Title
The prevention of HIV infection and the stimulation of cells latently infected with HIV

Author
Fischer, Kathryn Elyse

Publication Date
2020

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, MERCED

The prevention of HIV infection and the stimulation of cells latently infected with HIV

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Quantitative and Systems Biology

by

Kathryn Elyse Fischer

Committee in charge:

Professor Kirk Jensen, Chair of Advisory Committee

Professor Katrina Hoyer

Professor Maria Zoghbi

Professor Patricia LiWang, Supervisor

2020

The dissertation of Kathryn Elyse Fischer, titled, “The prevention of HIV infection and the stimulation of cells latently infected with HIV”, is approved, and is acceptable ion quality and form for publication on microfilm and electronically:

_____ Date _____

Professor Maria Zoghbi

_____ Date _____

Professor Katrina Hoyer

Supervisor _____ Date _____

Professor Patricia J. LiWang

Chair _____ Date _____

Professor Kirk Jensen

University of California, Merced

2020

Dedication

The journey is always made better by the company kept.

I cannot imagine a more loving, supportive or encouraging family than my mother Lynn, father Felix and brother Ryan. My family means more to me than I can express in words and have taught me more than I have space to list here, so I will limit myself to the lessons that were critical for my success in graduate school. Thank you to my mother for teaching me perseverance and that it is always okay to forge my own path. Thank you to my father for teaching me about quiet strength and the power of calm, analytical thought. Thank you to my brother for teaching me to look for the humor in every situation and that a smile always makes the darkest days better.

A gigantic thank you to all of my friends who were there, patiently waiting, for me when I was able to find time outside the lab. While all of my friends are amazing, there are a few incredible, inspiring women I must mention by name. Danielle, thank you for always taking the time to check in on me, you have always been a shining example of how to be a better friend and will always be a sister of my heart. Noelle, thank you for welcoming me so whole heartedly into the local theater family and lightening my stress by lending a sympathetic ear. Katy and Jessica, thank you for inspiring me and convincing me to continue to pursue creative outlets, outside of the lab, that kept me sane through the trials of research.

Last but not least, this dedication would not be complete without my partner Jon, whose love and support have sustained me through failed experiments, lack of sleep and grad school induced insanity, all while pursuing his own degree.
Jon, I love you more than I can say.

Table of Contents

I. List of Abbreviations

293-FT	Is a human embryonal kidney derived cell line transformed with the SZ40 large T antigen that allows for fast growth and high rates of transfection
AIDS	Acquired Immune Deficiency Syndrome
ARV	AIDS-associated retrovirus (an early name for HIV)
β-gal	β-galactosidase
βME	Beta-mercaptoethanol
bNAb	Broadly Neutralizing Antibodies
bp	Base Pair
cART	Combination Antiretroviral Therapy
CBD	Carbohydrate Binding Domain
CBS	Carbohydrate Binding Site
CCR5	C-C Chemokine receptor type 5 (originally referred to as fusin)
CD4	Cluster of Differentiation 4
CD8	Cluster of Differentiation 8
CDC	Center for Disease Control
CI	Combination Index
Conc.	Concentration
CPRG	Chlorophenol Red-β-D-galactopyranoside
CRS	Chemokine Recognition Site
CXCR4	C-X-C Chemokine receptor type 4 (originally referred to as fusin)
D1-D4	The four N terminal immunoglobulin domains of CD4. Numbered from N to C.
DART	Dual-Affinity Retargeting
DMEM	Dulbecco Modified Eagle Medium
DNA	Deoxynucleic Acid
dNTP	deoxyriboNucleotide TriPhosphate
EDTA	Ethylenediaminetetraacetic acid, also known as 2-[2-[bis(carboxymethyl)amino]ethyl-(carboxymethyl)amino]acetic acid

EK	Enterokinase
ELISA	Enzyme-Linked ImmunoSorbent Assay
ENV	HIV Envelope protein gp160
ER	Endoplasmic Reticulum
FACS	Fluorescence-activated cell sorting
FBS	Fetal Bovine Serum
Gag	Group-specific antigen
GFP	Green Fluorescent Protein
gp	Glycoprotein
Grft	Griffithsin
HCV	Hepatitis C Virus
HDAC	Histone Deacetylase
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HIV	Human Immunodeficiency Virus
HMT	Histone Methyltransferase
HPV	Human Papilloma Virus
HSV-2	Herpes Simplex Virus 2
HTLV	Human T-cell leukemia-lymphoma virus (an early name for HIV)
IC_{50}	Half maximal Inhibitory Concentration
IgG	Immunoglobulin class G1
IPTG	Isopropyl β-D-1-thiogalactopyranoside
JEV	Japanese Encephalitis Virus
J-Lat	Jurkat cells latently infected with an LTR linked GFP
JLTRG	Jurkat cell line with LTR linked GFP
JLTRG +R5	Jurkat cell line with LTR linked GFP with expression of CCR5
KCl	Potassium Chloride
LAV	Lymphadenopathy-associated virus (an early name for HIV)
LRA	Latency Reversal Agent
MAGI	HeLa cells with CD4 and LTR β-galactosidase

MERS-CoV	Middle Eastern Respiratory Syndrome - Coronavirus
MIP-1α	Macrophage Inflammatory Protein 1 alpha, officially referred to as CCL3 (C-C chemokine ligand type 3)
MIP-1β	Macrophage Inflammatory Protein 1 beta, officially referred to as CCL4 (C-C chemokine ligand type 4)
MMWR	Morbidity and Mortality Weekly Report
MVL	*Microcyctis vividis* lectin
NaCl	Sodium Chloride
NaOAc	Sodium Acetate
NaPi	Sodium Phosphate
NEB	New England Biolabs
NF	Nuclear Factor
NIAID	National Institute of Allergy and Infectious Diseases
NLS	Nuclear Localization Signal
NMR	Nuclear Magnetic Resonance
NNRTI	Non-Nucleoside Reverse Transcriptase Inhibitor
NRTI	Nucleoside Reverse Transcriptase Inhibitor
Oligo	Oligonucleotide
p24	HIV capsid protein, molecular weight of 24 kDa
PBS	Phosphate-buffered Saline
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
Pol	Polyprotein precursor to Reverse Transcriptase, Integrase and HIV Protease
pSG3$^{\Delta ENV}$	A proviral clone of HIV that contains a defective vpu gene and a premature stop codon in the ENV gene. Tends to delete viral sequence during propagation in E. coli[3]
PTD	Protein Transduction Domain
P-TEFb	Positive Transcription Elongation Factor b
Q	Quadrant
RANTES	Common name for CCL5 (C-C chemokine ligand type 5)

RT	Reverse Transcriptase
SARS-CoV	Sever Acute Respiratory Syndrome - Coronavirus
SDS	Sodium Dodecyl Sulfate
SIV	Simian Immunodeficiency Virus
SIVcpz	Simian Immunodeficiency Virus Chimpanzee
SIVcpzPtt	Simian Immunodeficiency Virus Chimpanzee *Pan troglodytes*
SIVgor	Simian Immunodeficiency Virus Gorilla
SIVsm	Simian Immunodeficiency Virus Sooty Mangabeys
T4	CD4 expressing T-cells
TAR	Transcription Response Element
Tat	Trans-Activator of Transcription, transcription regulatory protein
T-cell	Lymphocytes that are produced and/or processed in the Thymus
TCR	T cell Receptor
T/F	Transmitted/Founder strains of HIV
Tm	Melting Temperature
Tris	Tris(hydroxymethyl)aminomethane, also known as 2-Amino-2-(hydroxymethyl)-1,3-propanediol
Trx	Thioredoxin
TZM-bl	HeLa cell line, previously designated JC53-b1, that expresses CD4 & CCR5 with integrated copies of the luciferase and β-galactosidase genes under the control of the HIV-1 promoter.
UCLA	University of California, Los Angeles
UCM	University of California, Merced
V1-V5	Variable regions 1-5 of the HIV gp120 protein
vCCI	Viral C-C Chemokine Inhibitor
Wt	Wild type
WW	World Wide

II. List of Figures

III. List of Tables

IV. Acknowledgments

I would like to acknowledge and thank the following people for their technical assistance: Kimberly Nguyen, Sage Sorenson, James McIntyre, Robert Amidon, Julissa Garcia, Giselle San Roman and David Chen. Thank you to Arjan Bains for fighting through these tough projects alongside me. Thank you to Dr. Jessica Morgan for your guidance during your time at UC Merced. Thank you to Dr. Anna Nguyen, Dr. Li Zhang, Dr. Yongong Chang for training and guiding me in my early years in graduate school. Thank you to Dr. Shahar Sukenik and Karina Guadalupe for their assistance obtaining microscopy images of my cells. This work was funded by the National Institutes of Health 1R01AI112011.

My sincerest thanks to Drs. Patricia and Andy LiWang, as well as Dr. Katrina Hoyer, Dr. Maria Zoghbi and Dr. Kirk Jensen for their patience, wisdom and guidance throughout my graduate career.

V. Vita for Kathryn Elyse Fischer

Education:

2014-2020: Ph.D. & Masters University of California Merced, Quantitative and Systems Biology with research done under the supervision of Professor Patricia J. LiWang

- A biochemical approach to prevention of HIV infection and resurgence utilizing natural proteins.

2011-2013: Bachelor of Science, University of California Davis, Biological Sciences – Emphasis in Molecular and Cellular Biology with research done under the supervision of Professor Li Tian

- The study of the biofortification of pro-vitamin A in wheat to help combat Vitamin A deficiencies in 3rd world countries utilizing classic breeding techniques.

List of Publications:

Fischer, K. (2020) The prevention of HIV infection and the stimulation of cells latently infected with HIV: A dissertation submitted in partial satisfaction of the requirements for the degree of Doctor of Philosophy in Quantitative Systems and Biology. *Dissertation, University of California, Merced*, 1-201.

Fischer, K., Nguyen, K., LiWang, P.J. (2019) Griffithsin retains anti-HIV-1 potency with changes in gp120 glycosylation and complements broadly neutralizing antibodies PGT121 and PGT126. *Antimicrob. Agents Chemother. 64*:e01084-19

Qin X., **Fischer, K**., Yu, S., Dubcosvsky, J., Tian, L. (2016) Distinct expression and function of carotenoid metabolic genes and homoeologs in developing wheat grains. *BMC Plant Biol*. *16*, 155.

Qin X., **Fischer, K**., Dubcosvsky, J., Tian, L. (2016) Endosperm Carotenoid Concentrations in Wheat are Better Correlated with Transcript Levels than Enzyme Activities. *Crop Sci. 56*, 3173.

Awards and Fellowships:

Dean's Dissertation Fellowship Fall 2019

Quantitative and Systems Graduate Program Summer Fellowship 2015-2017 & 2019

VI. Abstract

The prevention of HIV infection and the stimulation of cells latently infected with HIV

Kathryn Elyse Fischer

Doctor of Philosophy

University of California, Merced

2020

Supervisor: Professor Patricia LiWang

Thirty-five years after its exposure, HIV is still a global health threat with millions of new infections annually. The goal of our research is to stop the global spread of HIV, first by preventing new HIV infections and then by working to cure existing infections. Chapter one reviews the history of HIV, the infection and transmission process and the current approaches for the prevention and cure of HIV so that the reader will have a thorough understanding of the molecular workings of HIV. I also address the origin, discovery and classification of HIV. In chapter 2 we investigated the anti-HIV lectin Griffithsin that inhibits HIV by binding to high mannose glycans on the viral envelope protein gp120. Mutational studies that removed or shifted glycans on gp120 showed that Griffithsin retains anti-HIV-1 potency despite changes in gp120 glycosylation. We found through HIV capture ELISAs that Griffithsin's ability to alter the structure of gp120, exposing the CD4 binding site, correlated with the presence of glycosylation at N295 only in B clade strains, not C clade strains. We further demonstrated that Griffithsin can rescue the activity of broadly neutralizing antibodies PGT121 and PGT126 in the event of loss or shift of glycosylation at N332, where these broadly neutralizing antibodies suffer a drastic loss of potency. Chapter 3 reviews the purification and testing of a chimeric protein composed of the HIV transcriptional activator protein Tat, and a variant of a human chemokine, 5P14-RANTES, which binds to CCR5, a receptor that HIV utilizes during the infection process. This chimeric protein was designed with the theoretical purpose of targeting to and then activating cells latently infected with HIV. We attempted to produce and purify the 5P14-Linker-Tat protein as a full-length construct and by linking the two separate proteins with a sortase enzyme. Regardless of the method used, the protein was difficult to produce in large quantities and the final product contained protein contaminants. Unfortunately, when we tested our construct (in both orientations, 5P14-Linker-Tat and Tat-Linker-5P14) using a MAGI colorimetric assay, we found that it did not show targeting to CCR5 cells and little to no activation of the HIV LTR linked GFP.

Chapter 1

Introduction

1.1 Significance of Research on HIV

Since the discovery of AIDS in 1981, millions of papers have been published and great strides have been made in understanding the origin, spread and infection process of HIV. Elucidation of the mechanism of infection has allowed researchers to develop drug regimens that significantly slow the progression of AIDS and even prevent transmission to patient's uninfected partners and children, allowing people living with HIV to lead relatively normal lives. Researchers continually strive to understand more about HIV's origins, its infection process and most importantly, how to prevent and cure HIV infection. Although we have learned a great deal about AIDS and HIV in the last thirty-five years, there is still no vaccine to prevent new infections and no cure for existing patients.

The benefits of HIV research extend beyond preventing and curing HIV and AIDS[4,5]. Research on HIV's degradation of the immune system has informed general immunology on the impacts of $CD4^+$ T cells in other infectious diseases[6,7]. The inflammation triggered by HIV has driven research into heart disease and other disorders triggered by persistent immune activation and inflammation[8–15]. Modified HIV virus has been used to genetically modify T cells to treat certain forms of leukemia and genetic immune disorders[16–18]. Protease, reverse transcriptase and integrase inhibitors developed to combat HIV equipped researchers to create drugs for hepatitis B & C (HBV & HCV)[19–22] and have advanced the targeted drug and vaccine development field as a whole[23–25]. While HIV vaccine development has not been successful to produce a viable preventative for HIV, the research into broadly neutralizing antibodies has been applicable towards autoimmune and chronic diseases such as Alzheimer's[26–33].

1.2 The History of HIV

1.2.1 Discovery of AIDS

AIDS was identified as a new disease in 1981, first by Dr. Gottlieb at the UCLA Medical Center[34,35], then 2 months later by Dr. Hymes et al from various medical facilities in New York[36]. Dr. Gottlieb published his results initially in the Morbidity and Mortality Weekly Report (MMWR), detailing the cases of five homosexual men who contracted *Pneumocystis* pneumonia, noting that the disease was normally only found in severely immunosuppressed patients and that there may a "cellular-immune dysfunction. . . that predisposes individuals to opportunistic infections"[34]. This was the first hypothesis that there was an immunodeficiency disease underlying the increased opportunistic infections seen in at risk populations such as homosexuals and intravenous drug users. Barely a month later the Center for Disease Control (CDC) put out another report detailing the cases of 26 homosexual men found in New York and Los Angeles that had various combinations of Karposi's Sarcoma, *Pneumocystis* Pneumonia and cytomegalovirus[37]. Hymes et al reported 8 patients, aged 27-45 years, who displayed highly aggressive forms of Kaposi's sarcoma. The high prevalence of Karposi's sarcoma in young, otherwise healthy people was highly unusual for several reasons. Firstly, Kaposi's sarcoma was rare

in the US, and generally found in patients in their 70's[36,38]. Secondly, the legions associated with Kaposi's sarcoma are normally found on patients legs, whereas the 8 men in Hymes's report presented lesions over their entire bodies[36,38]. Lastly, the normal survival period after diagnosis with Kaposi's sarcoma in the 1970-1980s was 8-13 years[39], but half of the patients whose cases were reported were dead within 20 months of diagnosis[36]. In September of 1982, a CDC MMWR coined the term Acquired Immune Deficiency Syndrome (AIDS) for the underlying disease causing a depletion of the immune system, resulting in secondary infections uncommonly seen in otherwise seemingly healthy individuals[40]. The report indicated that nearly 600 cases had been reported and of those, over 40% of those patient had died[40].

Though it was initially thought that AIDS was a gay-related disease, cases of AIDS were rapidly found in intravenous drug users (both male and female), hemophiliacs and even newborn children[40–46]. Infection studies in Macaques and surveillance data of people infected with AIDS demonstrated that AIDS could be transmitted through vaginal or anal sexual activity, blood transfusions, parenteral exposure via needles, prenatally and through breastmilk[40–49]. Within two years of the discovery of AIDS, there were over a thousand reported cases and nearly 700 AIDS related deaths, but the cause of this terrible disease had yet to be determined[45].

1.2.2 HIV Identified as Cause of AIDS

In 1983, there were 2 articles published in Science that claimed the isolation of the virus responsible for AIDS[50,51]. Dr. Barre-Sinoussi & Dr. Montagnier are credited with the discovery, having isolated a retrovirus from a patient with multiple lymphadenopathies who had sought treatment in December of the previous year[50]. This virus was named lymphadenopathy-associated virus (LAV). The other paper was published by Dr. Robert Gallo, who had discovered the first two human retroviruses, named Human T-cell leukemia-lymphoma virus (HTLV) I and II respectively[52]. Gallo's publication presented the theory that the virus responsible for AIDS was another HTLV and thus dubbed it HTLV-III. Antibodies against HTVL-III, that had no cross reactivity with HTLV-I or HTVL-II, were detected in patients diagnosed with AIDS, with none found in control populations[53–55]. This set of discoveries was enough for the research community to conclude that HTLV-III was the cause of AIDS. But not every AIDS patient showed reactivity to HTLV-III antibodies, some showed reactivity to antibodies against LAV and the newly introduced AIDS-associated retrovirus (ARV)[56,57]. Sequencing of all three viruses in 1984 revealed that they were simply different strains of the same virus[58–60]. In May 1986, the International Committee on the Taxonomy of Viruses termed this virus Human Immunodeficiency Virus (HIV)[61,62] in order to unify the previous titles given to the virus. HIV was then classified as a Lentivirus within the Retroviridae family due to its morphology, and genetics[62–65] (Figure 1.1). Lentiviruses are defined non-tumorigenic viruses who demonstrate progressive antigenic drift, have extend incubation periods that range from months to years and cause chronic, progressive and usually terminal diseases[16,65–68]. Twenty five years after the discovery of AIDS, Dr. Barre-Sinoussi and his college Dr. Montagnier were awarded the Nobel Prize in Medicine for the discovery of HIV[69,70].

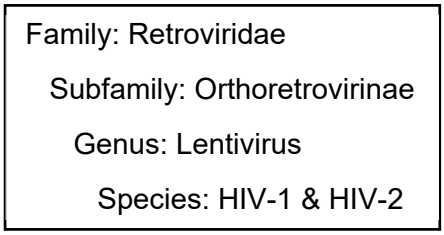


Figure 1.1 Taxonomy of HIV.

1.2.3 Classification of HIV

Currently, there are four major groups of HIV-1 have been classified: M for 'main'[71], O for 'outlier'[72–75], N for 'Non-M/non-O'[76] & P[77] (Figure 1.2). This currently utilized nomenclature was officiated by the theoretical division of the Los Alamos National Laboratory in 1992[78], with groups subdivided into subtypes/clades through phylogenetic differences between the HIV envelope and gag proteins[71,78]. As increasing amounts of sequencing data became available, variant strains of HIV exhibiting marked genetic divergence began to be discovered[72]. The high amount of genetic diversity in HIV comes from an error prone reverse transcriptase, high levels of persistent replication and frequent genomic recombination events[79–88]. Most variation occurs in the Env proteins, primarily gp120[89–93].

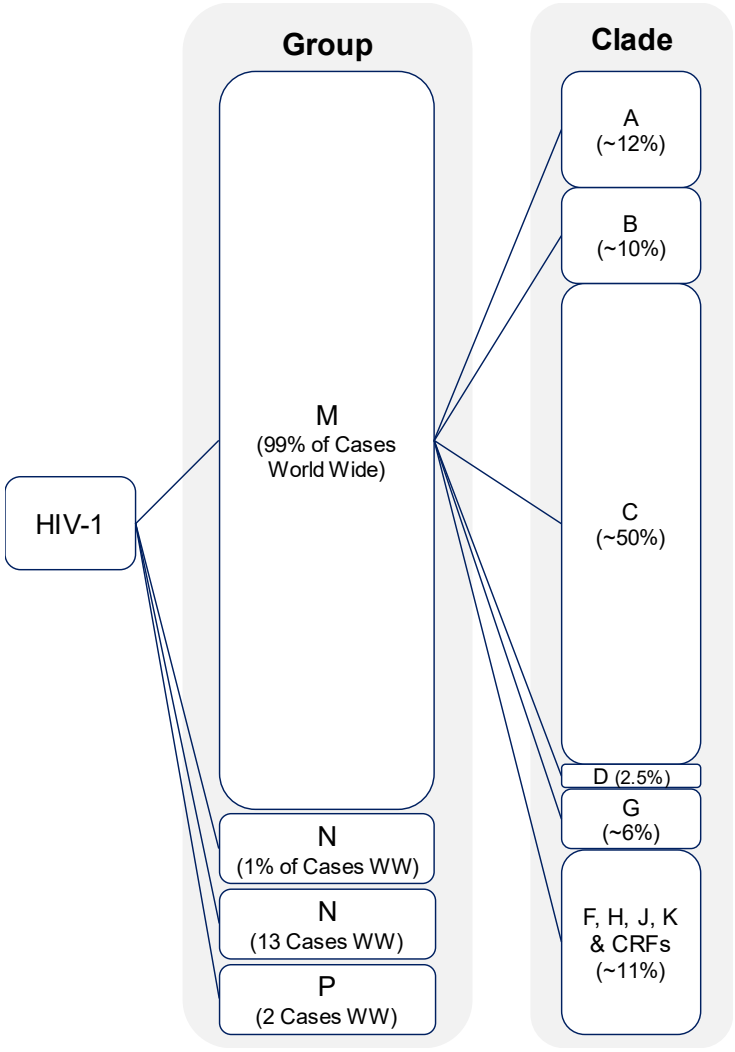


Figure 1.2 Classification of HIV-1.
A representation of the classification of HIV and the global prevalence of these groups and clades according to the most recent survey in 2013. WW stands for World-Wide.

In 1994, the two categories of isolates were ordered into Group M (for Main/Major), which included many subtypes/clades, and Group O (for Outgroup/outliers)[73–75,94,95]. Sequence analysis revealed that all HIV-O isolates seemed to be as diverse from each other as variation between M subgroups[95]. This evidence seems to indicate that the group O viruses are older than group M variants[95], however, group O infections currently only account for approximately one percent of HIV infections worldwide[96–98]. In 1998, group N variants were isolated from individuals in Cameroon[76] and currently only 13 group N infections have been reported[99,100]. HIV-1 group P was discovered most recently in 2009[77,101] and as of 2013, there are still only 2 confirmed cases involving HIV-P[77,102](Figure 1.2). The continued discovery of new groups of HIV indicates transfer events are still occurring, and infected

populations should be monitored carefully for the emergence of new HIV strains that may lead to medical and treatment complications.

Despite the large sequence divergence found in other groups, only the subtypes from group M are responsible for the pandemic spread of HIV[103,104] (Figure 1.2). M group subtypes/clades A,B,C,D & F were determined in 1993 by Louwagie et al by phylogenetic tree analysis of the complete gag protein sequence from 70 international HIV isolates[105], with four additional subtypes, G-K, determined in the following two years by various groups[89,106–108]. Clade C is responsible for approximately half of HIV infections with worldwide[109–113] with clades A and B accountable for 12% and roughly 10% repectively[109,111,112] (Figure 1.2). The HIV epidemic in the Americas is primarily caused by infection with B clade strains; HIV infections in Western Europe and Australia are also principally due to B clade virus[110,111,114,115]. C clade infections are predominate in Sub-Saharan Africa and India[110,111], which are also the regions with the highest levels of HIV infection[109,110,114,115]. The primary differences between B and C clade viruses are as follows: 1) B clades viruses are recognized more by broadly neutralizing antibodies (bNAbs) than C clade, 2) B clade viruses tend to transition to using the CXCR4 co-receptor (X4 tropism) in late stages, while C clade viruses consistently utilizes the CCR5 coreceptor (R5 tropism), 3) the variable loop 3 in gp120 has less variability in C clade than B clade viruses.

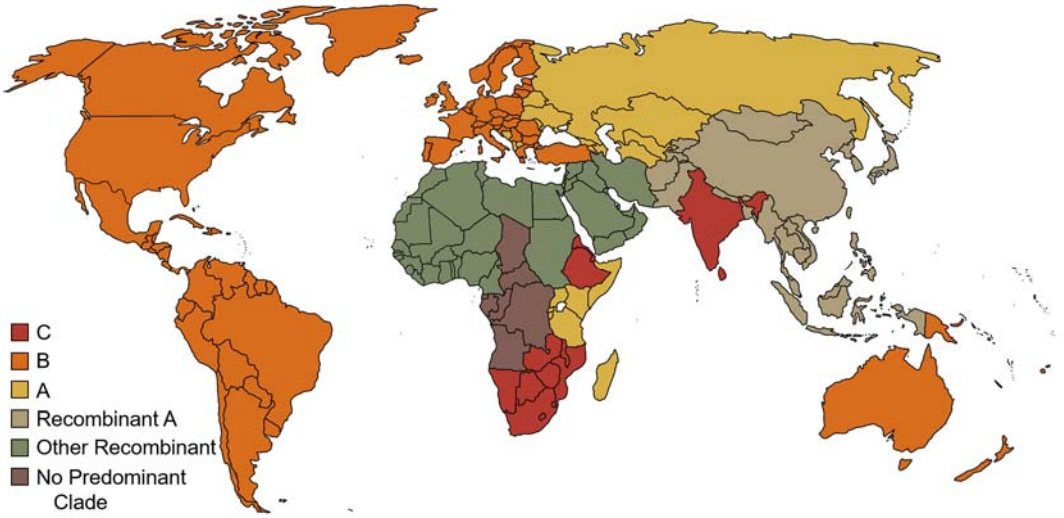


Figure 1.3 World map of predominate HIV clade in each region.
The predominate clade of HIV in each country is indicated using the color key shown in the figure. Predominance of a strain means that the indicated strain is the largest percentage, which is not necessarily larger than 50%. For the regions that have no predominate clade, this means that there was not a single clade that was significantly more present than the rest.

1.2.4 Origin of HIV

In 1983 an immunodeficiency syndrome similar to AIDS was discovered in captive Macaques[116,117]. Two years later, a virus similar to HTVL-III (HIV) was isolated and termed Simian T-cell Leukemia Virus, which was later changed to Simian Immunodeficiency Virus

(SIV)[118]. Within a few months of SIV's discovery, patients in Senegal, west Africa were found to have a new variant of HIV that was more closely related to SIV than HIV[119]. A year later, this new strain was classified as HIV-II (named LAV-II at the time)[120]. This discovery was the earliest indication of where the deadly AIDS epidemic had originated.

In 1989, an SIV found in Sooty mangabeys (SIVsm) was sequenced by the Johnson lab at Georgetown university and found to be closely related to HIV-2, suggesting that SIVsm was the origin of HIV-2[121]; this hypothesis was confirmed in the early 1990's through genetic characterization of the HIV sequences from various sooty mangabeys troops in West Africa[122,123]. Later that year, a strain of SIV linked to HIV-1 was discovered, SIVcpzPtt, that seemed to indicate the origin of HIV-1[124]. Isolated from wild-born chimpanzees that had no exposure to HIV or human blood products, the virus had more similar phylogenetic organization to HIV-1 than to any SIV and sparked the theory that HIV-1's origins were chimpanzee SIV (SIVcpz)[125]. Construction of phylogenetic trees of HIV and SIV confirmed that HIV-1 originated from SIVcpz; the virus was passed to humans from the common chimpanzee (*Pan troglodytes*)[126]. It is theorized that the disease was passed to humans when the chimps were hunted for food, and that the virus then evolved from SIV to HIV-1[126].

Sequencing by the Hahn lab at the University of Alabama indicated that sequences from the HIV groups M, N and O are similar to a single SIVcpz lineage[127], however phylogenetic analysis of groups M, N & O indicates that each group evolved from a unique transmission event[126,127]. The earliest evidence of human infection being a blood sample from an adult Bantu male who lived in the Belgian Congo (the Democratic Republic of Congo) in 1959[128]. Analysis of other tissue and blood samples from the same period not only confirmed the presence of HIV in 1959, but the level of diversification indicated that endemic HIV had already been circulating through west-central Africa for several decades[129]. It is quite possible that at the time AIDS was discovered to be a threat, hundreds of thousands of people were already infected with HIV[130]. The consensus is that the introduction of HIV-M into the human population occurred near the 1930's in Kinshasa, the capitol of the Democratic Republic of Congo[103,128,129,131,132]. Further research has since mapped the geographic distribution, diversity and endemicity of SIVcpz, which allowed researchers to pinpoint that HIV-1 groups M and N originated in southeastern and south central Cameroon respectively[133,134]. Group P variants are distinct from previously classified groups of HIV and are very closely related to SIV strains in Gorillas (SIVgor), leading to the conclusion that HIV-P originated from SIV from Western lowland gorillas of Cameroon [77,101,135] When scientists sought the primate reservoir for HIV-O they determined that while group O strains were still within the SIVcpzPtt radiation, they were more closely related to SIVgor[77,101,135]. There are still too many genetic differences between HIV-O and the simian viruses primate reservoir[101,135].

1.3 HIV Infection and Transmission

1.3.1 HIV Transmission

Since its discovery, 75 million people have become infected with HIV, with an estimated 1.7 million new infections in 2018[136]. Transmission, through any route, can be traced to founder strains (T/F) of virus that have signature characteristics including length of variable loops, level of glycosylation, utilization of CCR5 as a co-receptor, and

consensus sequences in HIV's Gag, Pol and Nef proteins (for descriptions of these proteins see section 1.3.3)[137,138]. The major routes of exposure, in order of prevalence, are male-male sexual contact, male-female sexual contact, and blood contact either via shared needles during intravenous drug use or mother to child placental contact. There can also be transmission between mother and child through genital secretions and breast milk. While blood transfusion has the highest risk per exposure, it accounts for less than 1% of transmissions (Table 1.1). Thanks to increasing awareness and availability of antiretroviral therapy, 92% globally of HIV positive pregnant women received combination antiretroviral therapy (cART) to prevent transmission to their child in 2018[136]. The risk of male-male transmission is the highest globally, estimated to be 28 times higher than the general population, with injectable drug users next with 22 times higher risk[139,140].

Table 1.1 Rates and prevalence of various routes of HIV exposure

Exposure Route	Transmission Medium	Risk per 10,00 exposures to an infected source[141]	Contributed to New HIV Diagnosis in the USA in 2018[142,143]	Contributed to New HIV Diagnosis Globally in 2018[140]
Sexual Contact				
Male- Male Intercourse	Receptive	138	24,909 (69% of all diagnosis)	17%
	Insertive	11		
Heterosexual Intercourse	Receptive	8	8,929 (24% of all diagnosis)	>50%
	Insertive	4		
Injection drug use	Blood, sharps	63	2,511 (7% of all diagnosis)	12%
Other				
Blood transfusion	Blood, sharps	9250	76 (<1% of all diagnosis)	
Perinatal exposure	Placenta	2260		

1.3.2 Discovery of the HIV Infection Process & HIV Structure

Studies of HTLV-I and HTLV-II set the stage for the discovery, isolation, and culturing of HIV[52,144]. Using immortalized T lymphocytes[50,51,54,55,145,146], researchers mimicked HTVL-I and HTVL-II, were culturing techniques[145,147–149], but inverted the normal ratio between T-helper cells and T-suppressor cells[35,150,151] to due the observation that AIDS patients exhibited a T-cell dysfunction that depleted T-helper cells. Trial and error determined that HIV propagated best in cells expressing the CD4 antigen[56,146,152–154],the majority of which where CD4 T lymphocytes (which were referred to as T4 T-cells at the time). Several groups established that CD4 was an essential receptor for HIV infection through competition assays with radiolabeled anti-CD4 antibodies[152,153].

CD4 (cluster of differentiation 4), originally called leu-3 and T4, was discovered in the 1970s[155–160]; it was renamed CD4 in 1984[161]. The structures of the first two extracellular domains of CD4 were solved in 1990[162], but the first crystal structures did not appear on the RCSB protein data bank until 1993, both the first two domains of human CD4[163] and the third and fourth of rat CD4[164]. CD4 is a glycoprotein found on T cells, macrophages, monocytes and dendritic cells. The first two domains of CD4 exhibit an immunoglobulin fold (Figure 1.6A). A NMR structure of the cytoplasmic domain of CD4 became available in 1996[165]. The transmembrane domain was not solved until 2010[166]. Key CD4 structures are summarized in Table Appendix A.1. The interactions between CD4 and HIV are discussed below.

In 1985, the NC1-Frederick Cancer Research Facility isolated the HIV gag and envelope proteins and characterized them for the first time[167]. They noted that the envelope gene generates a16 kDa glycosylated peptide (gp160), that was processed by proteolytic cleavage and glyosidic trimming[168] to form gp120 and other small gene products (later determined to be gp41[53]), and the gag gene produced two structural proteins dubbed p70 and p55 in accordance with their molecular weight[167]. In 1989, HIV's aspartyl protease was the first HIV protein crystal structure to be solved[169]. Researchers hoped that these, and subsequent studies of HIV structural proteins, would to help develop effective antibodies against HIV[53,167,170].

