population of SIVcpz envelope proteins. The TCM is present in a few SIVcpz species, but is largely not present in most published sequences. It is also important to note that SIVcpz causes a much more drastic disease in its natural host than other SIVs, notably SIVagm and SIVsmm [126, 341, 342]. SIVcpz results in CD4+ T cell depletion and studies suggest that naturally infected chimpanzees have a much higher mortality rate than their uninfected counterparts [343]. Based on sequence analysis, it seems like the twin cysteine motif was lost within the SIVcpz population, so it will be interesting to see if the presence of this motif correlates with changes in pathogenesis. It will also be important in future studies to examine how the twin-cysteine motif is involved in stabilization of the envelope trimer since the structures of SIV or HIV-2 envelope

are yet to be solved. According to our modelling, the twin-cysteines form a disulfide-bond that can stabilize this region in the V2 loop. It is also anticipated that this twin-cysteine loop can interact with the V1 loop, which is able to exert its effect for stabilization on the apex of the trimer. It seems possible that the alteration of the twin-cysteine motif may destabilize the whole envelope complex by disrupting the interaction of these two loops, causing a massive destabilization of the envelope complex. Interestingly, the SIV V1 loop has been shown to be immunodominant, and a linear neutralizing epitope in the SIV V1 loop region has been identified [344, 345]. Additionally, the interactions between V1 and V2 appear to be significant for virus genetic diversity in nature [346]. The lengths of the major loops (V1, V2 and V3) of HIV-1 are indispensable for envelope structural integrity and virus entry [239], and their variation are also found to be involved in viral envelope function and disease progression [347].

It will also be important in future studies to examine how the twin cysteine motif interacts with surrounding residues. Results described in chapter 3 of this work explain that the V2 and V3 loops fit together in a *key-in-lock* conformation, with hydrophobic residues of both the V2 and V3 loops forming a hydrophobic core at the center of the gp120 molecule. It seems possible that the alteration of the twin cysteine motif may destabilize the whole envelope complex by disrupting the interaction of these two loops, and in turn potentially the core of the gp120 molecule, causing a massive destabilization of the envelope complex. Further mutational studies should be performed in this region to verify the structural importance of the hydrophobic core.

FIGURE 4.1

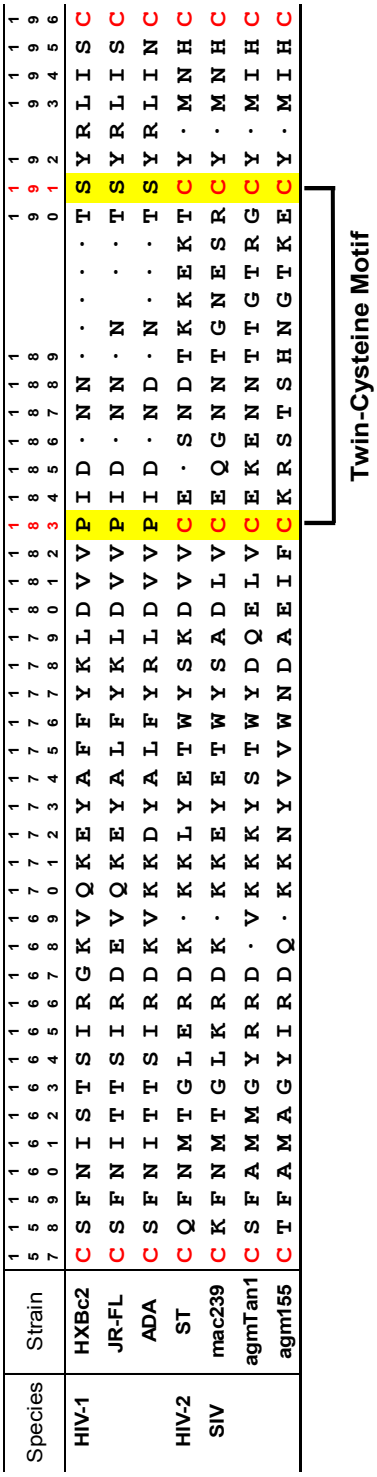
Species
Strain
HIV-1
HIV-2
SIV
HXBc2
JR-FL
ADA
ST
mac239
agmTan1
agm155
Twin-Cysteine Motif

FIGURE 4.1. Alignment of the primate lentiviral species V2 loops. This alignment shows the presence of the twin cysteine motif in SIV species and HIV-2 and the notable absence of the residues in HIV-1. The amino acids involved in the TCM are highlighted in yellow.

FIGURE 4.2

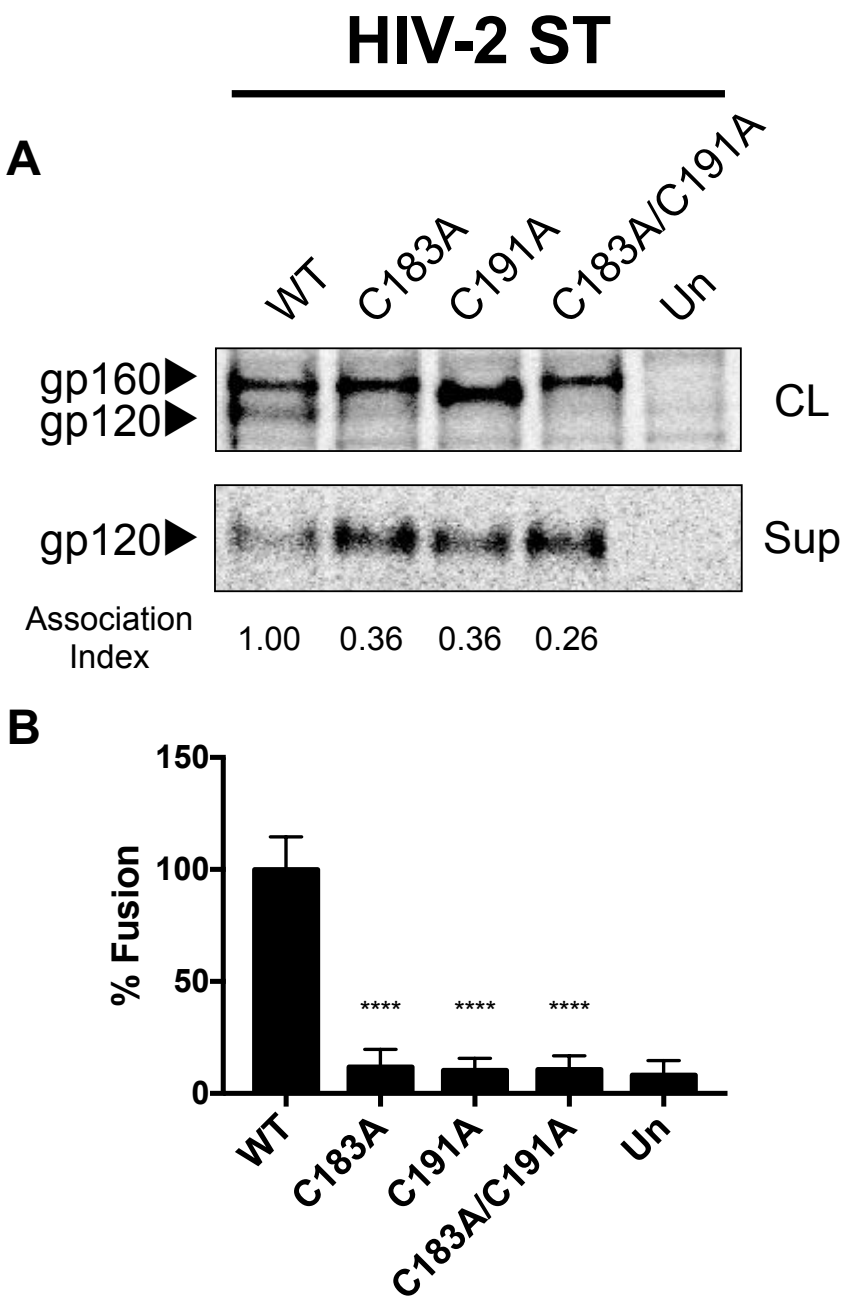
HIV-2 ST
A
WT
C183A
C191A
C183A/C191A
Un
gp160
gp120
CL
gp120
Sup
Association Index
1.00
0.36
0.36
0.26
B
% Fusion
150
100
50
0
****
****
****
WT
C183A
C191A
C183A/C191A
Un

FIGURE 4.2. Effects of the removal of twin cysteines from HIV-2 ST on envelope stability and envelope fusion. (A). The twin cysteine residues of HIV-2 ST were substituted for alanine residues by site directed mutagenesis individually (C183A, C191A) or in tandem (C183A/C191A) and a shedding assay were performed to determine the effect on envelope stability. The association index was calculated as previously described [293] and is shown below the shedding gel, values are representative of the average of three separate experiments. **(B).** The same mutants were used to perform a fusion assay (as described in the materials and methods as well as chapter 2), and the results are included as a percent of WT. Significance was determined by Student's T-test ($p \leq 0.05$).

FIGURE 4.3

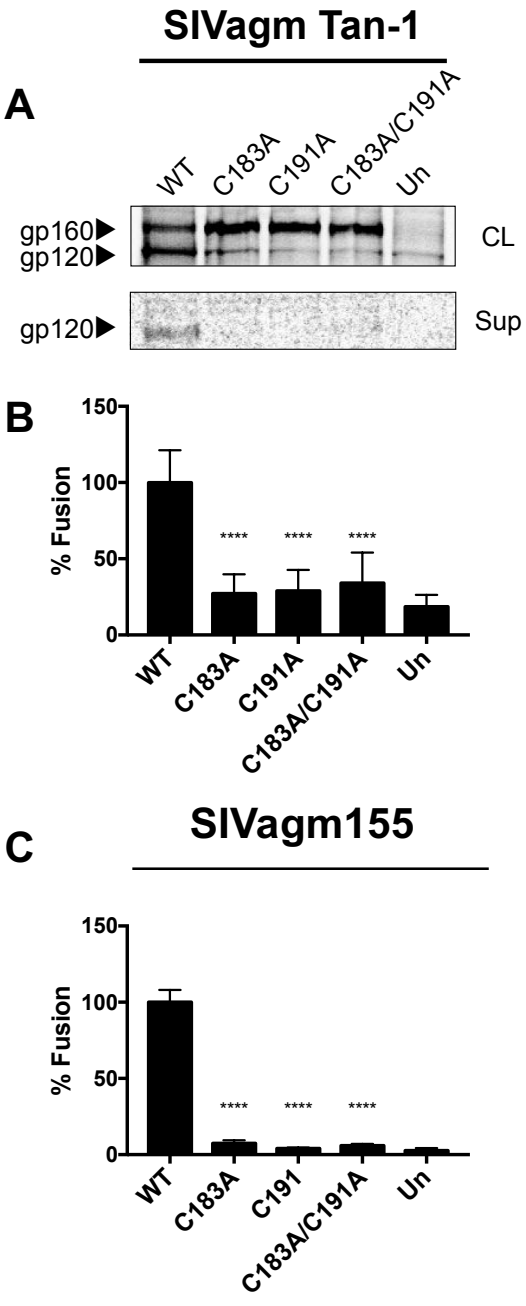
SIVagm Tan-1
A
WT
C183A
C191A
C183A/C191A
Un
gp160▶
gp120▶
CL
gp120▶
Sup
B
% Fusion
150
100
50
0
****
****
****
WT
C183A
C191A
C183A/C191A
Un
C
SIVagm155
% Fusion
150
100
50
0
****
****
****
WT
C183A
C191
C183A/C191A
Un

FIGURE 4.3. Effects of the removal of the twin cysteine motif from SIVagmTan-1 and SIVagm155 on envelope stability and envelope fusion. (A). Alanine substitutions were made to the twin cysteine motif individually (C183A, C191A) or in duplicate (C183A/C191A) and the mutants were used to express protein followed by metabolic labeling with ^{35}S and a shedding assay. Because gp120 was not detectable and is required to determine the association index, it was not calculated. The image shown is representative of three separate experiments. (B). The same SIVagmTan-1 mutations were used to perform a fusion assay. (C). Fusion assay was also performed on twin-cysteine removal mutants in an additional SIVagm strain, agm155. Experiments were performed three separate times in quadruplicate. Significance was determined by Student's t-test ($p \leq 0.05$).

FIGURE 4.4

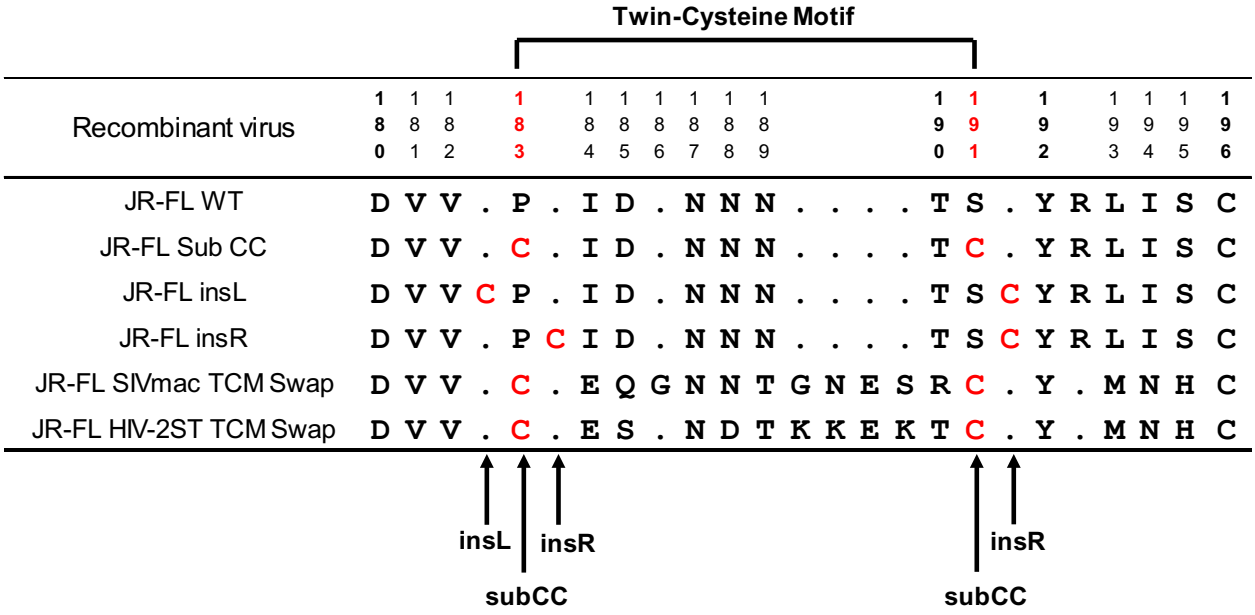
Twin-Cysteine Motif
Recombinant virus
180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196
JR-FL WT D V V . P . I D . N N N . . . . T S . Y R L I S C
JR-FL Sub CC D V V . C . I D . N N N . . . . T C . Y R L I S C
JR-FL insL D V V C P . I D . N N N . . . . T S C Y R L I S C
JR-FL insR D V V . P C I D . N N N . . . . T S C Y R L I S C
JR-FL SIVmac TCM Swap D V V . C . E Q G N N T G N E S R C . Y . M N H C
JR-FL HIV-2ST TCM Swap D V V . C . E S . N D T K K E K T C . Y . M N H C
insL
subCC
insR
insR
subCC

FIGURE 4.4. Diagram of the twin cysteine addition mutants in HIV-1. Alignment demonstrates the positioning of the twin cysteine mutations addition mutants in JR-FL and ADA (not shown). Numbering is based on the standard HXBc2 numbering system. SubCC, direct substitution of P183 and S191 with cysteine residues; insL, insertion of the cysteines to the right of 183 proline (P183) residue and between S191 and Y192; insR, insertion of cysteines to the left side of 183 proline (P) and between S191 and Y192; TCM swap, replaced the whole corresponding region from TCM element of other species.