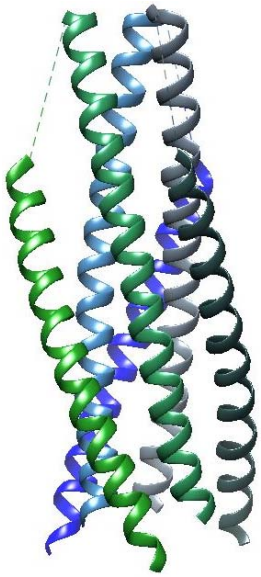


Figure 1.4 The fusion structure of gp41. Crystal structure of the gp41 trimer (PDB 1ENV) in its fusion form. Each monomer of gp41 is shown in a separate color, with the N terminal helix in a dimmer color (along the inside), and the C terminal helix in a brighter color (along the outside).

Meanwhile, the immunology branch of the CDC developed an immunoassay to detect LAV (HIV)[154], and confirmed that HIV tropism was primarily determined by the presence of the CD4 (T4) antigen/glycoprotein, which complexed with the envelope protein gp120[171,172]. Drs. Weiss and Axel proved that insertion of the T4 molecule into the membrane of non-target cells was sufficient for HIV infection to occur[173]. Along with CD4 T-lymphocytes, HIV was also found to infect macrophages, including those found in the central nervous system[174–176].

With the primary cell type and receptor located, researchers turned their efforts to elucidating the gp120-CD4 interaction and the envelope protein structure. In 1987, electron microscopy showed the morphology of HIV to be a 100nM enveloped virus covered in spikes made up of two glycoproteins, gp120 and gp41, with an outer shell composed of p17 proteins and a central core composed of p24 proteins surrounding two strands of RNA[62]. The genome is similar to other retroviruses in that it contains the genes gag, structural proteins, env, envelope proteins, and pol (a reverse transcriptase), but is more complex with an additional five proteins whose functions were not yet completely understood[62] and

are discussed further below. HIV's aspartyl protease was the first HIV protein structure to be published in 1989[169].

The envelope glycoproteins were found to have extensive genomic diversity[177], primarily in gp120; computer based analysis showed that while there were five highly variable regions in gp120, 80% of the gp41 residues were conserved[92]. The sequence of HIV strain HXB2, which was decided to be the reference strain for all other sequences and alignments, is shown in Figure 1.5. Research showed that proteolytic cleavage of gp160 into gp120 and gp41 was necessary for infectivity[178]. The fusion structure of the gp41 ectodomain was published in 1997[179], showing that the structure was very similar to the fusion domains of Influenza and Moloney murine leukemia virus: both termini of the alpha helical structure point towards the apex of the trimer(towards the host cell membrane) with the N terminal helix in the very center of the bundle (called the zipper domain), with the C terminal helix packed along the outside[179].

```
MRVKEKYQHLWRWGWRWGTMLLGMLMICSATEKLWVTVYYGVPVWKEATT  49
TLFCASDAKAYDTEVHNVWATHACVPTDPNPQEVVLVNVTENFNMWKNDM  99
VEQMHEDIISLWDQSLKPCVKLTPLCVSLKCTDLKNDTNTNSSSGRMIME 149
EKGEIKNCSFNISTSIRGKVQKEYAFFYKLDIIPIDNDTTSYKLTSCNTS 199
VITQACPKVSFEPIPIHYCAPAGFAILKCNNKTFNGTGPCTNVSTVQCTH 249
GIRPVVSTQLLLNGSLAEEEVVIRSVNFTDNAKTIIVQLNTSVEINCTRP 299
NNNTRKRIRIQRGPGRAFVTIGKIGNMRQAHCNISRAKWNNTLKQIASKL 349
REQFGNNKTIIFKQSSGGDPEIVTHSFNCGGEFFYCNSTQLFNSTWFNST 399
WSTEGSNNTEGSDTITLPCRIKQIINMWQKVGKAMYAPPISGQIRCSSNI 449
TGLLLTRDGGNSNNESEIFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPT 499
KAKRRVVQREKRAVGIGALFLGFLGAAGSTMGAASMTLTVQARQLLSGIV 549
QQQNNLLRAIEAQQHLLQLTVWGIKQLQARILAVERYLKDQQLLGIWGCS 599
GKLICTTAVPWNASWSNKSLEQIWNHTTWMEWDREINNYTSLIHSLIEES 649
QNQQEKNEQELLELDKWASLWNWFNITNWLWYIKLFIMIVGGLVGLRIVF 699
AVLSIVNRVRQGYSPLSFQTHLPTPRGPDRPEGIEEEGGERDRDRSIRLV 749
NGSLALIWDDLRSLCLFSYHRLRDLLLIVTRIVELLGRRGWEALKYWWNL 799
LQYWSQELKNSAVSLLNATAIAVAEGTDRVIEVVQGACRAIRHIPRRIRQ 849
GLERILL  856
```

Figure 1.5 Sequence of HIV strain HXB2 envelope protein.
The sequence for HXB2 envelope gp120 is in black, the sequence for gp41 is in grey. The variable loop regions are underlined, 1-red, 2-orange, 3-gold, 4-green, 5-blue. High mannose glycosites are highlighted in red, hybrid and complex glycosites are highlighted in yellow.

In 1990, Leonard et al characterized the 24 potential glycosylation sites (Figure 1.5) and assigned all nine intra-chain disulfide bonds for gp120[180], finding that disulfide bonds form loop structures and three of the loops define the highly variable regions 1-4 (V1-V4)[92,180] as seen in Figure 1.5, Figure 1.9 & Figure 1.11. The position of the cysteine residues are highly conserved, but only 13 of the 24 glycosylation sites are conserved[177].

With CD4 and gp120 both being glycosylated proteins, researchers questioned whether the glycosylation residues were involved in the CD4-gp120 interaction but Fenouillet et al confirmed that the glycosylation was not necessary for their interaction[181]. They stated that previous studies showing deglycosylated gp120 exhibiting a decreased binding capacity to CD4[176,182–184] were most likely due to disruption of gp120's conformation in relation to the use of SDS during the deglycosylation procedures[181]. Unifying these findings, Fenouillet hypothesized that the glycosylation residues may be involved in gp120 folding and or stabilization[181]. Gp120 mutational studies determined that the CD4 contact sites on gp120 included two hydrophobic regions, around T257 & W427, and two hydrophilic regions, around D386, E370 & D457[180,185–187]. The crystal structure of CD4 revealed that N terminal domain (D1) and the domain immediately adjacent to it (D2) were very similar in structure to immunoglobulin domains[162,188] (Figure 1.6). Analysis of the gp120-CD4 binding site suggested that the gp120 binding region would be a 25Å by 12Å groove or pocket[162,188]. Studies isolated the binding site to the D1 of CD4[185,189,190], locating specific contact sites,F43, K29, K46 and R59 , that are important for gp120-CD4 binding affinity[191,192]. It was hypothesized that the exposed phenylalanine fit into the

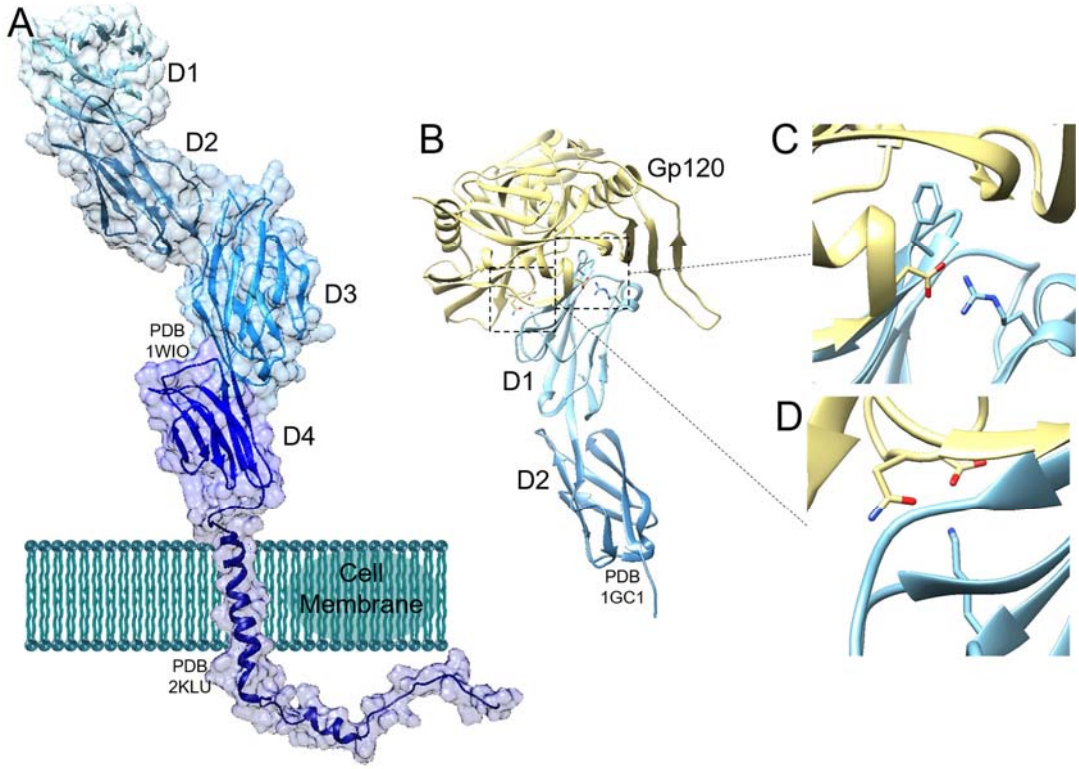


Figure 1.6 Structure of CD4 & its interaction with gp120.
A) shows the full structure of CD4. The cytoplasmic domain and transmembrane are shown in a dark navy blue (PDB 2KLU). Extracellular domains are shown in difference shades of blue to differentiate between the four domains, labeled D1-4 (PDB 1WIO). For extended information on the CD4 structure, please see Appendix A. **B-D**) show the interaction between a monomer of HIV gp120 and the D1 domain of CD4 (PDB 1GC1). **C)** shows CD4 F43 inserted into a hydrophobic pocket in gp120 and CD4 R59 electrostatically interacting with gp120 D368. **D)** shows CD4 K29 has possible electrostatic interactions with both D279 & N280 from gp120.

hydrophobic pocket in gp120 that was ringed by negatively charged amino acids[192]. A combination of mutational and kinetic studies confirmed that CD4 F43 was a critical residue for CD4-gp120 binding[193]. A slow 'on' rate that indicated a conformational adaptation likely occurs during CD4 binding to gp120[193]. The first published structure of the gp120 core bound to CD4 by Peter D. Kwong provided further evidence for this hypothesis in 1998, fully interpreting the F43 binding pocket and how its presence would be structurally unfeasible in the absence of CD4 binding[194]. An energetics study using Surface Plasmon Resonance by the Doyle lab in 2000 determined that the association rate constant between gp120 and CD4 was approximately 6 x 10^4 $M^{-1}S^{-1}$, consistent with a binding induced conformational change[195]. This conformational change resulted in the formation of what was termed the 'bridging sheet,' a four strand, anti-parallel β-sheet at the bottom (relative to the viral membrane) of gp120 (Figure 1.7 & Figure 1.11).

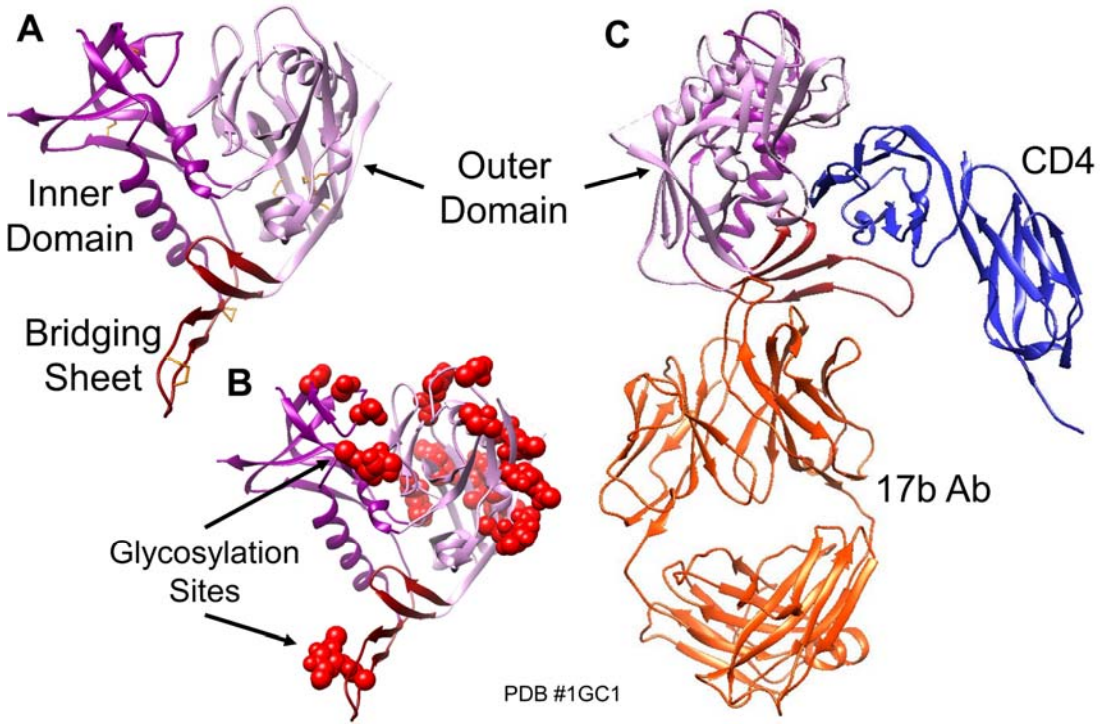


Figure 1.7 Structure of a gp120 monomer bound to CD4 and 17b showing the basic domains.
A) shows gp120 monomer PDB 1GC1 (bound to CD4 and 17b) with the outer domain colored light purple, the inner domain colored medium purple, the bridging sheet in dark red and the disulfide bonds shown in orange. The structure is oriented so that the top would be against the viral membrane and the bridging sheet would be pointed toward the host cell membrane. Only seven of the nine conserved disulfide bridges are present in this structure. The gp120 subunit is approximately 97Å long with a diameter of 66Å. **B)** shows the location of all the potential glycosylation sites as red balls. **C)** the gp120 structure has been rotated 90° to the left to show how the gp120 monomer binds to CD4, in blue, and the 17b Ab, in orange.

The gp120-CD4 interaction was well characterized by the beginning of the 1990's, and researchers were trying to determine the mechanism of HIV infection. It was originally thought that HIV was endocytosed into the cell after binding to CD4[196]. However, electron micrographs showed fusion of HIV virions to the host cell[197] and mutational studies that removed the cytoplasmic domain of CD4 (impairing the endocytosis process)[198,199] established that infection occurred through viral fusion to the host and not host endocytosis of the virion[197–199]. This coincided with previously published sequencing data that gp41 contained a fusion peptide, indicating gp41 was involved with viral fusion to the host cell[200]. Evidence seemed to indicate that after binding CD4, a conformational change of gp120 likely occurred to expose the fusigenic domain of gp41 to facilitate viral fusion to the host membrane[187,192,201–203]. The Sodroski lab found the first solid proof for the conformational change theory, showing that the binding of gp120 to CD4 induces epitopes for the antibodies 17b, 48d an A32, by shifting the V1 and V2 loops of gp120 using viral capture ELISAs with mutated HIV gp120[202](Figure 1.7).

While this simple model of fusion and infection was appealing, multiple groups had shown that the V3 region was a determinant for monocyte-macrophage tropism of HIV[204,205], and that there was a shift in HIV tropism from monocyte/macrophage tropic (M-tropic) to T-cell tropic (T-tropic) as the disease progressed[206]. The V3 loop was dubbed the "Neutralization Determinant" with a conserved Gly-Pro-Gly-Arg sequence that provided the binding site for antibodies able to block HIV infection[207–212] (Figure 1.12D). Researchers found that mutation of this conserved site abolished HIV infection and syncytium formation, leading them to incorrectly conclude that the V3 loop must be a fusion domain[213,214]. Around the same time, studies in non-human cells indicated that there were additional cell-type specific co-factors, aside from CD4, required for viral entry[215–218]. In 1995, Cocchi et al discovered that the chemokines RANTES, MIP-1 α, and MIP-1 β could inhibit HIV infection[219]. Then, within a year, the Berger group pieced those two ideas together and used functional complementary DNA cloning[220] to isolate a G-coupled protein receptor designated 'fusin'[221], officially discovering HIV-1's cofactor for viral entry. Several other groups proceeded to characterize fusin, which was later renamed CXCR4[222], and CCR5 as chemokine receptors and linking them to T-tropic and M-tropic strains respectively[223–228].

CCR5 was determined to be the co-receptor used in CD4 T-cell infection, and strains would adapt to use CXCR4 as CD4 T-cell counts declined[229–231]. Further research showed that CCR5 and CXCR4 bound primarily to the V3 loop of gp120, with epitopes that were induced by CD4 binding[223,232–235] (Figure 1.7 & Figure 1.11), and gp120 bound to the N terminal residues 13-21 of the chemokine receptors[236–240]. The first full structure of a chemokine receptor was of CXCR4 in 2010, bound to agonists IT1t and CVX15[241] (Figure 1.8A). This model provided support for the two-site model[242] of chemokine binding, which theorized that the N terminus the receptor reached up to bind to the globular core of the chemokine or HIV (chemokine recognition site 1, CRS1), and the N terminus of the chemokine (or V3 loop of HIV) would insert into a binding pocket formed by the seven transmembrane bundle of the receptor (chemokine recognition site 2, CRS2). It also gave the first structure showing how the receptor dimerizes in the presence of ligands (Figure 1.8B). The first full structure of CCR5, bound to Maraviroc, published two years later, continued to lend credence to the two-site model while also showing that Maraviroc is a noncompetitive inhibitor of HIV[243](Figure 1.8A). In 2018, a structure of CCR5 with a chemokine agonist, 5P7-CCL5, gave definitive evidence that the two-site model was correct (Figure 1.8C)[244].

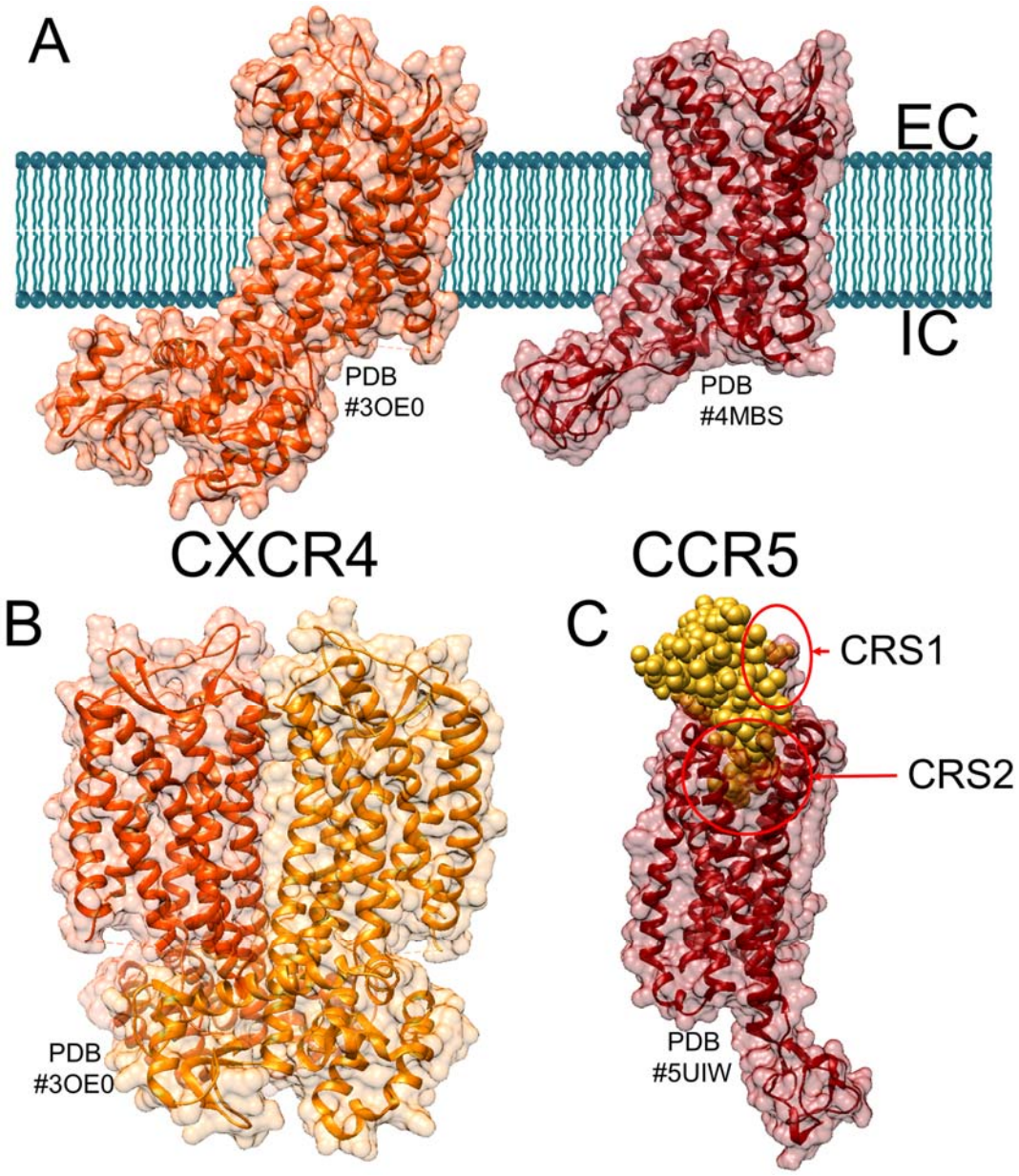


Figure 1.8 Crystal structures of HIV Coreceptors CXCR4 & CCR5.
A) Shows the monomeric structures of CXCR4[241] and CCR5[244] on the left and right respectively. **B)** Shows the dimeric form of CXCR4. C) Shows monomeric CCR5, in red, bound to chemokine agonist 5P7-CCL5, in yellow, demonstrating how ligands are bound by the N terminus of CCR5 (CRS1) and the transmembrane pocket (CRS2).

Table 1.2 Relevant structures of HIV co-receptors CCR5 and CXCR4

Structure Studied	Structural and Functional Characteristics	PDBs	Year
CCR5 N Terminus (7-15) modeled with gp120 and CD4	Modeling shows that CCR5 N terminus binds in a pocket of gp120 specialized for sulfotyrosine, causing the V3 loop to become a rigid β-hairpin	2RLL[245]	2007
CXCR4 Monomer and Dimer bound to IT1t & CVX15	Supports two-site model for receptor binding.	3ODU & 3OE0[241]	2010
CCR5 bound to Maraviroc	Full CCR5 monomeric structure, shows Maraviroc is a noncompetitive inhibitor of HIV	4MBS[243]	2013
CCR5 bound to 5P7-CCL5	Elucidates how ligands insert themselves into TM binding pocket, offers insight to gp120 binding	5UIW[244]	2017
CCR5 Bound to Agonists	Structure based drug design of CCR5 agonists, designed to surpass deficiencies in Maraviroc	6AKX & 6AKY[246]	2018
CCR5 bound to monomeric gp120 and CD4	V3 loop of gp120 binds in TM binding pocket but doesn't induce allosteric conformational changes that would propagate to gp41. N terminus of gp120 is flipped back, which may destabilize gp41 resulting in fusion	6MEO & 6MET[247]	2019

1.3.3 Current Understanding of HIV Structure & Infection Process

After the discovery of the co-receptor CCR5 and crystal structures of gp120 were published, the research curve around the HIV infection process was exponential. It was determined that the HIV envelope glycoproteins form a trimeric complex, with a trimer of gp120 shielding an inner trimer of gp41[248], with approximately 9-14 trimer spikes per virion. Interprotomer contacts between gp120 and gp41 occur primarily at the apex of the trimer, where the variable loops 1 through 3 form a cap, as well as near the viral surface[249,250]. Advances in HIV crystallography, including the development of SOSIP trimers[251–253](For more information on SOSIP please see Appendix A), lattice engineering[254–256] and crystallization chaperones[257], have allowed the crystallization of the HIV prefusion state, with full glycosylation, for multiple HIV clades.

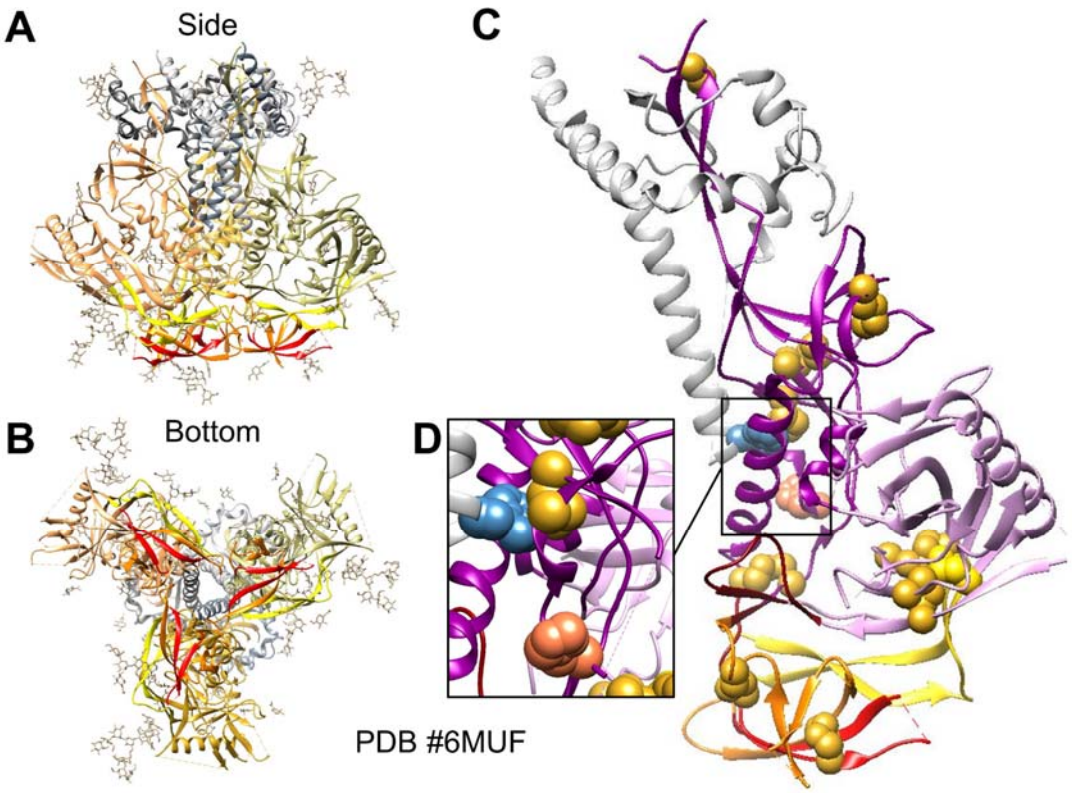


Figure 1.9 Structural elements of the prefusion, unbound trimer.
A fully glycosylated B clade SOSIP trimer (B41 SOSIP.664[557]) 3.4Å crystal structure, that is bound to Abs 3H109L and 35O22 (hidden), to show all the basic structures of the envelope proteins in the prefusion state. **A)** Shows the gp140 trimer from the side, oriented so the viral surface would be at the top and the bottom would be pointed towards the host cell. Each structural element is colored as follows: gp120 monomers in different shades of sand/khaki, gp41 monomers in different shades of grey, glycans in tan, V1 in red, V2 in orange and V3 in yellow. **B)** shows the trimer from the bottom view, or from the host cell's perspective, with the same coloration as A. **C)** Shows a single gp140 monomer. Each structural element is colored as follows: the gp41 monomer is grey, the inner domain of gp10 is medium purple, the outer domain of gp210 is light purple, the bridging sheet is dark red, the V1 loop is bright red, the V2 loop is orange, the V3 loop is yellow, disulfide bonds are indicated with golden spheres, H66 is indicated by salmon spheres and W571 is indicated with blue spheres. D) Shows a close-up view of gp120 H66 and gp41 W571, both of which influence conformational sampling.

The viral envelope undergoes a three step process of infection is classified as a type 1 viral membrane fusion[23], but is unique in that it requires two coreceptors. These steps are summarized in Figure 1.10. Step one is a fully reversible, CD4 binding induced, conformational change[258] that reveals epitopes on the V3 loop necessary for CCR5 binding[259–261]. Gp120 binds to CD4 with a kD of 22nM[195]. There is debate in the literature that suggests this conformational change may be independent of binding to CD4, that

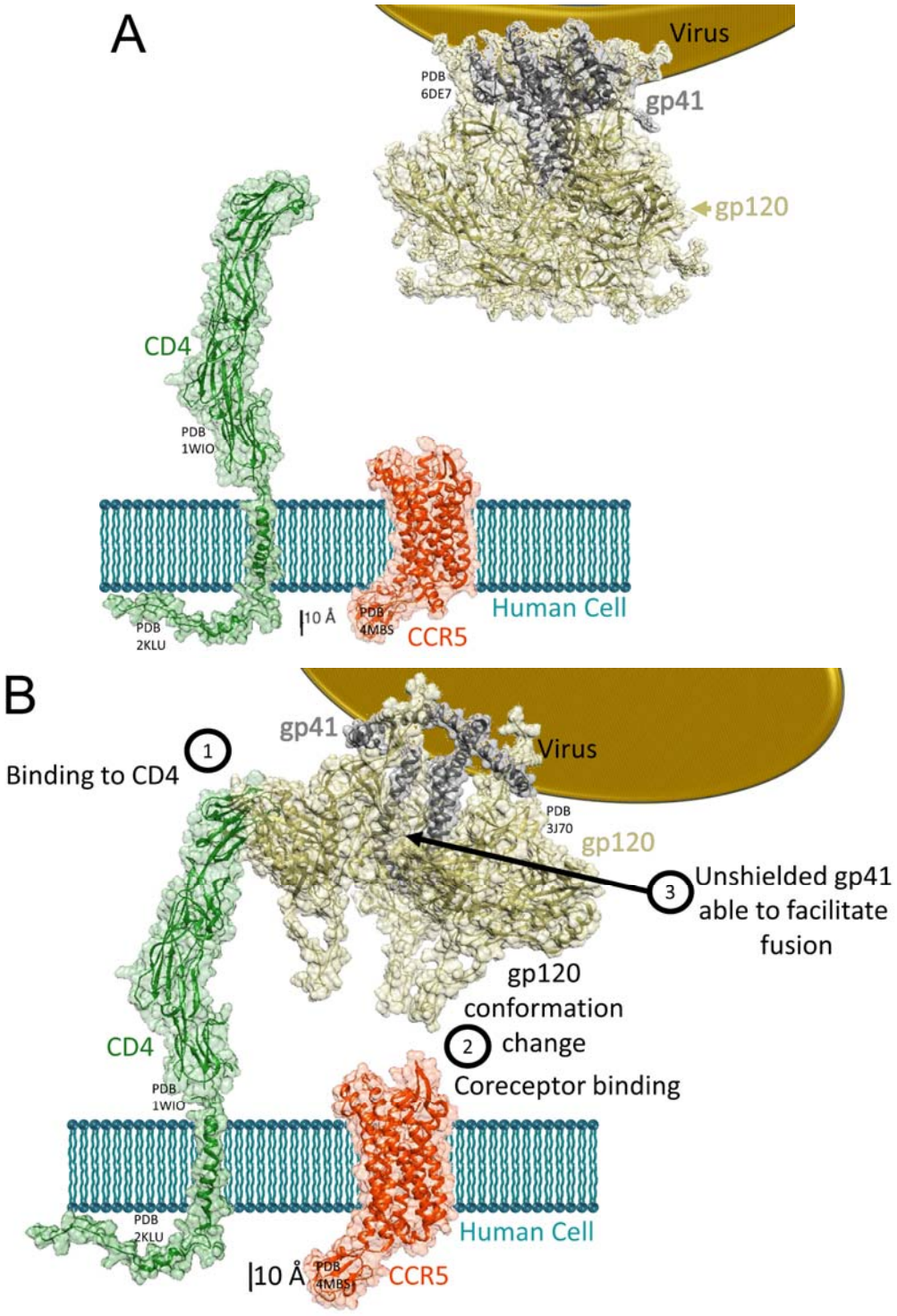
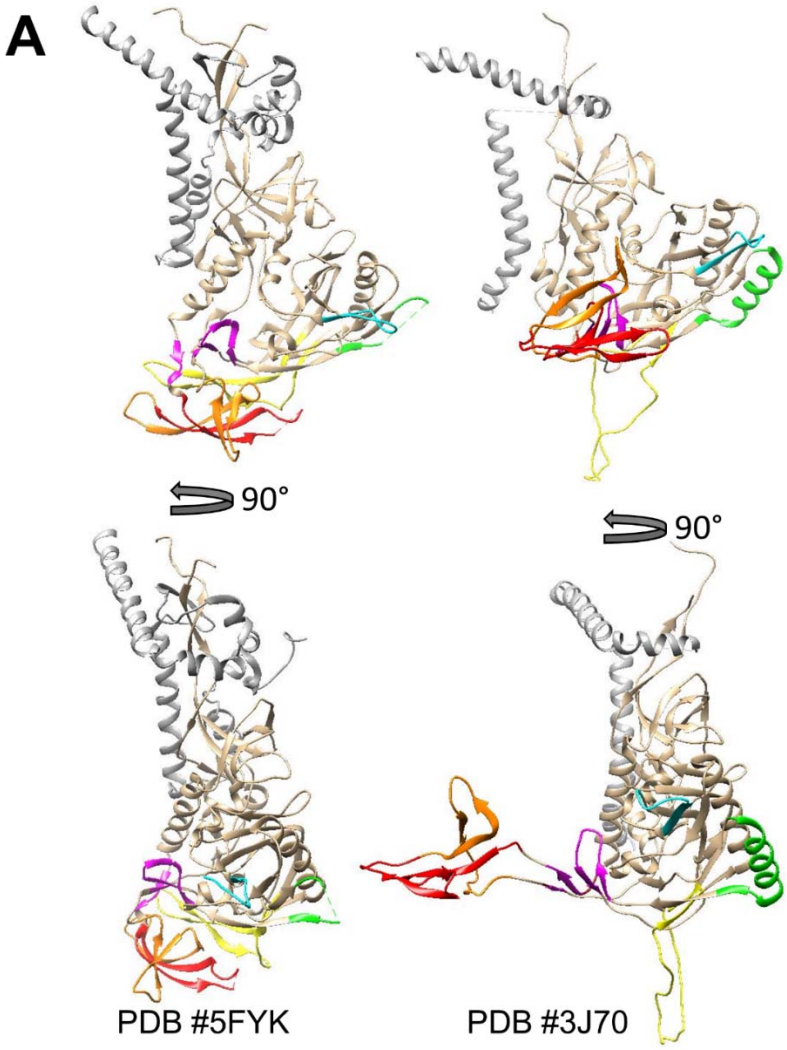


Figure 1.10 Structures showing the process of HIV infection.
A) Shows the structures of the unbound HIV envelope proteins (PDB 6DE7[575]), CD4 (PDBs 1WIO[901] & 2KLU[166]) and CCR5 (PDB 4MBS[243]) in appropriate scale to each other. **B)** Shows a summary of the HIV infection process with the structure of the envelope proteins aftergp120 has bound to CD4 (PDB 3J70[912]) with the full CD4 and CCR5 structures for comparison.

gp120 may sample several conformational states while unbound. It has been shown that conformational sampling is dependent on the presence of conserved residue H66 from gp120[262], which is located on the inner domain near gp120-gp41 interface (Figure 1.9) and variable loops 1-3[263]. Binding to CD4 prompts allosteric changes through the gp120 contacts with gp41 at W571 (Figure 1.9), stabilizing the conformational state with a fully formed bridging sheet and partial exposure of the V3 loop[195,262,264,265]. The V1/V2 and V3 loop separate away from the gp120 core, exposing the gp41 three-helix bundle and fusion elements[266,267] (Figure 1.11). The ability of CD4 to recruit the bridging sheet is a major determinant of R5 macrophage tropism[247,268–272]. CD4 normally serves as a co-receptor to the T cell receptor(TCR**)** during the T-cell interaction with antigen-presenting cells, part of an immunoglobulin super family involved in recognition events[193,273,274]. The cytoplasmic domain of CD4 is involved with signal transduction, associating with tyrosine kinase p56*lck*, which is implicated in T-cell activation[193,275,276].



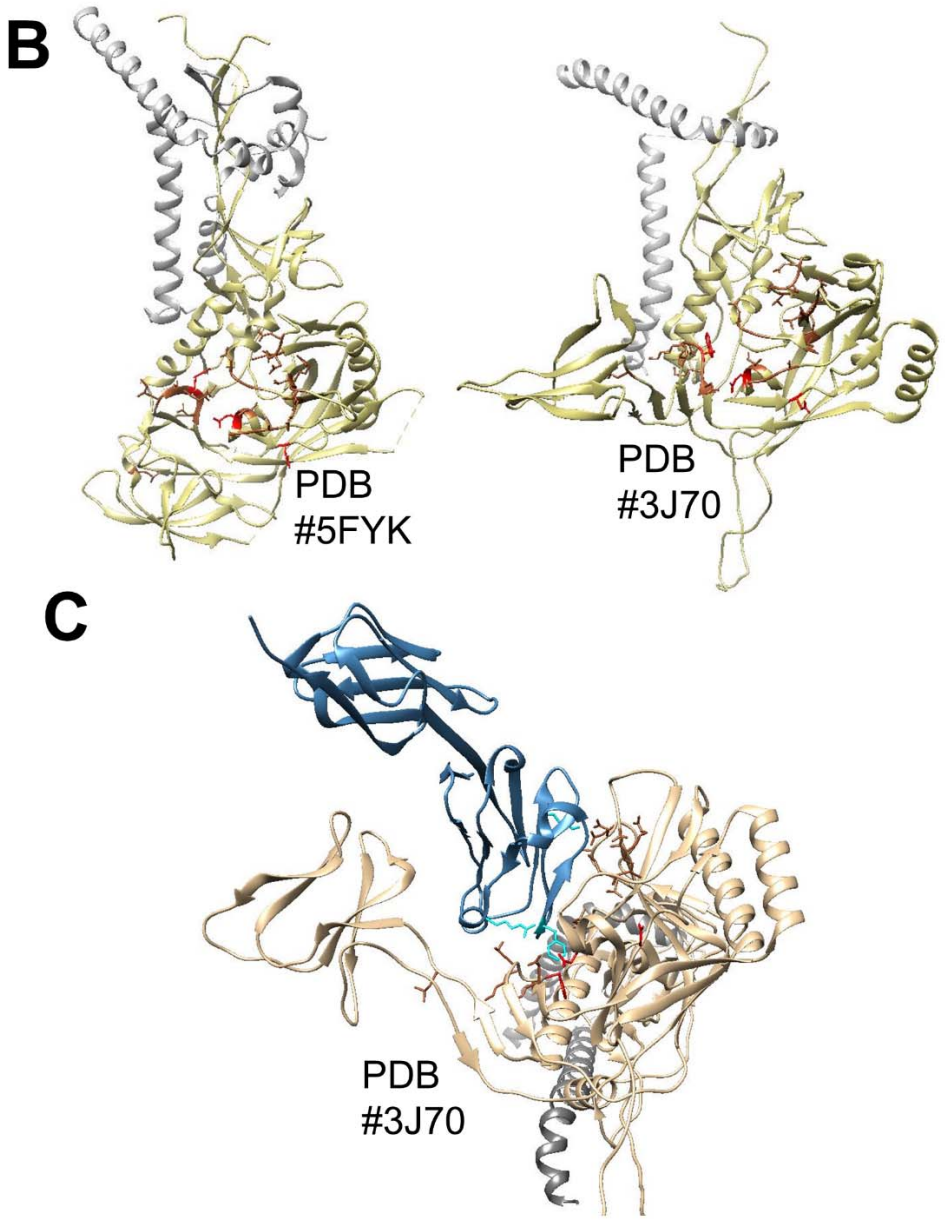


Figure 1.11 The conformational changes in the HIV envelope upon binding to CD4. This figure compares gp140 structures PDB 5FYK (a prefusion B clade SOSIP.664 bound to bNAbs PGT121, 35O22 & VRC01) and PDB 3J70[912] (a model based on EMD5020[282] and crystal structure 1GC1[194], which are both B clade structures – BaL and HXB2 respectively, to establish the structures of the full variable regions while in the CD4 and 17b bound state. **A)** Compares the locations and orientations of the variable loops and residues involved in the bridging sheet between the prefusion state (5FYK) and the CD4 bound state (3J70). The structures are colored as follows: gp41 in grey, gp120 in khaki, V1 in red, V2 in orange, V3 in yellow, V4 in green, V5 in blue, bridging sheet residues in purple. **B)** Shows the location of the residues in gp120 that contact CD4 in the prefusion(5FYK) and CD4 bound(3J70) states. The main three contacts are shown in red (D368,E370 & W527) [194], the other contacts in brown. **C)** Follows the same color scheme as B and includes CD4 in blue.

The second step of the initial infection process is gp120 binding to chemokine receptor CCR5 with a K_D of 4nM[258]. Binding to CCR5 is believed to occur in two stages, first sulfotyrosine residues on the N-terminus of the co-receptor binds to the base of the V3 loop near the bridging sheet, then the GPGR motif at the tip of the V3 loop inserts into the CRS2 of CCR5, contacting all seven of the transmembrane helices (Figure 1.12)[247]. Binding of sulfated tyrosines in the CCR5 N terminus converts its V3 from a flexible loop to a rigid Beta-hairpin., which facilitates interaction between the CCR5 extracellular loops and the tip of V3[247,277]. Binding to CCR5 was previously thought to prompt a second conformational change that exposes the fusigenic domain of gp41[277]; however, the

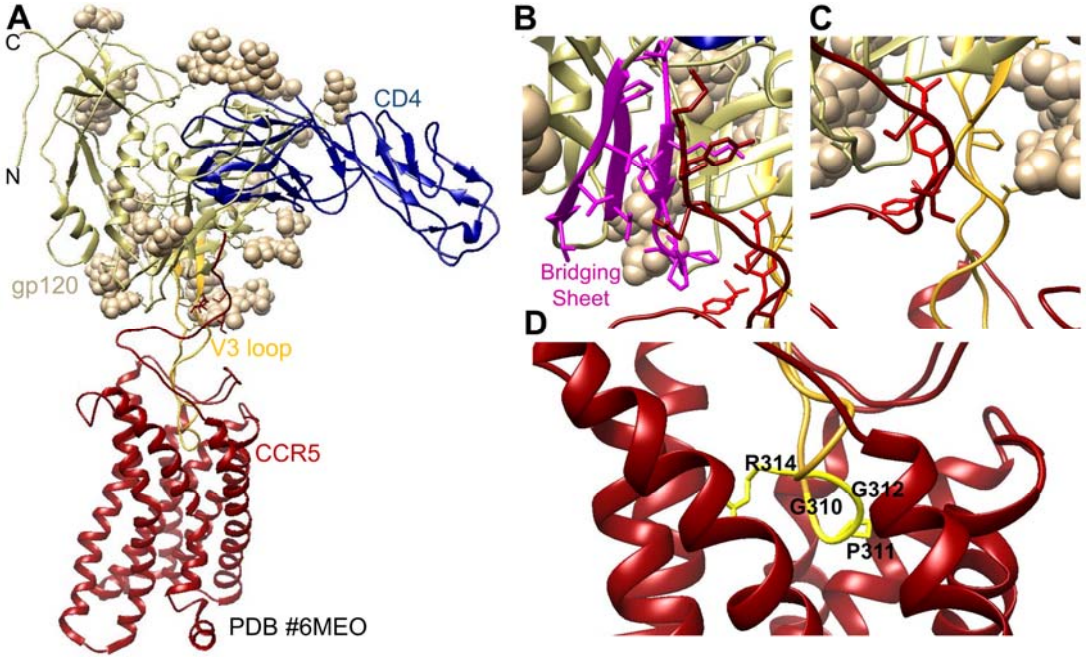


Figure 1.12 Structure of gp120 bound to CD4 and CCR5.
A) Shows the structure of gp120 in complex with CD4 and CCR5 (PDB 6MEO). The gp120 structure, in khaki, is strain 92BR020, clade B, with the V3 loop highlighted in gold. Glycans are shown as tan spheres, CD4 in blue and CCR5 in red. The N and C termini are labeled respectively. **B)** Shows the contacts the CCR5 N terminus (in dark red) makes with the gp120 bridging sheet (in purple). **C)** Shows the contact the CCR5 sulfotyrosines (in bright red) form with the root of the V3 loop (in gold). **D)** Shows the insertion of the GPRG motif at the tip of the V3 loop inserted into CCR5.

recently published structure with gp120 bound to CCR5 indicates that the only major changes that occur in gpp120 are that the N and C termini bend back to pack against the surface of gp120 and the V3 loop adjusts to fit into the CRS2 pocket[278]. Prior to CCR5 binding, the N and C termini of gp120 are surrounded by a '4-helix collar' (Figure 1.13). It has also been shown that the conformational changes induced by CD4 binding are relatively unstable and easily reversed unless gp120 binds to CCR5 in quick sucession[279,280].

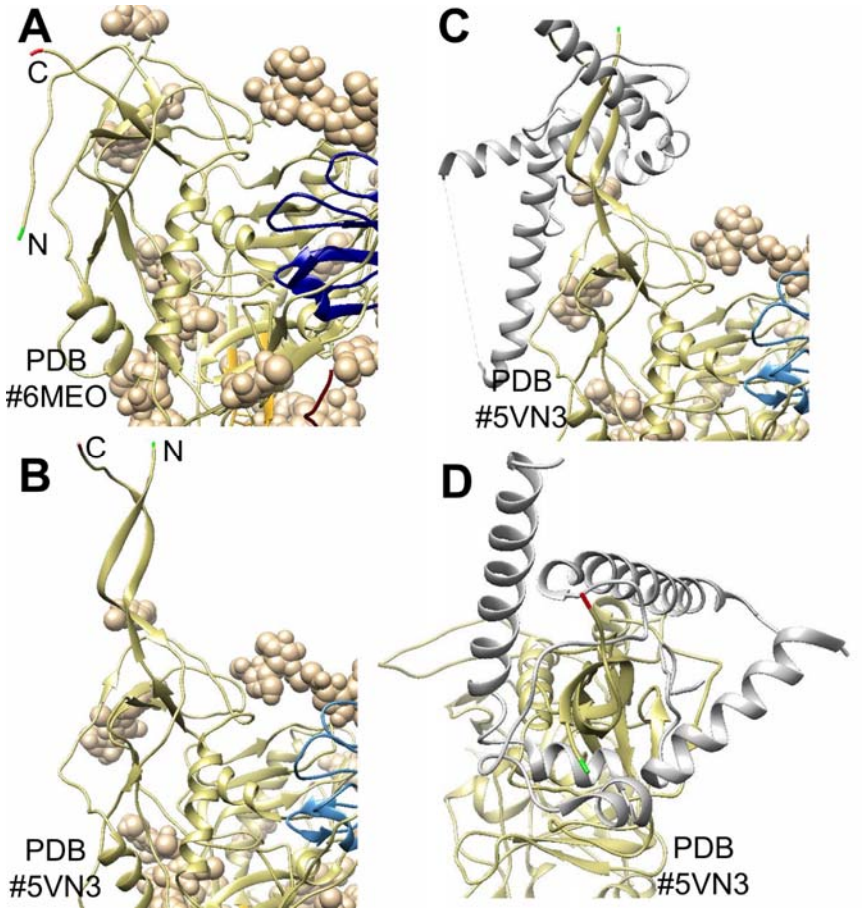


Figure 1.13 Comparison of the gp120 termini in unbound verses fully bound structures. **A)** Shows a gp120 monomer (PDB 6MEO) fully bound to CD4 and CCR5 with the N terminus labeled and highlighted in green and the C terminus labeled and highlighted in red. The gp120 is in khaki, glycans are tan spheres and the edge of CD4 is visible in blue. **B-D)** Show PDB 5VN3, a gp120 structure bound to sCD4 and Ab 17b in various orientations. The gp120 is colored khaki, with N and C termini labeled and highlighted in green and red respectively. The gp41, colored grey, has been hidden in **B** to reveal the location of the termini for comparison to the fully bound structure in **A.** CD4 is visible in blue in structures B-C. **D)** Is a top down (from the perspective of the viral membrane) view of the gp140 monomer to show the 4-helix collar formed by the gp41 around the termini of gp120 in the CD4 bound structure.

The exact trigger to expose the gp41 fusion domain is still unknown[281]. Previous structures have shown that the gp41 trimer becomes more compact after gp120 binds to CD4[282–285] (Figure 1.14), Researchers hypothesize that the conformational change of the gp120 termini, triggered by binding to CCR5, destabilizes the gp41, potentially initiating gp41 refolding to expose the fusion domain[278]. Binding to the coreceptor also brings the fusion peptide within 70Å of the host membrane, which is close enough for the fusion peptide to reach[278]. The prefusion structure of each gp41 monomer is a four-helix collar (Figure 1.13D, Figure 1.14C-D & Figure 1.15B), whereas the post fusion structure is composed of two helices[283]. The components of gp41 and its conformational changes are illustrated in Figure 1.15.

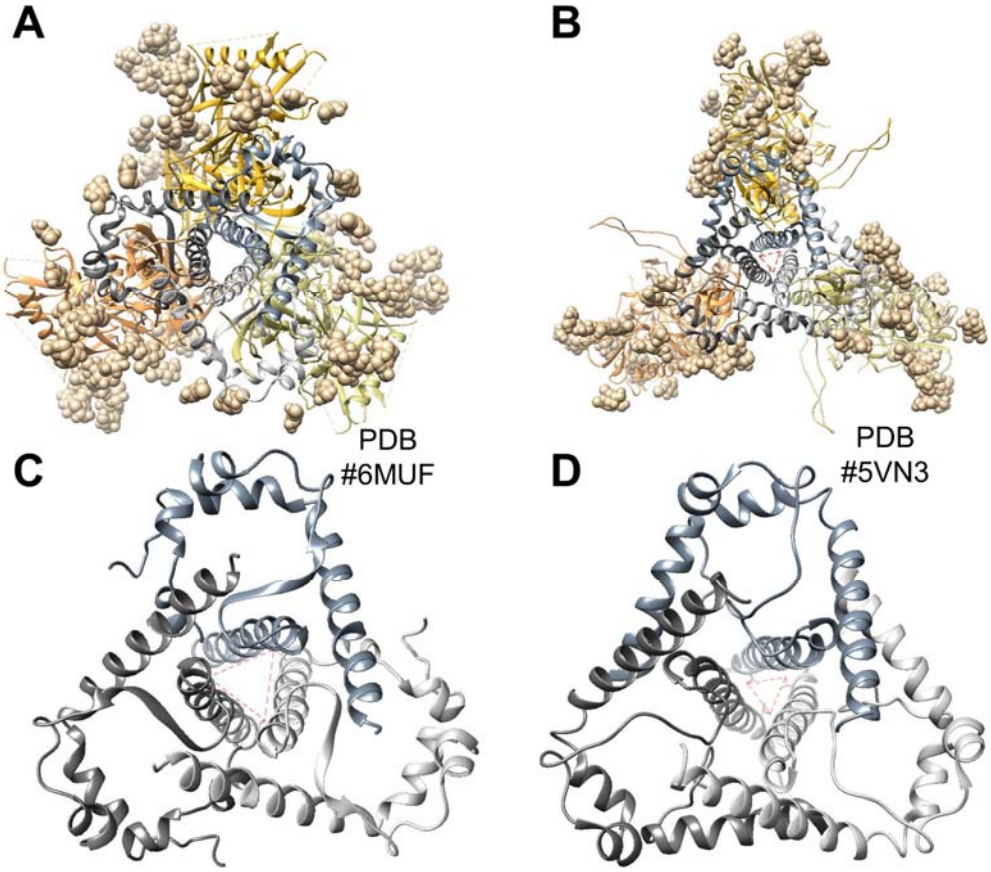


Figure 1.14 Comparison between unbound and CD4 bound gp41 trimer conformations. **A-D)** show HIV trimers with gp120 in various shades khaki and gp41 in various shades of grey. **A)** Shows the gp160 trimer (PDB 6MUF[254]) in the unbound state. This 'closed' trimer is approximately 105 Å high by 113 Å wide. **B)** Shows the gp160 trimer (PDB 5VN3[267]) in the CD4 bound conformation. This 'open' conformation is approximately 118 Å high x 152 Å wide, The V3 loop extending downwards causes the main difference in height. **C & D)** compare the gp41 trimer in the unbound (C) and CD4 bound (D) state. A faint red equilateral triangle is shown in the center of the helical bundle. It connects the oxygen atom of T569, near the apex of the trimer. The unbound conformation has a much wider aperture, with each side of the triangle measuring 12Å (C), compared to the 7Å per side of the CD4 bound structure (D).

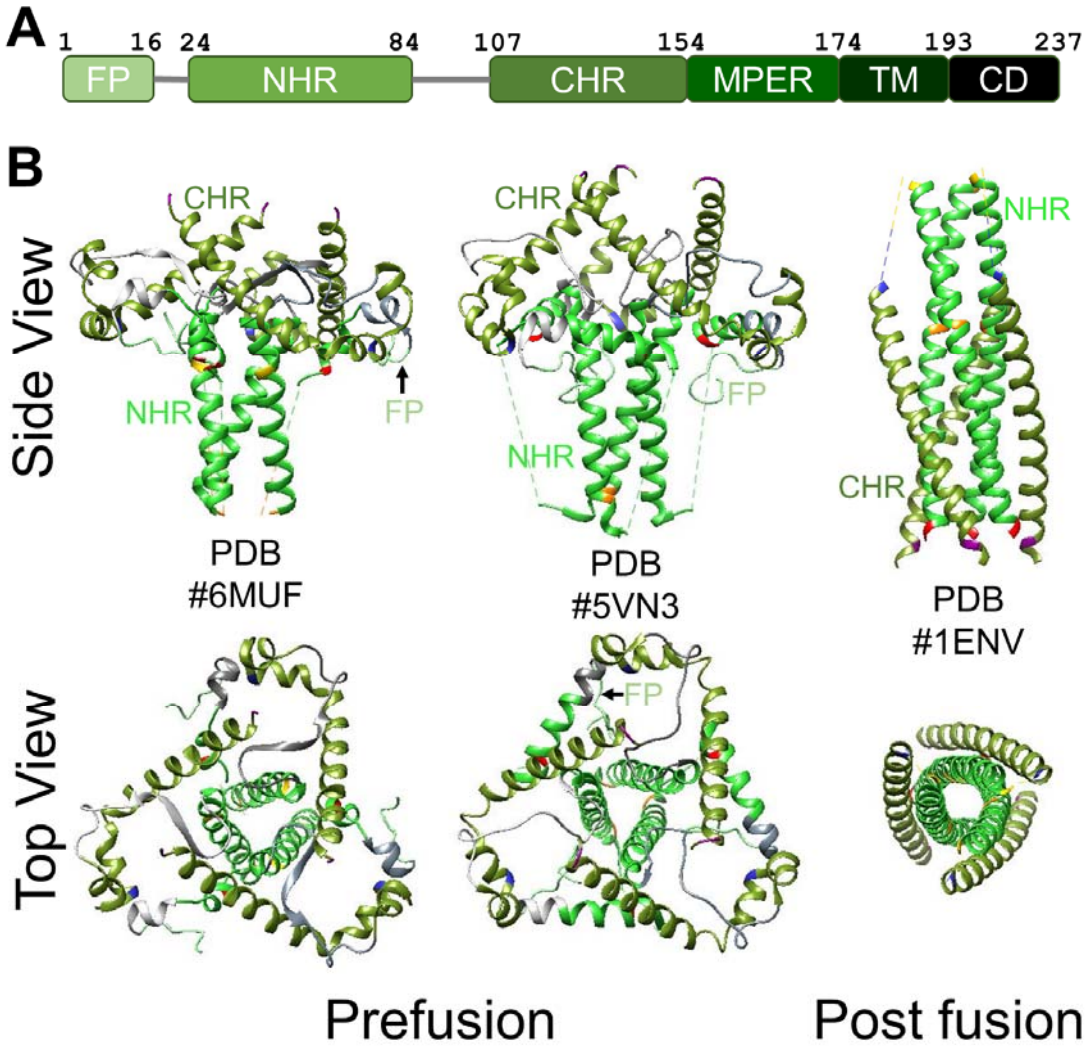


Figure 1.15 Structures of the gp41 trimer at various stages of infection.
A) Shows the general regions in the gp41 monomer. The abbreviations found in the figure are as follows: FP – Fusion peptide, NHR – N terminal Heptad Repeat, CHR – C terminal Heptad Repeat, MPER – Membrane-Proximal External Region, TM – Transmembrane domain, CD – Cytoplasmic Domain. **B)** Compares the gp41 trimer in the unbound (PDB 6MUF[254], residues 8-34 & 58-152) the CD4 bound (PDB 5VN3[267], residues 4-37 & 52-153) and post fusion (PDB 1ENV[179], residues 32-77 & 117-154) structures. The fusion peptide is colored in a sea foam green, the NHR is colored in bright green and the CHR is colored in an olive green. Corresponding residues are colored in each of the structures to orient the viewer between the conformational states. The side view structures are oriented so the top would be towards the viral surface and the bottom would point towards the host cell. In the post fusion structure, both membranes would be fused together at the bottom of the image since the fusion peptide (which would be inserted into the host membrane) is extended from the residue colored red, and the TM domain of gp41 would be in the extension of the structure after the residue in purple.

Once the gp41 fusion domain is exposed, it inserts into the host membrane. HIV's gp41 is defined as a class 1 viral fusion protein, meaning that in the post-fusion state the protein will form a six-helix bundle[286–289] with the N-terminal helices in a triple coiled-coil, and the C terminal helices packed in an anti-parallel orientation, into the grooves[23,290–293]. The formation of the six helix bundle draws the viral and host membranes together[23,294,295] (Figure 1.16). Research has proven that increased numbers of envelope spikes per virion does increase infectivity[296,297], and it is likely that multiple viral spikes are necessary to fully bring the membranes together and create a pore for the viral RNA and proteins to pass through into the host cell[280,298–300]. There is debate whether gp120 fully detaches or 'sheds' away at this point in order for the gp41 to fully undergo the conformational change to a six-helix bundle[268,301–304].

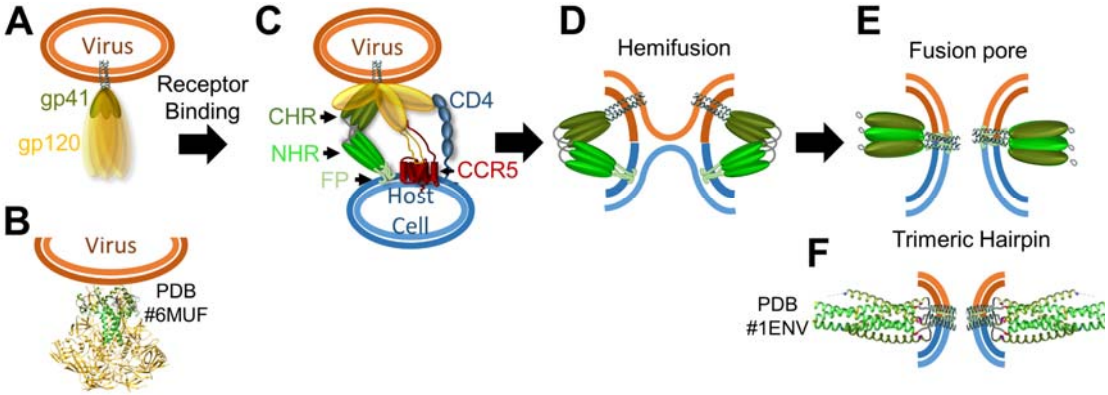


Figure 1.16 gp41 fusion to the host membrane and formation of the fusion pore. This figure walks through the steps of forming a fusion pore. The abbreviations are as follows: CHR – C terminal Helical Region, NHR – N terminal Helical Region, FP – Fusion Protein. **A)** Shows the original, unbound conformation of the HIV envelope proteins on the viral surface with gp41 colored green and gp120 colored gold. **B)** Shows a crystal structure of the envelope proteins in the conformation represented in part A (PDB 6MUF[254]), the glycans and stabilizing Abs have been hidden for clarity. **C)** Shows the theoretical conformation after gp20 has bound to both CD4 and CCR5 and the fusion peptide of gp41 has inserted into the host membrane. Currently, there are no crystal or NMR structures of this conformation. **D)** Shows the theoretical hemifusion state where the inner leaflet of the viral and host membranes has begun to fuse. **E)** Shows the formation of the fusion pore, where gp41 has formed a trimeric hairpin structure. F) Shows a crystal structure of gp41 in the trimeric hairpin conformation (PDB 1ENV[179]). This figure is modeled on Figure 1 from Melikyan's review, Klug et al and Hu et al., Figure 2 of Dom and More's review and Figure 5 from Liu et al. [286,913–916]

1.3.4 HIV replication in the host cell

Once the fusion pore has been created and the viral proteins have entered into the cell, HIV replication can be divided into four basic steps: 1) reverse transcription, 2) integration into the host genome, 3) protein production, and 4) viral assembly. Since the process of retroviral replication and the proteins of HIV have been thoroughly characterized in other reviews[305–310], they will only be briefly summarized here. **Error! Not a valid bookmark self-reference.** briefly describes all the relevant proteins and their function in the HIV lifecycle for quick reference. It should be noted that HIV-2 does not contain Vpu, it contains Vpx instead[305,307,311,312].

The first step of viral reproduction after the viral proteins have entered the host cell is reverse transcription of the viral RNA into double stranded DNA by the reverse transcriptase complex, which includes the viral proteins RT, RNase H, Pro, IN, MA, CA, NC, Vif, Tat, Nef and Vpr, with some host proteins as well. The exact composition of the complex may vary and is still under debate[313–317]. Once the dsDNA is produced, it is integrated into the host genome by the preintegration complex. The location of integration will affect the rate of transcription and integration into transcriptionally silent regions can lead to latent infection[318–323] (for more on the HIV latent reservoir, please see the section below). Integration into the host genome is key for viral persistence in all retroviruses[324–326].

The third step of viral reproduction, expression of the HIV genome, occurs in two stages, early and late, because HIV produces many polycistronic mRNAs, i.e. mRNAs that contain open reading frames for multiple proteins[305,327–329]. The fully spliced forms of the genes are transcribed in the early phase of gene expression, producing Rev, Nef and the long form of Tat. These proteins can even be transcribed before integration of viral DNA into the host genome. The late stage of gene expression is regulated by Rev and produces unspliced and incompletely spliced RNA. Expression of the unspliced mRNA produces Gag and Gag-Pol precursor proteins. The unsliced mRNA is also what will be packaged into new virions. Incompletely spliced mRNAs can produce Env, Vif, Vpu, Vpr and the shorter form of Tat. Many of the HIV viral proteins have additional functions to the ones listed in **Error! Not a valid bookmark self-reference.** that are dictated by alternate splicing and pairwise protein interactions[75,305,311,328,329].

The final step of viral reproduction is the formation of new viruses. Envelope proteins gp120 and gp41 are transported to the membrane through the secretory pathway after they are finished being processed in the Golgi[330–332]. Gag and Gag-Pol proteins bind to the plasma membrane and mediate the formation of the new capsid. PR cleaves Gag into its active forms of MA, CA, NC and p6 which trigger packaging of the RNA genome and assembly of the inner, canonical capsid[305,333–335]. HIV hijacks the host ESCRT pathway (endosomal sorting complex required for transport that is involved in cargo-recognition and membrane remodeling[336]) to complete the budding process[333].