FIGURE 4.5

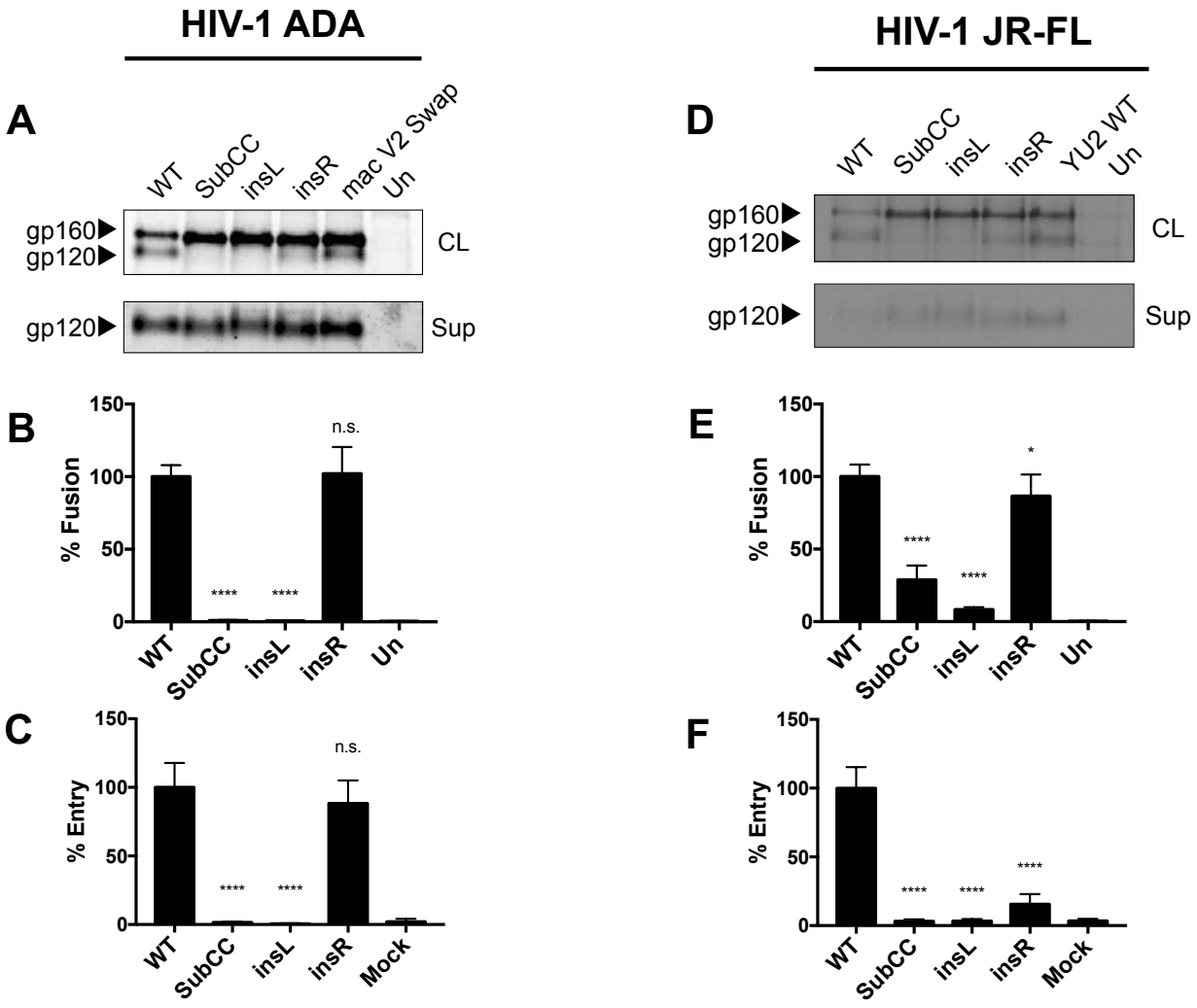
HIV-1 ADA
HIV-1 JR-FL
A
WT
SubCC
insL
insR
mac V2 Swap
Un
gp160
gp120
CL
Sup
D
YU2 WT
B
% Fusion
n.s.
****
E
*
C
% Entry
Mock
F

FIGURE 4.5. Addition of the twin-cysteine motif to HIV-1 strains in envelope processing and envelope function. (A-B). Twin cysteine mutations were added to the envelopes of HIV-1 strains ADA and JR-FL and the mutants were expressed in 293T cells followed by ^{35}S labeling and shedding assay. Because gp120 was undetectable in the CL fractions of most of the mutants, the association index was not calculated. Each gel is representative of three separate experiments. **(C-D).** Readdition mutations were used to conduct a fusion assay. Fusion of the mutants is shown as a percentage of the WT. Effector cells transfected with Tat only were used as a negative control for this experiment. **(E-F).** Mutants were used to produce single-round luciferase tagged viruses which were quantified by RT assay. Normalized quantities of virus based on RT value were used to infect TZM-bl cells. Cells were mock infected with media alone as a negative control for viral entry. Mutants are shown as a percentage of the WT virus. In the case of the fusion and entry assays, significance was determined using the Student's t-test ($p \leq 0.05$).

FIGURE 4.6

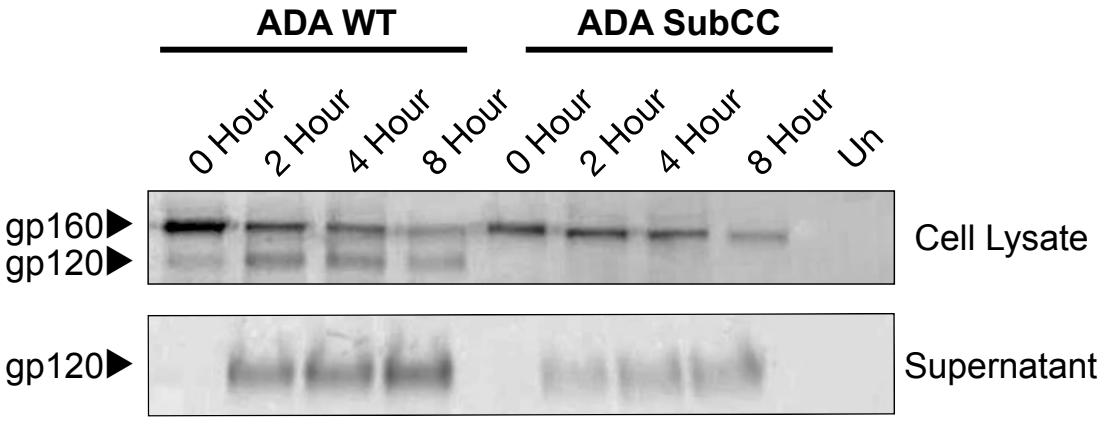
ADA WT
ADA SubCC
0 Hour
2 Hour
4 Hour
8 Hour
0 Hour
2 Hour
4 Hour
8 Hour
Un
gp160
gp120
Cell Lysate
gp120
Supernatant

FIGURE 4.6. Processing of gp120 in ADA twin cysteine substitution mutants. Pulse chase assay was performed to characterize the processing error seen in readdition mutants. Cells were transfected with ADA WT or ADA SubCC and the following day were pulse labeled with ^{35}S-cys-met containing media for 1 hour followed by replenishment with unlabeled complete media. After the pulse, the supernatants and cell lysate were harvested from each well separately at 0, 2, 4 and 8 hours' post followed by IP with HIV infected patient serum.

FIGURE 4.7

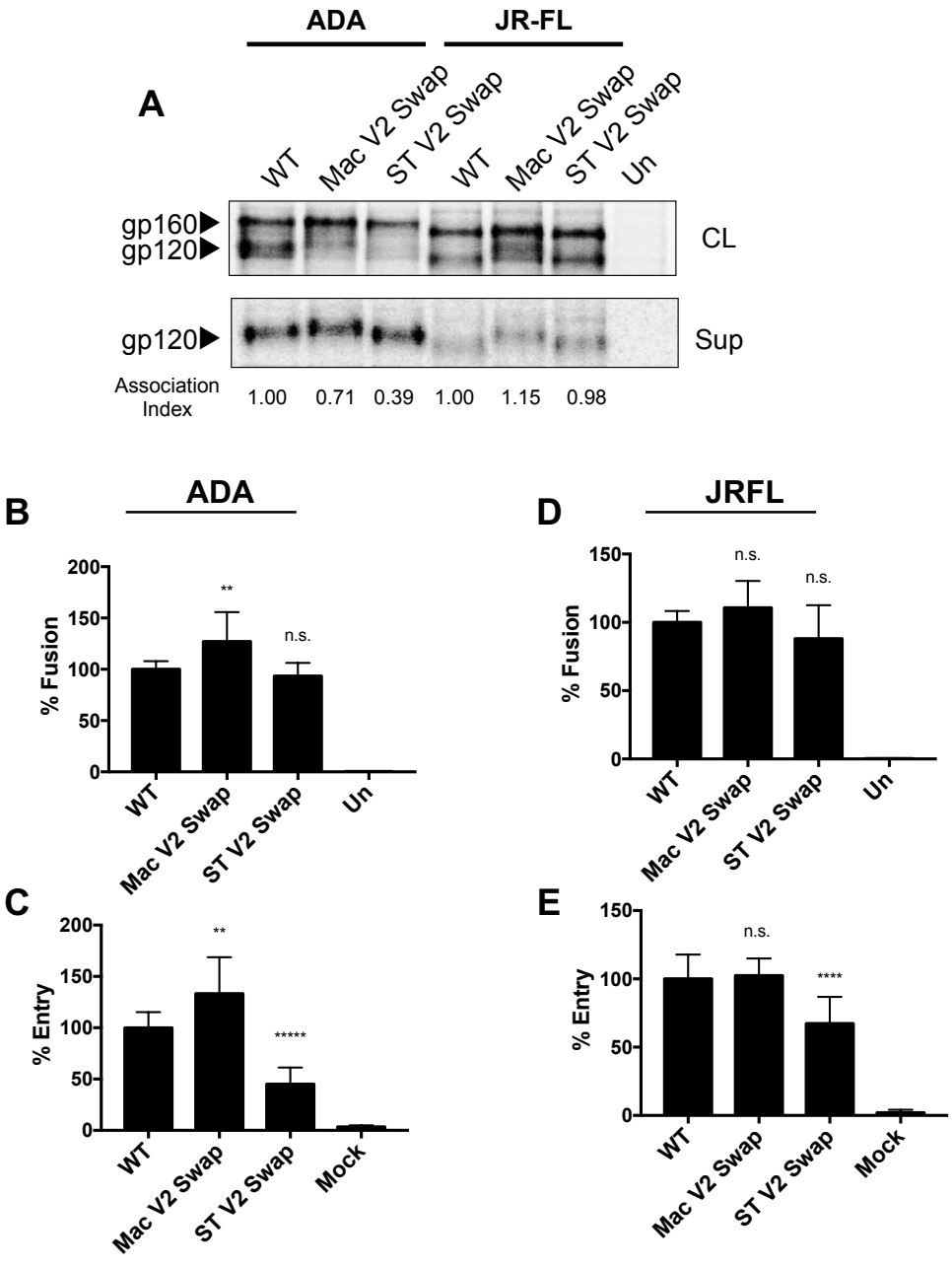
A
ADA
JR-FL
WT
Mac V2 Swap
ST V2 Swap
WT
Mac V2 Swap
ST V2 Swap
Un
gp160
gp120
CL
gp120
Sup
Association Index
1.00 0.71 0.39 1.00 1.15 0.98
B
ADA
% Fusion
200
150
100
50
0
**
n.s.
WT
Mac V2 Swap
ST V2 Swap
Un
C
% Entry
200
150
100
50
0
**
*****
WT
Mac V2 Swap
ST V2 Swap
Mock
D
JRFL
% Fusion
150
100
50
0
n.s.
n.s.
WT
Mac V2 Swap
ST V2 Swap
Un
E
% Entry
150
100
50
0
n.s.
****
WT
Mac V2 Swap
ST V2 Swap
Mock

FIGURE 4.7. Envelope gp120 shedding, fusion and viral infectivity in HIV-1 SIVmac239 and HIV-2 ST V2 Swap Mutants. (A). The shedding of the V2 swap mutants in both ADA and JRFL were investigated. The envelope constructs were expressed in 293T cells in a 6 well plate, were ^{35}S labeled and the supernatants and cell lysates from each well was harvested separately and immunoprecipitation with infected patient serum was performed on each fraction followed by analysis via SDS-PAGE. The association index is included below. **(B-C).** Following shedding analysis, the same envelope constructs (B- ADA, C-JRFL) were investigated for their ability to initiate cell-cell fusion. Mutant cell-cell fusion is expressed as a percentage of WT fusion. Cells expressing only Tat served as a negative control in this experiment. **(D-E).** V2 swap mutants were also used to generate single round, luciferase tagged HIV-1, were normalized by RT assay and used to infect TZM-bl cells. V2 swap envelope viral entry is expressed as a percentage of WT. Cells were mock infected with non-virus containing DMEM as a negative control in these experiments. In both the fusion and infectivity assay Student's T-test was used to determine statistical significance ($p \leq 0.05$).

FIGURE 4.8

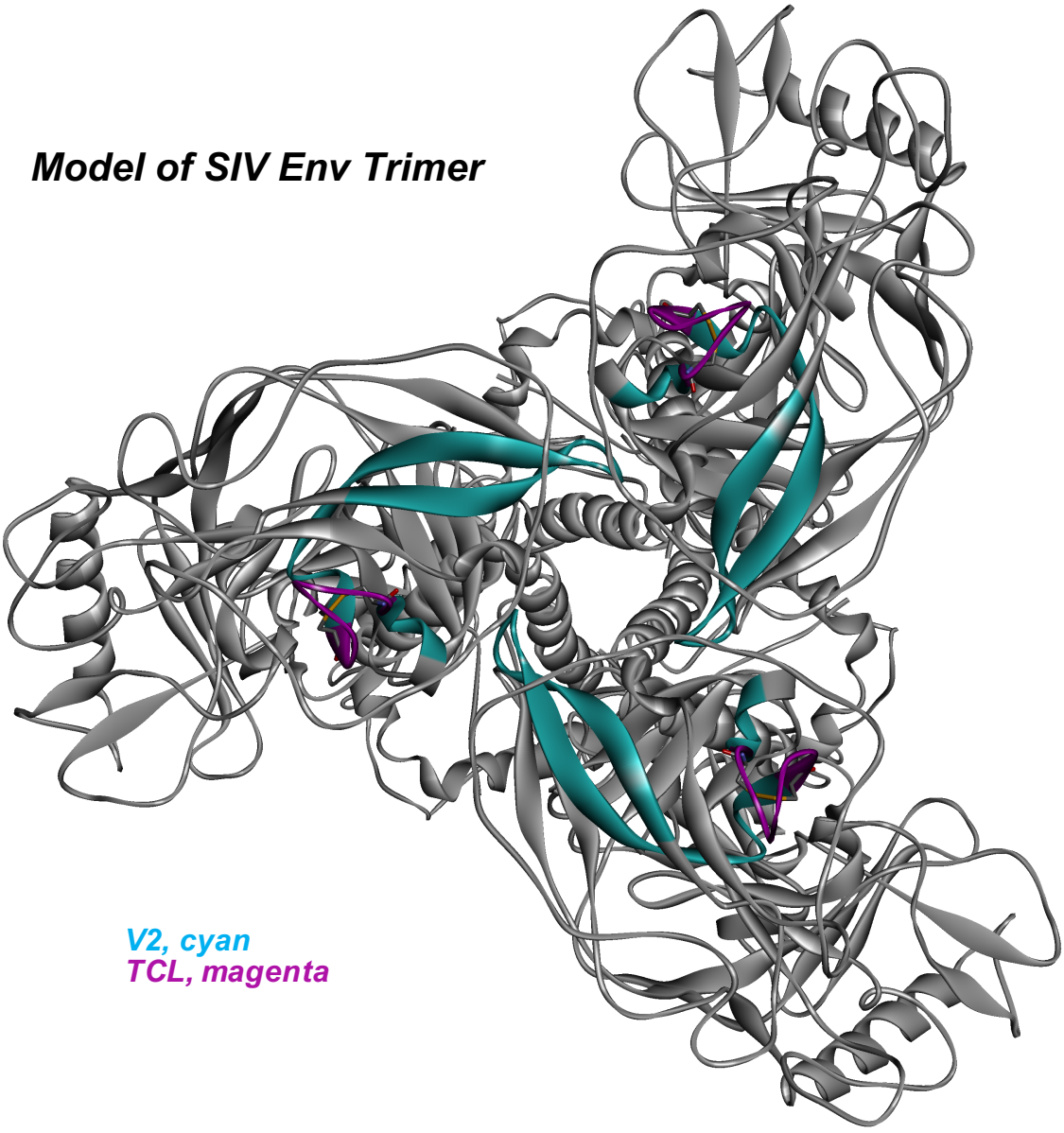

FIGURE 4.8. Molecular modeling of the SIV Envelope Trimer. Model of SIV envelope trimer from SIVmac251 (sequence accession number: M19499) based on the HIV-1 SOSIP structure (PDB: 4ZMJ) [1]. The V2 loop is colored in cyan and the twin-cysteine loop (TCL) in magenta.