Table 1.3 Summary of HIV proteins[305–310]

Category	Gene	Proteins	Function
Structural Proteins	Env	Gp160 > gp120 & gp41	Gp120 & gp41 form trimeric spikes on the viral surface that bind to the host cell receptors and initiate fusion to the host cell
	Gag	MA (matrix/p17), CA (capsid/ p24), NC (nucleocapsid/p9) & p6	These proteins make up the core of the viral capsid. MA forms the outer-most layer, underneath the lipid bilayer. CA forms the conical core of the capsid. The NC protein recognizes the packaging signal on HIV RNA and incorporates the RNA into the forming virions. P6 enables incorporation of Vpr into budding virions by facilitating interactions between the precursor Gag protein and Vpr.
Viral Enzymes	Pol	Pro (protease/p10), IN (integrase/p31), RT (reverse transcriptase/p50) & RNase H (p15)	Pro is an aspartyl protease that cleaves the Gag and Gal-Pol precursor proteins. RT uses the viral RNA to produce a copy of double stranded DNA. RNase H degrades the original RNA strand once the first strand of DNA has been made, permitting production of a complementary strand. IN is responsible for inserting the newly produced viral DNA into the host genome; this single enzyme integrates exonuclease, endonuclease and ligase functions.
Regulatory Proteins		Tat, 72aa & 101aa forms	Tat is a transcriptional activator that promotes the elongation of HIV mRNA by recruitment of serine kinase CDK9 using the cellular co-factor Cyclin T.
		Rev	Rev exports unspliced and incompletely spliced mRNA's out of the nucleus in the late phase of HIV gene expression to produce proteins such as Vpr, Vpu and Vif
Accessory Proteins		Vpu (Viral protein U)	Vpu induces degradation of CD4 that binds to envelope proteins in the process of virion assembly and enhances release of newly formed virions.
		Vpr (Viral protein R)	Vpr facilitates pre-integration complex localization to the nucleus and is associated with nuclear pores
		Vif (Viral infectivity factor)	Vif is essential for replication in specific cell lines including peripheral blood lymphocytes and macrophages and is incorporated into the virions of HIV to overcome antiviral factors produced by host cells.
		Nef (negative factor)	Nef downregulates the cell surface expression of CD4 and class I MHC, optimizes cellular conditions for viral production and increases HIV infectivity

1.3.5 HIV's use of glycosylation to evade the immune system

HIV uses glycosylation, oligosaccharides covalently bound to the protein surface, to shield itself from the host immune system[269,330,330,337–342]. There are five types of glycans: phosphoglycans, C-linked glycans, glypiation, O-linked and N-linked, with the final two being the most common[343,344]. HIV-1 is glycosylated with only N-linked glycans[330,345] and so that is the process that will be briefly explained here. Glycosylation is added to asparagine residues in the endoplasmic reticulum (ER) at the consensus sequence NXS/T, where X is any amino acid except for proline[346–349]. A preformed oligosaccharide precursor (Figure 1.17A) is transferred *en bloc* to the asparagine by olygosaccharyl transferase in the lumen of the ER[344]. As the protein passes through the ER, the glucose molecules are cleaved off. The glycan then passes on to the golgi through either the glucosidase-dependent and independent pathway, where it goes through a variable amount of processing that is dependent on the glycan accessibility to the enzymes[182,350–355]. High mannose glycans (Figure 1.17B) are the least processed glycans. Mannoses are

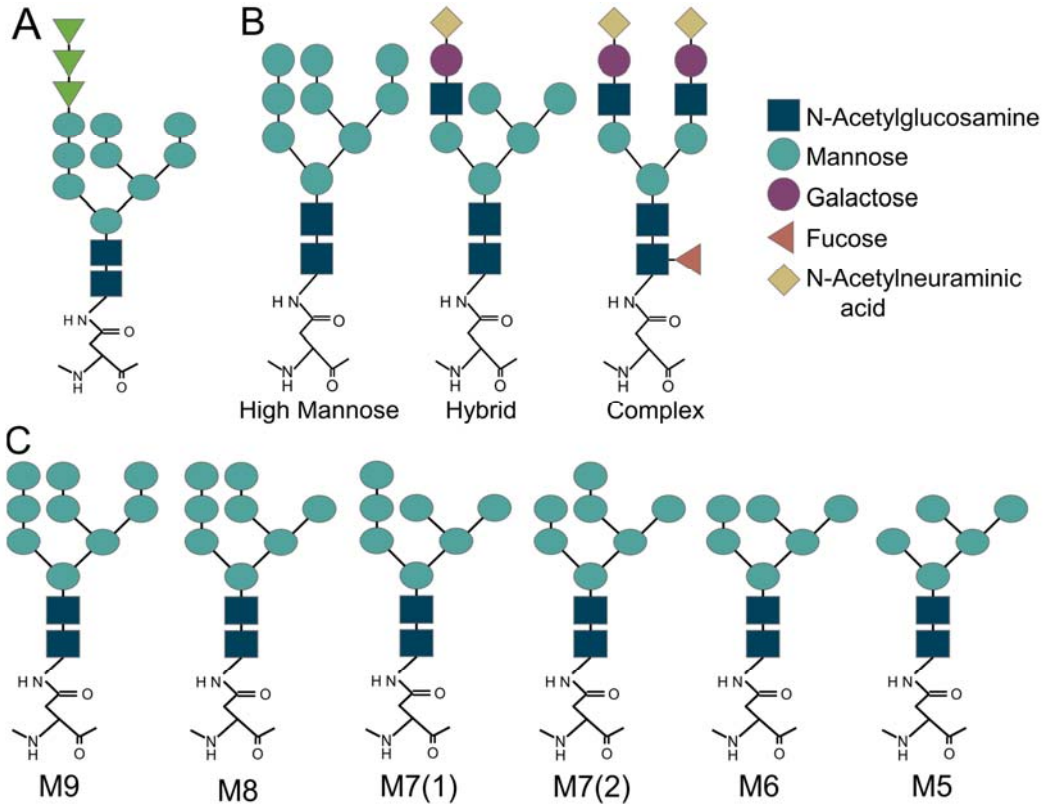


Figure 1.17 Glycan structures
A) shows the structure of the oligosaccharide precursor. **B)** shows the three main structures of glycans in order from least to most processed. **C)** shows how high mannose glycans are processed to remove mannose residues, starting with the least processed, M9, to the most processed, M5. It is possible to have further processing down to M4. Branch points are connected by α1-3 and α1-6 linkages, while linear connections are α1-2 linkages.

removed as glycans are more processed, and may eventually be replaced with other sugars including galactose, fucose, N-acetylglucosamine and N-Acetylneuraminic acid (Figure 1.17C). When only one branch of the glycan displays other sugars, it is classed as a hybrid glycan; when both of the branches contain variable sugars, the glycan is now termed complex (Figure 1.17C).

The glycan shield, interlinking oligosaccharides that composes approximately half of the HIV envelope, is used to circumvent the host immune system, shielding rapidly mutating epitopes on the envelope proteins gp120 and gp41[356,357]. Because the glycans are similar to those found on host proteins, the glycan shield makes most of the viral spike immunologically "silent"[358,359]. Even slight shifts in glycan location within the shield can completely disrupt binding of broadly neutralizing antibodies that include glycans in their target epitope[360–363]. Figure 1.18 shows the glycan shields of envelope proteins from four different clades. While the glycosylation pattern of each strain is variable there are regions where glycans are densely crowded which minimizes the processing of those glycans, resulting in a higher proportion of high mannose glycans in those regions[340,364–368]. There is one large region referred to as the mannose patch, that is retained across all clades, throughout disease progression and immune challenge[340,352,369–371].

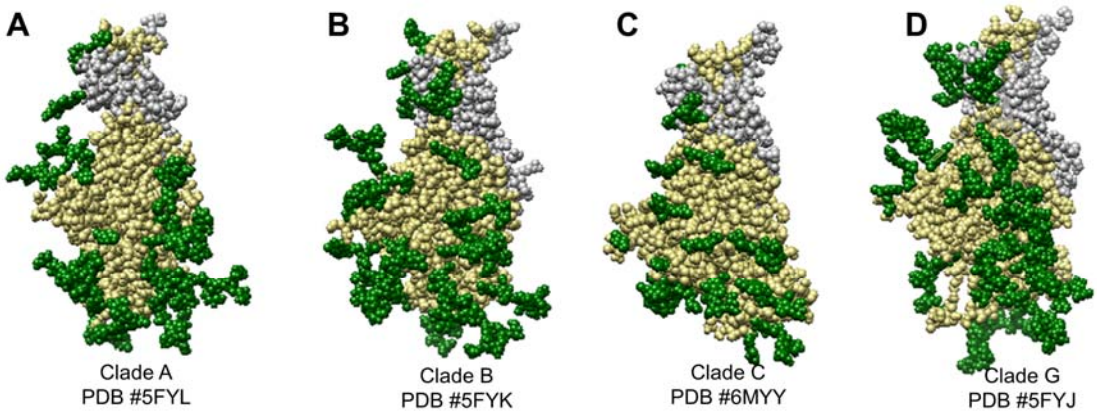


Figure 1.18 Glycan shields of various clades of HIV.
A-D) Show HIV envelope monomers with gp120 in khaki, gp41 in grey and glycans in green. The clade of each monomer is listed underneath the crystal structure.

Previous research showed that individual glycans have importance to the infectivity and folding of the virus[372]. Removal of glycan N262 (N262Q) resulted in complete loss of infectivity, and N>Q mutations of sites N197, N234, N241 and N289 yielded a significant decrease in viral infectivity[372,373]. This has since been disproven by the Varadarajan group, who demonstrated that it was the method of mutation, not the loss of glycosylation that interfered with viral folding and infectivity[364]. This confirms that the primary purpose of glycosylation of HIV envelope proteins is to create an evolving shield that protects the virus from the host immune system[337–339,374–377]. Specific glycans shield key epitopes on the gp120 surface. Glycosite N386 is key to protecting the CD4 binding site from antibodies, with assistance from sites N197 and N301[373,378]. There are antibodies that target the glycan shield (broadly neutralizing antibodies, bNAbs), rather than the protein epitopes. Extensive research has been done to understand the process necessary to illicit

bNAbs to develop an efficient vaccine and microbicides against HIV. BNAbs and their relation to HIV treatment are discussed further in a later section.

1.4 Microbicides and Prevention of HIV

1.4.1 Antiretroviral Therapy

Six years after the first AIDs patients were discovered, a therapeutic finally became available: AZT (also known as zidovudine)[379,380]. AZT is a nucleoside reverse transcriptase inhibitor (NRTIs), which inhibits reverse transcriptase. In the 30 years since AZT became available numerous other NRTI's, with fewer side effects, have been adopted. To minimize drug resistance, the current therapeutic regime, called combination antiretroviral therapy (cART) combines NRTIs with non-nucleoside reverse transcriptase inhibitors (NNRTIs), protease inhibitors, integrase inhibitors and infection/entry inhibitors (for a list of common drugs used in cART please see Table Appendix A.3). The development of entry inhibitors is arguably the most important since they prevent entry of the HIV virus into the cell, whereas the other classes of inhibitors only inhibit HIV replication after the virus has entered the cell. The first entry inhibitor developed was T-20 (aka Fuzeon or Enfruvitide)[381], which binds the CHR of gp41 to prevent fusion with the host membrane. In the nearly 20 years since the development of Enfurvitide, entry inhibitors have expanded to encompass three main classes of drugs – fusion inhibitors, co-receptor binding inhibitors and attachment inhibitors[382]. Less than a handful of entry inhibitors have been approved in the last 20 years, but there are many currently undergoing clinical trials[79,243,382–395]. Entry inhibitors are desperately needed, since they work differently from the other classes of drugs available, which are encountering viral resistance from overuse[86,229,312,389,396–402]. Many of the attachment inhibitors that are under development come from a single class of carbohydrate binding proteins[403–409].

1.4.2 Lectins

Lectins, originally called agglutinins, are a class of carbohydrate binding proteins that was discovered in extracts of castor bean seeds back in 1888 and immediately proved useful to the medical community for their ability to distinguish blood types[343]. Lectins have since been found in most prokaryotic and eukaryotic organisms, including humans. There is some evidence that the human immune system utilizes soluble and membrane-associated lectins to neutralize and degrade viruses[410,411], but unfortunately may viruses have adapted to exploit lectin receptors, such as DC-SIGN, to enter target cells, enabling viral spread[412–416].

In 1997, the lectin Cyanovirin-N was discovered to be protective against the latest medical epidemic - HIV. Isolated from cyanobacterium Nostoc ellipsosporum, this 11kDa monomeric protein has two symmetrically related domains, that each contain a CBS (carbohydrate binding site) that binds to an 1-2 oligomannose moiety, which are found on Man-8 and Man-9 glycans(Figure 1.17). Despite tight binding, and a broad specificity, safety concerns including mitogenic activity and stimulation of increased cytokine production[417] (disputed by some who continue to make advances on CV-N translational products[418–421]) left researchers looking for other lectins with fewer detrimental effects. Over the next several years, many other lectins with virucidal properties against HIV were discovered[403–409,418,422–439], and the search continues today[424,440,441]. Since lectins bind to

any glycan containing their specific carbohydrate, lectins have been shown to inhibit a broad range of other enveloped viruses, including Ebola, influenza[406,442], Coronaviruses[443,444], herpes[440,445], human papilloma virus (HPV)[446–448], hepatitis B and C[19–22,408,449–453], and Marburg[454,455]. The continue for lectins with antiviral properties

Lectins are differentiated by their carbohydrate recognizing domains (CBDs) with shallow pockets that recognize terminal groups on glycans. There are two primary folds that lectins with anti-HIV properties adopt: 1) a GNA-related lectin fold with β-prism II fold – three 4-stranded β-sheets with internal pseudo threefold symmetry, where the β-strands are **perpendicular** to the 3-fold axis (shown in Figure 1.19A) [456], and 2) a jacalin-like fold with β-prism I fold – three 4-stranded, anti-parallel β-sheets with internal pseudo threefold symmetry, where the β-strands are **parallel** to the 3-fold axis (shown in Figure 1.19B) [457]. Two outliers to these common folds are CV-N and MVL. Cyanovirin-N has a fold that is still formed from anti-parallel β-sheets but exhibits twofold symmetry instead of threefold symmetry (Figure 1.19C)[438]. *Microcystis viridis* lectin (MVL) is the only anti-HIV lectin that contains an alpha-helix integrated with an anti-parallel β-sheets within a twofold symmetrical domain (Figure 1.19D)[407].

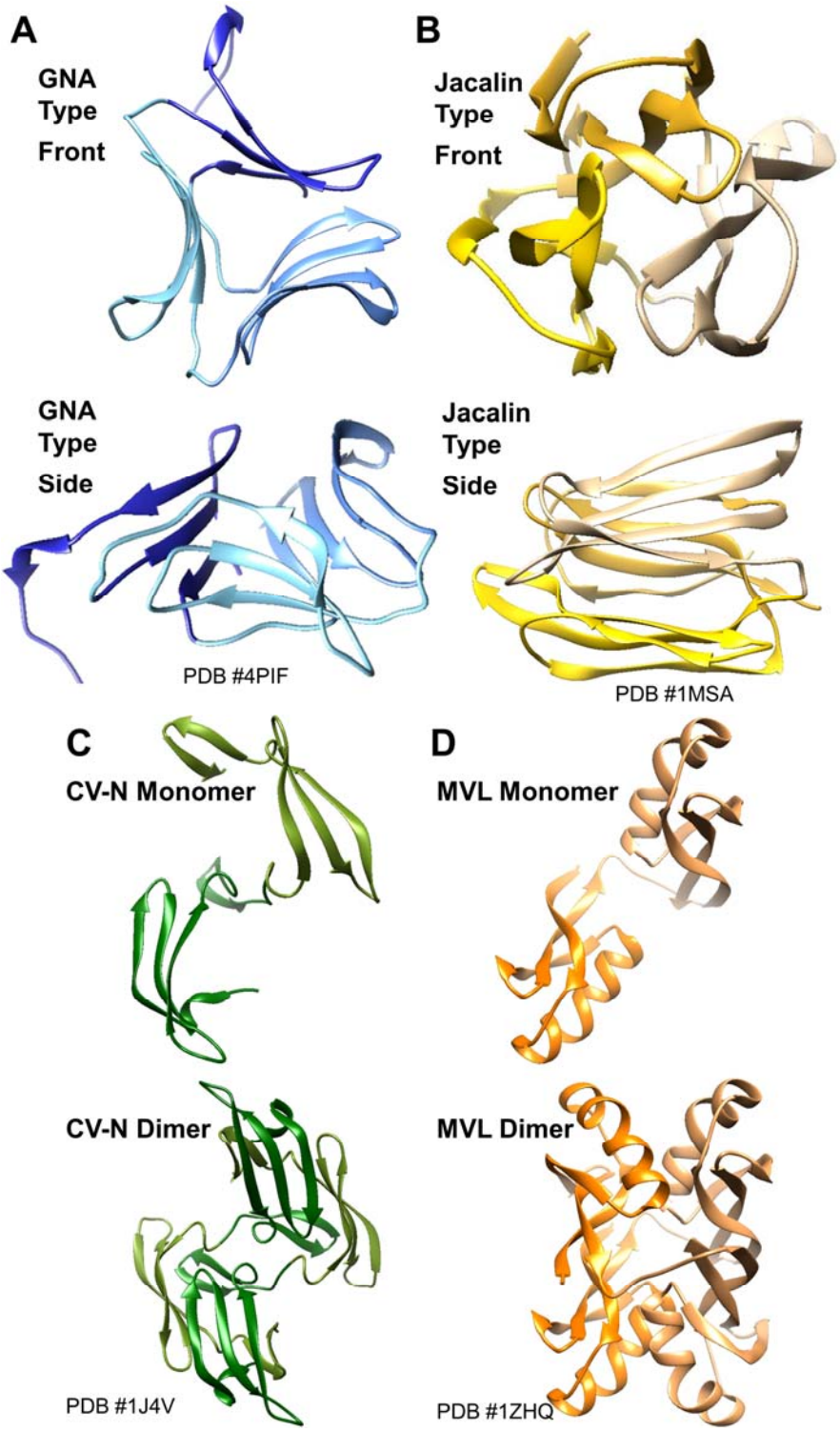


Figure 1.19 The Four Types of Lectin Structures. A) Shows the GNA type fold from the front and side, **B)** shows a Jacalin type fold from the front and side, **C)** shows a CV-N monomer and dimer, and **D)** shows an MVL monomer and dimer.

1.4.3 Griffithsin

The lectin Griffithsin (Grft) was discovered in 2004 by Mori et al. The protein was derived from the red algae, *Griffithsia* sp. (full classification shown in Figure 1.20)[458]. There are over 41 recognized species found around the world[459]. Griffithsia species had previously been studied for the biliprotein R-phycoerythrin, which can be used for natural food coloring and fluorescent tagging[460,461]. The Boyd group specialized in the discovery of novel therapeutic agents with anti-HIV activity from natural products[462]. Griffithsin was theorized to be a potential anti-HIV therapeutic based on the previous success of lectins Cyanovirin-N and scytovirin from cyanobacterium[463,464]: it proved to be the most potent anti-HIV lectin of all.

Empire	Eukaryota
Kingdom	Plantae
Subkingdom	Biliphyta
Phylum	Rhodophyta
Subphylum	Eurhodophytina
Class	Florideophyceae
Subclass	Rhodymeniophycidae
Order	Ceramiales
Family	Wrangliaceae
Tribe	Griffithsieae
Genus	Griffithsia

Figure 1.20 Taxonomy of Griffithsia Sp. Based on data from algaebase.org

Grft is a 121 amino acid protein with a non-canonical amino acid at position 31, that naturally coalesces into a domain swapped dimer (Figure 1.22)[465,466]. Based on a BLAST data base[467–472] sequence analysis, the highest sequence homology to Grft is 45.6% from a Jacalin-type lectin artocarpin[473]. What makes Grft a particularly strong microbial agent, is the presence of three sugar-binding sites on each monomer that form a nearly equilateral triangle (Figure 1.21); site 1 (M1 in Figure 1.21) is conserved in lectins, site 2 (M2) is found in only one other lectin, while site 3 (M3) is unique among Jacalin-like lectins[405,407,418,438]. The distance between the three sites is approximately 14 angstroms and the carbohydrate binding faces of each monomer are 50 angstroms apart and at a relative angle of ~160° to each other[474]. A summary of the main published structures of Grft are shown in Table 1.4 with a description of the Grft characteristics that each structure elucidated. The LiWang group and others have shown that the dimer and all six mannose binding sites are critical for Grft potency against HIV[475–477]. Like all lectins, Grft binds carbohydrates, with a particular affinity for mannose and oligomannose variants (Table 1.5)[429,477–479]. Grft's strong affinity for HIV gp120 contributes to IC_{50}'s against HIV in the high picomolar to low micromolar range[408,423,429,431,432,475,480–484].

While this lectin comes from algae, this lectin can be produced in large quantities from a variety of sources, including E. coli[466,485], rice endosperm[486], and tobacco (N. benthamiana)[487–490]. Grft has gone through pre-clinical safety testing in mice[429,491], rats[492], rabbits[489], macaques[448,493] and human explants[489,494–496] to confirm low to no toxicity, minimal inflammation, and no mitogenic activity. Griffithsin is currently undergoing human clinical trials under the supervision of Dr. Palmer at the University of Louisville[497].

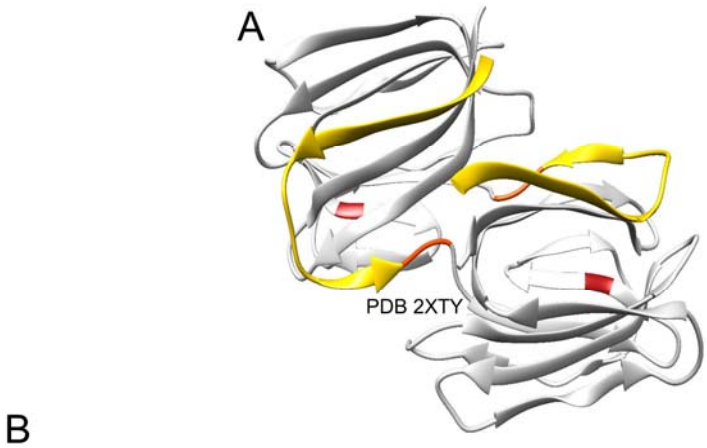


Figure 1.22 Crystal structure and Sequence of Grft domain swapped dimer.
A) The domain swapped dimer of Grft with swapped domains shown in yellow, hinge regions shown in orange and the non-standard amino acid (#31) is shown in red. PDB 2XTY. B) The sequence of Grft, with the non-standard amino acid marked X.

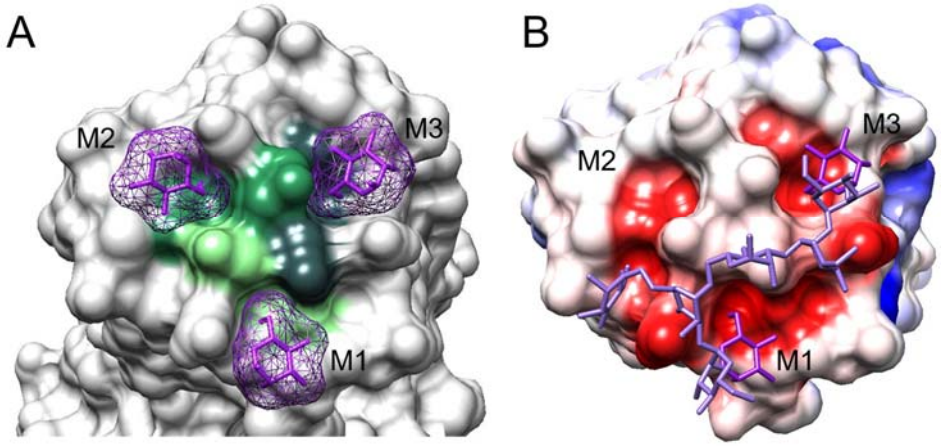


Figure 1.21 Mannose binding sites of Grft.
Grft has three sugar binding sites, designated M1-3. **A** identifies the three binding pockets in progressively darker shades of green to show the critical amino acids in sites M1 through M3 binding to the mannose residues, which are shown in purple (PDB 2GUD). **B** shows a Grft monomer binding to a high mannose glycan. The surface of Grft has been colored according to its electrochemical potential as calculated by Coulomb's law (using the Coulombic surface coloring tool in UCSF Chimera[917,918], https://www.cgl.ucsf.edu/chimera/docs/ContributedSoftware/coulombic/coulombic.html) with red indicating negative charge and blue indicating positive charge. The glycan has been colored in shades of purple, with the darker shade to differentiate the mannose residues that fit specifically into the binding pockets (PDB 3LL2).

Table 1.4 Summary of Grft Structural Studies

Structure Studied	Structural and Functional Characteristics	PDBs
Grft alone & with mannose	Jacalin related lectin, domain swapped dimer with a β-prism motif and three mannose-binding sites pre monomer	Alone 2GTY[498] Mannose 2GUC, 2GUD[498]
Monomeric Grft alone & with nona-manoside	Binding of monomeric Grft to two different nona-mannosides	Monomer alone 3LKY,3LL0,3LL1[478] Nona-manoside 3LL2[478]
Grft with glucose or N-acetylglucosamine	Glucose and N-acetylglucosamine are bound by Grft in a similar way to mannose	Glucose 2NUO[465] N-acetylglucosamine 2GUE[498] 2NU5[465]
Grft with 1 → 6α – Mannobiose or maltose	The binding of maltose is weaker than the binding of mannose. The binding mechanism of 1 → 6α – Mannobiose is similar to mannose	1→ 6α-Mannobiose 2HYQ[479] Maltose 2HYR[479]

Table 1.5 Grft Affinities to Sugars

	Affinity	
	Grft Dimer	**Grft Monomer**
Maltose	394 uM[479]	Not published
Mannose	102 uM[479]	Not published
1 → 6α - Mannobiose	83 uM[479]	85 uM[478]
Mannopentose	19 uM[478]	15 uM[478]
Nonamannoside	500 pM[478]	510 pM[478]
gp120	8-69 pM[429,477,478]	112-386 pM[477,478]

Grft's carbohydrate binding activity also allows potent activity against other enveloped viruses including severe acute respiratory syndrome-Coronavirus (SARS-CoV), Hepatitis C virus (HCV), herpes simplex virus 2 (HSV-2), Japanese encephalitis virus (JEV), human papillomavirus (HPV), Middle East respiratory syndrome-Coronavirus (MERS-CoV)[20,429,443,447,499,500] and Nipah virus[501].

Microbicides such as Griffithsin (Grft) can be used in topical or oral solutions to prevent HIV infection. Unfortunately, there is not a one size fits all prevention plan. A recent clinical trial of tenofovir gel showed that the vaginal microbiome has significant effects on the efficacy of the microbicide[502]. This highlights the need for a diverse compliment of microbicides that can be used alone or in combination to prevent HIV infection. Grft is one of the most promising microbicides currently undergoing clinical trials. Griffithsin has been tested in both quick and long-term release formulations. Subcutaneous injection of Grft was bioavailable and detectable at therapeutic levels up to 96 hours post injection[492]. Bioavailability refers to the ability and rate at which a drug (in this case a protein) is absorbed into the circulatory system and detectable in the serum. Oral dosages of Grft were not bioavailable but were able to prevent rectal transmission of HIV-1[492]. Formulations encapsulating Grft in silk disks[496] and electrospun fibers[503–505] have been used for both short term and long-term release[496,503–507]. Silk formulations have been used to help stabilize other microbicides, but Grft is already highly stable, with a melting temperature of 78.8°C[474].

1.4.4 Broadly Neutralizing Antibodies and Vaccines

While there has been progress in developing drugs that can prevent initial infection, there still has not been any success in developing a vaccine to prevent HIV infection. The current, optimistically, anticipated date for a widely available vaccine is 2030[508–512]. There are those who believe it may take another 20-30 years before a vaccine is widely available[513–515]. Since no vaccine confers 100% protection, there must be a balance between the effectiveness of the vaccine (and the number of different vaccines required to cover all the different clades epidemic in different regions) with the cost of production and delivery of vaccinations in areas where HIV is epidemic[516–518].

The purpose of a vaccine is to provide an antigenic template for the immune system to create antibodies before the host is challenged with the real virus, but the rapid mutation rate of HIV makes it extremely difficult for the human immune system to produce new neutralizing antibodies that cover the breadth of HIV strains fast enough to clear the infection[30,80,374,519–524]. Many of the vaccine developers are trying to elicit broadly neutralizing antibodies(bNAbs) that recognize regions of the viral envelope that are more highly conserved[24,522,525–537] (for a hierarchy of antibody types and functionality see box on next page), which are only naturally made by less than 20% of people infected with HIV[538–545]. A sequential immunization strategy seems to be necessary to induce bNAbs, using structure guided techniques to steer B cell maturation from the production of germline target immunogens to mature bNAbs[546–550]. Traditional vaccines use neutralized viral proteins to elicit neutralizing antibodies and stimulate CD8 T cell responses[551–555]. To try and guide what antibodies are elicited, researchers developed modified, soluble HIV envelope trimers called SOSIP trimers[250,267,281,342,364,534,556–565] (for more information on SOSIP trimers please see Appendix A). These targeted responses can be guided by covering undesirable epitopes with glycosylation, referred to as glycan masking[342,566–569]. Improvement of protein imaging, primarily in electron microscopy, has rapidly improved

vaccine development by allowing researchers to determine the epitopes of the antibodies induced by vaccine trials[256,526,534,561,570–575]. The antibody neutralization response is the only reliable indication of vaccine efficacy in macaque models[576,577]. Many vaccines have proceeded to clinical trials[30,521,522,578–580].

Classes of Antibodies[919–928]

IgM – the first antibodies made to fight infectious antigens but is temporary and only found in lymph fluid and blood, large with 10 antigen binding sites

IgA – expressed in mucosal tissues, are a first line of defense

IgE – are involved in allergies and defense against protozoan parasites

IgD – signal activation of B cells

IgG – Most common, produced by B cells as a part of the adaptive immune system, highly optimized for neutralization activity against infectious antigens, found in all bodily fluids, four types – IgG1-IgG4

Antibody Functionality[528,929,930]

Binding antibodies, usually produced by the innate immune system, can attach to viruses or other antigens but cannot inhibit virus infectivity

Neutralizing antibodies are antibodies that bind to specific regions and disable the virus, are highly specific

Broadly neutralizing antibodies can neutralize many genetic variants of a virus, so far these have only been identified against HIV

Passive antibodies are doses of antibodies that have been introduced as a treatment, not produced by the patient's immune system

Figure 1.23 Antibody terminology

Broadly neutralizing antibodies are also being tested for use as passive antibodies in combinational therapies, both preventative and therapeutic drug regimes[581–590]. While the bNAbs primarily function like normal antibodies, drawing CD8 T cells to eliminate the virus, there are some broadly neutralizing antibodies that act as attachment inhibitors[337,591–593]. For example, VRC01 locks the virus in the closed state, preventing binding of gp120 to CD4[594–601].

1.5 HIV Latency and the Eradication of HIV

1.5.1 The Challenge of Curing HIV

Even years after taking antiretroviral therapy and having an undetectable viral load, there is essentially always a rebound of viral load after cessation of treatment, indicating that there are cells harboring viral DNA even during cART[322,602–604]. Life-long cART treatment extends patients' life expectancy, but is expensive, not universally available, requires life-long adherence, and confers negative side effects such as hyperlipidemia, bone loss, cardiovascular complications, and neuropathy [32,136,395,397,605–609]. Even patients

who have faithfully adhered to anti-retroviral therapy for years still harbor infected cells that don't actively expressing the HIV genome, which are therefore unable to be detected and removed by the host immune system[610,611].These cells are referred to as the "latent reservoir"[602,604,612–614].

As a retrovirus, the HIV virus permanently integrates its genome into the host cell's DNA during the viral replication process [323,602,615,616]. After release of viral proteins into the cell, the pre-integration complex (PIC) is formed from the positive stranded viral RNA, and HIV proteins, including reverse transcriptase and integrase. In the cytoplasm, RT is primed with a host tRNA and then the viral RNA is reverse transcribed first into a complementary negative strand DNA. Before the second, positive strand of DNA is transcribed, RNAse H cleaves the RNA/DNA complex to degrade the RNA and host tRNA primer until only a small section is left. This small piece is used as a primer to generate the positive strand of DNA, resulting in the final double stranded DNA, which is referred to as the provirus[313,315–317,617]. The provirus is then transported to the nucleus by the PIC where integrase facilitates integration of the proviral DNA into the host genome[323–325,618,619]. HIV seems to preferentially integrate into transcriptionally active regions, but the mechanism by which the location of integration along the chromosome is still being determined[319,321,325,326,620,621]. Once the host enzymes re-seal the chromosome with the HIV DNA now inserted, the host cell usually begins producing infective virions. However, some cells enter a non-reproductive state of infection. These transcriptionally silent cells form the HIV reservoir.

HIV reverse transcriptase is highly error prone, which accounts for HIV's rapid mutation rate and is also thought to be a common cause of latency[87]. The errors accrued in the reverse transcription process can lead to the formation of defective proviruses, most commonly via APOBEC3G-induced hypermutation[622–624]. While many of these defective proviruses are non-inducible during single round T cell activation, they still contain replication competent virus. Non-inducible virus may have also integrated into a transcriptionally silent, heterochromatin region of the host genome[321,616,623,625–628]. Activation of latent reservoirs appears to be stochastic, which increases the barrier to a cure[620,623,629–632]. Infected cells can proliferate without activating viral replication, which contributes to the stability of the latent reservoir[616,631,633,634].

The latent reservoir is comprised of many different cell types that contain transcriptionally silent virus. The most common are memory CD4 T cells, including central memory, effector memory, migratory memory, transitional memory, tissue resident memory, stem cell memory and terminally differentiated cells[635–638]. Of these subsets, central memory CD4 T cells are considered the most predominant and important[635,636,638]. Other cell types believed to contribute to the latent reservoir are macrophages (both monocyte derived and tissue resident), dendritic cells, epithelial cells, astrocytes and fibrocytes[175,412,414,608,610,635–657]. There are also tissues, such as the central nervous system, the gastrointestinal tract, and lymphoid tissues, that comprise a anatomical reservoir where low levels of viral replication can occur, which may be inaccessible to host immune CD8 T killer cells and cART drug regimes[635,658–664].

There are several strategies that have been attempted to locate and remove cells hiding HIV as a functional cure. The two main routes that have been pursued are a "Locking" the cells and "Kick and Kill" (sometimes also called block & lock and shock & kill respectively). Locking cells involves the use of genomic editing techniques to prevent any

latent HIV from being reactivated[320,626,627,650,665–668]. The kick/shock and kill technique involves stimulating cell populations most likely to be latently infected, inducing expression of HIV proteins then supporting the host immune system to remove the infected cells, sometimes with additional anti-retroviral agents[321,608,627,669–674]. Locking techniques and development have been thoroughly reviewed[320,608,626,627,632,665,674–676] and will not be discussed further here.

Shock and kill methods have shown promise in cell[660,677–681] and animal models[653,667,682–688] but, more targeted approaches for shock and kill methods are needed since many of the previously tested strategies are very broadly acting with toxic side effects[608,689–692]. Latency reversal agents (LRAs) include Histone Methyltransferase (HMT) inhibitors[693], Toll-like receptor 7 agonists[694–698], histone deacetylase (HDAC) inhibitors[626,670,672,699–702], and activators of the nuclear factor-κB (NF-κB) pathway [687,688,703]. HDAC and HMT inhibitors loosen the interactions between histones and DNA, unwinding the DNA and allowing transcription and production of viral proteins[704]. NFκB is a protein complex that controls transcription of DNA[705] , so activation generally up-regulates transcription in a cell, increasing the chance of viral expression.

The most natural and potent enhancer of HIV transcription, that would not be expected to induce transcription of non-HIV DNA, is the HIV protein Tat. The Tat protein is responsible for ensuring viral transcription during the viral lifecycle. Tat stimulates transcription by binding to the transactivation response element (TAR) in developing RNA transcripts and recruiting Cyclin T1 and cyclin-dependent kinase 9, which compose the positive transcription elongation factor b (P-TEFb) and will promote transcriptional elongation by recruiting and phosphorylating RNA polymerase II, releasing the elongation inhibitor complexes termed NELF and DSIF[706–715]. Tat has also been shown to bind the upstream promoter regions[716]. Research has shown that insufficient Tat transactivation activity can result in latency[620,652,717–720]. Tat inhibitors are being investigated as a means to lock latently infected cells[666,668,721]. Tat has also been investigated as a shock agent[718,719,722–725], but issues with cytotoxicity[718,726,727] highlight a need for a method to target Tat and other LRAs to cells most likely to be infected with HIV.

Chapter 2

Griffithsin retains anti-HIV-1 potency with changes in gp120 glycosylation and complements broadly neutralizing antibodies PGT121 and PGT126

2.1 Abstract

Griffithsin (Grft) is an antiviral lectin that has been shown to potently inhibit HIV-1 by binding high mannose N-linked glycosylation sites on HIV-1 gp120. A key factor for Grft potency is glycosylation at N295 of gp120, which is directly adjacent to N332, a target glycan for an entire class of broadly neutralizing antibodies (bNAbs). Here we unify previous work on the importance of other glycans to Grft potency against HIV-1 and Grft's role in mediating conformational change of gp120 by mutating nearly every glycosylation site in gp120 individually and in combination. In addition to significant loss of Grft activity by removal of glycosylation at N295, glycan absence at N332 or N448 was found to have moderate effects on Grft potency. Interestingly, in the absence of N295, Grft effectiveness could be improved by mutation that results in the glycan at N448 shifting to N446, indicating that the importance of individual glycans may be related to their effect on glycosylation density. Grft's ability to alter the structure of gp120, exposing the CD4 binding site, correlated with the presence of glycosylation at N295 only in B clade strains, not C clade strains. We further demonstrate that Grft can rescue the activity of bNAbs PGT121 and PGT126 in the event of loss or shift of glycosylation at N332, where the bNAbs suffer a drastic loss of potency. Despite targeting the same region, Grft in combination with PGT121 and PGT126 produced additive effects. This indicates Grft could be an important combinational therapeutic.

2.2 Introduction

HIV-1 infects about 2 million people per year, predominantly in developing areas such as sub-Saharan Africa[728]. No vaccine is yet available, and existing prevention methods such as daily oral PrEP are problematic in these regions in terms of availability and user acceptability[729]. Microbicides that inhibit HIV-1 infection represent a promising strategy for prevention of HIV-1, and formulations are envisioned as insertable items that stop the sexual transmission of the virus. For instance, intravaginal rings that release small molecules or antibodies are being developed[730–733] and several proteins have recently been encapsulated into silk fibroin films that could potentially be used to prevent infection[496,605].

Several major types of proteins have been shown to be highly potent HIV-1 inhibitors, including broadly neutralizing antibodies[538,587,734,735], CCR5 binding proteins[219,384,387,736], and lectins[403,418,466,737]. Each of these types of protein has moved toward pre-clinical or clinical trials with some success[493,586,589,599,738–740].

Griffithsin (Grft) is among the most promising and potent potential microbicidal proteins. This lectin can be produced inexpensively in large quantities from tobacco leaves[487–489,741] , *E. coli*[485] and rice endosperm[486], and has been shown to have low/no toxicity, a melting temperature of 78.8°C, stability at pH 4-8, resistance to proteolytic degradation, conservation of potency after incubation in up to 50°C, and safety in mice

and macaques when applied topically, injected, or ingested [446,448,478,484,489,491–493,496,742–745]. Multiple groups have begun in-human trials [497,746]. It is arguably the most potent lectin HIV-1 inhibitor, showing nanomolar to sub-nanomolar efficacy against a wide range of HIV-1 strains[466,489,745] and has synergistic activity with currently used HIV-1 antiretroviral drugs including tenofovir, maraviroc, enfuvirtide and broadly neutralizing antibody (bNAb) VRC01[494,747]. Further, due to its ability to bind glycosylated viral surfaces, Grft has been shown to inhibit other viruses such as severe acute respiratory syndrome-Coronavirus (SARS-CoV), Hepatitis C virus (HCV), herpes simplex virus 2 (HSV-2), Japanese encephalitis virus (JEV), human papillomavirus (HPV), Middle East respiratory syndrome-Coronavirus (MERS-CoV) as well as HIV-1 and HIV-2[20,429,443,447,499,500].

Grft is a 121x2 amino acid domain swapped dimer that has three saccharide binding sites per monomer (Figure 1.22, Figure 2.1), and binds N-linked high-mannose glycans, such as Man-9, on viral surfaces with very high affinity[465,466,476,478,498]. It has been shown that both subunits of the Grft dimer are required for potent inhibition of HIV-1 despite tight binding (pM K_D) by the individual monomeric subunits to glycosylated gp120[477,478]. This seeming disconnect between affinity and inhibitory potency has led to the suggestion that while Grft may inhibit HIV-1 infection in a general way by simply binding to any high mannose site(s) on gp120, Grft may be most effective when it binds to specific regions or can cross-link between particular high mannose sites on gp120, possibly causing (or preventing) a conformational change in gp120[433,477,480]. Further insight into Grft's mechanism has come from cryo-EM studies, where the Bewley group has suggested that Grft can crosslink two separate viruses as part of its inhibition[433,748].

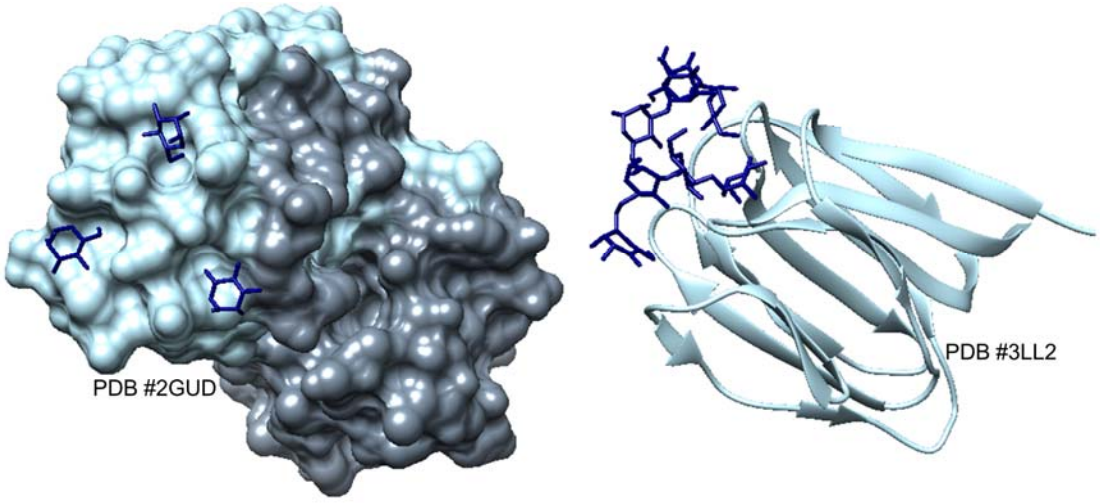


Figure 2.1 Griffithsin bound to Mannose and Nonamannoside
On the left is the dimeric structure of Grft, with one monomer in blue, the other in grey. The structure on the right shows the backbone of a Grft monomer. The three sugar binding sites per monomer are shown binding individual mannose molecules in the structure on the left, and a nonammanoside in the structure on the right, with the sugar moieties shown in navy blue. Figure 1.21 shows these interactions in greater detail.

Gp120 is found on the HIV-1 surface as a trimer[570,749], with each monomeric unit having about a dozen relatively conserved high mannose glycans that can potentially be

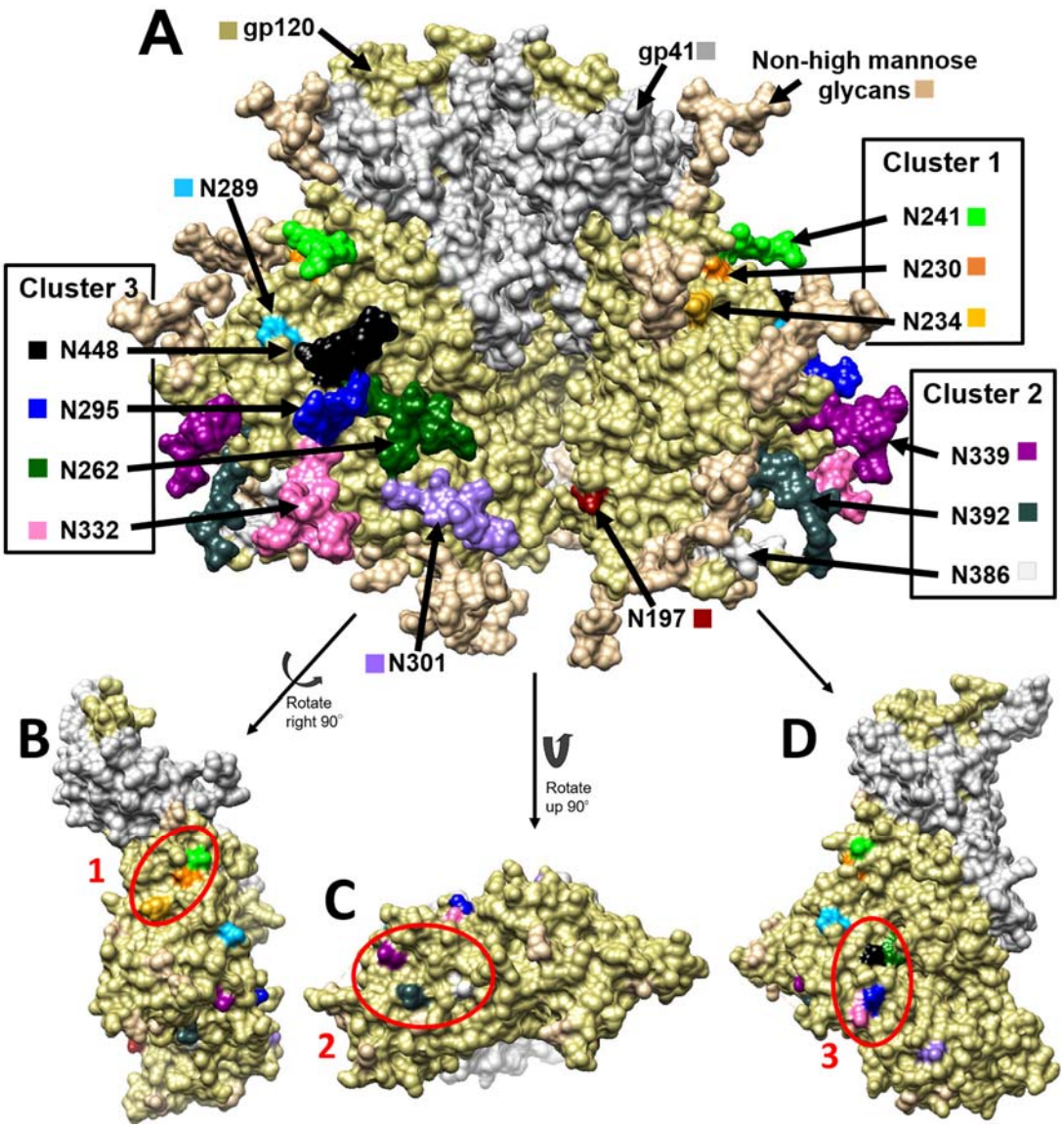


Figure 2.2 Structure of HIV-1 gp160 unbound trimer showing full glycoylation and glycan clusters.

A) Structure of the HIV gp160 trimer in the unbound, "closed" state, i.e. not bound to CD4 or CCR5. This structure of clade B JR-FL includes glycosylation (PDB 5FYK)[367] and was prepared using UCSF Chimera[917,918]. Gp120 is shown in khaki, gp41 in white. B-D Gp120 monomers with glycosylation hidden to show underlying amino acids that fall into clusters. Glycosylation sites (Asn) are shown in a variety of colors as labeled in the diagram. B) Cluster 1 encompasses the glycans at N230 (Orange), N234 (Yellow), & N241 (Green). C) Cluster 2 encompasses the glycans at N339 (Dark Magenta), N386 (Dim Grey) & N392 (Dark Slate Grey). D) Cluster 3 encompasses the glycans at N295 (Blue), N332 (Hot Pink), N262 (Dark Green), & N448 (Black). The V3 loop is highlighted in yellow to demonstrate the proximity of glycans at N295 and N332.

bound by Grft[180,368,750] (Figure 2.2A). The high mannose glycans group together to form 3 main clusters, defined by Balzarini et al.[430], as shown in Figure 2.2B-D; cluster 1 is composed of glycosites N230, N234 & N241, cluster 2 is comprised of glycosites N339, N386 & N392, and cluster 3 contains glycosites N295, N262, N332 & N448. The exact glycosylation pattern varies by strain; relevant strains are shown in Figure 2.3. The gp120 glycosylation site(s) utilized by Grft have been studied by several groups; the consensus is that glycosylation at N295 of gp120 (in cluster 3) is key to Grft potency against HIV-1. Several groups showed a correlation between the presence of the glycosylation sites at N234 and N295 with very high potency of Grft against several HIV-1 strains[474,495]. Huang and colleagues, working with different viral strains, confirmed the importance of glycosylation at N295 of gp120, but showed little effect at position N234 in the potency of Grft[482], instead suggesting glycosylation at N448 was important for Grft potency[482,483]. Cluster 3 has structural significance sitting at the head of the V3 loop, which interacts with the chemokine receptor on the host cell during HIV-1 infection. When Grft binds to cluster 3, it may inhibit the interaction between the V3 loop of gp120 and the chemokine receptor by forming a physical blockade or by impeding the conformational change necessary to expose the V3 loop in the open conformation of the HIV-1 trimer[277,282,751,752]. Further, cluster 3 includes glycosite N262 which cannot be removed without complete loss of viral infectivity[372].

Glycosylation of HIV-1 gp120 has become particularly noteworthy due to several highly potent broadly neutralizing antibodies (bNAbs) that target glycosylation in gp120 and are being considered for clinical trials[577,589,753–756]. One class of these antibodies target the glycan at N332, which is found in cluster 3 next to N295. Loss or shift of this glycan to N334 upon repetitive bNAb challenge is detrimental to inhibition by these bNAbs[362,363,734,757–759]. BNAbs that bind in this area, such as PGT121, have a range of effectiveness against these strains [362,363,537,758,760,761]. It has been reported that Grft is synergistic with the bNAb VRC01, that targets the gp120 CD4 binding site[747]. However, it is not known if Grft can be used in conjunction with bNAbs that target glycans on gp120 because Grft may compete with these proteins for binding (Figure 2.2).

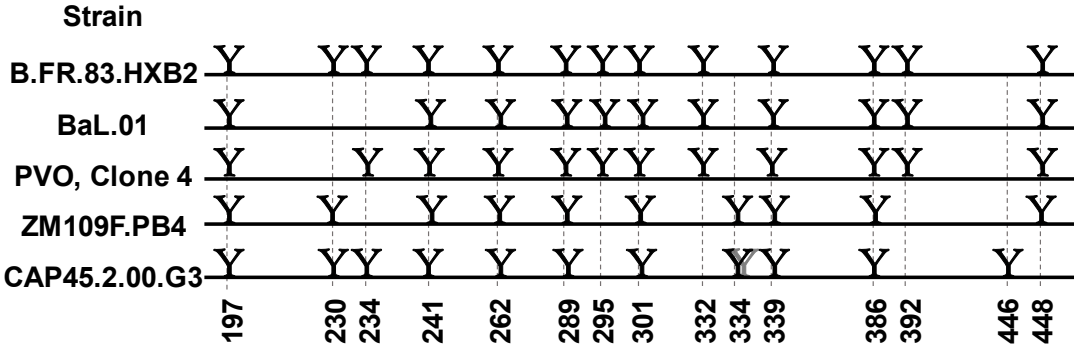


Figure 2.3 High mannose glycosylation patterns of HIV strains BaL.01, PVO.4, ZM109 & CAP45.

"Y" indicates the presence of a likely high-mannose site. The glycan on CAP45 shown in grey is at position 335 and is most likely another high mannose glycan. Cluster 3 is composed of glycans at N262, N295, N332, and N448. For a full alignment of the sequences see Supplemental Figure 1. High mannose glycosylation sites were identified through sequence analysis of each strain. Sequences can be found at www.hiv.lanl.gov.

We sought to clarify Grft's usage of glycosylation sites by extensively mutating gp120 from several strains of HIV-1 having distinct properties: two B clade strains, (PVO.4 & BaL.01) that are very sensitive to Grft and have essentially all high mannose sites present, a C clade strain (ZM109F.PB4) that is less sensitive to Grft and lacks several glycosylation sites, and a C clade strain (CAP45) that is sensitive to Grft but does not have a glycosite at 295. We specifically sought to elucidate the effect shifting glycans had on Grft potency. The sum of these mutants is the largest collection of mutants in a single paper to be tested against Grft. We demonstrate Grft's continued effectiveness in viral strains where the glycosite at N332 and/or N448 have been shifted, and that Grft is able to remain effective against strains that are resistant to bNAbs PGT121 and PGT126. This suggests that Grft could be valuable if used in combination with bNAbs that are more sensitive to such glycosylation changes. We also demonstrate that Grft-mediated conformational change of gp120 appears to be dependent on the presence of the N295 glycosite in clade B strains but not clade C strains, indicating distinct mechanisms of Grft binding to gp120 in different HIV-1 clades and therefore potentially distinct mechanisms of HIV-1 inhibition.

2.3 Results

2.3.1 Effect of glycosylation at each site in gp120 on Grft sensitivity in PVO.4 Mutants

The glycosylation patterns of several strains of HIV-1 are shown in Figure 2.3. We chose HIV-1 strain PVO.4 to resolve the effect of glycosylation placement on gp120 in relation to Grft sensitivity. This clade B strain is very sensitive to Grft, having a sub-nM IC_{50} of 0.2 ±0.09 nM, and is also glycosylated at essentially all the sites suggested to be important in Grft inhibition[431,482,483,495]. In our experiments, each individual high-mannose site in PVO.4 gp120 (except N262 due to its importance for structural stability[762] *It has since been shown that all glycans can be removed without loss of viral infectivity[364], see Chapter 4 for further information) was removed by mutating the saccharide-containing Asn to the closely related Gln, which is not glycosylated by cellular machinery. The single high mannose site that is not naturally found in PVO.4 (Asn 230), was also mutated to add glycosylation. The sensitivity to Grft inhibition for each PVO.4 variant was measured in single round inhibition assays and directly compared to wild type virus. As shown in Table 2.1, most individual mutations had little effect on the ability of Grft to inhibit PVO.4, except for mutation N295Q in gp120, which led to an 11-fold loss of sensitivity. To a lesser extent (approximately 5-fold), the mutations N448Q and N332Q showed a loss of sensitivity to Grft upon point mutation.

To determine the combined effect of each of these potentially important glycosylation sites, a double mutation was made in every combination of the sites N295, N448, and N332, as well as with the site at N234, which had previously been observed to correlate with Grft inhibition[431,495]. Each of these double deletions showed a larger effect on the ability of Grft to inhibit PVO.4, except N234Q/N295Q, which was effectively unchanged. The largest effect was caused by the N295Q/N332Q deletion, which showed an IC_{50} 43-fold higher than Grft against the wild type virus (Table 2.1 and Figure 2.4). In general, each combination that included loss of the glycosylation site at 295 showed a relatively high loss of sensitivity to Grft (Table 2.1 and Figure 2.4), with each lost glycosite at positions 332 and 448 each providing some additive effect. Mutations to remove all three of the proximal glycosylation sites (N295Q/N332Q/ N448Q) in HIV-1 gp120 were made in PVO.4 (Table 2.1 & Figure 2.4), resulting in an 88-fold loss of sensitivity to Grft. In addition, a triple mutant removing the site at 234, as well as 295 and 448 (N234Q/N295Q/ N448Q) showed a 28-fold decrease in sensitivity (Table 2.1). Removal of more than three glycosylation sites from gp120 resulted in increasingly poor inhibition by Grft. Deletion of the gp120 glycosylation site 301, in addition to the existing N295Q/N332Q/ N448Q triple mutation, resulted in a 55-fold difference in inhibition compared to wild type. The removal of the glycan at position 386, in conjunction with triple mutation, revealed a 129-fold difference from wild type. Combining these mutations (N295Q/N301Q/ N332Q/N386Q/ N448Q), led to a 126-fold difference from wild type. Continuing in a serial fashion, glycans at Asn392 and then Asn339 were removed, resulting in IC_{50}'s increasing up to 500 and 1000-fold respectively (Figure 2.5). However, mutations beyond the triple mutation N295Q/N332Q/ N448Q had a severe decrease in infectivity as well as loss of sensitivity to Grft, so the results must be interpreted with caution (Figure 2.5). The low infectivity of these latter mutants made it unfeasible to continue with serial deletions.

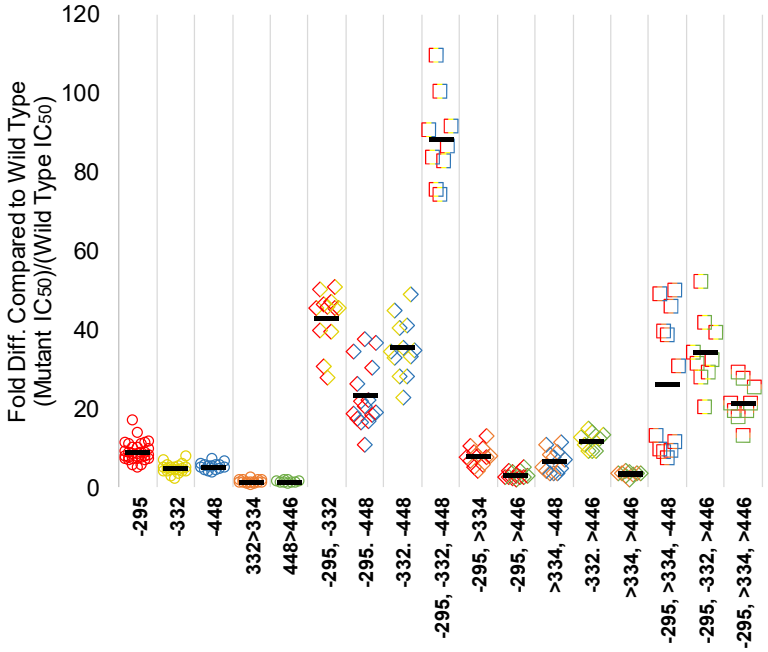


Figure 2.4 Grft inhibition against PVO.4 gp120 mutants. Each individual mutation is represented by a color: -N295 –Red, -N332 –Yellow, -N448 –Blue, N332>N334 –Orange, N448>N446 –Green. A minus sign (-) indicates that glycan has been removed. An arrow (>) indicates that glycan has been shifted. These colors are combined to represent combinational mutants. Circles indicate single mutants; diamonds indicate double mutants and squares indicate triple mutants. Glycan deletion mutants were made by mutating the Asn to Gln. For statistical analysis see Table 2.2

Table 2.1 Inhibition by Grft against HIV gp120 variants.

Strain	Mutation	FD[a]	S.E[b]	P[c]
PVO.4	Wild Type	IC_{50} = 0.2 nM	S.E 0.09	
PVO.4 Single Mutants	-197	2.7	0.58	1.6E-06
	+230	2.1	0.56	0.005
	-234	2.2	0.58	3.7E-06
	-241	3.8	0.91	7.2E-09
	-289	2.8	0.79	0.001
	-295	11.1	5.83	< 2.2e-16
	-301	1.4	0.32	6.6E-04
	-332	4.6	1.22	1.4E-08
	-339	2.8	0.59	3.5E-11
	-386	1.5	0.73	0.5
	-392	1.7	0.35	4.5E-05
	-448	5.1	0.91	9.7E-12
PVO.4 Double & Triple Mutants	-234,-295	9.8	3.2	1.7E-07
	-234,-448	6.4	2.4	9.2E-10
	-295, -332	42.9	6.9	2.2E-16
	-295, -448	23.2	7.8	2.7E-10
	-332, -448	35.5	7.1	9.4E-07
	-234,-295,-448	28.2	5.0	3.15E-13
	-295, -332, -448	88.2	10.7	8.1E-14
PVO.4 Shift of glycos. site Mutants	332 > 334	1.3	0.49	0.2
	448 > 446	1.7	0.79	0.04
	-295, 332 > 334	8.0	2.3	8.1E-09
	-295, 448 > 446	3.2	0.9	6.6E-09
	332 > 334, -448	6.6	2.7	1.1E-07
	-332, 448 > 446	10.2	3.3	1.4E-11
	332 > 334, 448 > 446	5.4	3.7	8.8E-05
	-295, 332 > 334, -448	21.2	18.0	3.8E-05
	-295, -332, 448 > 446	29.4	11.5	4.3E-09
	-295, 332 >334, 448 >446	21.5	4.7	2.8E-06
BaL.01	Wild Type	IC_{50} = 0.26 nM	S.E = 0.08	
	-295	14.1	1.9	1.0E-10
	-332	2.1	0.6	0.002
	-448	1.0	0.3	0.57
ZM109	Wild Type	IC_{50} = 26.9 nM	S.E = 6.2	
	+234	10.2	2.6	1.0E-11
	+295	52.9	12.6	8.1E-13
	+332	1.0	0.3	0.8
	-334	1.4	0.2	0.005
	+446	1.2	0.1	0.03
	-448	4.3	1.2	1.4E-11
	+234, +295	157	66	2.6E-16
	+234, +446	20.6	6.3	1.1E-11
CAP45	Wild Type	IC_{50} = 0.54	S.E = 0.24	
	-234	1.2	0.3	0.4
	+295	3.2	1.2	7.3E-10
	+332	3.2	1.2	4.7E-08
	-334	2.3	0.9	1.1E-03
	-446	2.0	0.9	1.5E-03
	+448	1.3	0.6	0.1

[a]FD is fold difference. Red text indicates that Grft is less effective at inhibiting the mutant virus compared to the wild type. Green text indicates the mutant is more sensitive to Grft than wild type. [b]S.E is standard error. [c]P values were determined by an ANOVA with a Bonferroni correction.

Table 2.2 Pairwise statistical analysis of the inhibition of Grft on various PVO.4 mutants.

	-295	-332	-448	>334	>446	-295 -332	-295 -448	-332 -448	-295 -332 -448	-295 >334	-295 >446	>334 -448	-332 >446	>334 >446	-295 >334 -448	-295 -332 >446	-295 >334 >446
Wild Type	<2 e-16	1.4 e-08	9.7 e-12	2.2 e-01	4.4 e-02	<2 e-16	2.7 e-10	9.4 e-07	8.1 e-14	8.1 e-09	6.6 e-09	1.1 e-07	1.4 e-11	8.8 e-05	3.8 e-05	4.3 e-09	2.8 e-06
-295 >334 >446	2.6 e-05	1.1 e-08	2.7 e-07	5.8 e-11	7.6 e-09	3.4 e-11	1	9.9 e-05	<2 e-16	2.2 e-04	2.7 e-08	2.0 e-05	0.14	4.4 e-07	1	2.6 e-03	
-295 -332 >446	<2 e-16	<2 e-16	<2 e-16	<2 e-16	<2 e-16	0.25	5 e-3	1	<2 e-16	3.6 e-16	<2 e-16	<2 e-16	8.0 e-11	<2 e-16	0.4		
-295 >334 -448	4.1 e-12	4.0 e-16	1.0 e-13	<2 e-16	1.2 e-14	3.6 e-08	1	0.04	<2 e-16	2.0 e-09	1.5 e-14	7.7 e-11	5.6 e-05	1.9 e-12			
>334 >446	1	1	1	1	1	<2 e-16	1.6 e-1	<2 e-16	<2 e-16	1	1	1	0.6				
-332 >446	1	0.5	1	0.01	0.05	<2 e-16	3 e-3	1.9 e-13	<2 e-16	1	0.3	1					
>334 -448	1	1	1	1	1	<2 e-16	7.7 e-9	<2 e-16	<2 e-16	1	1						
-295 >446	1	1	1	1	1	<2 e-16	1.4 e-12	<2 e-16	<2 e-16	1							
-295 >334	1	1	1	0.7	1	<2 e-16	1.9 e-7	<2 e-16	<2 e-16								
-295 -332 -448	<2 e-16	<2 e-16	<2 e-16	<2 e-16	<2 e-16	<2 e-16	<2 e-16	<2 e-16									
-332 -448	<2 e-16	<2 e-16	<2 e-16	<2 e-16	<2 e-16	0.49	1 e-4										
-295 -448	5.8 e-10	3.9 e-14	1.1 e-11	<2 e-16	1.0 e-12	3.5 e-12											
-295 -332	<2 e-16	<2 e-16	<2 e-16	<2 e-16	<2 e-16												
-448 >446	0.2	1	1	1													
-332 >334	0.02	1	1														
-448	1	1															
-332	1																

Heat map of Anova P values

<2e-16
2e-16 > 1e-8
1e-8 > 1e-5
1e-5 > 1e-3
1e-3 > 1e-2
1e-2 > 5e-2
>5e-2

p values for an Anova with Bonferroni correction analysis for Grft IC_{50} of PVO.4 mutants divided by Grft IC_{50} of PVO.4 wild type. A minus sign (-) indicates that glycan has been removed. An arrow (>) indicates that glycan has been shifted. Glycan deletion mutants were made by mutating the Asn to Gln.

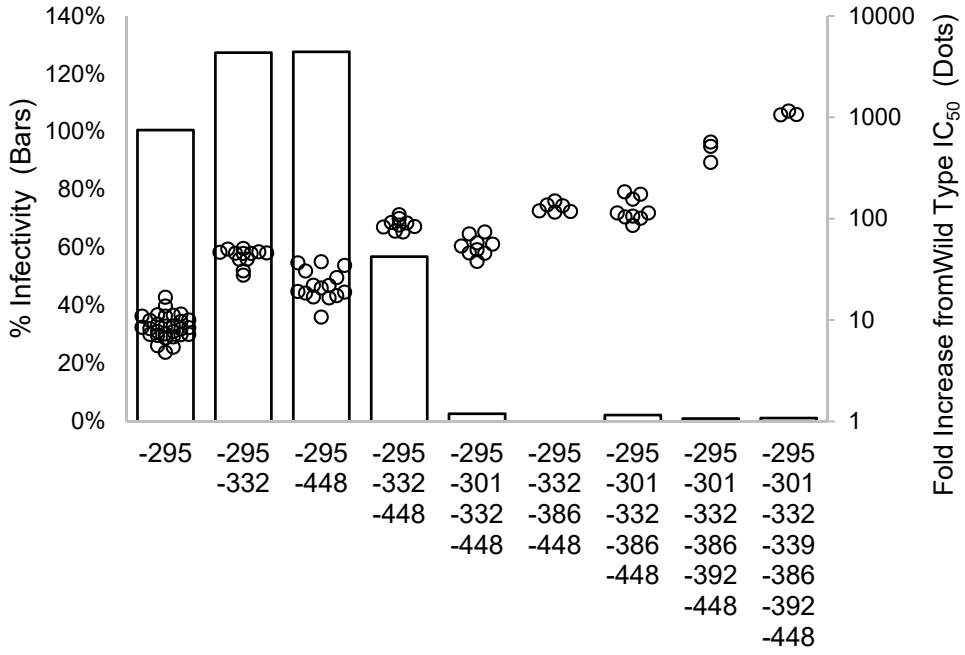


Figure 2.5 Serial deletion of glycosylation sites of gp120 degrades HIV-1 infectivity. Mutants of HIV strain PVO.4, showing an increase in IC_{50}, but a drastic degradation of infectivity, as glycan sites are removed. IC_{50} values were determined using HIV neutralization assays, with a wild type control on each plate. The wild type IC_{50} was then divided by the mutant IC_{50} to determine the fold increase. Infectivity was determined using a p24 ELISA in combination with a single-cycle neutralization assay to determine the infectivity of each mutant relative to the PVO.4 wild type virus.

It has been previously shown that glycan density affects the degree of enzymatic processing during protein glycosylation in the endoplasmic reticulum[332,352,368,376,750,763,764] which in turn can effect lectin epitopes[422,765]. This means that glycan clustering creates regions of mostly unprocessed glycan, Man-9 and Man-8 glycans. All possible combinational mutations of cluster 3 had been made and tested, so combination mutants of the other two clusters (defined by Balzarini et al[430]) were tested to determine whether the glycan clustering was contributing to Grft potency. Cluster 1 is defined to contain glycosite N230, which the strain PVO.4 lacks. To compensate, glycosylation was returned to site 230 to fully form cluster 1, Table 2.3 shows the mutant combinations removing glycans from clusters 1 and 2 that were tested, most showed little to no difference from wild type.

Table 2.3 Inhibition by Grft against HIV gp120 cluster mutants.

Cluster	Mutations	FD[a]	SE[b]
1	+230, -234	3.0	0.55
	-234, -241	2.9	1.2
2	-339, -386	1.7	0.22
	-339, -392	1.8	0.59
	-386, -392	1.2	0.13
	-339, -386, -392	1.8	0.42

[a]FD is fold difference. Red text indicates that Grft is less effective at inhibiting the mutant virus compared to the wild type. Green text indicates the mutant is more sensitive to Grft than wild type. [b]S.E is standard error.