FIGURE 4.9

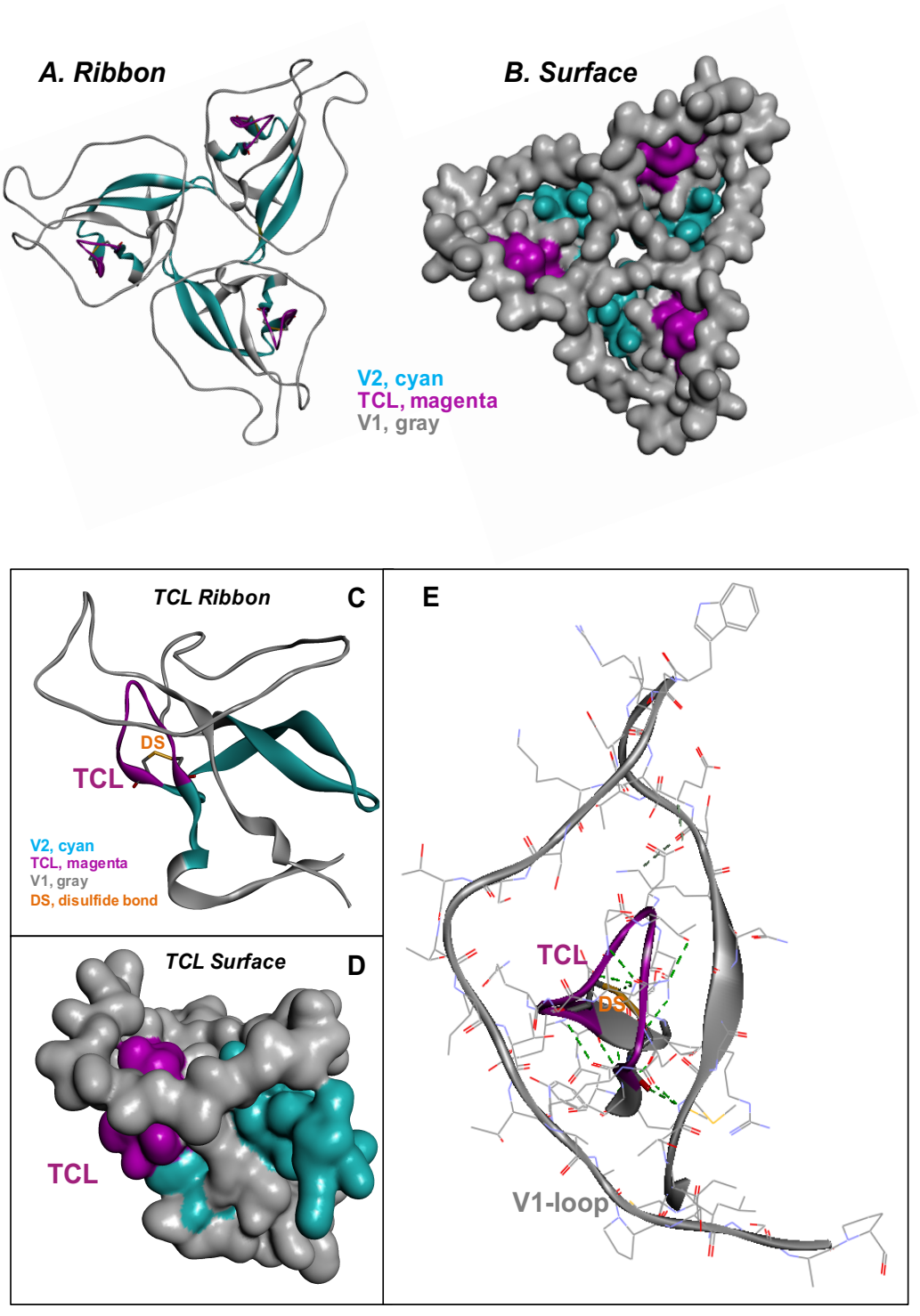
A. Ribbon
B. Surface
V2, cyan
TCL, magenta
V1, gray
TCL Ribbon
C
DS
TCL
V2, cyan
TCL, magenta
V1, gray
DS, disulfide bond
TCL Surface
D
TCL
E
TCL
DS
V1-loop

FIGURE 4.9. Molecular modeling of the twin-cysteine loop in SIVmac251 envelope. (A-B) The relationships of V1, V2 and TCM in the envelope trimer apex based on the model of the SIV envelope trimer of SIVmac251 depicted in figure 9. **(C-E)** The twin-cysteine loop (TCL), disulfide-bond and the interactions between TCM and V1 loop. The V2 loop is depicted in cyan, the V1 loop is depicted in gray. The TCL is shown in magenta and the disulfide bond formed by the TCM is shown in orange.

TABLE 4.1

	Association Index *a*	Processing Index *b*	Cell-cell Fusion (%) *c*	Infectivity (%) *d*
HIV-2 ST				
WT	1	1	100	n.d.
C183A	0.36	0.63	11.82	n.d.
C191A	0.36	0.36	10.26	n.d.
C183A/C191A	0.26	0.69	10.73	n.d.
SIVagmTan1				
WT	1	1	100	n.d.
C183A	0.89	0.18	27.24	n.d.
C191A	0.73	0.24	28.96	n.d.
C183A/C191A	0.69	0.19	34.11	n.d.
SIVagm155				
WT	n.d.	n.d.	100	n.d.
C183A	n.d.	n.d.	7.45	n.d.
C191A	n.d.	n.d.	4.06	n.d.
C183A/C191A	n.d.	n.d.	5.98	n.d.
HIV-1 ADA				
WT	1	1	100	100
SubCC	0.4	0.38	0.96	3.21
insL	0.31	0.42	0.66	3.27
insR	0.67	0.65	102.08	15.62
Mac V2 Swap	0.7	0.58	127.11	133.27
ST V2 Swap	0.38	0.63	93.44	45
HIV-1 JRFL				
WT	1	1	100	100
SubCC	0.1	0.36	28.85	1.58
insL	0.21	0.47	8.28	0.61
insR	0.49	0.87	86.56	88.28
Mac V2 Swap	1.16	1.12	110.63	102.44
ST V2 Swap	0.94	1.02	88	67.27

TABLE 4.1. Phenotypes of primate lentivirus twin cysteine mutants.

a. gp120 shedding assays were performed by immunoprecipitating 35S labeled gp120 from the cell lysate and supernatant fractions of WT or mutant env expressing 293T cells. Bands for gp120 in the cell lysate and supernatant were quantified and the association index was calculated using the equation ([mutant $gp120_{cell}$ * WT $gp120_{supernatant}$] / [mutant $gp120_{supernatant}$ * WT $gp120_{cell}$]). A lower association index indicates a lower level of gp120 association with the cells. Each association index value is the representative average of three separate experiments.

b. gp120 processing index was calculated from gp120 shedding gels. 293T cells expressing 35S labeled WT or mutant env protein were analyzed by SDS-PAGE and the bands for gp160 in the cells, gp120 in the cells and gp120 in the supernatant were quantified. The processing index was then calculated using these values using the equation, ([total $gp120_{mutant}$ * $gp160_{WT}$] / [total $gp120_{WT}$ * $gp160_{mutant}$]). Each processing index value is the representative average of three separate experiments.

c. Cell-cell fusion was calculated by cotransfecting 293T effector cells to express WT or mutant env and HIV-1 tat. These cells were counted and coincubated with target TZMbl cells at a ratio of 1:1. TZMbl cells stably express CD4 and CCR5 as well as the gene for luciferase under a Tat-inducible promoter. Following coincubation, luciferase values were determined. WT cells were set to 100% fusion, and the mutants were expressed as a percentage. Each value is the mean of three separate experiments performed in triplicate.

d. Viral infectivity values were determined by producing single round, replication deficient viruses expressing WT or mutant envelope proteins. These viruses were quantified based on reverse transcriptase activity, and normalized volumes were used to transduce TZMbl cells. Luciferase values were measured, and the WT virus values were set to 100%, and mutants were normalized to these values. Each value represented in the table is the representative mean of three separate experiments performed with quadruplicate replicate values.

CHAPTER 5- CONCLUDING REMARKS AND FUTURE DIRECTIONS

Despite tireless efforts and decades of research, HIV continues to be a global health crisis, with 36.7 million people living with the virus, 1.8 million new infections and 1.0 million AIDS related deaths in 2016. Due to the use of antiretroviral therapy (ART) and education about HIV/AIDS, new infections and death rates are declining, but the effects of this virus are still staggering [8]. Despite the advent of more effective antiretroviral drugs and effective pre-exposure prophylactic drugs, an effective vaccine is critical to preventing new infections and to stopping the HIV epidemic. The HIV envelope protein, as the primary antigen of the virus is a prime candidate for the development of a vaccine, and has been the subject of extensive vaccine trials and studies, though these trials have shown limited success [247, 273, 278]. One of the targets of many recent vaccine studies is the generation of broadly neutralizing antibodies, many of which target quaternary, trimer-specific epitopes [244, 274, 277, 330]. An important idea that currently permeates the literature is that soluble trimeric envelope immunogens will elicit the production of these antibodies more effectively than monomeric envelope [273, 278]. This is supported by the fact that envelope gp120 oligomers induce a more potent immune response than monomers [271, 276, 348].To this end, several groups have attempted to use the SOSIP soluble envelope trimer as an immunogen and preliminary studies have resulted in the promising production of antibodies, some of which are trimer specific, though a consistent broadly neutralizing heterologous response is yet to be seen, so more research must be done [219, 220, 257]. Advancements in structural biology have yielded a near atomic structure of a soluble envelope

trimer, which has furthered our understanding of the assembly and structure of this trimer, including confirmation that the variable loops are packaged into the apex of the trimer as a part of the trimer association domain [1-3, 275]. Trimer based immunogens and trimer-specific epitope based vaccines are promising in the generation of a sterilizing vaccine, and will likely include features of the trimer association domain, including the V2 and V3 loops [4]. To successfully generate these immunogens, it will be crucial to understand the features of the envelope protein that actively contribute to the stability of the envelope trimer complex.

The aim of this work was to further investigate the role of the variable regions V3 and V2 in envelope trimer stability by focusing on two motifs: the hydrophobic patch in the V3 loop, which is conserved in all PLVs; and the Twin Cysteine motif (TCM) that is found in the V2 loop of SIVs and HIV-2, but is notably absent from HIV-1. We sought to investigate the roles of these specific amino acids by performing mutational studies that remove these motifs when present, and introduce them when absent, followed by observation of envelope stability and functionality. Our hypothesis that the V3 loop hydrophobic patch is conserved in function was largely supported by our results. Our hypothesis that the twin cysteine motif in the V2 loop contributes to envelope trimer stability was supported and revealed an additional crucial role for this motif in appropriate envelope processing. We also sought to add this motif back to HIV-1 to determine whether this motif could stabilize the envelope trimer in HIV-1 and again found that this region is critical for envelope functionality and processing. Altogether, the molecular data generated in these studies strongly supports the

structural models published. This demonstrates that the variable loops, which were often disregarded in early structural studies, are indeed important structural components of the PLV envelope trimer.

The V3 loop has been shown previously to be involved in trimer stability in HIV-1 subtype B species [5]. In chapter 3, we expanded on this knowledge, demonstrating the same results in subtype C HIV-1, but also in an HIV-2 strain as well as in SIVs from two species. We show that the length of V3 loop is important across PLV species, and insertions to the V3 loop cause a massive destabilization of the envelope complex resulting in high amounts of shedding of gp120 into the media. The hydrophobic patch of the V3 loop is conserved in sequence in subtype C HIV-1, HIV-2 and SIVs, and we have shown that it is also conserved in function. Disruption of the residues involved in this motif by site directed mutagenesis causes a decrease in envelope stability, evidenced by increased shedding of gp120 into the media. This decrease in envelope trimer stability correlated with a decrease in envelope functionality evidenced by cell fusion and viral entry. The hydrophobicity and space filling nature of the residues is especially important. Substitution of the residues in this patch with hydrophilic and small neutral amino acids resulted in the most drastic decrease in envelope stability, while substitution with alternative hydrophobic residues partially rescued this reduced stability. Altogether, these results indicate that the hydrophobic patch has been conserved through the evolutionary history of HIV. The hydrophobic patch represents an important conserved motif within the variable loop, in addition to the previously characterized conserved features [234]. Part of

this dissertation is the exploration of a phenomenon present in HIV-1 subtype B and subtype C. Because of the differences in pathogenesis, transmission and antigenic capabilities shown in subtype C, and because of the high prevalence of this subtype, it is paramount that more studies explore the differences in subtype-C viruses. A successful anti-HIV vaccine or novel viral inhibitor will need to be functional against all viral subtypes and strains to be truly effective.

We propose that the hydrophobic patch plays an important role in trimer stability, but also in functionality in both the CD4 bound and unbound conformations. Based on modeling, in the unbound conformation, the V3 loop hydrophobic patch is clearly positioned within a hydrophobic cavity in the V2 loop, which cradles and supports the V3 loop. This interaction helps to stabilize the envelope trimer (Figure 3.8). When insertions are made to the length of the V3 loop, this initiates a crash of the V3 loop with the V2 loop, where the increased length of the V3 loop presses against the V2 loop, destabilizing the trimer. Also, because this interaction is reliant on hydrophobic interactions, mutation of these hydrophobic residues to hydrophilic or neutral amino acids reduces this interaction, also destabilizing the trimer. In the CD4 bound conformation, the V3 loop is rearranged and exposed outwardly where it participates in forming the CCR5 binding site. Contact with the V3 loop is dependent on a specific V3 loop length, which is supported by the fact that the length of the V3 loop is very well conserved at approximately 35 residues. The V3 loop has been demonstrated to contact CCR5 via residues both in the base and crown, and indicate that the length of the V3 loop is crucial to this function

[349-351]. The exact nature of the contacts between the V3 loop and CCR5 are still unknown since a complex structure of gp120 bound to CCR5 or CXCR4 is unavailable, so additional structural studies will be required to fully characterize this interaction.

Chapter 4 describes the role of the TCM in envelope trimer stability. Our data clearly demonstrates that this motif is not evolutionarily conserved in HIV-1 from SIV (Figure 4.1). Disruption of the TCM in PLVs in the SIVsm lineage clearly cause a reduction in envelope trimer stability that correlated with a decreased functionality, which is demonstrated by a reduction in cell-cell fusion. Interestingly, in SIVagm species, this motif is conserved, but the removal of this motif results in an unexpected change in envelope processing, as indicated by a stronger presence of uncleaved gp160 in the cellular fraction, with no detectable gp120 protein in either the cells or the supernatant. Similarly, the addition of the twin cysteine motif back to HIV-1 species also resulted in a similar processing phenotype, though it appears that any processed gp120 is immediately shed into the media. As expected based on this envelope stability profile and lack of gp120 at the cell surface, cell-cell fusion and viral entry were abolished by these mutations. The exact mechanism by which processing is altered is unknown, but is unexpected as furin proteases cleave gp160 at a conserved Lys/Arg-X-Lys/Arg-Arg motif located at the junction of gp120 and gp41, approximately amino acid positions 516 [42, 339]. This process could be affected either directly via forming a barrier to the cleavage site, or indirectly by blocking the furin protease accessibility. It should be investigated further through additional

processing studies and additional mutational configurations. The release of high resolution structures of the trimer, and the experimental data that we gathered suggest that a bond is indeed formed by the TCM within the envelope monomer. The TCM appears to contact the V1 loop to stabilize the envelope trimer structure. Removal of this bond from some PLV species may also change the conformation of envelope and leading to masking of the cleavage site, thus altering the processing of the envelope protein in SIVsyk/SIVagm lineage. Additional experiments should be performed in SIVcpz species, as this is the only PLV species where the TCM is present in some strains and absent in others. It is likely that the TCM was lost in the process of evolution from SIVcpz to HIV-1, and this would be an interesting problem to explore further. Regardless of the phenotype, we have shown clearly that the TCM is critical for the stability and functionality of the envelope protein of PLVs, though its exact role differs within PLV lineages.