2.3.2 BaL.01 Mutants

Several point mutations to gp120 in the commonly used HIV-1 strain BaL.01 were also carried out to remove the glycosylation at the glycosylation cluster 3[430]. As expected, removal of glycosylation at 295 showed the largest effect. Little effect was seen when individually removing either the site at 332 or 448 (Table 2.1). This differs from the results of Huang et al[482] who observed a loss of potency of Grft when gp120 was mutated to remove glycosylation at position 448 in this strain. This disparity may be due to differences in IC_{50} assay technique or potentially media composition (Appendix B).

2.3.3 Shifting gp120 glycosylation sites around N295 has little effect on sensitivity to Grft

The N332 glycosylation site was further explored due to the importance of this glycan as a target for bNAbs[362,363,537,758,760,761]. As shown in Table 2.1, Grft maintains potency against POV.4 when glycosite N332 is shifted to position 334. An analysis of thousands of published HIV-1 strains (245 A clade, 2231 B clade and 1449 C clade strains) in the HIV-1 sequence database was carried out, looking at the conservation of each high mannose glycan and the frequency of shifts occurring at each site(Figure 2.6). The glycans in cluster 3 showed the most movement in their location. Approximately 12% of all HIV-1 strains have a glycan located at N334 (referred to here as a N332>N334 shift; Figure 2.6). The other most common relocation is found primarily in C clade strains: 26% of C clade strains do not have glycosylation at 448, and 72% of those exhibit a glycan at N446 instead (referred to here as a N448>N446 shift; Figure 2.6C). Since removal of glycosite N448 in strain PVO.4 showed a clear decrease in Grft potency, we evaluated the N448>N446 shift as well. Table 2.1 shows that shifting the glycosylation site in PVO.4 (N448>N446) did not have a significant effect on Grft sensitivity.

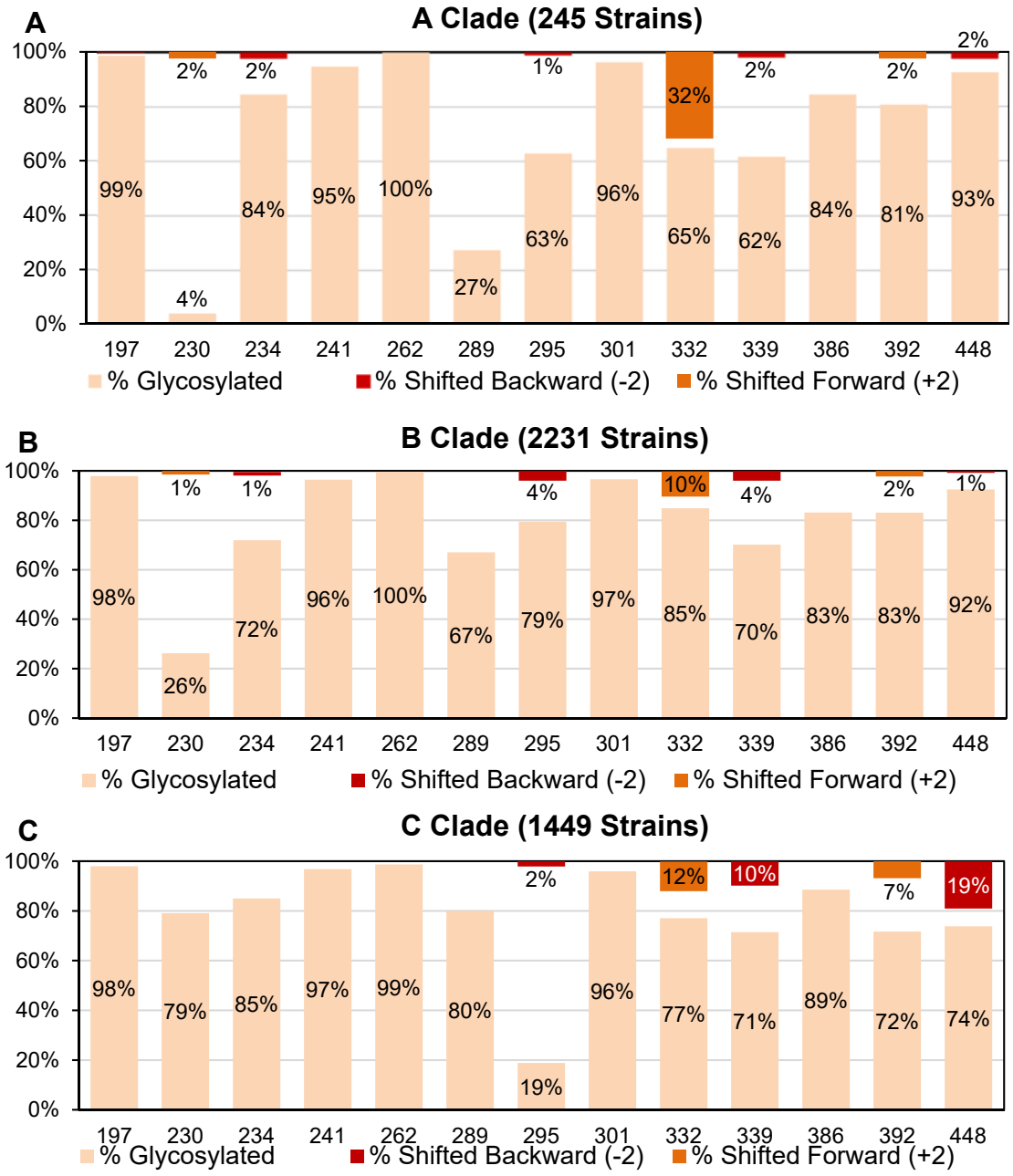


Figure 2.6 Analysis of the frequency of glycan shifts.
All published HIV strains in the https://www.hiv.lanl.gov HIV sequence database as of about March, 2016 were analyzed to determine the conservation of each high mannose glycan and the frequency of shifts in sequence (e.g. from N332 to N334 or from N448 to N446) occurring at each site. **A)** shows clade A. **B)** shows clade B. **C)** shows clade C.

We then investigated individual shift mutations (N332>N334; N448>N446) in combination with glycan deletions at other sites in cluster 3, N295Q, N332Q & N448Q (see Table 2.1 for IC_{50} and Figure 2.4 for graphical representation, Table 2.2 for statistical analysis). In addition, two mutants were generated that had both shifts, one in the presence, the other in the absence of a glycan at N295 (Table 2.1 and Figure 2.4). Overall, shifting of N332>N334 had little effect when added to any of the variant viruses. In general, the N448>N446 shift also had little effect, with the notable exception that this shift increased the sensitivity of the virus to Grft (lower IC_{50}) when the virus also lacked the glycan at 295.

2.3.4 Effect of glycosylation on HIV-1 strain ZM109

While Grft has been shown to be effective against strains from all clades tested, it has shown variable effectiveness against strains from clade C. This has been ascribed to differences in glycosylation patterns across clades, particularly, the lack of glycosylation at site 295 for 79% of clade C strains (Figure 2.6). ZM109 has lower sensitivity to Grft than PVO.4, with an IC_{50} of 27 ±6 nM. This strain lacks the glycosylation site at positions 234 and 295 and also exhibits the N332>N334 shift in glycosylation[766] (Figure 2.3). To investigate the role of glycosylation in sensitivity to Grft, several mutations were carried out, focusing on the critical glycosylation cluster comprised of N295/N33(2)4/N448. As shown in Table 2.1, placing a glycan at position 332 or removing one at 334 has little effect on sensitivity to Grft, while adding a glycosylation site at position 295 dramatically improved sensitivity to Grft, with a 53-fold reduction in IC_{50}.

Interestingly, when the glycosylation site in ZM109 gp120 at position N448 is shifted to 446 in the virus (naturally without glycosylation at 295), a 4-fold increase in sensitivity to Grft is observed (Table 2.1). Similar to the results with strain PVO.4, this indicates that a shift of glycosylation to the 446 position may be able to partially rescue the loss of the glycosylation site at 295 in terms of Grft action (Figure 2.7).

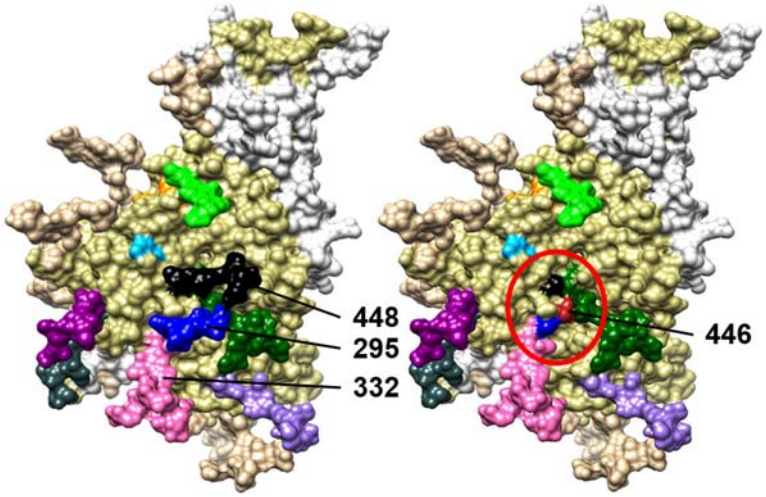


Figure 2.7 Monomer of gp120 highlighting location of position 446 relative to position 295.
A visual representation of how the shift of a glycan from N448 to N446 in the absence of a glycan at N295 may lead to higher glycan density and therefore a better binding environment for Grft. Left: gp120 monomer with relevant glycan sites in cluster 3 (N295, N332, N448) indicated. Right: Gp120 with amino acid at 446 (red) to show its proximity to cluster 3 (circled in red) if it is mutated to Asn to become a glycosite. PDB 5FYK[367] modeled using UCSF Chimera[917,918].

The glycosylation site at position N234 in gp120 has also been described as important in some strains, despite this site being in cluster 1 and distal to the critical cluster 3 (Figure 2.2)[430,495]. Although ZM109 was not mentioned in that study, when a glycosylation site is added at site 234, a 10-fold increase in sensitivity to Grft was observed. Most striking, when glycosylation is enabled at both 234 and 295, a 157-fold increase in Grft sensitivity was observed (Table 2.1). C clade strains seem to have much more highly processed glycans, minimizing the amount of Mannose residues accessible to Grft[340]. Returning glycosylation to key sites on C clade viruses may decrease the processing of glycans in that area, allowing Grft better potency by increasing the number of Mannose residues available.

2.3.5 Effect of glycosylation on gp120 in CAP45.G3J, a Grft-sensitive strain that naturally lacks glycosylation at position 295

Grft is highly potent across many strains of HIV-1, including some that do naturally lack the glycosylation site at 295. One such strain is CAP45.G3J (which will be referred to here as CAP45), a clade C virus that lacks glycosylation sites at 295, 332, 392, and yet is inhibited by Grft with an IC_{50} of 0.54 nM ± 0.23 nM. To investigate the role of glycosylation in such a strain, we made mutations in the important glycan cluster 3. When the glycosylation site at N295 was added back to this gp120, sensitivity to Grft was moderately improved, by factor of 3.2-fold (Table 2.1). The other two glycans in cluster 3 are naturally in the shifted positions, N334 and N446 (Figure 2.3). Returning the glycan at N334 back to its canonical site at N332 or deleting it entirely also moderately decreased the IC_{50}. Movement of a glycosite to N448 had no difference from wild type, whereas deletion of the glycan in that location entirely was one of the few mutations that caused a slight increase in Grft IC_{50}.

2.3.6 Grft is compatible with Broadly Neutralizing Antibodies Targeting Glycan at gp120 N332

We next investigated the potential of Grft as a combinational therapy with bNAbs. It has previously been shown that Grft is synergistic with the bNAb VRC01 which targets a separate area of gp120 (the CD4 binding site) so we sought to determine if Grft is compatible with another major class of bNAbs which target the glycan at N332 and may therefore be antagonistic with Grft. We chose PGT121 because it is in clinical trials[756,767]. As shown in Figure 2.8, in strains where Grft is relatively weak, PGT121 and PGT126 are potent, and vice versa. We then analyzed HIV-1 neutralization activity of Grft in combination with the two bNAbs in a constant 1:1 IC_{50} ratio. Grft was able to rescue activity when bNAbs became ineffective after the loss or shift of the glycan at N332 (Figure 2.9B-C, Figure 2.10B-D). Combinations range from synergistic to mildly antagonistic with combination index (CI) values ranging from 0.64 to 1.48 (Figure 2.9 & Figure 2.10). CI values were undeterminable when the bNAb IC_{50} was greater than 60nM, usually in the abcense of a glycan at position 332.

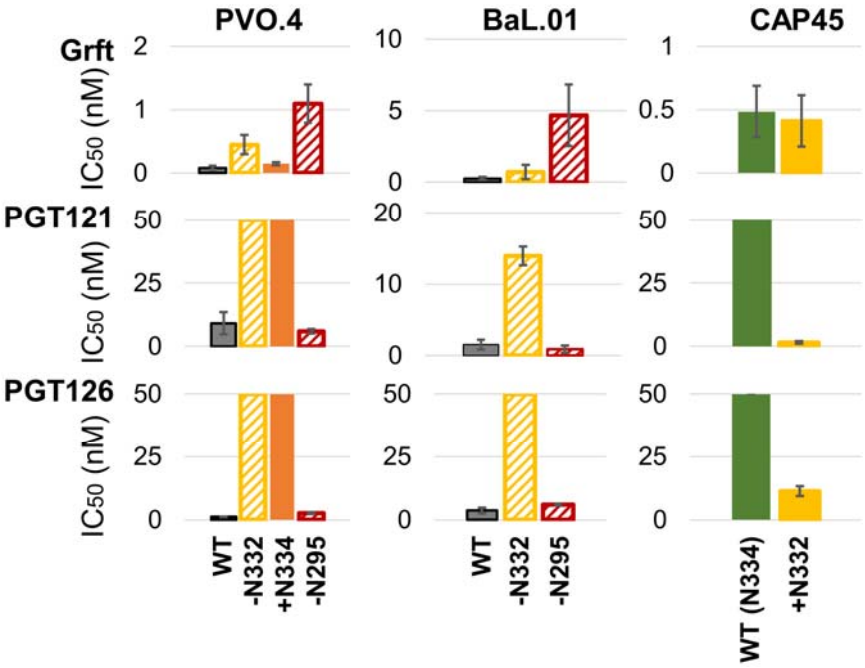


Figure 2.8 Grft and bNAbs PGT121 and PGT126 have complementary activity against a range of viral variants.
IC_{50} values for Grft, PGT121 and PGT126 in single round infection assays against various strains and mutants involving glycosylation at N332 and N295. CAP45 Wt naturally has N334, therefore +N332 indicates a shift of the glycan to N332. Solid bars indicate the presence of the glycan, slashed bars indicate the absence of the glycan (also labeled on the x-axis, a (+) indicates presence, a (-) indicates absence of a glycan at that site).

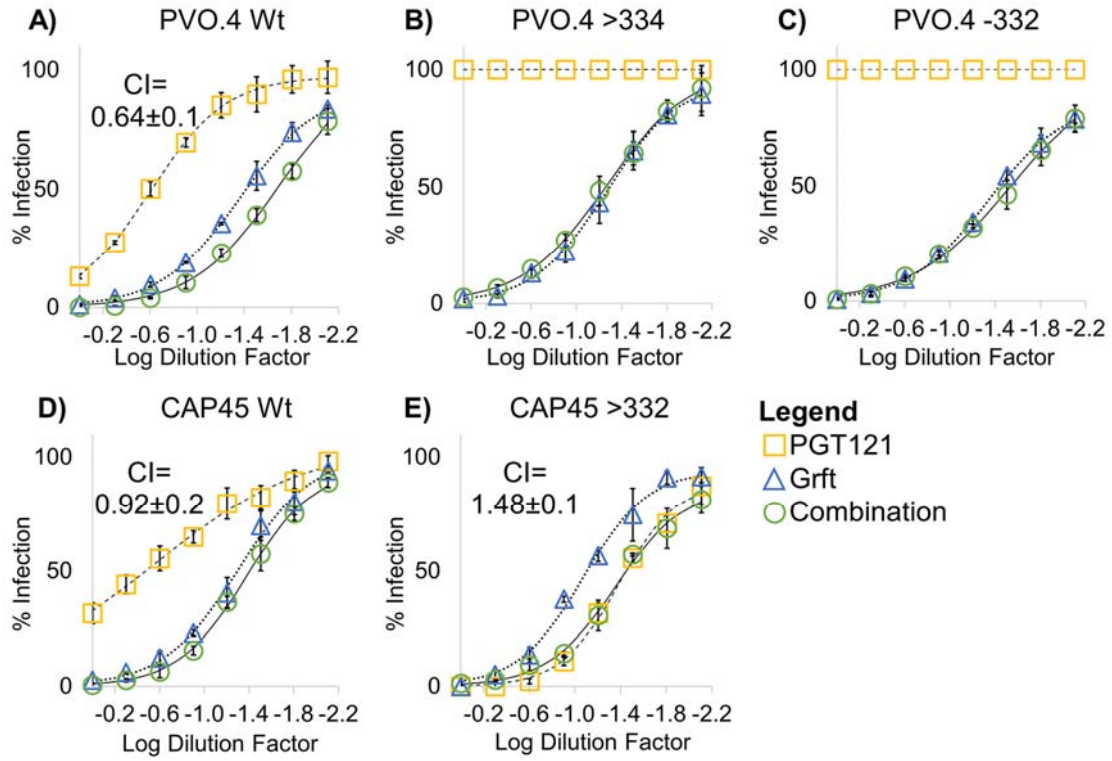


Figure 2.9 Combination of Grft & PGT121 against select strains and mutations of HIV. A-E show dose response curves of Grft and PGT121 singly and in combination against strains/ mutations of HIV to demonstrate Grft's ability to rescue in the absence or shift of a glycan at N332 without antagonistic effects. Combinations were tested by dilution of Grft and PGT121 at a fixed 1:1 ratio proportional to their IC_{50}'s. Relatively high starting concentrations were serially diluted to get a dose response curve. Starting concentrations of Grft were as follows: 2 nM for PVO.4 Wild Type (Wt) and PVO.4 with glycan shift to 334 (>334), 12 nM for PVO.4 with deletion of glycan at 332 (-332), 5.4 nM for CAP45 Wild Type (Wt) and 3.4 nM for CAP45 with a glycan shifted back to 332 (>332). Starting concentrations of PGT121 were as follows: 60 nM for all PVO.4 variants, 107 nM for CAP45 Wild Type and 20 nM for CAP45 with a glycan shifted back to 332.

Table 2.4 IC_{50}'s and CI values of Grft in combination with PGT121

		Grft			PGT121			
Strain	Mutation	Initial Conc. (nM)	IC_{50}	IC_{50}' [a]	Initial Conc. (nM)	IC_{50}	IC_{50}' [a]	CI Value
PVO.4	Wt	2	0.08 ±0.01	0.04 ±0.004	60	15.1 ±1.3	1.3 ±0.03	0.64 ±0.12
	>334	2	0.11 ±0.03	0.12 ±0.04	60	>60	3.57 ±1.1	
	-332	12	0.46 ±0.04	0.34 ±0.05	60	>60	1.69 ±0.26	
Cap45	Wt	5.4	0.28 ±0.03	0.23 ±0.05	107	39.8 ±13	4.5 ±0.94	0.92 ±0.18
	>332	3.4	0.29 ±0.02	0.13 ±0.01	20	0.77 ±0.06	0.79 ±0.06	1.48 ±0.10

[a] IC_{50}' is the IC_{50} of the inhibitor indicated in the presence of both inhibitors

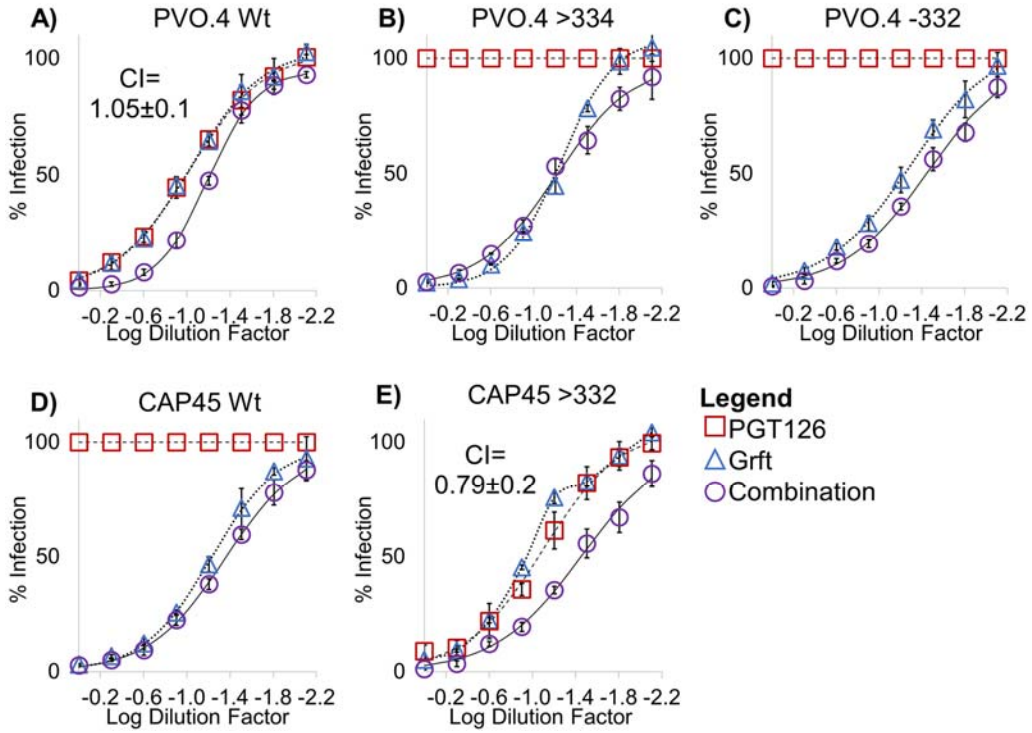


Figure 2.10 Combination of Grft & PGT126 against select strains and mutations of HIV. A-E show dose response curves of Grft and PGT126 singly and in combination against strains/mutations of HIV to demonstrate Grft's ability to rescue in the absence or shift of a glycan at N332 without antagonistic effects. Combinations were tested by dilution of Grft and PGT121 at a fixed 1:1 ratio proportional to their IC_{50}'s. Relatively high starting concentrations were serially diluted to get a dose response curve. Starting concentrations of Grft were as follows: 1 nM for PVO.4 Wild Type (Wt), 2 nM for PVO.4 with glycan shift to 334 (>334), 12 nM for PVO.4 with deletion of glycan at 332 (-332), 5.4 nM for CAP45 Wild Type (Wt) and 3.4 nM for CAP45 with a glycan shifted back to 332 (>332). Starting concentrations of PGT121 were as follows: 11 nM for PVO.4 Wild Type, 60 nM for PVO.4 with the glycan shifted to 334 or removed from 332, 60 nM for CAP45 Wild Type and 120 nM for CAP45 with a glycan shifted back to 332.

Table 2.5 IC_{50}'s and CI values of Grft in combination with PGT126

		Grft			PGT126			
Strain	Mutation	Initial Conc. (nM)	IC_{50}	IC_{50}' [a]	Initial Conc. (nM)	IC_{50}	IC_{50}' [a]	CI Value
PVO.4	Wild Type	1	0.11 ±0.01	0.06 ±0.002	11	1.2 ±0.12	0.66 ±0.02	1.05 ±0.09
	>334	2	0.12 ±0.01	0.14 ±0.02	60	>60	4.2 ±0.47	
	-332	12	0.70 ±0.08	0.49 ±0.08	60	>60	2.5 ±0.41	
Cap45	Wild Type	5.4	0.31 ±0.03	0.24 ±0.01	60	>60	2.7 ±0.14	
	>332	3.4	0.39 ±0.04	0.14 ±0.02	120	10.8 ±2.1	4.8 ±0.86	0.79 ±0.23

[a]IC_{50}' is the IC_{50} of the inhibitor indicated in the presence of both inhibitors

2.3.7 Conformational Changes Induced by Griffithsin

It was previously demonstrated that the naturally occurring presence of glycosite 295 in some strains of gp120 seemed to be a strong determinant in the ability of Grft to induce a conformational change in gp120, exposing the CD4 binding site, and that this correlated well with inhibition potency [477,480]. These experiments used virion capture by a conformationally sensitive antibody (b12) after incubation of the virion with Grft. We sought to confirm this finding using the same virion capture assay against wild type and mutated viruses (Figure 2.11). B clades strains PVO.4 and BaL.01 showed a significant decrease in CD4 site exposure upon incubation with Grft when glycosite N295 was removed (Figure 2.11). However, C clade viruses did not show a correlation between the presence of glycosite N295 in gp120 and exposure of the CD4 binding site upon incubation with Grft (Figure 2.11) despite the previous results showing that the addition of glycosite N295 did increase sensitivity of Grft to varying degrees in the C clade strains (Table 2.1).

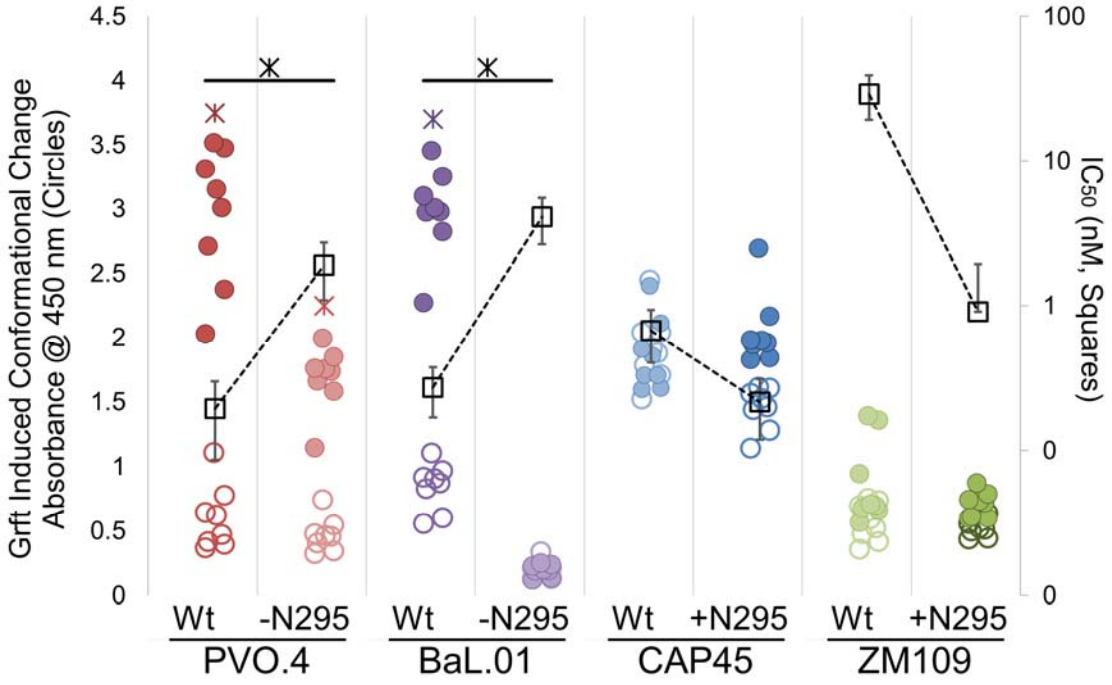


Figure 2.11 Grft mediates CD4 binding site exposure in B clade viruses.
Virion capture assay using mAb b12 to capture HIV virions with an exposed CD4 binding site. Filled circles indicate samples with 4nM Grft. Open circles indicate samples with 0 nM Grft. PVO.4 and Bal.01 naturally have N295, while CAP45 and ZM109 do not. Colored asterisks, (found over PVO.4 WT, PVO.4 -295, BaL.01 Wt, & Cap45 +295) indicate a comparison of filled versus open circles: that the mean of the samples with 4nM Grft was significantly ($p < 0.0001$) different than the samples with no Grft. Black asterisks indicate a comparison between the WT and mutant of a single strain, showing where the WT and mutant samples at 4nM Grft for a single strain were significantly different from each other ($p < 0.0001$). Significance was calculated using a one-way Anova with a post-hoc pairwise T-Test & Bonferroni correction. Squares indicate the IC_{50} of each strain and mutant, with dashed lines connecting the wild type and mutant of each strain to show how the absence of glycosylation at N295 results in a higher IC_{50}(lower Grft potency).

2.4 Discussion

In the fight against HIV-1, the use of the same drug(s) in both HIV-1 prevention and treatment leads to concerns about developing viral resistance. Therefore, it is critical to develop other means of HIV-1 prevention, that can be cheaply produced[485–487,489,741] with a sustained release, heat stable, formulation[496,730–733,768,769] so that they can be easily distributed in sub-Saharan Africa. The lectin protein Grft is a strong candidate for such a prevention technique. A successful microbicide must be effective against a wide range of HIV-1 strains, so it is imperative to understand the elements of Grft inhibition. Grft potency is independent of co-receptor usage by the virus; however, Grft's direct binding to high mannose glycans on gp120 necessitates the determination of which glycosylation sites are involved in Grft inhibition, with the goal of optimizing Grft usage either alone or in combination with other inhibitors such as bNAbs that have separate modes of action[494,770].

It has previously been reported that glycosylation on gp120 at N295 is critical for the potency of Grft and other lectins against HIV-1[431,474,482,495,743,771,772]. However, there is disagreement about the importance of glycosylation at other sites, with some reports demonstrating that the sites N448, N339, and N234 are also important for Grft inhibition [474,482,483,495]. Our work provides detailed and comprehensive data on the effect of glycosylation at various sites on gp120 and its relation to the potency of Grft. In so doing, we also gain insight into the mechanism of Grft action against HIV-1 entry. In the "typical" strain PVO.4 each high-mannose site was individually eliminated through mutation. It was found that the most important individual site was, as expected, N295; moderate effects were also seen upon individual elimination of glycosites N332 and N448 of gp120. Grft sensitivity to mutations at other high-mannose sites diminished with distance from the center of the high mannose patch at cluster 3 (Table 2.1).

Combinations of mutations at this glycan cluster 3 site drastically impair Grft potency, with the triple mutant N295Q/N332Q/N448Q showing an 88-fold reduction in sensitivity. The effectiveness of Grft was also tested against two strains that naturally lack glycosylation at the gp120 295 position, one of which is relatively insensitive to Grft (ZM109) and the other of which is quite sensitive to Grft (CAP45). It was found that adding a glycosylation site at position 295 did increase the sensitivity of both strains to Grft, although the effect was much larger for ZM109 (Table 2.1). The return of glycosite N234 to ZM109 interestingly resulted in a 10-fold improvement to Grft potency, but removal of the same site from PVO.4 showed little to no effect, even in combinatorial mutations.

These results suggest that Grft potency and the importance of glycosylation at N295 is at least partially due to the density of the glycan cluster. The location of N295 places it near the center of the high mannose patch[340,369]. Previous research has shown that the N295 site in gp120 stabilizes the high mannose patch and that the loss of this glycosylation site leads to increased enzymatic processing of the glycans in the area[352,556](Figure 2.7). In PVO.4, cluster 3 forms the core of the high mannose patch. In other strains, the high mannose patch may vary in size and location due to the different glycosylation patterns, which would explain disparate results by other researchers since glycan locations in each strain would have slightly different impacts on glycan processing based on gp120 conformation. Thermodynamic studies with Grft have shown that Grft binds more tightly to a saccharide that is closely related to Man-9 than to one that more resembles a processed site such as Man-5[478](Figure 1.17). Also, this glycan is placed at the base of the V3 loop which means that binding to this site would likely allow Grft to

inhibit gp120 interaction with CCR5, possibly through hindrance of gp120 conformational change, or CCR5-V3 loop interaction. This reasoning is corroborated by the rescue of Grft potency by the glycan shift from N448 to N446, in the absence of N295, in strains PVO.4 and ZM109 (Table 2.1 & Figure 2.4). The presence of a glycan at N446 may fill the hole in the high mannose patch created by loss of N295 and so facilitate Grft binding in a similar orientation and location to when a glycan is present at N295 (Figure 2.7). For CAP45, which is sensitive to Grft without glycosylation at 295, this strain may retain a high-density glycan patch in cluster 3 by virtue of having an additional, non-conserved glycan in the region at N335 (Figure 2.3). This additional glycan may help to maintain the density of the glycan patch, thereby minimizing glycan processing. Further mutational studies in conjunction with analysis of glycan composition by mass spectrometry would substantiate these hypotheses.

The glycosylation in HIV-1 gp120 at site N332 is particularly important in HIV-1 prevention because it provides the binding site for several highly potent bNAbs, which have been proposed as possible microbicides or injectable therapies[760,761,773]. In some viral strains, the natural glycosite is at N332, while others show a glycan at N334 instead. A range of effectiveness is observed for bNAbs that bind this area, with the antibodies retaining potency against some strains while losing efficacy against others[759,773,774]. We show that Grft is largely insensitive to shifts in the glycosite at either N332 or N448, indicating that Grft could be used to complement bNAbs as part of an HIV-1 prevention plan (Table 2.1 & Figure 2.4). Indeed, we found that bNAbs complement the activity of Grft against a variety of strains: the combination of PGT121 or PGT126 with Grft resulted in potent inhibition of HIV-1 regardless of the presence of glycosylation at site N295 or N332, making such a combination valuable for inhibiting a wide range of viral strains. Grft was able to rescue inhibition when the bNAbs lost potency after the shift or removal of the glycan at N332, and the bNAbs have minimal sensitivity to the loss of N295(Figure 2.7). Despite their preference for neighboring glycans (and therefore the possibility of antagonism between Grft and the bNAbs), Grft displayed primarily additive effects when combined with these bNAbs(Figure 2.8, Table 2.4, Figure 2.9 and Table 2.5). There was only one instance (the return of a glycan at N332 in CAP45) where Grft and PGT121 showed mildly antagonistic effects when combined(Figure 2.9E and Table 2.4). This effect was of little concern considering that the mutation of a virus to improve inhibitor potency is highly unlikely. Since only the full removal (of primarily N295), and not the shifting of glycans, inhibits Grft potency, Grft's addition to any combinational therapy appears to be consistently beneficial as Grft can rescue in instances where shifting glycans negate antibody potency. It has previously been shown that viral escape from Grft primarily requires complete removal of glycans[431], which would increase viral susceptibility to the immune system.

Grft is able to induce a conformational change in gp120, increasing gp120 binding to the conformational sensitive antibody b12, that generally is correlated to inhibitory potency[477,480]. In the current work, we studied this conformational change by making mutations to a set of viruses, allowing direct comparison in the presence or absence of particular glycosites in each virus rather than comparing separate strains and inferring a correlation. We show that for clade B strains, the conformational change mediated by Grft is dependent on the presence of the glycosite at 295 and is correlated with inhibitory potency. However, the clade C strains CAP45 and ZM109 do not show a significant difference in conformational change regardless of the presence or absence of glycosite

295, although they have different properties. CAP45 may be in a b12-binding conformation regardless of the presence of Grft (Figure 2.11); and ZM109 shows little conformational change regardless of the presence of the N295 site, but the presence of N295 dramatically enhances its sensitivity to Grft. Taken together, these data suggest that Grft is capable of binding clade B viruses differently than clade C viruses. However, the link between conformational change and inhibition of B clade viruses may be an instance of correlation and not causation.

Overall, this work delineates several modes of interaction between Grft and gp120 and shows that vulnerabilities in Grft inhibition are complementary to vulnerabilities in some bNAbs that bind to glycosylation on gp120, such as PGT121 and PGT126. A combination of Grft with either bNAb has additive effects and Grft can fully rescue inhibition in the presence of glycan shifts in gp120 that inactivate these bNAbs. Therefore, Grft is a suitable candidate for combinational therapies with bNAbs.

2.5 Materials & Methods

2.5.1 Griffithsin Production & Purification

Plasmids containing the Grft gene with an associated N-terminal histidine tag were transformed into *Escherichia coli* BL21 (DE3) (Novagen) competent cells and expressed in minimal media with $^{15}NH_4Cl$ as the sole nitrogen source. After induction with isopropyl β - D- 1- thiogalactopyranoside (IPTG), the cells were incubated at 22°C for 16 h. Once harvested, pelleted cells were resuspended in 6 M guanidine hydrochloride, 200 mM NaCl, and 50 mM Tris (pH 8.0), which allowed complete solubilization of proteins from both the inclusion and the supernatant upon cell disruption. The resulting solution was French pressed three times at 1,000 psi, then centrifuged at 27,000 x g for 1 h. The supernatant was loaded onto a Nickel chelating column (Qiagen, Hilden, Germany) equilibrated with the same re-suspension buffer. Flow through was collected and re-run over the column. The column was washed with the re-suspension buffer, then a wash buffer with 6 M guanidine hydrochloride, 200 mM NaCl and 20 mM NaPi (pH 7.2). Grft was eluted using 6 M guanidine hydrochloride, 200 mM NaCl, and 20mM NaOAc (pH 4.0), then added dropwise to a refolding buffer (200 mM NaCl, 1 mM EDTA, 550 mM arginine hydrochloride, and 60 mM Tris, pH 8.0) and allowed to stir overnight at 4°C. The solution was then dialyzed in 4L 200 mM NaCl, and Tris (pH 8.0) for 5 hours at 4°C, then transferred to new dialysis buffer (4L) to stir at 4°C overnight. The protein solution was then transferred to 4 L of a low salt dialysis buffer (80 mM NaCl, 20 mM Tris, pH8.0) for 5 h. The protein was finally purified on a C_4 reverse-phase chromatography column (Vydac, Hesperia, CA) and lyophilized in a Labconco freeze-dry system for long term storage. All samples were analyzed by SDS-Page gel at each step of purification to confirm the proper size of the Grft construct. NMR spectroscopy was used to determine that Griffithsin was properly folded (Figure 2.13), as described previously by the LiWang group[477].

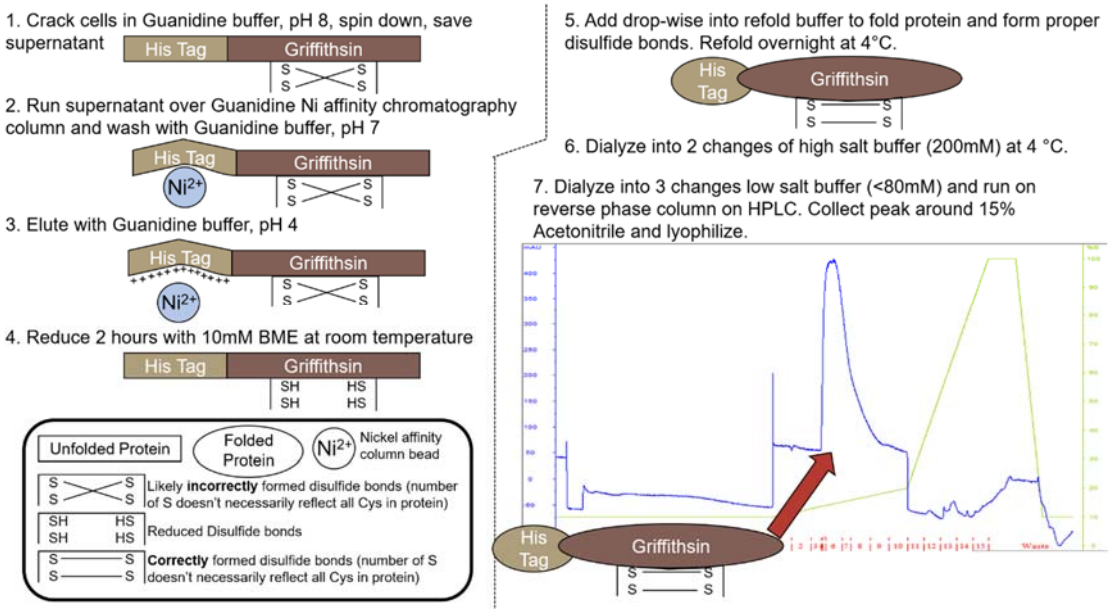


Figure 2.12 Overview of Grft purification.
A general overview of the purification of Grft. Tan sections represent the His tag, used in Ni column purification, brown sections represent the Grft protein itself. Steps with rectangular shapes are where the protein is not properly folded yet, meanwhile steps with oval shapes should have properly folded protein. A full legend is boxed in the lower left corner.

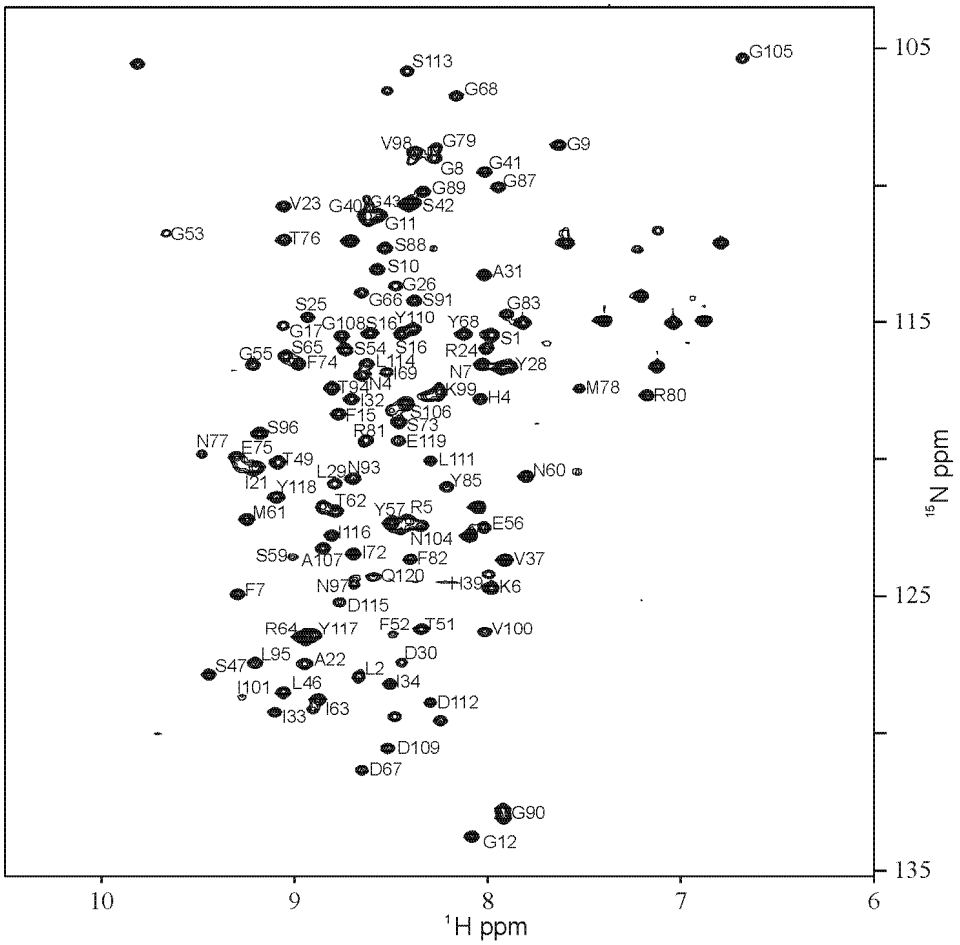


Figure 2.13 NMR of Grft
Assigned ^{15}N-^{1}H HSQC NMR specturm of Grft wild type at pH 7 in 20mM NaPi showing that the protein is completely folded. Protein was produced and NMR run by Kathryn Fischer with assistance from Anna Nguyen. Assignment was done by overlay with spectrum from Xue 2012 using Adobe Illustrator.

2.5.2 Viruses and Reagents

Viral plasmids containing the ENV gene from HIV were obtained from the NIH AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH, as follows: PVO, clone 4, SVPB11 (referred to as PVO.4) from Dr. David Montefiori and Dr. Feng Gao[775], ZM109F.PB4, SVPC13 (referred to as ZM109) from Drs. C.A. Derdeyn and E. Hunter[776], CAP45.2.00.G3, SVPC16 (referred to as CAP45) from Drs. L. Morris, K. Mlisana and D. Montefiori[777], HIV-1 clone Bal.01, Cat# 11445 (referred to as Bal.01) from Dr. Mascola[778] and pSG3$^{\Delta Env\ 3}$ from Drs. John C. Kappes and Xiaoyun Wu[339,399]. Sequencing primers for all viruses were generated using the Integrated DNA Technologies OligoAnalyzer (https://www.idtdna.com/calc/analyzer) and ordered from ELIM Biopharm (http://www.elimbio.com/index.htm). The sequencing primers were designed at 400-800 base pair intervals along the gene for good coverage when sequencing. (For future sequencing primer design, the spacing can be at a minimum of 700 base pairs since the sequencing consistently had reads of over 800 base pairs). Sequencing was done by the

UC Berkeley DNA Sequencing Facility. TZM-bl cells were obtained thought the NIH AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH from Dr. John C. Kappes, Dr. Xiaoyun Wu and Tranzyme Inc[399,779–782]. 239-FT cells were a kind gift from Dr. Jennifer Manilay and were originally obtained from Invitrogen[1,783–785]. Both cell lines were cultured in DMEM containing 10% FBS, 25mM HEPES with Penicillin/streptomycin or Geneticin respectively. Anti-HIV-1 gp120 Monoclonal (IgG1 b12) was obtained through the NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH from Dr. Dennis Burton and Carlos Barbas[786–789].

2.5.3 Generation of Mutant Envelope Pseudotype Virus DNA

All strains were aligned with strain HXB2 for standard numbering of glycan locations (Figure 2.15). Mutational primers were designed using a combination of 1) Integrated DNA Technologies OligoAnalyzer (https://www.idtdna.com /calc/analyzer) with the parameters as follows: 0.25uM Oligo Conc., 1.5mM Mg^{2+}Conc., 0nM dNTPs, and 2) New England Biolabs Inc. Tm Calculator (http://tmcalculator.neb.com/#!/) with the parameters set as follows: Product group – Phusion, Polymerase/kit – Phusion High-Fidelity DNA Polymerase (HF Buffer). Primers were designed to have the following properties: a) the mutation(s) centered in the primer as much as possible, b) primers should begin and end with a C or G nucleotide, c) primers should be at least 18bp long, d) annealing temperature should be between 50-59°C, e) the hairpin melting temperature is at least 15°C lower than the annealing temperature, and f) the self-dimerization free energy should be as positive as possible, meaning as high a kcal/mol value as possible, and no lower than -8 kcal/mol if it can be avoided. All mutants were generated by 2 step PCR

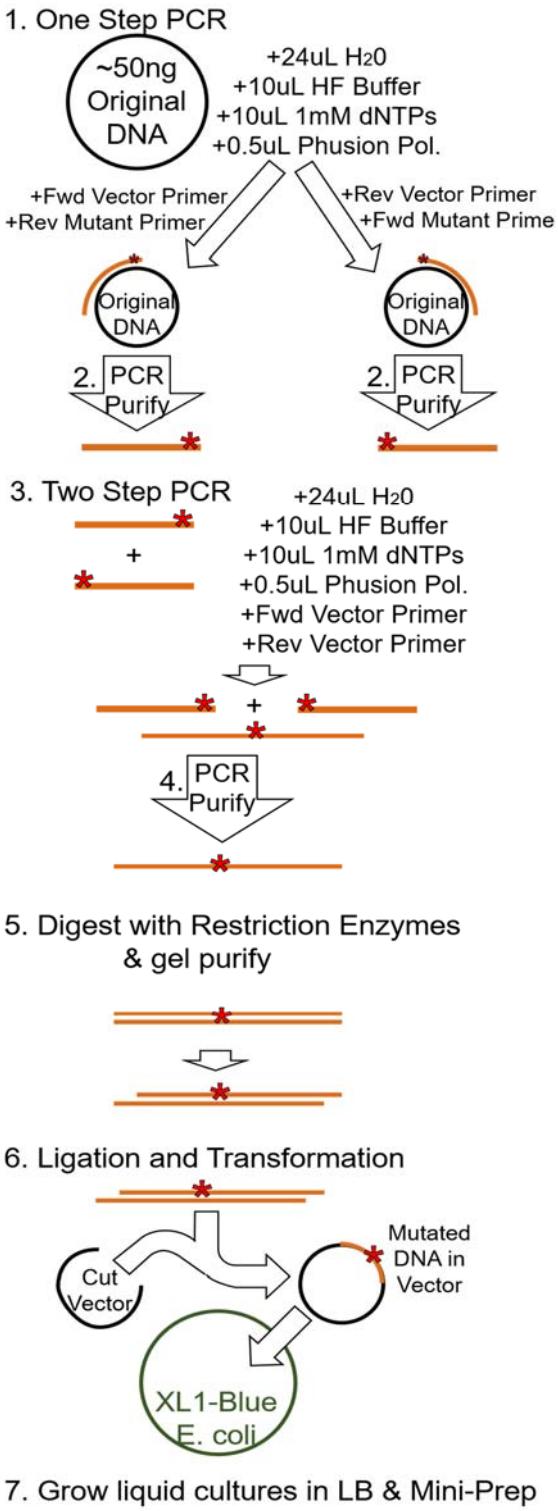


Figure 2.14 Flowchart for the generation of mutant envelope Pseudotype virus DNA.

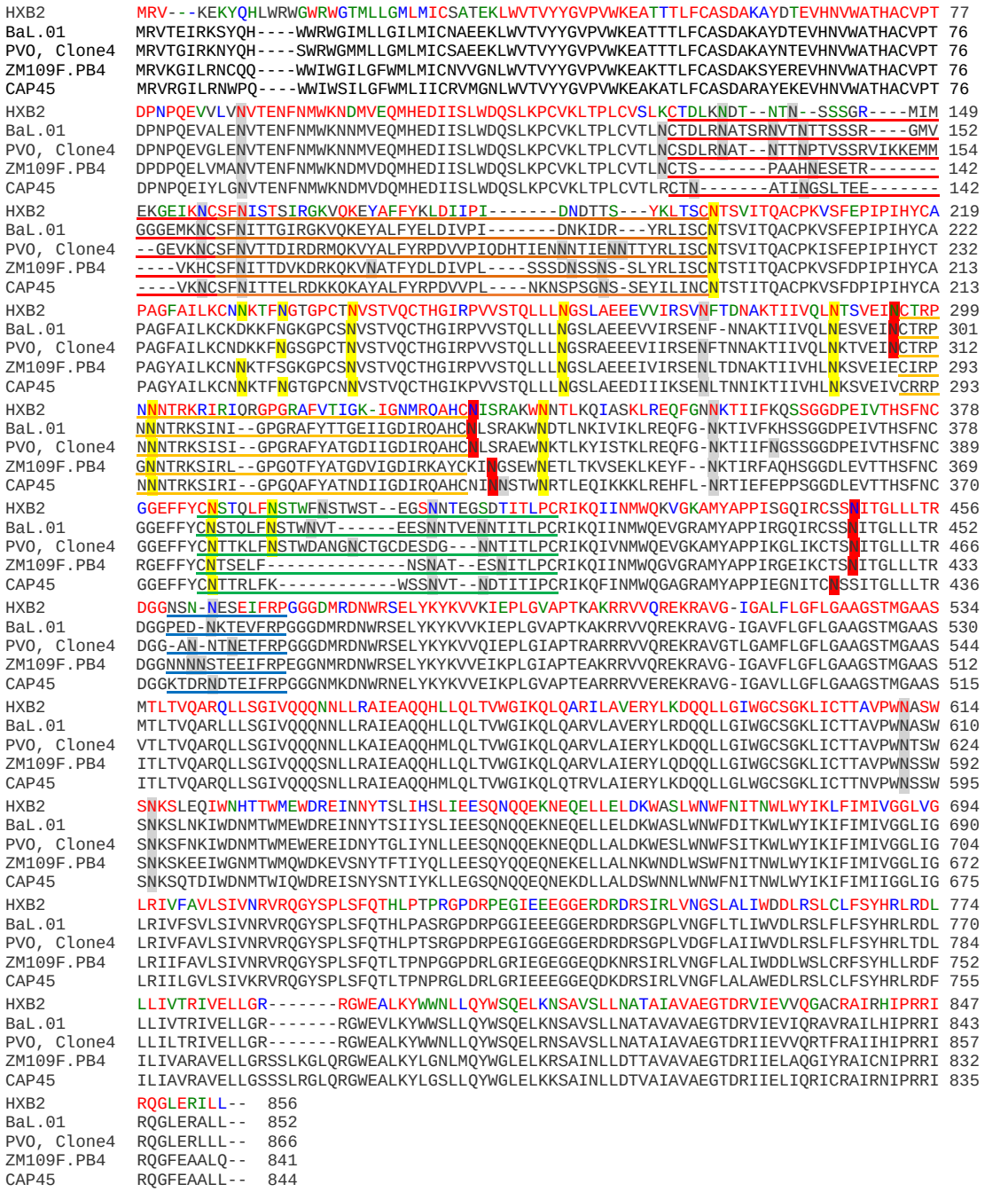


Figure 2.15 Alignment of HIV-1 strains used for mutational studies.
Sequences are from https://www.hiv.lanl.gov/content/sequence/VIRALIGN/viralign.html. Colored text indicates alignment identity: red for 100%, blue for 99-75%, green for 74-50%, black for 50%. Highlighted residues indicate potential glycan locations: yellow indicates conserved high mannose glycans, red indicates a high mannose glycan that was subject to extensive mutation (namely N295, N332(4) and N448(6), all other potential glycans were highlighted in grey. The variable loops of gp120 are underlined in bold, V1 in red, V2 in orange, V3 in yellow, V4 in Green, V5 in Blue.

using reagents from New England Biolabs (NEB) Inc (MA, USA). Step one of the PCR used Phusion polymerase to generate fragments of the gene, overlapping at the site of the mutational primer, using the protocol suggested by New England Biolabs (Figure 2.16) and outlined in Figure 2.14. Step two of the PCR combined the two fragments without primers for 10 cycles, then, an additional 30 cycles were run under normal conditions (end primers of the gene were added) suggested in the NEB protocol (Figure 2.16). Both steps of PCR were purified using the GeneJET PCR Purification Kit (Thermo Scientific, MA, USA). The second step PCR was digested then ligated into either pCR3.1 or pcDNA3.1 vectors using enzymes from New England Biolabs. The vectors with mutated DNA were then transformed into XL1-Blues cells and allowed to grow 16-18 hours. Cultures were grown at 37°C 220rpm for 14-16 hours, then purified using a Qiagen Mini Prep kit (MD, USA). The presence of mutations was confirmed by DNA sequencing.

Reaction Setup:

Component	50uL Rxn	[Final]
Water	to 50 uL	
5X Phusion HF Buffer	10 uL	1X
10mM dNTPs	1 uL	200 uM
10uM Fwd Primer	2.5 uL	0.5 uM
10uM Rev Primer	2.5 uL	0.5 uM
Template DNA	Variable	<250 ng
Phusion Polymerase	0.5 uL	1 unit

Thermocycling Conditions:

Step	Temp	Time
Initial Denaturation	98°C	30 sec
25-35 Cycles	98°C 45-72°C 72°C	5-10 sec 10-30 sec 15-30 sec per kb
Final Extension	72°C	5-10 min
Hold	4-10°C	

Figure 2.16 NEB recommended Protocol for PCR with Phusion Polymerase
A brief overview of the recommended parameters for PCR with Phusion polymerase as seen on the NEB protocol at https://www.neb.com/protocols.

2.5.4 Pseudo-Virus Production

293-Ft[1,2] cells were plated in 10cm dishes to an approximate 30-40% confluency. The cells should be passaged more than 6 times before this transfection process. When dish confluency reached 60-80%, usually in about 24 hours, cells were co-transfected with 10ng desired envelope plasmid and 10ng pSG3$^{\Delta Env}$ plasmid using X-tremeGENE HP DNA Transfection Reagent (Roche/Sigma-Aldrich, St. Louis, MO, USA). The transfection cocktail was prepared by adding 5ng of the desired envelope plasmid, 5ng pSG3$^{\Delta Env}$ plasmid[3] and 1mL of Opti-MEM to each of two 1.5mL low binding tube, then tipping to mix. Then, 30uL of X-tremeGENE transfection reagent was added to the tube, being careful not to touch the pipette tip to the wall of the tube. The solution was mixed by tipping and then left to rest. After fifteen minutes, the transfection cocktail was dripped into the dish containing 293 cells in a spiral pattern to distribute the mixture evenly. Transfected cells grown 24-48 h depending on the strain. Supernatant was collected in 15 mL tubes and centrifuged (Precision 100 centrifuge, 1100rpm, 2 minutes), then filtered through 0.45 um membranes to remove cellular debris. Viral stocks were stored in aliquots ranging from 200uL to 1.5mL, depending on the expected infectivity of the virus, in low binding tubes at -80°C. The tops of the tubes were labeled with the strain, mutation, and, if space was available, the date of collection. The side of the tubes were labeled with my initials (KF), the full date of collection, the strain, and the mutation of the virus.

2.5.5 Pseudo-Virus Quantification of p24 Concentration & Relative Infectivity

The amount of p24 was measured using the Human Immunodeficiency Virus type 1 (HIV-1) p24/Capsid Protein p24 ELISA Kit from Sinobiological (Beijing, China) according to the manufacturer's instructions. Virus samples were diluted to a final volume of 225uL and then 25ul of 5% Triton X-100 was added to samples. Samples were then vortexed for 20 seconds before being incubated at 37°C for 30 minutes. Results were calculated using a four-parameter logistic curve on myassays.com.

2.5.6 Single-cycle Neutralization Assay (TZM-bl assay)

On day one, 10^4 TZM-bl cells were seeded into a 96-well plate and allowed to incubate at 37°C for 18-20 hours. On day two, the media was changed and hour and 40 minutes before the time of the assay to a final volume of 50 uL per well. A dilution series of Grft, tailored to the strain and mutation of virus to be added, was made in phosphate buffered saline (PBS) and 20ul was added to wells (or 20uL of PBS for controls). 30uL of single round virus (at a dilution that would result in a 570 nm read at approximately 40 minutes) was added 20 minutes later, with three rows of the plate used for the mutated virus and three rows for the wild type of the virus being tested. The wells on the edge of the plate were not used due to evaporation effects. On day 3, after 18-20 hours, the media was changed, and the cells allowed to incubate for 48 hours. On day five, the cells were lysed using 0.5% NP-40 in PBS and a substrate solution was added (20 mM KCl, 100 mM 2-Mercaptoethanol (βME) and 8.3 mM Chlorophenol red-β-D-galactopyranoside (CPRG) (CalBiochem) in PBS). The absorbance signal was measured at 570 nm and 630 nm when the 570 nm reading was approximately 2. The ratio of the signal at 570 nm over 630 nm was calculated and compared to the controls to determine percent infection for each well. The fifty percent inhibitory concentration was determined using a linear equation fitted between the two data points surrounding 50% inhibition for each row and the three values were averaged for a final IC_{50}. All mutants were tested in triplicate three times or more. For an example plate see Figure 2.18, for an example of calculations see Appendix B. The data was plotted using Microsoft Excel. Statistical analysis, an ANOVA with a Bonferroni correction, was done using R software[790].

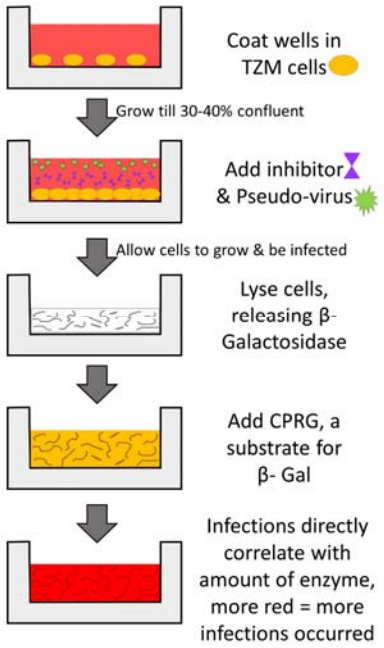


Figure 2.17 Overview of a single-cycle neutralization assay
A general representation of a single-cycle neutralization assay. Pink solution indicates DMEM media, white solution is a lysis solution of 0.5% NP-40, yellow solution is a solution with CPRG and red solution indicates that β-galactosidase has cleaved the CPRG to form a red product.

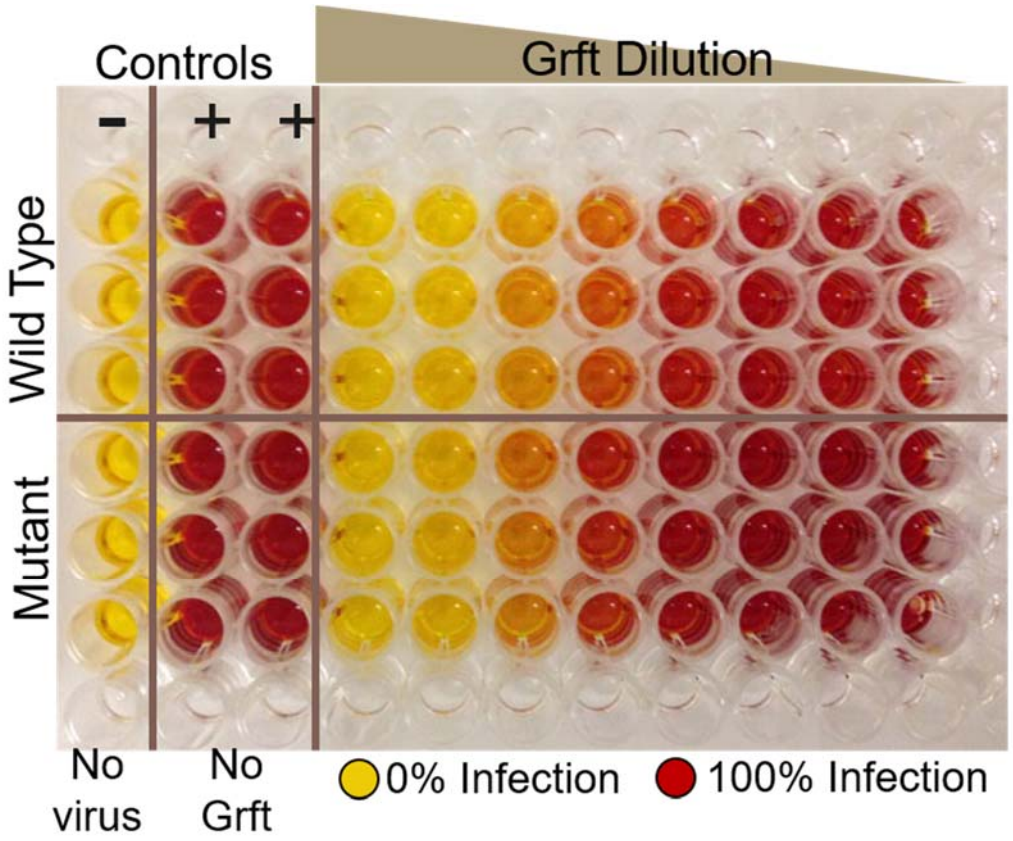


Figure 2.18 Example plate of a single-cycle neutralization assay.
This plate is an example of the single-cycle neutralization assays that were run to compare Grft potency against wild type and mutated viruses. In column 1, the negative control, for 0% HIV infection, remains yellow in color since only PBS was added during the infection stage. Columns two and three contain the positive control for 100% HIV infection, turning a vibrant red and dictating the optimum time to read the assay (when the A570 is approximately 2). The remaining columns contained a uniform amount of pseudovirus in each well but a decreasing concentration of Grft, which results in a sigmoidal distribution of infection, from fully inhibited virus, to fully infected cells as evidenced by the color gradient from yellow to red.

2.5.7 Synergy Tests

For synergy studies, combinations of Grft with bNAbs PGT121 and PGT126 were tested by a dilution series at a fixed 1:1 ratio proportional to their IC50's before being added to a standard TZM neutralization assay. For example, if the IC_{50} of Grft was 1nM and the IC_{50} of the Ab was 0.5nM, the starting concentrations would be 50nM and 25nM respectively. For these experiments, the IC_{50} values of each inhibitor were independently determined using single cycle neutralization assays within a month of the synergy experiment taking place. These IC_{50} values may differ from the average IC_{50} presented in Table 2.1 due to a change in media composition (Appendix B) Combinational index (CI) was calculated using CompuSyn software[791] based on the equations of Chou and Talalay

seen in Figure 1[792]. The data was plotted using Microsoft Excel with curve fits generated in MatLab (Appendix B).

2.5.8 Virion Capture ELISA

Exposure of the gp120 CD4 binding site was measured as described previously by the Morris and LiWang groups[477,495]. Briefly, high binding 96-well flat-bottomed plates (Thermo Scientific, MA, USA) were coated overnight at 4°C with 1ug/well of mAb b12 in PBS. On day two, the plate was washed and blocked. Pseudovirus samples were preincubated with Grft at 4nM and 20 nM and a media only control for 1 hour at 37°C, then added to the b12-coated plate and incubated another 2 hours at 37°C. Then the plate was washed, and bound virus was lysed with 150uL 0.5% Triton X-100. 100uL of virus lysate was transferred to a Human Immunodeficiency Virus type 1 (HIV-1) p24/Capsid Protein p24 ELISA Kit from Sinobiological (Beijing, China) and tested according to the manufacturer's instructions.

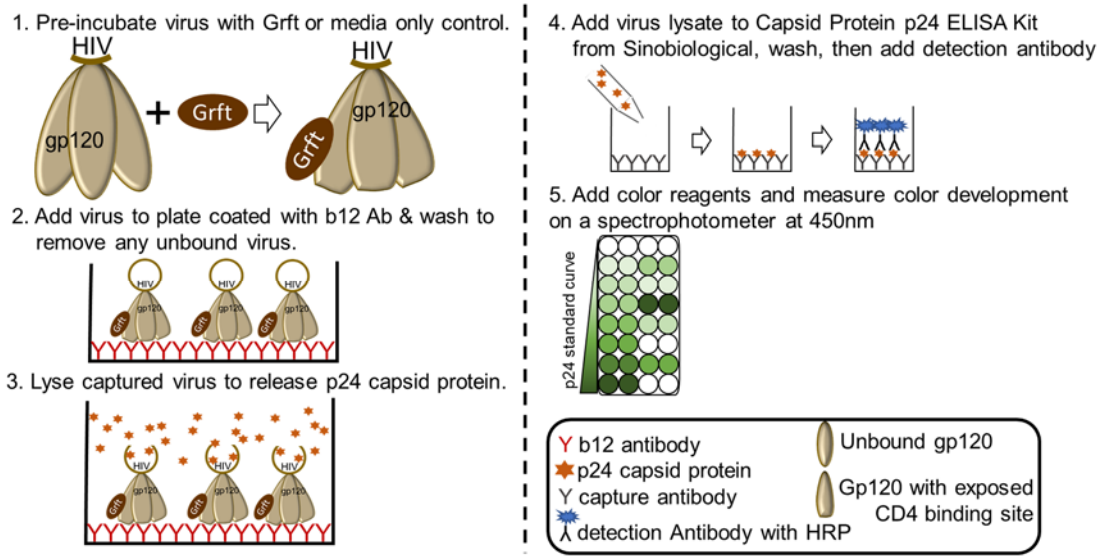


Figure 2.19 Overview of virion capture ELISA
A simplified representation of a virion capture ELISA using a p24 ELISA kit. Pre-incubation of the virus (tan) with Grft (brown) can, in certain cases, induce a conformational change that exposes the CD4 binding site. If this site is exposed, it will bind to the b12 antibody coated onto the first plate (shown as a red Y). A detergent is used to lyse the virus and release the p24 capsid protein (shown as orange stars). The lysate was then transferred to the Sinobiological kit, where the p24 was detected with an antibody conjugated to HRP. Addition of the color reagents yielded color development at 450nm.