Structural analysis of the HIV envelope trimer has shown that the V2 and V3 loops are packaged together in apex of the trimer, and that the V3 loop is cradled by the V1/V2 loop, forming the trimer association domain (Figure 3.7) [1-3, 211]. Despite contributing to the formation of the trimer association domain, there have been relatively few studies in the literature that demonstrate the V2 and V3 loops contribution to trimer association aside from previous work by our laboratory [5, 6]. The V2 and V3 loops sit atop a larger conglomerate of hydrophobic residues that we have preliminary coined the hydrophobic core. Molecular modeling has elucidated that the hydrophobicity of these residues is crucial to maintain gp120

structure and stability, and a future direction for this research would be to perform similar mutational studies with these residues to validate our modeling, and to further explore the interaction of the variable loops located in the apex of the trimer with more deeply buried hydrophobic residues in the gp120 core.

In conclusion, the work shown in this dissertation has further characterized the role of the variable loops in the stability of the envelope trimer of primate lentiviruses. The studies presented here are basic research into the overall structure of the envelope protein of primate lentiviruses and the role of the variable loops in maintaining these structures. While the goal of this work was not to directly create an HIV immunogen, the knowledge garnered from these studies could be used to aid in the stabilization of the HIV envelope trimer for this purpose. This work took a broad approach to this problem by exploring amino acid motifs that were thought to play a role in envelope stability, by performing mutational studies and molecular analysis in subtype C HIV-1, HIV-2 and SIV species to further the understanding of the PLV envelope structure and function. Because a neutralizing vaccine against HIV is still unavailable, understanding the biology of the envelope protein, the primary immunogen on the virus, is crucial. This work contributes to the understanding of the stabilization of the envelope trimer and will hopefully aid in the development of epitope and trimer-based vaccines and antiviral therapies.

APPENDIX A- Primers for ZM249 V3 Loop Mutagenesis

Env and Mutation	Direction	Primer Sequence (5'-3')
HIV-1 ZM249M_B10		
303 G Insert	F	cgacccaataataatacaggaagaaaaagtataaggata
	R	tatccttatacttttttcttcctgtattattattgggtcg
I307A	F	cccaataataatacaagaaaaagtgcaaggataggaccaggacaaacattc
	R	gaatgtttgtcctggtcctatccttgcacttttttcttgtattattattggg
I307E	F	aataatacaagaaaaagtgaaaggataggaccaggacaa
	R	ttgtcctggtcctatcctttcacttttttcttgtattatt
I307L	F	cccaataataatacaagaaaaagtttaaggataggaccaggacaaacattc
	R	gaatgtttgtcctggtcctatccttaaacttttttcttgtattattattgggg
I309A	F	ccaataataatacaagaaaaagtataagggcaggaccaggacaaacattctatgcaacagg
	R	cctgttgcatagaatgtttgtcctggtcctgcccttatacttttttcttgtattattattgg
I309E	F	ccaataataatacaagaaaaagtataagggaaggaccaggacaaacattctatgcaacagg
	R	cctgttgcatagaatgtttgtcctggtccttcccttatacttttttcttgtattattattgg
I309L	F	acaagaaaaagtataaggctaggaccaggacaaacattc
	R	gaatgtttgtcctggtcctagccttatacttttttcttgt
F317A	F	ggataggaccaggacaaacagcctatgcaacaggagaaataatagg
	R	ctatttttcctattatttctcctgttgcataggctgtttgtcctggtcctatccttatacttttttc
F317E	F	ataggaccaggacaaacagagctatgcaacaggagaaat
	R	atttctcctgttgcatagctctgtttgtcctggtcctat
F317L	F	ggataggaccaggacaaacattgtatgcaacaggagaaataatagg
	R	cttatttttcctattatttctcctgttgcatacaatgtttgtcctggtcctatccttatacttttttc

APPENDIX B- Primers for 1084i V3 Loop Mutagenesis

Env and Mutation	Direction	Primer Sequence (5'-3')
HIV-1 1084i		
303 G Insert	F	aggcccaacaataatacaggtaggaaaagtatgaggatag
	R	ctatcctcatacttttcctacctgtattattgttgggcct
322 G Insert	F	ttctatgcaacaggagaaggtataataggaaatataaga
	R	tcttatatttcctattataccttctcctgttgcatagaa
M307A	F	aataatacaaggaaaagtgcaaggataggaccaggacaa
	R	ttgtcctggtcctatccttgcacttttccttgtattatt
M307E	F	aataatacaaggaaaagtgagaggataggaccaggacaa
	R	ttgtcctggtcctatcctctcacttttccttgtattatt
M307L	F	aataatacaaggaaaagtcttaggataggaccaggacaa
	R	ttgtcctggtcctatcctaagacttttccttgtattatt
I309A	F	acaaggaaaagtatgagggcaggaccaggacaagcattc
	R	gaatgcttgtcctggtcctgccctcatacttttccttgt
I309E	F	acaaggaaaagtatgagggagggaccaggacaagcattc
	R	gaatgcttgtcctggtccctccctcatacttttccttgt
I309L	F	acaaggaaaagtatgaggcttggaccaggacaagcattc
	R	gaatgcttgtcctggtccaagcctcatacttttccttgt
F317A	F	ataggaccaggacaagcagcatatgcaacaggagaaata
	R	tatttctcctgttgcatatgctgcttgtcctggtcctat
F317E	F	ataggaccaggacaagcagagtatgcaacaggagaaata
	R	tatttctcctgttgcatactctgcttgtcctggtcctat
F317L	F	ataggaccaggacaagcactttatgcaacaggagaaata
	R	tatttctcctgttgcataaagtgcttgtcctggtcctat

APPENDIX C- Primers for HIV-2 ST V3 Loop Mutagenesis

Env and Mutation	Direction	Primer Sequence (5'-3')
HIV-2 ST		
303 G Insert	F	aagaggccaggaaacaagacagggttgtaccaataacactcatg
	R	catgagtgttattggtacaaccctgtcttgtttcctggcctctt
V307A	F	ggaaacaagacagttgtaccagcaacactcatgtcaggttagtg
	R	cactaccctgacatgagtgttgctggtacaactgtctgtttcc
I309A	F	aagacagttgtaccaataacagccatgtcagggttagtgtttcac
	R	gtgaaacactaaccctgacatggctgttattggtacaactgtctt
F317A	F	cactcatgtcagggttagtggctcactcccagccaatcaatag
	R	ctattgattggctgggagtgagccactaaccctgacatgagtg

APPENDIX D- Primers for SIVagmTan-1 V3 Mutagenesis

Env and Mutation	Direction	Primer Sequence (5'-3')
SIVagmTan-1		
303 G Insert	F	agaccaggaaataagacagggtcttaccagtaaccatc
	R	gatggttactggtaagacccctgtcttatttcctggtct
V307A	F	aataagacagtcttaccagcaaccatcatggcaggtctc
	R	gagacctgccatgatggttgctggtaagactgtcttatt
I309A	F	acagtcttaccagtaaccgccatggcaggtctcgtgttc
	R	gaacacgagacctgccatggcggttactggtaagactgt
F317A	F	atcatggcaggtctcgtggcccactcacaacaatataat
	R	attatattgttgtgagtgggccacgagacctgccatgat

APPENDIX E- Primers for SIVmac239 V3 Mutagenesis

Env and Mutation	Direction	Primer Sequence (5'-3')
SIVmac239		
303 G Insert	F	agaccaggaaataagacaggggttttaccagtcaccatt
	R	aatggtgactggtaaaaccctgtcttatttcctggtct
V307A	F	aataagacagtttaccagccaccattatgtctggattg
	R	caatccagacataatggtggctggtaaaactgtcttatt
I309A	F	caggaaataagacagtttaccagtcaccgctatgtctggattggtttccactcacaacc
	R	ggttgtgagtggaaaaccaatccagacatagcggtgactggtaaaactgtcttatttcctg
F317A	F	cagtcaccattatgtctggattggttgcccactcacaacccaatcaatgataggcc
	R	ggcctatcattgattggttgtgagtgggcaacccaatccagacataatggtgactg

APPENDIX F- Primers for HIV-1 V2 Loop Twin Cysteine Addition Mutations

Env and Mutation	Direction	Primer Sequence (5'-3')
HIV-1 Twin Cysteine Addition		
ADA Substitute Twin Cysteine	F	agacttgatgtagtatgtatagataatgataatacttgctataggttgataaat
	R	atttatcaacctatagcaagtattatcattatctatacatactactacatcaagtct
ADA Insert Cysteine 183 Left	F	atagacttgatgtagttgtccaatagataatgataat
	R	attatcattatctattggacaactacatcaagtctat
ADA Insert Cysteine 183 Right	F	gacttgatgtagtaccatgtatagataatgataatgat
	R	atcattatcattatctatacatggtactactacatcaagtc
ADA Insert Cysteine 191	F	gataatgataatactagctgttataggttgataaattgt
	R	acaatttatcaacctataacagctagtattatcattatc
JRFL Substitute Twin Cysteine	F	aaacttgatgtagtatgtatagataataataatacctgctataggttgataag
	R	cttatcaacctatagcaggtattattattatctatacatactacatcaagttt
JRFL Insert Cysteine 183 Left	F	aaacttgatgtagtatgtccaatagataataataata
	R	tattattattatctattggacatactactacatcaagttt
JRFL Insert Cysteine 183 Right	F	aaacttgatgtagtaccatgtatagataataataata
	R	tattattattatctatacatggtactactacatcaagttt
JRFL Insert Cysteine 191	F	gataataataataccagctgttataggttgataagttgt
	R	acaacttatcaacctataacagctggtattattattatc

APPENDIX G- Primers for HIV-1 V2 Swap Mutations

Env and Mutation	Direction	Primer Sequence (5'-3')
HIV-1 V2 Swap		
Subtype B 5' Amp (Primer A)	F	ggccggtacctgtgtggaaagaagcaaccaccac
ADA 3' Amp (Primer D)	R	ggccggatccgttcactaatcgaatggatctgtc
JRFL 3' Amp (Primer D)	R	ggccggatccgttcactaatcgtccggatctgtc
ADA SIVmac V2 Swap 3' Reverse (Primer B)	F	tctactttcattaccagtgttattccttgttcacatactacatcaagtctataaaa
ADA SIVmac V2 Swap 5' Forward (Primer C)	R	aacactggtaatgaaagtagatgttacatgaaccactgtaatacctcaaccattaca
JRFL SIVmac V2 Swap 3' Reverse (Primer B)	F	tctactttcattaccagtgttattccttgttcacatactacatcaagtttataaaa
JRFL SIVmac V2 Swap 5' Forward (Primer C)	R	aacactggtaatgaaagtagatgttacatgaaccactgtgacacctcagtcattaca
ADA HIV2 ST V2 Swap 3' Reverse (Primer B)	F	gtaacatgtttctctttcttggtgtcatttgattcacatactacatcaagtctataaaa
ADA HIV2 ST V2 Swap 5' Reverse (Primer C)	R	gacaccaagaaagagaaaacatgttacatgaaccactgcaatacctcaaccattacacag
JRFL HIV2 ST V2 Swap 3' Reverse (Primer B)	F	gtaacatgtttctctttcttggtgtcatttgattcacatactacatcaagtttataaaa
JRFL HIV2 ST V2 Swap 5' Reverse (Primer C)	R	gacaccaagaaagagaaaacatgttacatgaaccactgcgacacctcagtcattacacag

APPENDIX H- Primers for Twin Cysteine Removal from HIV-2 and SIV Species

Env and Mutation	Direction	Primer Sequence (5'-3')
Non-HIV-1 Twin Cysteine Removal		
HIV-2 ST C183A	F	tactcaaaagatgtagtcgctgaatcaaatgacaccaag
	R	cttggtgtcatttgattcagcgactacatctttgagta
HIV-2 ST C191A	F	accaagaaagagaaaacagcttacatgaaccactgcaac
	R	gttgcagtggttcatgtaagctgtttctctttcttggt
SIVagmTan-1 C183A	F	cttggtatgatcaggagttggtagctgagaaggaaaacaacaccacaggc
	R	gcctgtggtgttgttttccttctcagctaccaactcctgatcataccaag
SIVagmTan-1 C191A	F	caacaccacaggcacgagaggtgcttacatgattcactgcaacgactctg
	R	cagagtcgttgcagtgaatcatgtaagcacctctcgtgcctgtggtgttg
SIVagm155 C183A	F	ggaatgatgcagaaatctttgctaagcgtagtacatcgcataatagg
	R	cctattatgcgatgtactacgcttagcaaagatttctgcatcattcc
SIVagm155 C191A	F	cataataggacaaaagaggcctatatgatccactgtaatg
	R	cattacagtggatcatataggcctctttgtcctattatg

APPENDIX I- Primers for PLV Envelope Gene Amplification

PLV Species	Direction	Primer Sequence (5'-3')
Envelope Gene Amplification		
HIV-2 ST	F	gcgcgccttaaggccgccaccatgtggtaggaatcaactatttgttgcc
	R	gcgcgcggtacctcacaggagggcgatttctgcccc
SIVagmTan-1	F	ggccggtaccgccgccaccatgggaccattaaggggaaaaggggtattattagtaatttggg
	R	ggccgaattcctagttgaaccaccgttccagttcttgtcgcattcttcttgg
SIVagm155	F	ggccctcgaggccgccaccatggcaaagttcttaggaatttttatag
	R	ggcctctagactaatgcgtagatttgctttg
HIV-1 ZM249M_B10	F	ggccaagcttgccgccaccatgagtgatggggatactg
	R	ggccgaattcttatagcaaagctctttc

BIBLIOGRAPHY

1. Kwon, Y.D., et al., *Crystal structure, conformational fixation and entry-related interactions of mature ligand-free HIV-1 Env.* Nat Struct Mol Biol, 2015. **22**(7): p. 522-31.

2. Lyumkis, D., et al., *Cryo-EM structure of a fully glycosylated soluble cleaved HIV-1 envelope trimer.* Science, 2013. **342**(6165): p. 1484-90.

3. Julien, J.P., et al., *Crystal structure of a soluble cleaved HIV-1 envelope trimer.* Science, 2013. **342**(6165): p. 1477-83.

4. Xiang, S.H., *Recent advances on the use of structural biology for the design of novel envelope immunogens of HIV-1.* Curr HIV Res, 2013. **11**(6): p. 464-72.

5. Xiang, S.H., et al., *A V3 loop-dependent gp120 element disrupted by CD4 binding stabilizes the human immunodeficiency virus envelope glycoprotein trimer.* J Virol, 2010. **84**(7): p. 3147-61.

6. Bohl, C., et al., *A twin-cysteine motif in the V2 region of gp120 is associated with SIV envelope trimer stabilization.* PLoS One, 2013. **8**(7): p. e69406.

7. (WHO), W.H.O. *HIV/AIDS Data and Statistics*. 2016 [cited 2018; Available from: http://www.who.int/hiv/en/.

8. UNAIDS, *UNAIDS Data 2017*. 2017, Joint United Nations Programme on HIV/AIDS (UNAIDS).

9. Barre-Sinoussi, F., et al., *Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS).* Science, 1983. **220**(4599): p. 868-71.

10. Gallo, R.C., et al., *Isolation of human T-cell leukemia virus in acquired immune deficiency syndrome (AIDS).* Science, 1983. **220**(4599): p. 865-7.

11. Sharp, P.M. and B.H. Hahn, *Origins of HIV and the AIDS pandemic.* Cold Spring Harb Perspect Med, 2011. **1**(1): p. a006841.

12. Muriaux, D. and J.L. Darlix, *Properties and functions of the nucleocapsid protein in virus assembly.* RNA Biol, 2010. **7**(6): p. 744-53.

13. Briggs, J.A., et al., *Structural organization of authentic, mature HIV-1 virions and cores.* EMBO J, 2003. **22**(7): p. 1707-15.

14. Freed, E.O., *HIV-1 gag proteins: diverse functions in the virus life cycle.* Virology, 1998. **251**(1): p. 1-15.

15. Tedbury, P.R. and E.O. Freed, *The role of matrix in HIV-1 envelope glycoprotein incorporation.* Trends Microbiol, 2014. **22**(7): p. 372-8.

16. Hill, M., G. Tachedjian, and J. Mak, *The packaging and maturation of the HIV-1 Pol proteins.* Curr HIV Res, 2005. **3**(1): p. 73-85.

17. Zhu, P., et al., *Electron tomography analysis of envelope glycoprotein trimers on HIV and simian immunodeficiency virus virions.* Proc Natl Acad Sci U S A, 2003. **100**(26): p. 15812-7.

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

19. Laboratory, L.A.N. *HIV-1 Gene Map*. 2017 Jan 24, 2017 [cited 2018 March 31]; Available from: https://www.hiv.lanl.gov/content/sequence/HIV/MAP/landmark.html.

20. Freed, E.O. and M.A. Martin, *Human Immunodeficiency Viruses: Replication*, in *Field's Virology*, D.M. Knipe and P.M. Howley, Editors. 2013, Lippincott, Williams & Wilkins: Philadelphia, PA. p. 1502-1560.

21. Marcello, A., et al., *Nuclear organization and the control of HIV-1 transcription.* Gene, 2004. **326**: p. 1-11.

22. Hu, W.S. and S.H. Hughes, *HIV-1 reverse transcription.* Cold Spring Harb Perspect Med, 2012. **2**(10).

23. Brin, E. and J. Leis, *HIV-1 integrase interaction with U3 and U5 terminal sequences in vitro defined using substrates with random sequences.* J Biol Chem, 2002. **277**(21): p. 18357-64.

24. Hindmarsh, P. and J. Leis, *Retroviral DNA integration.* Microbiol Mol Biol Rev, 1999. **63**(4): p. 836-43, table of contents.

25. Mondor, I., S. Ugolini, and Q.J. Sattentau, *Human immunodeficiency virus type 1 attachment to HeLa CD4 cells is CD4 independent and gp120 dependent and requires cell surface heparans.* J Virol, 1998. **72**(5): p. 3623-34.

26. Arthos, J., et al., *HIV-1 envelope protein binds to and signals through integrin alpha4beta7, the gut mucosal homing receptor for peripheral T cells.* Nat Immunol, 2008. **9**(3): p. 301-9.

27. Geijtenbeek, T.B., et al., *DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells.* Cell, 2000. **100**(5): p. 587-97.

28. Kwong, P.D., et al., *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.

29. Miyauchi, K., et al., *HIV enters cells via endocytosis and dynamin-dependent fusion with endosomes.* Cell, 2009. **137**(3): p. 433-44.

30. Stein, B.S., et al., *pH-independent HIV entry into CD4-positive T cells via virus envelope fusion to the plasma membrane.* Cell, 1987. **49**(5): p. 659-68.

31. Wilen, C.B., J.C. Tilton, and R.W. Doms, *HIV: cell binding and entry.* Cold Spring Harb Perspect Med, 2012. **2**(8).

32. Shaharabany, M. and A. Hizi, *The DNA-dependent and RNA-dependent DNA polymerase activities of the reverse transcriptases of human immunodeficiency viruses types 1 and 2.* AIDS Res Hum Retroviruses, 1991. **7**(11): p. 883-8.

33. Luo, G.X. and J. Taylor, *Template switching by reverse transcriptase during DNA synthesis.* J Virol, 1990. **64**(9): p. 4321-8.

34. di Marzo Veronese, F., et al., *Characterization of highly immunogenic p66/p51 as the reverse transcriptase of HTLV-III/LAV.* Science, 1986. **231**(4743): p. 1289-91.

35. Karageorgos, L., P. Li, and C. Burrell, *Characterization of HIV replication complexes early after cell-to-cell infection.* AIDS Res Hum Retroviruses, 1993. **9**(9): p. 817-23.

36. Heinzinger, N.K., et al., *The Vpr protein of human immunodeficiency virus type 1 influences nuclear localization of viral nucleic acids in nondividing host cells.* Proc Natl Acad Sci U S A, 1994. **91**(15): p. 7311-5.

37. Bushman, F.D., T. Fujiwara, and R. Craigie, *Retroviral DNA integration directed by HIV integration protein in vitro.* Science, 1990. **249**(4976): p. 1555-8.

38. Brown, P.O., et al., *Retroviral integration: structure of the initial covalent product and its precursor, and a role for the viral IN protein.* Proc Natl Acad Sci U S A, 1989. **86**(8): p. 2525-9.

39. Felber, B.K., et al., *rev protein of human immunodeficiency virus type 1 affects the stability and transport of the viral mRNA.* Proc Natl Acad Sci U S A, 1989. **86**(5): p. 1495-9.

40. Sodroski, J., et al., *A second post-transcriptional trans-activator gene required for HTLV-III replication.* Nature, 1986. **321**(6068): p. 412-7.

41. Yu, X.F., et al., *Role of the C terminus Gag protein in human immunodeficiency virus type 1 virion assembly and maturation.* J Gen Virol, 1995. **76 (Pt 12)**: p. 3171-9.

42. Checkley, M.A., B.G. Luttge, and E.O. Freed, *HIV-1 envelope glycoprotein biosynthesis, trafficking, and incorporation.* J Mol Biol, 2011. **410**(4): p. 582-608.

43. Dilley, K.A., et al., *Interactions between HIV-1 Gag and Viral RNA Genome Enhance Virion Assembly.* J Virol, 2017. **91**(16).

44. Paillart, J.C., et al., *Dimerization of retroviral RNA genomes: an inseparable pair.* Nat Rev Microbiol, 2004. **2**(6): p. 461-72.

45. Morita, E., et al., *ESCRT-III protein requirements for HIV-1 budding.* Cell Host Microbe, 2011. **9**(3): p. 235-242.

46. Zhao, G., et al., *Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics.* Nature, 2013. **497**(7451): p. 643-6.

47. Kohl, N.E., et al., *Active human immunodeficiency virus protease is required for viral infectivity.* Proc Natl Acad Sci U S A, 1988. **85**(13): p. 4686-90.

48. Bryant, M. and L. Ratner, *Myristoylation-dependent replication and assembly of human immunodeficiency virus 1.* Proc Natl Acad Sci U S A, 1990. **87**(2): p. 523-7.

49. Gottlinger, H.G., J.G. Sodroski, and W.A. Haseltine, *Role of capsid precursor processing and myristoylation in morphogenesis and infectivity of human immunodeficiency virus type 1.* Proc Natl Acad Sci U S A, 1989. **86**(15): p. 5781-5.

50. Gheysen, D., et al., *Assembly and release of HIV-1 precursor Pr55gag virus-like particles from recombinant baculovirus-infected insect cells.* Cell, 1989. **59**(1): p. 103-12.

51. Mark-Danieli, M., et al., *Single point mutations in the zinc finger motifs of the human immunodeficiency virus type 1 nucleocapsid alter RNA binding specificities of the gag protein and enhance packaging and infectivity.* J Virol, 2005. **79**(12): p. 7756-67.

52. Amarasinghe, G.K., et al., *NMR structure of the HIV-1 nucleocapsid protein bound to stem-loop SL2 of the psi-RNA packaging signal. Implications for genome recognition.* J Mol Biol, 2000. **301**(2): p. 491-511.

53. Harrison, G.P. and A.M. Lever, *The human immunodeficiency virus type 1 packaging signal and major splice donor region have a conserved stable secondary structure.* J Virol, 1992. **66**(7): p. 4144-53.

54. Kondo, E., et al., *The p6gag domain of human immunodeficiency virus type 1 is sufficient for the incorporation of Vpr into heterologous viral particles.* J Virol, 1995. **69**(5): p. 2759-64.

55. Jenkins, Y., et al., *Biochemical analyses of the interactions between human immunodeficiency virus type 1 Vpr and p6(Gag).* J Virol, 2001. **75**(21): p. 10537-42.

56. Gottlinger, H.G., et al., *Effect of mutations affecting the p6 gag protein on human immunodeficiency virus particle release.* Proc Natl Acad Sci U S A, 1991. **88**(8): p. 3195-9.

57. Shehu-Xhilaga, M., S.M. Crowe, and J. Mak, *Maintenance of the Gag/Gag-Pol ratio is important for human immunodeficiency virus type 1 RNA dimerization and viral infectivity.* J Virol, 2001. **75**(4): p. 1834-41.

58. Jacks, T., et al., *Characterization of ribosomal frameshifting in HIV-1 gag-pol expression.* Nature, 1988. **331**(6153): p. 280-3.

59. Lv, Z., Y. Chu, and Y. Wang, *HIV protease inhibitors: a review of molecular selectivity and toxicity.* HIV AIDS (Auckl), 2015. **7**: p. 95-104.

60. Kurt Yilmaz, N., R. Swanstrom, and C.A. Schiffer, *Improving Viral Protease Inhibitors to Counter Drug Resistance.* Trends Microbiol, 2016. **24**(7): p. 547-557.

61. Fu, T.B. and J. Taylor, *When retroviral reverse transcriptases reach the end of their RNA templates.* J Virol, 1992. **66**(7): p. 4271-8.

62. Sarafianos, S.G., et al., *Structure and function of HIV-1 reverse transcriptase: molecular mechanisms of polymerization and inhibition.* J Mol Biol, 2009. **385**(3): p. 693-713.

63. Jacobo-Molina, A., et al., *Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 A resolution shows bent DNA.* Proc Natl Acad Sci U S A, 1993. **90**(13): p. 6320-4.

64. Kohlstaedt, L.A., et al., *Crystal structure at 3.5 A resolution of HIV-1 reverse transcriptase complexed with an inhibitor.* Science, 1992. **256**(5065): p. 1783-90.

65. Preston, B.D., B.J. Poiesz, and L.A. Loeb, *Fidelity of HIV-1 reverse transcriptase.* Science, 1988. **242**(4882): p. 1168-71.

66. De Clercq, E., *The nucleoside reverse transcriptase inhibitors, nonnucleoside reverse transcriptase inhibitors, and protease inhibitors in the treatment of HIV infections (AIDS).* Adv Pharmacol, 2013. **67**: p. 317-58.

67. Usach, I., V. Melis, and J.E. Peris, *Non-nucleoside reverse transcriptase inhibitors: a review on pharmacokinetics, pharmacodynamics, safety and tolerability.* J Int AIDS Soc, 2013. **16**: p. 1-14.

68. Craigie, R., *The molecular biology of HIV integrase.* Future Virol, 2012. **7**(7): p. 679-686.

69. Yoder, K.E. and F.D. Bushman, *Repair of gaps in retroviral DNA integration intermediates.* J Virol, 2000. **74**(23): p. 11191-200.

70. Moulard, M., et al., *Processing and routage of HIV glycoproteins by furin to the cell surface.* Virus Res, 1999. **60**(1): p. 55-65.

71. Lasky, L.A., et al., *Neutralization of the AIDS retrovirus by antibodies to a recombinant envelope glycoprotein.* Science, 1986. **233**(4760): p. 209-12.

72. Wei, X., et al., *Antibody neutralization and escape by HIV-1.* Nature, 2003. **422**(6929): p. 307-12.

73. Wyatt, R. and J. Sodroski, *The HIV-1 envelope glycoproteins: fusogens, antigens, and immunogens.* Science, 1998. **280**(5371): p. 1884-8.

74. Barin, F., et al., *Virus envelope protein of HTLV-III represents major target antigen for antibodies in AIDS patients.* Science, 1985. **228**(4703): p. 1094-6.

75. Weiss, R.A., *HIV receptors and cellular tropism.* IUBMB Life, 2002. **53**(4-5): p. 201-5.

76. Dayton, A.I., et al., *The trans-activator gene of the human T cell lymphotropic virus type III is required for replication.* Cell, 1986. **44**(6): p. 941-7.

77. Fisher, A.G., et al., *The trans-activator gene of HTLV-III is essential for virus replication.* Nature, 1986. **320**(6060): p. 367-71.

78. Bartz, S.R. and M. Emerman, *Human immunodeficiency virus type 1 Tat induces apoptosis and increases sensitivity to apoptotic signals by up-regulating FLICE/caspase-8.* J Virol, 1999. **73**(3): p. 1956-63.

79. Ensoli, B., et al., *Tat protein of HIV-1 stimulates growth of cells derived from Kaposi's sarcoma lesions of AIDS patients.* Nature, 1990. **345**(6270): p. 84-6.

80. Clark, E., B. Nava, and M. Caputi, *Tat is a multifunctional viral protein that modulates cellular gene expression and functions.* Oncotarget, 2017. **8**(16): p. 27569-27581.

81. Hammarskjold, M.L., et al., *Regulation of human immunodeficiency virus env expression by the rev gene product.* J Virol, 1989. **63**(5): p. 1959-66.

82. Fischer, U., et al., *The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs.* Cell, 1995. **82**(3): p. 475-83.

83. Conticello, S.G., R.S. Harris, and M.S. Neuberger, *The Vif protein of HIV triggers degradation of the human antiretroviral DNA deaminase APOBEC3G.* Curr Biol, 2003. **13**(22): p. 2009-13.

84. Feng, Y., et al., *Suppression of APOBEC3-mediated restriction of HIV-1 by Vif.* Front Microbiol, 2014. **5**: p. 450.

85. Simon, J.H., et al., *Evidence for a newly discovered cellular anti-HIV-1 phenotype.* Nat Med, 1998. **4**(12): p. 1397-400.

86. Madani, N. and D. Kabat, *An endogenous inhibitor of human immunodeficiency virus in human lymphocytes is overcome by the viral Vif protein.* J Virol, 1998. **72**(12): p. 10251-5.

87. Sheehy, A.M., et al., *Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein.* Nature, 2002. **418**(6898): p. 646-50.

88. Batisse, J., et al., *The role of Vif oligomerization and RNA chaperone activity in HIV-1 replication.* Virus Res, 2012. **169**(2): p. 361-76.

89. Hoglund, S., et al., *Role of vif during packing of the core of HIV-1.* Virology, 1994. **201**(2): p. 349-55.

90. Cohen, E.A., et al., *Identification of HIV-1 vpr product and function.* J Acquir Immune Defic Syndr, 1990. **3**(1): p. 11-8.

91. Yuan, X., et al., *Human immunodeficiency virus vpr gene encodes a virion-associated protein.* AIDS Res Hum Retroviruses, 1990. **6**(11): p. 1265-71.

92. Kogan, M. and J. Rappaport, *HIV-1 accessory protein Vpr: relevance in the pathogenesis of HIV and potential for therapeutic intervention.* Retrovirology, 2011. **8**: p. 25.

93. Bachand, F., et al., *Incorporation of Vpr into human immunodeficiency virus type 1 requires a direct interaction with the p6 domain of the p55 gag precursor.* J Biol Chem, 1999. **274**(13): p. 9083-91.

94. Popov, S., et al., *Viral protein R regulates nuclear import of the HIV-1 pre-integration complex.* EMBO J, 1998. **17**(4): p. 909-17.

95. Popov, S., et al., *Viral protein R regulates docking of the HIV-1 preintegration complex to the nuclear pore complex.* J Biol Chem, 1998. **273**(21): p. 13347-52.

96. Agostini, I., et al., *Phosphorylation of Vpr regulates HIV type 1 nuclear import and macrophage infection.* AIDS Res Hum Retroviruses, 2002. **18**(4): p. 283-8.

97. Wang, L., et al., *Interaction of virion protein Vpr of human immunodeficiency virus type 1 with cellular transcription factor Sp1 and trans-activation of viral long terminal repeat.* J Biol Chem, 1995. **270**(43): p. 25564-9.

98. Agostini, I., et al., *The human immunodeficiency virus type 1 Vpr transactivator: cooperation with promoter-bound activator domains and binding to TFIIB.* J Mol Biol, 1996. **261**(5): p. 599-606.

99. Groschel, B. and F. Bushman, *Cell cycle arrest in G2/M promotes early steps of infection by human immunodeficiency virus.* J Virol, 2005. **79**(9): p. 5695-704.

100. Goh, W.C., N. Manel, and M. Emerman, *The human immunodeficiency virus Vpr protein binds Cdc25C: implications for G2 arrest.* Virology, 2004. **318**(1): p. 337-49.

101. Goh, W.C., et al., *HIV-1 Vpr increases viral expression by manipulation of the cell cycle: a mechanism for selection of Vpr in vivo.* Nat Med, 1998. **4**(1): p. 65-71.

102. Tiganos, E., et al., *Structural and functional analysis of the membrane-spanning domain of the human immunodeficiency virus type 1 Vpu protein.* Virology, 1998. **251**(1): p. 96-107.

103. Andrew, A. and K. Strebel, *HIV-1 Vpu targets cell surface markers CD4 and BST-2 through distinct mechanisms.* Mol Aspects Med, 2010. **31**(5): p. 407-17.

104. Willey, R.L., et al., *Human immunodeficiency virus type 1 Vpu protein induces rapid degradation of CD4.* J Virol, 1992. **66**(12): p. 7193-200.

105. Margottin, F., et al., *A novel human WD protein, h-beta TrCp, that interacts with HIV-1 Vpu connects CD4 to the ER degradation pathway through an F-box motif.* Mol Cell, 1998. **1**(4): p. 565-74.

106. Lama, J., A. Mangasarian, and D. Trono, *Cell-surface expression of CD4 reduces HIV-1 infectivity by blocking Env incorporation in a Nef- and Vpu-inhibitable manner.* Curr Biol, 1999. **9**(12): p. 622-31.

107. Neil, S.J., T. Zang, and P.D. Bieniasz, *Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu.* Nature, 2008. **451**(7177): p. 425-30.

108. Douglas, J.L., et al., *Vpu directs the degradation of the human immunodeficiency virus restriction factor BST-2/Tetherin via a {beta}TrCP-dependent mechanism.* J Virol, 2009. **83**(16): p. 7931-47.

109. Mangeat, B., et al., *HIV-1 Vpu neutralizes the antiviral factor Tetherin/BST-2 by binding it and directing its beta-TrCP2-dependent degradation.* PLoS Pathog, 2009. **5**(9): p. e1000574.

110. Mitchell, R.S., et al., *Vpu antagonizes BST-2-mediated restriction of HIV-1 release via beta-TrCP and endo-lysosomal trafficking.* PLoS Pathog, 2009. **5**(5): p. e1000450.

111. Kestler, H.W., 3rd, et al., *Importance of the nef gene for maintenance of high virus loads and for development of AIDS.* Cell, 1991. **65**(4): p. 651-62.

112. Jin, Y.J., et al., *HIV Nef-mediated CD4 down-regulation is adaptor protein complex 2 dependent.* J Immunol, 2005. **175**(5): p. 3157-64.

113. Gondim, M.V., et al., *AP-2 Is the Crucial Clathrin Adaptor Protein for CD4 Downmodulation by HIV-1 Nef in Infected Primary CD4+ T Cells.* J Virol, 2015. **89**(24): p. 12518-24.

114. Sol-Foulon, N., et al., *HIV-1 Nef-induced upregulation of DC-SIGN in dendritic cells promotes lymphocyte clustering and viral spread.* Immunity, 2002. **16**(1): p. 145-55.

115. Swigut, T., N. Shohdy, and J. Skowronski, *Mechanism for down-regulation of CD28 by Nef.* EMBO J, 2001. **20**(7): p. 1593-604.

116. El-Far, M., et al., *Down-regulation of CTLA-4 by HIV-1 Nef protein.* PLoS One, 2013. **8**(1): p. e54295.

117. Schwartz, O., et al., *Endocytosis of major histocompatibility complex class I molecules is induced by the HIV-1 Nef protein.* Nat Med, 1996. **2**(3): p. 338-42.

118. Foster, J.L. and J.V. Garcia, *HIV-1 Nef: at the crossroads.* Retrovirology, 2008. **5**: p. 84.

119. Moarefi, I., et al., *Activation of the Src-family tyrosine kinase Hck by SH3 domain displacement.* Nature, 1997. **385**(6617): p. 650-3.

120. Arora, V.K., et al., *Lentivirus Nef specifically activates Pak2.* J Virol, 2000. **74**(23): p. 11081-7.

121. Hahn, B.H., et al., *AIDS as a zoonosis: scientific and public health implications.* Science, 2000. **287**(5453): p. 607-14.

122. Gao, F., et al., *Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes.* Nature, 1999. **397**(6718): p. 436-41.

123. Apetrei, C., D.L. Robertson, and P.A. Marx, *The history of SIVS and AIDS: epidemiology, phylogeny and biology of isolates from naturally SIV infected non-human primates (NHP) in Africa.* Front Biosci, 2004. **9**: p. 225-54.

124. Ma, D., et al., *SIVagm infection in wild African green monkeys from South Africa: epidemiology, natural history, and evolutionary considerations.* PLoS Pathog, 2013. **9**(1): p. e1003011.

125. Keele, B.F., et al., *Chimpanzee reservoirs of pandemic and nonpandemic HIV-1.* Science, 2006. **313**(5786): p. 523-6.

126. Keele, B.F., et al., *Increased mortality and AIDS-like immunopathology in wild chimpanzees infected with SIVcpz.* Nature, 2009. **460**(7254): p. 515-9.

127. Sharp, P.M., G.M. Shaw, and B.H. Hahn, *Simian immunodeficiency virus infection of chimpanzees.* J Virol, 2005. **79**(7): p. 3891-902.

128. Hirsch, V.M. and P.R. Johnson, *Pathogenic diversity of simian immunodeficiency viruses.* Virus Res, 1994. **32**(2): p. 183-203.

129. Pandrea, I.V., et al., *Acute loss of intestinal CD4+ T cells is not predictive of simian immunodeficiency virus virulence.* J Immunol, 2007. **179**(5): p. 3035-46.

130. Souquiere, S., et al., *Simian immunodeficiency virus types 1 and 2 (SIV mnd 1 and 2) have different pathogenic potentials in rhesus macaques upon experimental cross-species transmission.* J Gen Virol, 2009. **90**(Pt 2): p. 488-99.

131. Beer, B.E., et al., *Immunodeficiency in the absence of high viral load in pig-tailed macaques infected with Simian immunodeficiency virus SIVsun or SIVlhoest.* J Virol, 2005. **79**(22): p. 14044-56.

132. Chakrabarti, L.A., et al., *Normal T-cell turnover in sooty mangabeys harboring active simian immunodeficiency virus infection.* J Virol, 2000. **74**(3): p. 1209-23.

133. Santiago, M.L., et al., *Simian immunodeficiency virus infection in free-ranging sooty mangabeys (Cercocebus atys atys) from the Tai Forest, Cote d'Ivoire: implications for the origin of epidemic human immunodeficiency virus type 2.* J Virol, 2005. **79**(19): p. 12515-27.

134. Apetrei, C., et al., *Kuru experiments triggered the emergence of pathogenic SIVmac.* AIDS, 2006. **20**(3): p. 317-21.

135. Apetrei, C., et al., *Molecular epidemiology of simian immunodeficiency virus SIVsm in U.S. primate centers unravels the origin of SIVmac and SIVstm.* J Virol, 2005. **79**(14): p. 8991-9005.

136. Gardner, M., *A historical perspective: Simian AIDS-an accidental windfall.* J Med Primatol, 2016. **45**(5): p. 212-4.

137. Bailes, E., et al., *Hybrid origin of SIV in chimpanzees.* Science, 2003. **300**(5626): p. 1713.

138. McCutchan, F.E., *Global epidemiology of HIV.* J Med Virol, 2006. **78 Suppl 1**: p. S7-S12.

139. Peeters, M., et al., *Geographical distribution of HIV-1 group O viruses in Africa.* AIDS, 1997. **11**(4): p. 493-8.

140. Ayouba, A., et al., *HIV-1 group O infection in Cameroon, 1986 to 1998.* Emerg Infect Dis, 2001. **7**(3): p. 466-7.

141. Vessiere, A., et al., *Diagnosis and monitoring of HIV-1 group O-infected patients in Cameroun.* J Acquir Immune Defic Syndr, 2010. **53**(1): p. 107-10.

142. Villabona-Arenas, C.J., et al., *HIV-1 group O infection in Cameroon from 2006 to 2013: Prevalence, genetic diversity, evolution and public health challenges.* Infect Genet Evol, 2015. **36**: p. 210-216.

143. Rodgers, M.A., et al., *Identification of rare HIV-1 Group N, HBV AE, and HTLV-3 strains in rural South Cameroon.* Virology, 2017. **504**: p. 141-151.

144. Delaugerre, C., et al., *HIV-1 group N: travelling beyond Cameroon.* Lancet, 2011. **378**(9806): p. 1894.

145. Ayouba, A., et al., *HIV-1 group N among HIV-1-seropositive individuals in Cameroon.* AIDS, 2000. **14**(16): p. 2623-5.

146. Vallari, A., et al., *Confirmation of putative HIV-1 group P in Cameroon.* J Virol, 2011. **85**(3): p. 1403-7.

147. Roques, P., et al., *Phylogenetic characteristics of three new HIV-1 N strains and implications for the origin of group N.* AIDS, 2004. **18**(10): p. 1371-81.

148. Simon, F., et al., *Identification of a new human immunodeficiency virus type 1 distinct from group M and group O.* Nat Med, 1998. **4**(9): p. 1032-7.

149. D'Arc, M., et al., *Origin of the HIV-1 group O epidemic in western lowland gorillas.* Proc Natl Acad Sci U S A, 2015. **112**(11): p. E1343-52.

150. Van Heuverswyn, F., et al., *Human immunodeficiency viruses: SIV infection in wild gorillas.* Nature, 2006. **444**(7116): p. 164.

151. Plantier, J.C., et al., *A new human immunodeficiency virus derived from gorillas.* Nat Med, 2009. **15**(8): p. 871-2.

152. Hemelaar, J., *The origin and diversity of the HIV-1 pandemic.* Trends Mol Med, 2012. **18**(3): p. 182-92.

153. Korber, B., et al., *Timing the ancestor of the HIV-1 pandemic strains.* Science, 2000. **288**(5472): p. 1789-96.

154. Robertson, D.L., et al., *HIV-1 nomenclature proposal.* Science, 2000. **288**(5463): p. 55-6.

155. Hemelaar, J., et al., *Global trends in molecular epidemiology of HIV-1 during 2000-2007.* AIDS, 2011. **25**(5): p. 679-89.

156. John-Stewart, G.C., et al., *Subtype C Is associated with increased vaginal shedding of HIV-1.* J Infect Dis, 2005. **192**(3): p. 492-6.

157. Paxton, W.A., *Why Are Some HIV-1 Subtypes More "Wimpy" at Causing Disease?* EBioMedicine, 2016. **13**: p. 27-28.

158. Venner, C.M., et al., *Infecting HIV-1 Subtype Predicts Disease Progression in Women of Sub-Saharan Africa.* EBioMedicine, 2016. **13**: p. 305-314.

159. Peeters, M., et al., *Evidence for differences in MT2 cell tropism according to genetic subtypes of HIV-1: syncytium-inducing variants seem rare among subtype C HIV-1 viruses.* J Acquir Immune Defic Syndr Hum Retrovirol, 1999. **20**(2): p. 115-21.

160. Renjifo, B., et al., *Preferential in-utero transmission of HIV-1 subtype C as compared to HIV-1 subtype A or D.* AIDS, 2004. **18**(12): p. 1629-36.

161. Ball, S.C., et al., *Comparing the ex vivo fitness of CCR5-tropic human immunodeficiency virus type 1 isolates of subtypes B and C.* J Virol, 2003. **77**(2): p. 1021-38.

162. Iordanskiy, S., et al., *Subtype-associated differences in HIV-1 reverse transcription affect the viral replication.* Retrovirology, 2010. **7**: p. 85.

163. Chen, Z., et al., *Genetic characterization of new West African simian immunodeficiency virus SIVsm: geographic clustering of household-derived SIV strains with human immunodeficiency virus type 2 subtypes and genetically diverse viruses from a single feral sooty mangabey troop.* J Virol, 1996. **70**(6): p. 3617-27.

164. Ayouba, A., et al., *Evidence for continuing cross-species transmission of SIVsmm to humans: characterization of a new HIV-2 lineage in rural Cote d'Ivoire.* AIDS, 2013. **27**(15): p. 2488-91.

165. de Silva, T.I., M. Cotten, and S.L. Rowland-Jones, *HIV-2: the forgotten AIDS virus.* Trends Microbiol, 2008. **16**(12): p. 588-95.

166. Lemey, P., et al., *Tracing the origin and history of the HIV-2 epidemic.* Proc Natl Acad Sci U S A, 2003. **100**(11): p. 6588-92.

167. Pieniazek, D., et al., *Predominance of human immunodeficiency virus type 2 subtype B in Abidjan, Ivory Coast.* AIDS Res Hum Retroviruses, 1999. **15**(6): p. 603-8.

168. Peeters, M., C. Toure-Kane, and J.N. Nkengasong, *Genetic diversity of HIV in Africa: impact on diagnosis, treatment, vaccine development and trials.* AIDS, 2003. **17**(18): p. 2547-60.

169. Smith, S.M., et al., *Isolation of a new HIV-2 group in the US.* Retrovirology, 2008. **5**: p. 103.

170. Marlink, R., et al., *Reduced rate of disease development after HIV-2 infection as compared to HIV-1.* Science, 1994. **265**(5178): p. 1587-90.

171. Gueudin, M., et al., *Differences in proviral DNA load between HIV-1- and HIV-2-infected patients.* AIDS, 2008. **22**(2): p. 211-5.

172. MacNeil, A., et al., *Direct evidence of lower viral replication rates in vivo in human immunodeficiency virus type 2 (HIV-2) infection than in HIV-1 infection.* J Virol, 2007. **81**(10): p. 5325-30.

173. Bjorling, E., et al., *Autologous neutralizing antibodies prevail in HIV-2 but not in HIV-1 infection.* Virology, 1993. **193**(1): p. 528-30.

174. Shi, Y., et al., *Evolution of human immunodeficiency virus type 2 coreceptor usage, autologous neutralization, envelope sequence and glycosylation.* J Gen Virol, 2005. **86**(Pt 12): p. 3385-96.

175. Rodriguez, S.K., et al., *Comparison of heterologous neutralizing antibody responses of human immunodeficiency virus type 1 (HIV-1)- and HIV-2-infected Senegalese patients: distinct patterns of breadth and magnitude distinguish HIV-1 and HIV-2 infections.* J Virol, 2007. **81**(10): p. 5331-8.

176. Silvestri, G., et al., *Divergent host responses during primary simian immunodeficiency virus SIVsm infection of natural sooty mangabey and nonnatural rhesus macaque hosts.* J Virol, 2005. **79**(7): p. 4043-54.

177. Ling, B., et al., *Classic AIDS in a sooty mangabey after an 18-year natural infection.* J Virol, 2004. **78**(16): p. 8902-8.

178. Caffrey, M., *HIV envelope: challenges and opportunities for development of entry inhibitors.* Trends Microbiol, 2011. **19**(4): p. 191-7.

179. Vergne, L.P., M., *The challenge of HIV sequence diversity in the envelope glycoprotein*, in *Entry Inhibitors in HIV Therapy*, C.A.D. Jacqueline D. Reeves, Editor. 2007, Birkhauser. p. 17-30.

180. Matthews, T., et al., *Enfuvirtide: the first therapy to inhibit the entry of HIV-1 into host CD4 lymphocytes.* Nat Rev Drug Discov, 2004. **3**(3): p. 215-25.

181. Westby, M. and E. van der Ryst, *CCR5 antagonists: host-targeted antiviral agents for the treatment of HIV infection, 4 years on.* Antivir Chem Chemother, 2010. **20**(5): p. 179-92.

182. Tilton, J.C. and R.W. Doms, *Entry inhibitors in the treatment of HIV-1 infection.* Antiviral Res, 2010. **85**(1): p. 91-100.

183. Patel, M., et al., *Cell-surface heparan sulfate proteoglycan mediates HIV-1 infection of T-cell lines.* AIDS Res Hum Retroviruses, 1993. **9**(2): p. 167-74.

184. McDougal, J.S., et al., *Binding of HTLV-III/LAV to T4+ T cells by a complex of the 110K viral protein and the T4 molecule.* Science, 1986. **231**(4736): p. 382-5.

185. McDougal, J.S., et al., *The T4 glycoprotein is a cell-surface receptor for the AIDS virus.* Cold Spring Harb Symp Quant Biol, 1986. **51 Pt 2**: p. 703-11.

186. Maddon, P.J., et al., *The T4 gene encodes the AIDS virus receptor and is expressed in the immune system and the brain.* Cell, 1986. **47**(3): p. 333-48.

187. Liu, J., et al., *Molecular architecture of native HIV-1 gp120 trimers.* Nature, 2008. **455**(7209): p. 109-13.

188. Biscone, M.J., et al., *Functional impact of HIV coreceptor-binding site mutations.* Virology, 2006. **351**(1): p. 226-36.

189. Rizzuto, C. and J. Sodroski, *Fine definition of a conserved CCR5-binding region on the human immunodeficiency virus type 1 glycoprotein 120.* AIDS Res Hum Retroviruses, 2000. **16**(8): p. 741-9.

190. Rizzuto, C.D., et al., *A conserved HIV gp120 glycoprotein structure involved in chemokine receptor binding.* Science, 1998. **280**(5371): p. 1949-53.

191. Berger, E.A., et al., *A new classification for HIV-1.* Nature, 1998. **391**(6664): p. 240.

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

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

194. Buzon, V., et al., *Crystal structure of HIV-1 gp41 including both fusion peptide and membrane proximal external regions.* PLoS Pathog, 2010. **6**(5): p. e1000880.

195. Cilliers, T., et al., *Use of alternate coreceptors on primary cells by two HIV-1 isolates.* Virology, 2005. **339**(1): p. 136-44.

196. Lackner, A.A., M.M. Lederman, and B. Rodriguez, *HIV pathogenesis: the host.* Cold Spring Harb Perspect Med, 2012. **2**(9): p. a007005.

197. Schuitemaker, H., A.B. van 't Wout, and P. Lusso, *Clinical significance of HIV-1 coreceptor usage.* J Transl Med, 2011. **9 Suppl 1**: p. S5.

198. Manches, O., D. Frleta, and N. Bhardwaj, *Dendritic cells in progression and pathology of HIV infection.* Trends Immunol, 2014. **35**(3): p. 114-22.

199. Smed-Sorensen, A., et al., *Differential susceptibility to human immunodeficiency virus type 1 infection of myeloid and plasmacytoid dendritic cells.* J Virol, 2005. **79**(14): p. 8861-9.

200. Piguet, V. and R.M. Steinman, *The interaction of HIV with dendritic cells: outcomes and pathways.* Trends Immunol, 2007. **28**(11): p. 503-10.

201. Wu, L. and V.N. KewalRamani, *Dendritic-cell interactions with HIV: infection and viral dissemination.* Nat Rev Immunol, 2006. **6**(11): p. 859-68.

202. Koppensteiner, H., R. Brack-Werner, and M. Schindler, *Macrophages and their relevance in Human Immunodeficiency Virus Type I infection.* Retrovirology, 2012. **9**: p. 82.

203. Koenig, S., et al., *Detection of AIDS virus in macrophages in brain tissue from AIDS patients with encephalopathy.* Science, 1986. **233**(4768): p. 1089-93.

204. Groot, F., S. Welsch, and Q.J. Sattentau, *Efficient HIV-1 transmission from macrophages to T cells across transient virological synapses.* Blood, 2008. **111**(9): p. 4660-3.

205. Kawamura, T., et al., *R5 HIV productively infects Langerhans cells, and infection levels are regulated by compound CCR5 polymorphisms.* Proc Natl Acad Sci U S A, 2003. **100**(14): p. 8401-6.

206. Keele, B.F., et al., *Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection.* Proc Natl Acad Sci U S A, 2008. **105**(21): p. 7552-7.

207. Connor, R.I., et al., *Change in coreceptor use correlates with disease progression in HIV-1--infected individuals.* J Exp Med, 1997. **185**(4): p. 621-8.

208. Castellano, P., L. Prevedel, and E.A. Eugenin, *HIV-infected macrophages and microglia that survive acute infection become viral reservoirs by a mechanism involving Bim.* Sci Rep, 2017. **7**(1): p. 12866.

209. Bhattacharya, A., *Protein structures: Structures of desire.* Nature, 2009. **459**(7243): p. 24-7.

210. Mao, Y., et al., *Subunit organization of the membrane-bound HIV-1 envelope glycoprotein trimer.* Nat Struct Mol Biol, 2012. **19**(9): p. 893-9.

211. Mao, Y., et al., *Molecular architecture of the uncleaved HIV-1 envelope glycoprotein trimer.* Proc Natl Acad Sci U S A, 2013. **110**(30): p. 12438-43.

212. Subramaniam, S., *Structure of trimeric HIV-1 envelope glycoproteins.* Proc Natl Acad Sci U S A, 2013. **110**(45): p. E4172-4.

213. Mao, Y., L.R. Castillo-Menendez, and J.G. Sodroski, *Reply to Subramaniam, van Heel, and Henderson: Validity of the cryo-electron microscopy structures of the HIV-1 envelope glycoprotein complex.* Proc Natl Acad Sci U S A, 2013. **110**(45): p. E4178-82.

214. van Heel, M., *Finding trimeric HIV-1 envelope glycoproteins in random noise.* Proc Natl Acad Sci U S A, 2013. **110**(45): p. E4175-7.

215. Henderson, R., *Avoiding the pitfalls of single particle cryo-electron microscopy: Einstein from noise.* Proc Natl Acad Sci U S A, 2013. **110**(45): p. 18037-41.

216. Beddows, S., et al., *Evaluating the immunogenicity of a disulfide-stabilized, cleaved, trimeric form of the envelope glycoprotein complex of human immunodeficiency virus type 1.* J Virol, 2005. **79**(14): p. 8812-27.

217. Kang, Y.K., et al., *Structural and immunogenicity studies of a cleaved, stabilized envelope trimer derived from subtype A HIV-1.* Vaccine, 2009. **27**(37): p. 5120-32.

218. McCoy, L.E., et al., *Holes in the Glycan Shield of the Native HIV Envelope Are a Target of Trimer-Elicited Neutralizing Antibodies.* Cell Rep, 2016. **16**(9): p. 2327-38.

219. Voss, J.E., et al., *Elicitation of Neutralizing Antibodies Targeting the V2 Apex of the HIV Envelope Trimer in a Wild-Type Animal Model.* Cell Rep, 2017. **21**(1): p. 222-235.

220. Sanders, R.W., et al., *HIV-1 VACCINES. HIV-1 neutralizing antibodies induced by native-like envelope trimers.* Science, 2015. **349**(6244): p. aac4223.

221. Stewart-Jones, G.B., et al., *Trimeric HIV-1-Env Structures Define Glycan Shields from Clades A, B, and G.* Cell, 2016. **165**(4): p. 813-26.

222. Gristick, H.B., et al., *Natively glycosylated HIV-1 Env structure reveals new mode for antibody recognition of the CD4-binding site.* Nat Struct Mol Biol, 2016. **23**(10): p. 906-915.

223. Lee, J.H., G. Ozorowski, and A.B. Ward, *Cryo-EM structure of a native, fully glycosylated, cleaved HIV-1 envelope trimer.* Science, 2016. **351**(6277): p. 1043-8.

224. Sullivan, N., et al., *Effect of amino acid changes in the V1/V2 region of the human immunodeficiency virus type 1 gp120 glycoprotein on subunit association, syncytium formation, and recognition by a neutralizing antibody.* J Virol, 1993. **67**(6): p. 3674-9.

225. Walker, L.M., et al., *Broad neutralization coverage of HIV by multiple highly potent antibodies.* Nature, 2011. **477**(7365): p. 466-70.

226. Chojnacki, J., et al., *Maturation-dependent HIV-1 surface protein redistribution revealed by fluorescence nanoscopy.* Science, 2012. **338**(6106): p. 524-8.

227. Brandenberg, O.F., et al., *The HIV-1 Entry Process: A Stoichiometric View.* Trends Microbiol, 2015. **23**(12): p. 763-774.

228. Magnus, C., et al., *Estimating the stoichiometry of human immunodeficiency virus entry.* J Virol, 2009. **83**(3): p. 1523-31.

229. Herrera, C., et al., *Dominant-negative effect of hetero-oligomerization on the function of the human immunodeficiency virus type 1 envelope glycoprotein complex.* Virology, 2006. **351**(1): p. 121-32.

230. Klasse, P.J., *Modeling how many envelope glycoprotein trimers per virion participate in human immunodeficiency virus infectivity and its neutralization by antibody.* Virology, 2007. **369**(2): p. 245-62.

231. Yang, X., et al., *Stoichiometry of envelope glycoprotein trimers in the entry of human immunodeficiency virus type 1.* J Virol, 2005. **79**(19): p. 12132-47.

232. Sougrat, R., et al., *Electron tomography of the contact between T cells and SIV/HIV-1: implications for viral entry.* PLoS Pathog, 2007. **3**(5): p. e63.

233. Starcich, B.R., et al., *Identification and characterization of conserved and variable regions in the envelope gene of HTLV-III/LAV, the retrovirus of AIDS.* Cell, 1986. **45**(5): p. 637-48.

234. Jiang, X., et al., *Conserved structural elements in the V3 crown of HIV-1 gp120.* Nat Struct Mol Biol, 2010. **17**(8): p. 955-61.

235. Chen, B., et al., *Structure of an unliganded simian immunodeficiency virus gp120 core.* Nature, 2005. **433**(7028): p. 834-41.

236. Zolla-Pazner, S., et al., *Structure/Function Studies Involving the V3 Region of the HIV-1 Envelope Delineate Multiple Factors That Affect Neutralization Sensitivity.* J Virol, 2016. **90**(2): p. 636-49.

237. Powell, R.L.R., et al., *Plasticity and Epitope Exposure of the HIV-1 Envelope Trimer.* J Virol, 2017. **91**(17).

238. Zolla-Pazner, S. and T. Cardozo, *Structure-function relationships of HIV-1 envelope sequence-variable regions refocus vaccine design.* Nat Rev Immunol, 2010. **10**(7): p. 527-35.

239. Yuan, T., J. Li, and M.Y. Zhang, *HIV-1 envelope glycoprotein variable loops are indispensable for envelope structural integrity and virus entry.* PLoS One, 2013. **8**(8): p. e69789.

240. Wyatt, R., et al., *Functional and immunologic characterization of human immunodeficiency virus type 1 envelope glycoproteins containing deletions of the major variable regions.* J Virol, 1993. **67**(8): p. 4557-65.

241. Cao, J., et al., *Replication and neutralization of human immunodeficiency virus type 1 lacking the V1 and V2 variable loops of the gp120 envelope glycoprotein.* J Virol, 1997. **71**(12): p. 9808-12.

242. Perez, L.G., et al., *Envelope glycoprotein binding to the integrin alpha4beta7 is not a general property of most HIV-1 strains.* J Virol, 2014. **88**(18): p. 10767-77.

243. Walker, L.M., et al., *Broad and potent neutralizing antibodies from an African donor reveal a new HIV-1 vaccine target.* Science, 2009. **326**(5950): p. 285-9.

244. Gorny, M.K., et al., *Identification of a new quaternary neutralizing epitope on human immunodeficiency virus type 1 virus particles.* J Virol, 2005. **79**(8): p. 5232-7.

245. O'Connell, R.J., J.H. Kim, and J.L. Excler, *The HIV-1 gp120 V1V2 loop: structure, function and importance for vaccine development.* Expert Rev Vaccines, 2014. **13**(12): p. 1489-500.

246. Haynes, B.F., et al., *Immune-correlates analysis of an HIV-1 vaccine efficacy trial.* N Engl J Med, 2012. **366**(14): p. 1275-86.

247. Karasavvas, N., et al., *The Thai Phase III HIV Type 1 Vaccine trial (RV144) regimen induces antibodies that target conserved regions within the V2 loop of gp120.* AIDS Res Hum Retroviruses, 2012. **28**(11): p. 1444-57.

248. Gu, L., et al., *Adenoviral vectors elicit humoral immunity against variable loop 2 of clade C HIV-1 gp120 via "Antigen Capsid-Incorporation" strategy.* Virology, 2016. **487**: p. 75-84.

249. LaRosa, G.J., et al., *Conserved sequence and structural elements in the HIV-1 principal neutralizing determinant.* Science, 1990. **249**(4971): p. 932-5.

250. Chandramouli, B., et al., *Importance of V3 loop flexibility and net charge in the context of co-receptor recognition. A molecular dynamics study on HIV gp120.* J Biomol Struct Dyn, 2012. **29**(5): p. 879-91.

251. Watabe, T., et al., *Fold recognition of the human immunodeficiency virus type 1 V3 loop and flexibility of its crown structure during the course of adaptation to a host.* Genetics, 2006. **172**(3): p. 1385-96.

252. Sander, O., et al., *Structural descriptors of gp120 V3 loop for the prediction of HIV-1 coreceptor usage.* PLoS Comput Biol, 2007. **3**(3): p. e58.

253. Chiou, S.H., et al., *Studies on the role of the V3 loop in human immunodeficiency virus type 1 envelope glycoprotein function.* AIDS Res Hum Retroviruses, 1992. **8**(9): p. 1611-8.

254. Davis, K.L., et al., *High titer HIV-1 V3-specific antibodies with broad reactivity but low neutralizing potency in acute infection and following vaccination.* Virology, 2009. **387**(2): p. 414-26.

255. Gorny, M.K., et al., *The v3 loop is accessible on the surface of most human immunodeficiency virus type 1 primary isolates and serves as a neutralization epitope.* J Virol, 2004. **78**(5): p. 2394-404.

256. Hioe, C.E., et al., *Anti-V3 monoclonal antibodies display broad neutralizing activities against multiple HIV-1 subtypes.* PLoS One, 2010. **5**(4): p. e10254.

257. de Taeye, S.W., et al., *Immunogenicity of Stabilized HIV-1 Envelope Trimers with Reduced Exposure of Non-neutralizing Epitopes.* Cell, 2015. **163**(7): p. 1702-15.

258. de Taeye, S.W., et al., *Stabilization of the gp120 V3 loop through hydrophobic interactions reduces the immunodominant V3-directed non-neutralizing response to HIV-1 envelope trimers.* J Biol Chem, 2018. **293**(5): p. 1688-1701.

259. Gilbert, P.B., et al., *Correlation between immunologic responses to a recombinant glycoprotein 120 vaccine and incidence of HIV-1 infection in a phase 3 HIV-1 preventive vaccine trial.* J Infect Dis, 2005. **191**(5): p. 666-77.

260. Flynn, N.M., et al., *Placebo-controlled phase 3 trial of a recombinant glycoprotein 120 vaccine to prevent HIV-1 infection.* J Infect Dis, 2005. **191**(5): p. 654-65.

261. Pitisuttithum, P., et al., *Randomized, double-blind, placebo-controlled efficacy trial of a bivalent recombinant glycoprotein 120 HIV-1 vaccine*

among injection drug users in Bangkok, Thailand. J Infect Dis, 2006. **194**(12): p. 1661-71.

262. Gray, G.E., et al., *Safety and efficacy of the HVTN 503/Phambili study of a clade-B-based HIV-1 vaccine in South Africa: a double-blind, randomised, placebo-controlled test-of-concept phase 2b study.* Lancet Infect Dis, 2011. **11**(7): p. 507-15.

263. Buchbinder, S.P., et al., *Efficacy assessment of a cell-mediated immunity HIV-1 vaccine (the Step Study): a double-blind, randomised, placebo-controlled, test-of-concept trial.* Lancet, 2008. **372**(9653): p. 1881-1893.

264. Rerks-Ngarm, S., et al., *Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand.* N Engl J Med, 2009. **361**(23): p. 2209-20.

265. Jones, A.T., et al., *A Trimeric HIV-1 Envelope gp120 Immunogen Induces Potent and Broad Anti-V1V2 Loop Antibodies against HIV-1 in Rabbits and Rhesus Macaques.* J Virol, 2018. **92**(5).

266. Hammer, S.M., et al., *Efficacy trial of a DNA/rAd5 HIV-1 preventive vaccine.* N Engl J Med, 2013. **369**(22): p. 2083-92.

267. Hu, D., et al., *Tryptophan 375 stabilizes the outer-domain core of gp120 for HIV vaccine immunogen design.* Vaccine, 2017. **35**(23): p. 3067-3075.

268. Torrents de la Pena, A., et al., *Immunogenicity in Rabbits of HIV-1 SOSIP Trimers from Clades A, B, and C, Given Individually, Sequentially, or in Combination.* J Virol, 2018. **92**(8).

269. Klasse, P.J., et al., *Sequential and Simultaneous Immunization of Rabbits with HIV-1 Envelope Glycoprotein SOSIP.664 Trimers from Clades A, B and C.* PLoS Pathog, 2016. **12**(9): p. e1005864.

270. Ringe, R.P., et al., *Reducing V3 Antigenicity and Immunogenicity on Soluble, Native-Like HIV-1 Env SOSIP Trimers.* J Virol, 2017. **91**(15).

271. Forsell, M.N., W.R. Schief, and R.T. Wyatt, *Immunogenicity of HIV-1 envelope glycoprotein oligomers.* Curr Opin HIV AIDS, 2009. **4**(5): p. 380-7.

272. Chuang, G.Y., et al., *Structure-Based Design of a Soluble Prefusion-Closed HIV-1 Env Trimer with Reduced CD4 Affinity and Improved Immunogenicity.* J Virol, 2017. **91**(10).

273. Sanders, R.W. and J.P. Moore, *Native-like Env trimers as a platform for HIV-1 vaccine design.* Immunol Rev, 2017. **275**(1): p. 161-182.

274. Julien, J.P., et al., *Asymmetric recognition of the HIV-1 trimer by broadly neutralizing antibody PG9.* Proc Natl Acad Sci U S A, 2013. **110**(11): p. 4351-6.

275. Pancera, M., et al., *Structure and immune recognition of trimeric pre-fusion HIV-1 Env.* Nature, 2014. **514**(7523): p. 455-61.

276. Kovacs, J.M., et al., *HIV-1 envelope trimer elicits more potent neutralizing antibody responses than monomeric gp120.* Proc Natl Acad Sci U S A, 2012. **109**(30): p. 12111-6.

277. Cheng, C., et al., *Immunogenicity of a Prefusion HIV-1 Envelope Trimer in Complex with a Quaternary-Structure-Specific Antibody.* J Virol, 2015. **90**(6): p. 2740-55.

278. Medina-Ramirez, M., R.W. Sanders, and Q.J. Sattentau, *Stabilized HIV-1 envelope glycoprotein trimers for vaccine use.* Curr Opin HIV AIDS, 2017. **12**(3): p. 241-249.

279. Karlsson Hedestam, G.B., et al., *The challenges of eliciting neutralizing antibodies to HIV-1 and to influenza virus.* Nat Rev Microbiol, 2008. **6**(2): p. 143-55.

280. Graham, F.L., et al., *Characteristics of a human cell line transformed by DNA from human adenovirus type 5.* J Gen Virol, 1977. **36**(1): p. 59-74.

281. Wei, X., et al., *Emergence of resistant human immunodeficiency virus type 1 in patients receiving fusion inhibitor (T-20) monotherapy.* Antimicrob Agents Chemother, 2002. **46**(6): p. 1896-905.

282. Derdeyn, C.A., et al., *Sensitivity of human immunodeficiency virus type 1 to the fusion inhibitor T-20 is modulated by coreceptor specificity defined by the V3 loop of gp120.* J Virol, 2000. **74**(18): p. 8358-67.

283. Platt, E.J., et al., *Effects of CCR5 and CD4 cell surface concentrations on infections by macrophagetropic isolates of human immunodeficiency virus type 1.* J Virol, 1998. **72**(4): p. 2855-64.

284. Montefiori, D.C., *Measuring HIV neutralization in a luciferase reporter gene assay.* Methods Mol Biol, 2009. **485**: p. 395-405.

285. Nelson-Rees, W.A., et al., *Source, alterations, characteristics and use of a new dog cell line (Cf2Th).* In Vitro, 1976. **12**(10): p. 665-9.

286. Kong, L.I., et al., *West African HIV-2-related human retrovirus with attenuated cytopathicity.* Science, 1988. **240**(4858): p. 1525-9.

287. Kumar, P., et al., *Molecular characterization of an attenuated human immunodeficiency virus type 2 isolate.* J Virol, 1990. **64**(2): p. 890-901.

288. Soares, M.A., et al., *A full-length and replication-competent proviral clone of SIVAGM from tantalus monkeys.* Virology, 1997. **228**(2): p. 394-9.

289. Johnson, P.R., et al., *Simian immunodeficiency viruses from African green monkeys display unusual genetic diversity.* J Virol, 1990. **64**(3): p. 1086-92.

290. Ivey-Hoyle, M., et al., *Envelope glycoproteins from biologically diverse isolates of immunodeficiency viruses have widely different affinities for CD4.* Proc Natl Acad Sci U S A, 1991. **88**(2): p. 512-6.

291. Reimann, K.A., et al., *Immunopathogenic events in acute infection of rhesus monkeys with simian immunodeficiency virus of macaques.* J Virol, 1994. **68**(4): p. 2362-70.

292. Mulligan, M.J., et al., *Human immunodeficiency virus type 2 envelope glycoprotein: differential CD4 interactions of soluble gp120 versus the assembled envelope complex.* Virology, 1992. **187**(1): p. 233-41.

293. Finzi, A., et al., *Topological layers in the HIV-1 gp120 inner domain regulate gp41 interaction and CD4-triggered conformational transitions.* Mol Cell, 2010. **37**(5): p. 656-67.

294. Parolin, C., et al., *Use of cis- and trans-acting viral regulatory sequences to improve expression of human immunodeficiency virus vectors in human lymphocytes.* Virology, 1996. **222**(2): p. 415-22.

295. Helseth, E., et al., *Rapid complementation assays measuring replicative potential of human immunodeficiency virus type 1 envelope glycoprotein mutants.* J Virol, 1990. **64**(5): p. 2416-20.

296. Moore, J.P., *Simple methods for monitoring HIV-1 and HIV-2 gp120 binding to soluble CD4 by enzyme-linked immunosorbent assay: HIV-2 has a 25-fold lower affinity than HIV-1 for soluble CD4.* AIDS, 1990. **4**(4): p. 297-305.

297. Clapham, P.R., et al., *Soluble CD4 blocks the infectivity of diverse strains of HIV and SIV for T cells and monocytes but not for brain and muscle cells.* Nature, 1989. **337**(6205): p. 368-70.

298. Bahraoui, E., et al., *Study of the interaction of HIV-1 and HIV-2 envelope glycoproteins with the CD4 receptor and role of N-glycans.* AIDS Res Hum Retroviruses, 1992. **8**(5): p. 565-73.

299. Hill, C.M., et al., *Envelope glycoproteins from human immunodeficiency virus types 1 and 2 and simian immunodeficiency virus can use human CCR5 as a coreceptor for viral entry and make direct CD4-dependent interactions with this chemokine receptor.* J Virol, 1997. **71**(9): p. 6296-304.

300. Trkola, A., et al., *CD4-dependent, antibody-sensitive interactions between HIV-1 and its co-receptor CCR-5.* Nature, 1996. **384**(6605): p. 184-7.

301. Wu, L., et al., *CD4-induced interaction of primary HIV-1 gp120 glycoproteins with the chemokine receptor CCR-5.* Nature, 1996. **384**(6605): p. 179-83.

302. Basmaciogullari, S., et al., *Identification of conserved and variable structures in the human immunodeficiency virus gp120 glycoprotein of importance for CXCR4 binding.* J Virol, 2002. **76**(21): p. 10791-800.

303. Shioda, T., J.A. Levy, and C. Cheng-Mayer, *Small amino acid changes in the V3 hypervariable region of gp120 can affect the T-cell-line and macrophage tropism of human immunodeficiency virus type 1.* Proc Natl Acad Sci U S A, 1992. **89**(20): p. 9434-8.

304. Huang, C.C., et al., *Structure of a V3-containing HIV-1 gp120 core.* Science, 2005. **310**(5750): p. 1025-8.

305. Cashin, K., et al., *CoRSeqV3-C: a novel HIV-1 subtype C specific V3 sequence based coreceptor usage prediction algorithm.* Retrovirology, 2013. **10**: p. 24.

306. Hartley, O., et al., *V3: HIV's switch-hitter.* AIDS Res Hum Retroviruses, 2005. **21**(2): p. 171-89.

307. Foda, M., S. Harada, and Y. Maeda, *Role of V3 independent domains on a dualtropic human immunodeficiency virus type 1 (HIV-1) envelope gp120 in CCR5 coreceptor utilization and viral infectivity.* Microbiol Immunol, 2001. **45**(7): p. 521-30.

308. Cocchi, F., et al., *The V3 domain of the HIV-1 gp120 envelope glycoprotein is critical for chemokine-mediated blockade of infection.* Nat Med, 1996. **2**(11): p. 1244-7.

309. Cardozo, T., et al., *Structural basis for coreceptor selectivity by the HIV type 1 V3 loop.* AIDS Res Hum Retroviruses, 2007. **23**(3): p. 415-26.

310. Buonaguro, L., M.L. Tornesello, and F.M. Buonaguro, *Human immunodeficiency virus type 1 subtype distribution in the worldwide epidemic: pathogenetic and therapeutic implications.* J Virol, 2007. **81**(19): p. 10209-19.

311. Taylor, B.S., et al., *The challenge of HIV-1 subtype diversity.* N Engl J Med, 2008. **358**(15): p. 1590-602.

312. Shet, A., P. Nagaraja, and N.M. Dixit, *Viral Decay Dynamics and Mathematical Modeling of Treatment Response: Evidence of Lower in vivo Fitness of HIV-1 Subtype C.* J Acquir Immune Defic Syndr, 2016. **73**(3): p. 245-251.

313. Zolla-Pazner, S., *Improving on nature: focusing the immune response on the V3 loop.* Hum Antibodies, 2005. **14**(3-4): p. 69-72.

314. White-Scharf, M.E., et al., *Broadly neutralizing monoclonal antibodies to the V3 region of HIV-1 can be elicited by peptide immunization.* Virology, 1993. **192**(1): p. 197-206.

315. Krachmarov, C.P., et al., *V3-specific polyclonal antibodies affinity purified from sera of infected humans effectively neutralize primary isolates of human immunodeficiency virus type 1.* AIDS Res Hum Retroviruses, 2001. **17**(18): p. 1737-48.

316. Xiang, S.H., et al., *Characterization of a dual-tropic human immunodeficiency virus (HIV-1) strain derived from the prototypical X4 isolate HXBc2.* Virology, 2013. **438**(1): p. 5-13.

317. Huang, C.C., et al., *Structures of the CCR5 N terminus and of a tyrosine-sulfated antibody with HIV-1 gp120 and CD4.* Science, 2007. **317**(5846): p. 1930-4.

318. Cimbro, R., et al., *Tyrosine sulfation in the second variable loop (V2) of HIV-1 gp120 stabilizes V2-V3 interaction and modulates neutralization sensitivity.* Proc Natl Acad Sci U S A, 2014. **111**(8): p. 3152-7.

319. Wu, L., et al., *Interaction of chemokine receptor CCR5 with its ligands: multiple domains for HIV-1 gp120 binding and a single domain for chemokine binding.* J Exp Med, 1997. **186**(8): p. 1373-81.

320. Maeda, K., et al., *Involvement of the second extracellular loop and transmembrane residues of CCR5 in inhibitor binding and HIV-1 fusion: insights into the mechanism of allosteric inhibition.* J Mol Biol, 2008. **381**(4): p. 956-74.

321. Abayev, M., et al., *An extended CCR5 ECL2 peptide forms a helix that binds HIV-1 gp120 through non-specific hydrophobic interactions.* FEBS J, 2015. **282**(10): p. 1906-1921.

322. Korber B., F.B.T., Kuiken C., Pillai S.K., Sodroski J.G., *Numbering Positions in HIV Relative to HXB2CG.* The Human Retroviruses and AIDS

1998 Compendium Los Alamos National Laboratories, 1998: p. III-102- III-111.

323. Mourez, T., F. Simon, and J.C. Plantier, *Non-M variants of human immunodeficiency virus type 1.* Clin Microbiol Rev, 2013. **26**(3): p. 448-61.

324. Gao, F., et al., *Human infection by genetically diverse SIVSM-related HIV-2 in west Africa.* Nature, 1992. **358**(6386): p. 495-9.

325. de Silva, T.I., et al., *Potent autologous and heterologous neutralizing antibody responses occur in HIV-2 infection across a broad range of infection outcomes.* J Virol, 2012. **86**(2): p. 930-46.

326. Kong, R., et al., *Broad and potent neutralizing antibody responses elicited in natural HIV-2 infection.* J Virol, 2012. **86**(2): p. 947-60.

327. Ozkaya Sahin, G., et al., *Potent intratype neutralizing activity distinguishes human immunodeficiency virus type 2 (HIV-2) from HIV-1.* J Virol, 2012. **86**(2): p. 961-71.

328. Lee, J.H., et al., *A Broadly Neutralizing Antibody Targets the Dynamic HIV Envelope Trimer Apex via a Long, Rigidified, and Anionic beta-Hairpin Structure.* Immunity, 2017. **46**(4): p. 690-702.

329. Blattner, C., et al., *Structural delineation of a quaternary, cleavage-dependent epitope at the gp41-gp120 interface on intact HIV-1 Env trimers.* Immunity, 2014. **40**(5): p. 669-80.

330. Derking, R., et al., *Comprehensive antigenic map of a cleaved soluble HIV-1 envelope trimer.* PLoS Pathog, 2015. **11**(3): p. e1004767.

331. Huang, J., et al., *Broad and potent HIV-1 neutralization by a human antibody that binds the gp41-gp120 interface.* Nature, 2014. **515**(7525): p. 138-42.

332. Barnes, C.O., et al., *Structural characterization of a highly-potent V3-glycan broadly neutralizing antibody bound to natively-glycosylated HIV-1 envelope.* Nat Commun, 2018. **9**(1): p. 1251.

333. Willey, R.L., et al., *Identification of conserved and divergent domains within the envelope gene of the acquired immunodeficiency syndrome retrovirus.* Proc Natl Acad Sci U S A, 1986. **83**(14): p. 5038-42.

334. Catasti, P., et al., *Local and global structural properties of the HIV-MN V3 loop.* J Biol Chem, 1995. **270**(5): p. 2224-32.

335. Torres, O.B., et al., *Heroin-HIV-1 (H2) vaccine: induction of dual immunologic effects with a heroin hapten-conjugate and an HIV-1*

envelope V2 peptide with liposomal lipid A as an adjuvant. NPJ Vaccines, 2017. **2**: p. 13.

336. Mansfield, K.G., et al., *Origins of simian immunodeficiency virus infection in macaques at the New England Regional Primate Research Center.* J Med Primatol, 1995. **24**(3): p. 116-22.

337. Hirsch, V.M., et al., *An African primate lentivirus (SIVsm) closely related to HIV-2.* Nature, 1989. **339**(6223): p. 389-92.

338. Courgnaud, V., et al., *Characterization of a novel simian immunodeficiency virus from guereza colobus monkeys (Colobus guereza) in Cameroon: a new lineage in the nonhuman primate lentivirus family.* J Virol, 2001. **75**(2): p. 857-66.

339. Freed, E.O., D.J. Myers, and R. Risser, *Mutational analysis of the cleavage sequence of the human immunodeficiency virus type 1 envelope glycoprotein precursor gp160.* J Virol, 1989. **63**(11): p. 4670-5.

340. McCune, J.M., et al., *Endoproteolytic cleavage of gp160 is required for the activation of human immunodeficiency virus.* Cell, 1988. **53**(1): p. 55-67.

341. Rudicell, R.S., et al., *Impact of simian immunodeficiency virus infection on chimpanzee population dynamics.* PLoS Pathog, 2010. **6**(9): p. e1001116.

342. Rudicell, R.S., et al., *High prevalence of simian immunodeficiency virus infection in a community of savanna chimpanzees.* J Virol, 2011. **85**(19): p. 9918-28.

343. Etienne, L., et al., *Characterization of a new simian immunodeficiency virus strain in a naturally infected Pan troglodytes troglodytes chimpanzee with AIDS related symptoms.* Retrovirology, 2011. **8**: p. 4.

344. Jurkiewicz, E., et al., *Identification of the V1 region as a linear neutralizing epitope of the simian immunodeficiency virus SIVmac envelope glycoprotein.* J Virol, 1997. **71**(12): p. 9475-81.

345. Petry, H., et al., *Naturally occurring V1-env region variants mediate simian immunodeficiency virus SIVmac escape from high-titer neutralizing antibodies induced by a protective subunit vaccine.* J Virol, 2000. **74**(23): p. 11145-52.

346. Ondoa, P., et al., *Genetic variability of the V1 and V2 env domains of SIVcpz-ant and neutralization pattern of plasma viruses in a chimpanzee infected naturally.* J Med Virol, 2001. **65**(4): p. 765-76.

347. Curlin, M.E., et al., *HIV-1 envelope subregion length variation during disease progression.* PLoS Pathog, 2011. **6**(12): p. e1001228.

348. Yang, X., R. Wyatt, and J. Sodroski, *Improved elicitation of neutralizing antibodies against primary human immunodeficiency viruses by soluble stabilized envelope glycoprotein trimers.* J Virol, 2001. **75**(3): p. 1165-71.

349. Dogo-Isonagie, C., et al., *Peptides from second extracellular loop of C-C chemokine receptor type 5 (CCR5) inhibit diverse strains of HIV-1.* J Biol Chem, 2012. **287**(18): p. 15076-86.

350. Bobyk, K.D., et al., *Design of HIV Coreceptor Derived Peptides That Inhibit Viral Entry at Submicromolar Concentrations.* Mol Pharm, 2017. **14**(8): p. 2681-2689.

351. Cormier, E.G. and T. Dragic, *The crown and stem of the V3 loop play distinct roles in human immunodeficiency virus type 1 envelope glycoprotein interactions with the CCR5 coreceptor.* J Virol, 2002. **76**(17): p. 8953-7.