Chapter 3

Targeted reactivation of the latent HIV reservoir

3.1 Introduction

Currently, over 35 million people world-wide are infected with HIV[136]. While current antiretroviral therapies have significantly increased life expectancy, decreased symptoms, and prevent transmission, there is no cure[136,397,515,605,632,667,793]. The HIV reservoir, a population of cells quiescently infected with HIV, is the barrier to a functional cure. Even though patients on combination antiretroviral therapy (cART) do not have detectable virus, interruption of treatment leads to re-emergence of viral replication from latently infected reservoirs that have existed since the first week of infection[323,602,616,717,794,795]. As a retrovirus, the viral DNA is irreversibly integrated into the host cell as part of the replication process as described in the section on replication above(1.3.4)[323,602,615,616]. It seems to be preferentially integrated into transcriptionally active regions, although the exact locations are still unclear[318,325,619,796]. Location of viral integration during infection may influence latency, with the HIV becoming transcriptionally silent instead of producing infective virions[319,320,323,326,670,797–800]. Cells found to be harboring latent HIV are fairly diverse: peripheral blood monocytes, macrophages, dendritic cells, microglial cells, astrocytes and hematopoietic stem cells[523,603,646,693,801–803]. One of the most prominent latent cellular populations are CD4 memory cells[633,651,804,805], which are long lived and have a low turnover rate[323,614,643,794,801,806–810]. Latency may also be induced by replication of virus in deep tissues, such as lymphatic tissues, that are not accessible to CD8 killer cells or therapeutics[608,689,794,802,811]. A further complication in the reactivation of latent viruses is that latently infected cells reactivate stochastically[620,630,678,810,812].

3.1.1 Reversal of latency

A cure for HIV is impossible until the HIV reservoir is fully understood and either removed or permanently silenced so that the ability to produce new virions is completely abrogated. Various strategies have been proposed to eradicate the latently infected cells: “shock and kill”, “locking” cells in latency and dual-affinity retargeting (DART)[608,626,627,665,668,670,671,673,675,813]. The earliest and most investigated HIV elimination strategies are the “shock and kill”, which usually consist of a drug/compound that activates transcription of the dormant HIV genome (the “shock), yielding the expression of viral proteins that would be recognizable by the immune system, and potentially additional targeted drugs, that could then “kill” the infected cell[627,670,671,673,675]. The process of “locking” cells into a latent state, usually through gene therapy or tat inhibitors, theoretically prevents any future transcription of the HIV genome. Locking techniques and development has been thoroughly reviewed[320,608,626,627,632,665,674–676] and will not be discussed further here.

Shock and kill methods have shown promise to transiently increase viral expression in cell[660,677–681] and animal models[653,667,682–688] but, there is no visible evidence that the viral reservoir has been reduced, let alone eliminated. More targeted approaches for shock and kill methods are needed since many of the previously tested strategies are very broadly acting with toxic side effects[608,689–692]. Latency reversal agents (LRAs) include Histone Methyltransferase (HMT) inhibitors[693], Toll-like receptor 7 agonists[694–698], histone deacetylase (HDAC) inhibitors[626,670,672,699–702], and activators of the nuclear factor-κB (NF-

κB) pathway [687,688,703]. HDAC and HMT inhibitors loosen the interactions between histones and DNA, unwinding the DNA and allowing transcription and production of viral proteins[704]. NFκB is a protein complex that controls transcription of DNA[705] , so activation generally up-regulates transcription in a cell, increasing the chance of viral expression.

The most natural and potent enhancer of HIV transcription, that should only induce activation of HIV DNA, is the HIV protein Tat. The Tat protein is responsible for ensuring viral transcription during the viral lifecycle. Tat stimulates transcription by binding to the transactivation response element (TAR) in developing RNA transcripts and recruiting Cyclin T1 and cyclin-dependent kinase 9, which compose the positive transcription elongation factor b (P-TEFb) and will promote transcriptional elongation by recruiting and phosphorylating RNA polymerase II, releasing the elongation inhibitor complexes termed NELF and DSIF[706–715]. Tat has also been shown to bind the upstream promoter regions[716]. Research has shown that insufficient Tat transactivation activity can result in latency[620,652,717–720]. Tat inhibitors are being investigated as a means to lock latently infected cells[666,668,721]. Tat has also been investigated as a shock agent[718,719,722–725], but issues with cytotoxicity[718,726,727] highlight a need for a method to target Tat and other LRAs to cells most likely to be infected with HIV. Geng et al generated a mutated Tat to minimize its toxicity while maintaining activity as an HIV transcriptional activator in various latency models[718].

3.1.2 Targeting the latent reservoir

Immune cells are primarily recognized and classified based on the receptors found on the cell surface[814]. Therefore, binding cellular receptors either used in HIV entry or other receptors displayed on HIV latent reservoir populations could be a viable strategy for selective targeting. Unfortunately, while researchers previously suggested that CD2[815] and CD32a[816–818] were biomarkers of latently infected cells, it has since been disproven[819–821], indicating that there is no single cellular receptor associated with latency. The natural diversity and robustness of chemokines and chemokine receptors utilized by the immune system present an ideal system[822]. An extensive library of modified chemokine ligands with various binding and receptor-activation properties[823,824], offers options to target combinations or individual cell types.

Previous work has indicated that it may be possible to link a targeting molecule or peptide to an LRA, toxin (such as the KLA peptide[825–827]), genomic editing proteins (such as CRISPR[677–681,828]) or latency locking molecule (such as Diehydro-Cortistatin A[666–668,721]) to create a chimeric protein that would be able to selectively target cells most likely to be latently infected with HIV and remove HIV latency through transcriptional activation of HIV provirus, cell death, removing the provirus from the genome or permanent genomic silencing of the provirus respectively. Our chosen route was to link Tat to variants of the chemokine RANTES, which binds several chemokine receptors, including the HIV co-receptor CCR5.

Tat is a small HIV protein responsible for naturally activating transcription of HIV proviral DNA[718,719,722–725], that is one of the few protein-based LRAs shown to reactivate HIV in latently infected cells, making cellular expression feasible for production, unlike the small molecule toxins, LRAs and latency locking molecules previously listed. While Tat has been show to exhibit cytotoxic effects[726,727,829,830], it may be possible with attenuated dosages and mutations to minimize toxicity[718], which is important, as some of the cells that are entered by the construct may not contain HIV proviral DNA. Benefits of using Tat

include Tat's natural ability to escape endosomes through the process of pinocytosis through the protein transduction domain (PTD)[831–835], which would aid the construct in exiting the endosome produced by a chemokine targeting-protein interacting with its receptor, and Tat's Nuclear Localization Sequence (NLS), allowing for localization into the nucleus once the cell is entered[832,836–841] (Figure 3.1). NMR studies show Tat to be structurally flexible, lacking most secondary structure[707,842] (Figure 3.1A), even when bound to other proteins[709,843] (Figure 3.1B).

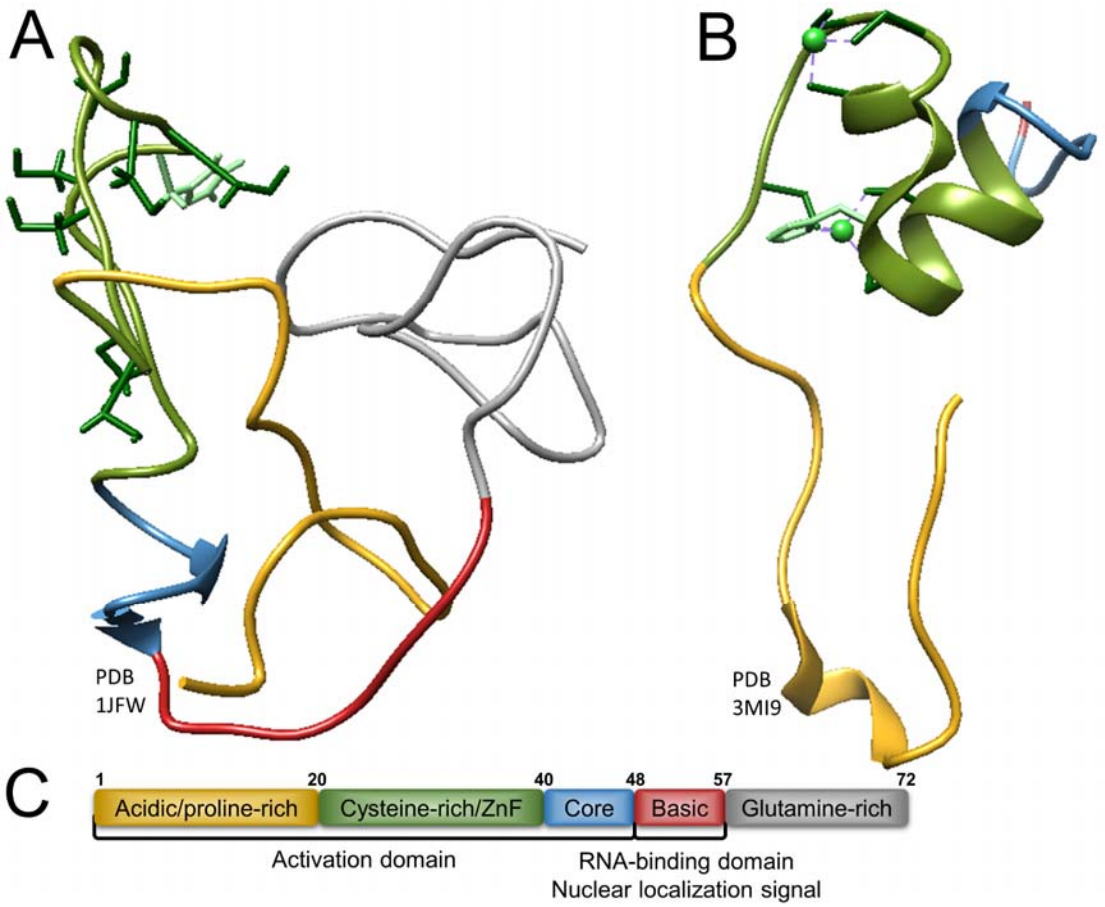


Figure 3.1 Basic structure and sequence of Tat.
The NMR structures of Tat are shown with each region color coded, with each region indicated in C. A) shows the NMR structure of unbound Tat (PDB 1JFW[842]). B) shows the NMR structure of Tat bound to p-TEFb (PBD 3MI9[843]). The residues involved in binding to zinc (shown in spring green) are the seven cysteine residues are colored dark green and His 33 is colored light green. C) shows the basic sequence of Tat, indicating the various regions and domains of Tat. This figure is based on Figure 2A of Tahirov et al 2010[843].

The chemokine RANTES was chosen for initial studies of a chemokine-based targeting system because 1) it binds to CCR5, which is one of the co-receptors of HIV[225,657,844–848], and 2) many variants of RANTES with various receptor binding properties have already been produced and characterized[244,824,849–853]. A variant known as 5P14-RANTES has been shown to bind CCR5 tightly and cause internalization of the CCR5

receptor without causing signaling of the associated G-proteins, which would result in calcium flux from cellular stores[823]. According to Kraft 2001, the CCR5 receptor is internalized within 30 minutes after exposure to a chemokine (125pM). Phosphorylation of the chemokine N terminus is essential for internalization[854]. CCR5 internalization and recycling back to the cell surface can occur one of two ways[855]: 1) Arrestin binds to the receptor, which then localizes to clathrin coated pits or 2) the Caveolae pathway, which is independent of clathrin-coated pits. In previous research receptors returned to the membrane after 120 minutes after being incubated with 50nM chemokine for 1 hour, but then chemokine was removed by washing with fresh medium[854].

We hypothesize that 5P14-RANTES will induce internalization into the cell, bringing Tat with it. The PTD region of Tat would then allow escape from the endosome, and the NLS would stimulate transport of the chimeric protein (5P14-Linker-Tat) to the nucleus. Once in the nucleus, the Tat would bind to the TAR of nascent HIV RNA and promote transcription (Figure 3.2).

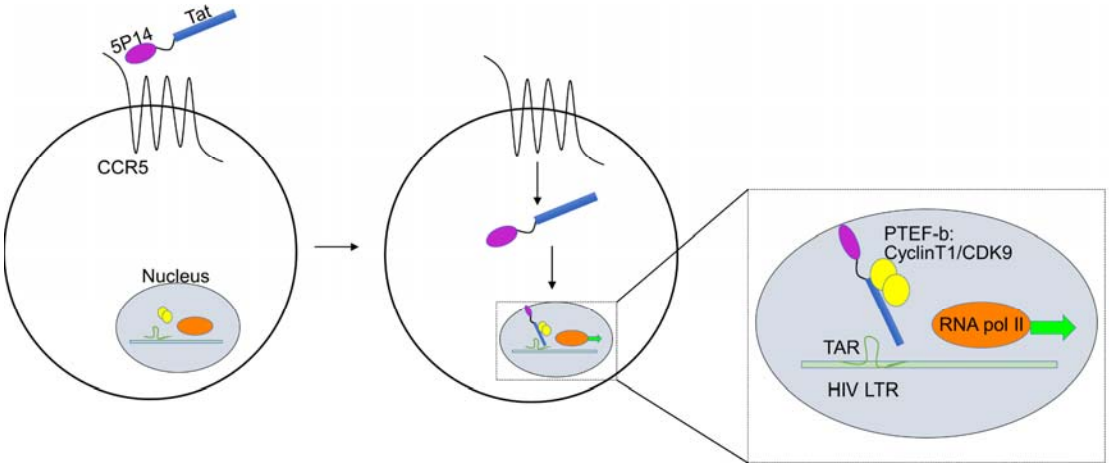


Figure 3.2 Theoretical function of 5P14-Linker-Tat chimeric protein.
5P14-RANTES would bind to the receptor CCR5, triggering internalization of the receptor with the chimeric protein (5P14-Linker-Tat) attached. Tat would allow escape from the endosome and subsequent transport to the nucleus where it would stimulate expression of proteins linked to the HIV LTR.

3.2 Materials

3.2.1 DNA

Mutational primers for all proteins were generated using a combination of 1) Integrated DNA Technologies OligoAnalyzer (https://www.idtdna.com /calc/analyzer) with the parameters as follows: 0.25uM Oligo Conc., 1.5mM Mg^{2+}Conc., 0nM dNTPs, and 2) New England Biolabs Inc. Tm Calculator (http://tmcalculator.neb.com/#!/) with the parameters set as follows: Product group – Phusion, Polymerase/kit – Phusion High-Fidelity DNA Polymerase (HF Buffer). Primers were designed to have the following properties: a) the mutation(s) centered in the primer as much as possible, b) primers should begin and end with a C or G nucleotide, c) primers should be at least 18bp long, d) annealing temperature should be between 50-59°C, e) the hairpin melting temperature

is at least 15°C lower than the annealing temperature, and f) the self-dimerization free energy should be as positive as possible, meaning as high a kcal/mol value as possible, and no lower than -8 kcal/mol if it can be avoided. Primers were ordered from ELIM Biopharm (http://www.elimbio.com/index.htm). Sequencing was done by the UC Berkeley DNA Sequencing Facility utilizing T7 promoter and terminator sequencing primers.

3.2.2 Cells and Reagents

All cell lines were obtained through the NIH AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH as follows: TZM-bl cells (Cat# 8129) from Dr. John C. Kappes, Dr. Xiaoyun Wu and Tranzyme Inc[399,779–782], Jurkat LTR-GFP Cells (JLTRG)(Cat #11587) and Jurkat LTR-GFP CCR5+ Cells (JLTRG-R5)(Cat #11586) from Dr. Olaf Kutsch[856,857], J-Lat Full Length Clone 8.4 cells (JLat 8.4 cells) from Dr. Eric Verdin[322], HeLa-CD4-LTR-β-gal cells (MAGI cells) from Dr. Michael Emerman[628], and MAGI+CCR5 from Dr. Julie Overbaugh[858]. The following mouse anti-human antibodies were used to characterize the presences of CCR5 on the cell lines above: CD195 APC Clone 2D7/CCR5 (BD Biosciences), anti-CCR5 PE Clone 45531 (R&D Systems), anti-CCR5 APC Clone CTC5 (R&D Systems), CD195 APC-Cy7 (J418F1, Biolegend).

3.3 Methods and results

Dr. Anna Nguyen had begun development and initial testing of a chimeric protein that could selectively enter cells most likely to be latently infected with HIV and activate the viral genome so that the cells could be targeted for elimination. A protocol was developed to produce a folded 5P14-Linker-Tat chimera using a Sumo fusion tag that was cleaved off with ULP1 after the refolding step. 5P14 is a variant of the RANTES chemokine that binds to the receptor CCR5, which is co-receptor for HIV infection. This particular variant was chosen because it causes internalization of the receptor, without activation of the inflammatory signaling pathway, which would draw more vulnerable cells to the area of infection. Tat is an HIV transcription factor that activates production of viral proteins. Unfortunately, high quantities of Tat are cytotoxic and therefore Tat alone cannot be used as the “kick” component of a therapeutic regime.

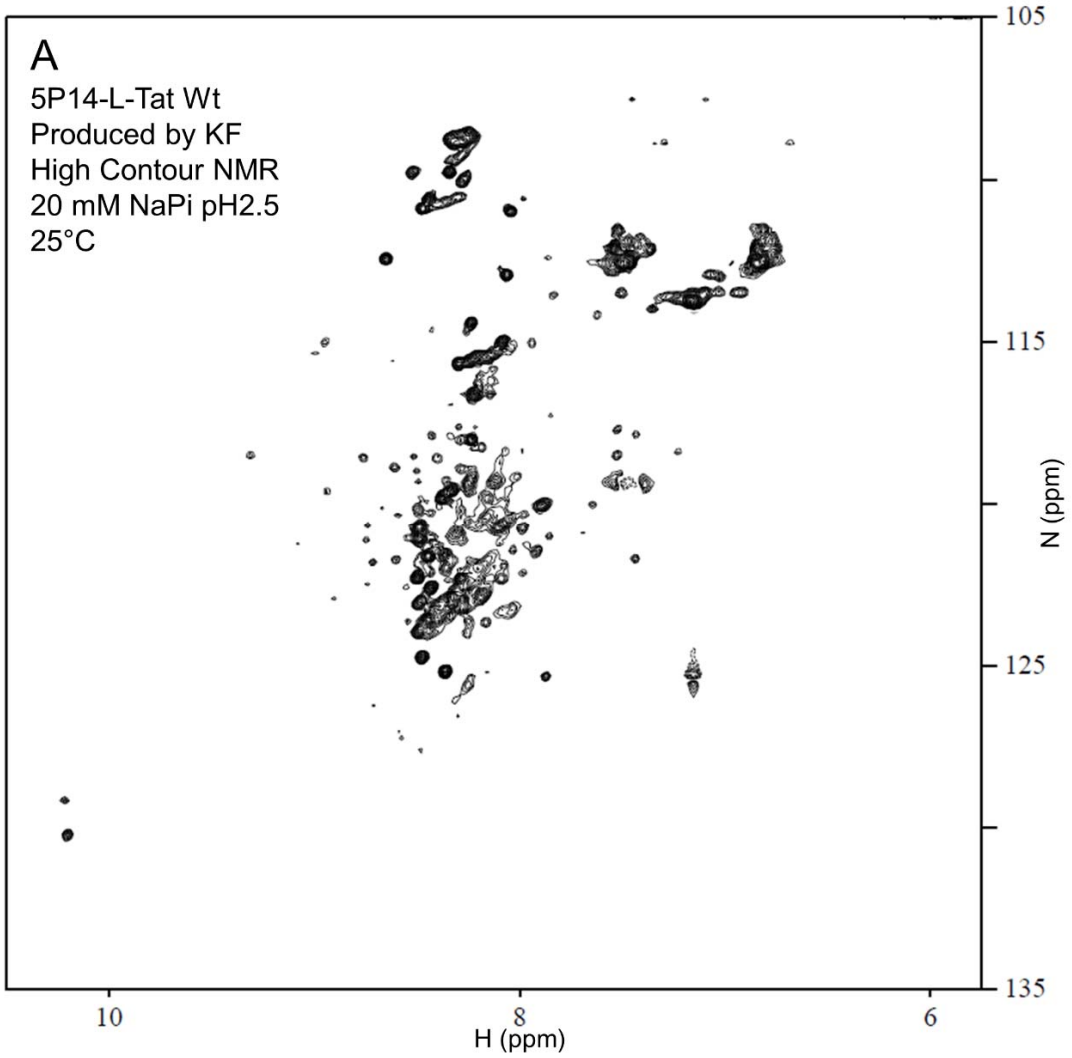
The primary struggle after finalizing the chimera refold conditions was low yield. The protocol was particularly exacting because of the dual nature of the chimera. 5P14-RANTES is a CC chemokine, which means it has two disulfide bonds that are essential for a proper fold and function. Tat, on the other hand, contains seven free cysteine residues that function best in a reduced state. The initial challenge was to produce large quantities of folded protein, both the wild type and three new mutants designed to modulate the function or cytotoxicity of Tat. I would then test these proteins in a cellular assay to determine whether the chimera could selectively activate cells containing an HIV LTR-linked GFP reporter and bearing a CCR5 receptor, using free Tat and 5P14-RANTES as positive and negative controls respectively. Other conditions that would be taken into consideration were the levels of cell death and ease of entry into the cell, which may be affected by the order of the chimeric proteins since the N-terminal portion of Tat contains the nuclear localization signal (Figure 3.1).

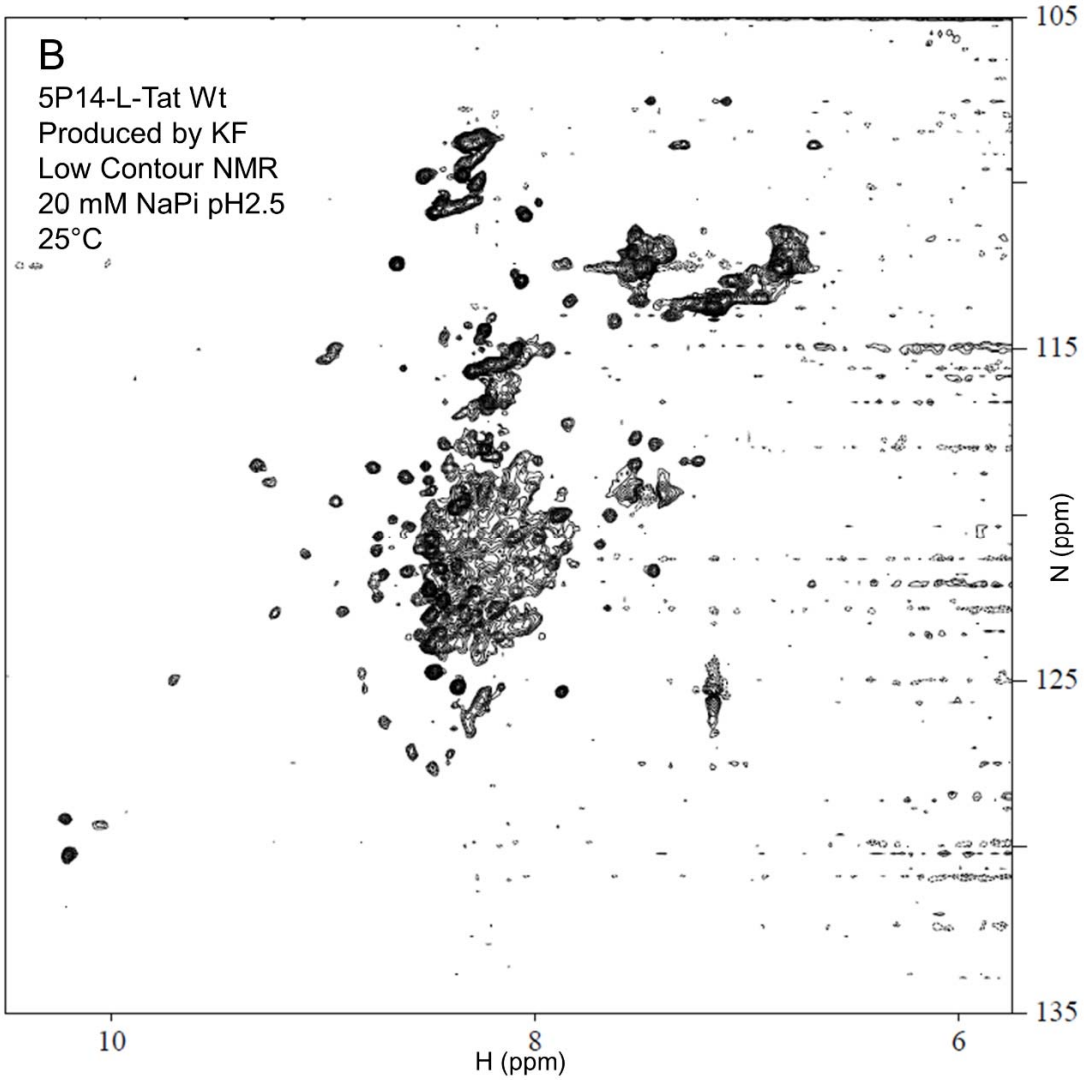
3.3.1 Correction of 5P14-Linker-Tat DNA from a random Tat sequence to Tat_{IIIb}

Soon after I inherited the project, I discovered that the protein sequence did not match the proposed sequence. According to Dr. Nguyen's records, the proposed sequence for Tat was supposed to be the HIV strain IIIb sequence but instead was an unidentified HIV sequence. In order to convert the sequence to IIIb, the following point mutations were made using two-step PCR and confirmed by sequencing: T39I, T42N & S77P. For two-step PCR protocol, see section 2.5.3, Figure 2.14 and Figure 2.16.

3.3.2 Protein Production of the Full Length 5P14-Linker-Tat Chimeric Protein

Once expression of the 5P14-Linker-Tat, with IIIb wild type Tat, chimeric protein had been re-established, large scale production began. Production of the updated chimeric protein construct Sumo-5P14-Linker-Tat_{IIIb} was done using Dr. Nguyen's protocol, as seen in her doctoral thesis, page 82[859]. Yield was still dismal however and large solution volumes resulted in lengthy and tedious purification stages that flirted with unfolding, misfolding or conglomeration of the protein due to extended exposure to room temperature conditions. A single preparation of protein had a high enough yield to be tested via HSQC NMR (Figure 3.3 A & B). An overlay with Anna Nguyen's HSQC indicates that my protein has folded elements, but is not completely folded (Figure 3.3C). Despite recommendations to both increase the scale of the initial bacterial culture, while also minimizing the overall time of the protocol, subsequent purification attempts failed due to combinations of poor protein yields and misfolding of the chimera.


A
5P14-L-Tat Wt
Produced by KF
High Contour NMR
20 mM NaPi pH2.5
25°C
105
115
125
135
N (ppm)
10
8
6
H (ppm)


B
5P14-L-Tat Wt
Produced by KF
Low Contour NMR
20 mM NaPi pH2.5
25°C
105
115
125
135
N (ppm)
10
8
6
H (ppm)

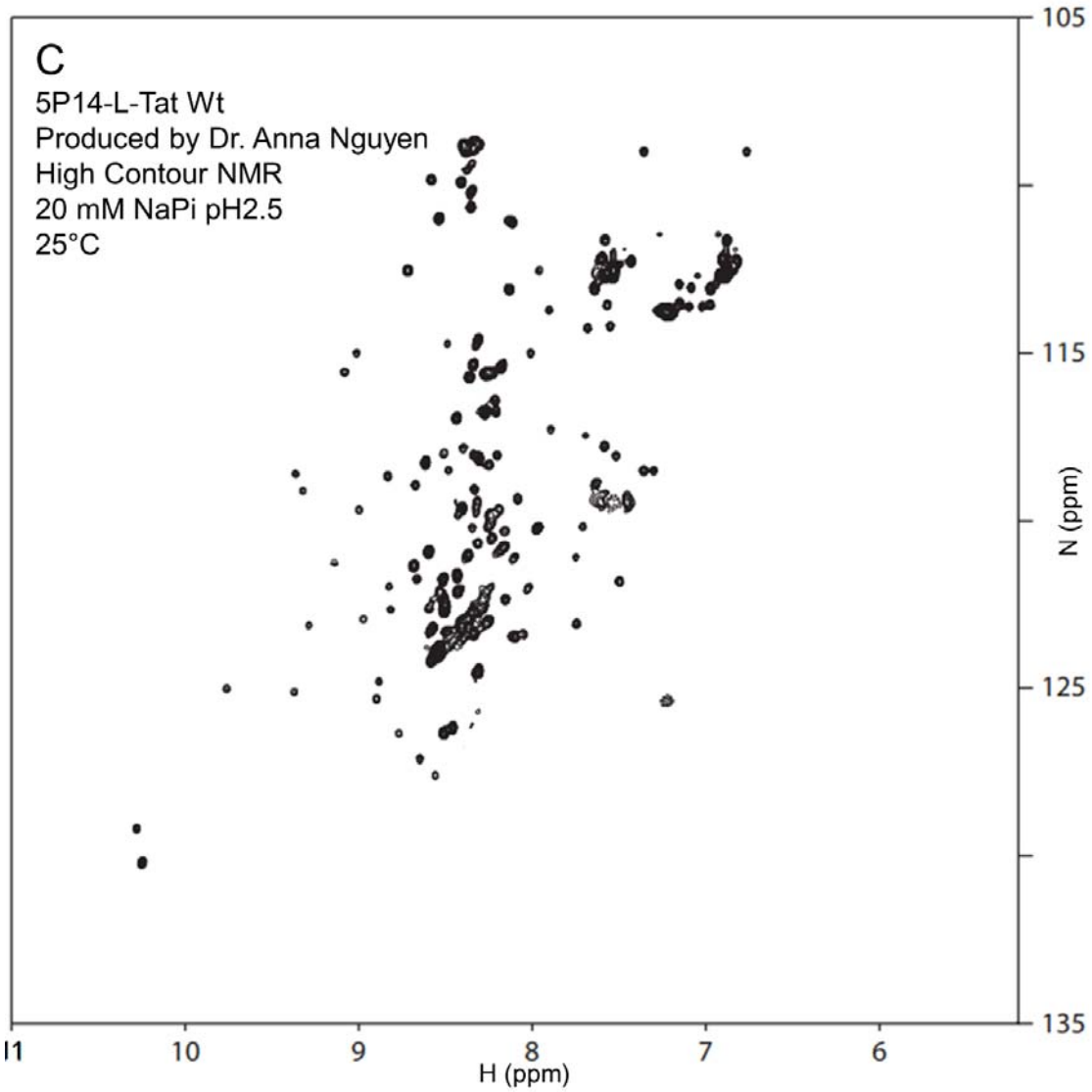


Figure 3.3 HSQC NMR Spectra of 5P14-Linker-Tat produced as full length chimeric protein.

^{15}N-Labeled HSQC spectrums of 5P14-Linker-Tat run at 25°C in 20 mM NaPi Buffer pH 2.5. **A & B)** show the NMR spectra of the 5P14-Linker-Tat protein I produced at high and low contours respectively. The spectra are largely unfolded, as exhibited by the smear of peaks around 8.2 ppm. The sample appears to have a small amount of folded 5P14-RANTES, shown by the few peaks above 9ppm in approximately the same position observed for the correctly folded protein shown in C. **C)** shows the NMR spectra of the 5P14-Linker-Tat protein produced by Dr. Anna Nguyen. This protein is similar to the known spectrum of 5P14-RANTES (Figure 3.10A), particularly the peaks above 9ppm. The several peaks around 8.2 ppm likely result from disordered Tat.

3.3.3 Production of 5P14-Linker-Tat Using Sortase Enzyme to Link Separately Prepared Proteins

To overcome the difficulties of producing the full length chimeric protein, I proposed to produce the 5P14-RANTES and Tat separately, using the enzyme Sortase, a prokaryotic enzyme that recognizes a c-terminal signal sequence (LPXTG) and covalently attaches it to an N-terminal poly-glycine sequence, to link the two proteins together after each had been appropriately folded and purified. The individual production of 5P14-RANTES and Tat would significantly shorten protein production time, simplify refolding conditions and increase final yields of each protein. This would also ease alteration of the chimera to test Tat mutants, various chemokines and protein order within the chimera.

An N-terminal poly-glycine, or a C-terminal LPMTG-6His sequences were added to both Tat_{IIIb} and 5P14-RANTES using single step PCR. The use of Met in position 3 of the sortase signal was chosen based on research by Clancy et. al[860]. For basic PCR protocol, see section 2.5.3, Figure 2.14 and Figure 2.16. The production of 5P14-RANTES followed protocols previously established by in the lab[384–386,496,506,736,850]. The purification scheme is briefly described below and illustrated in Figure 3.4.

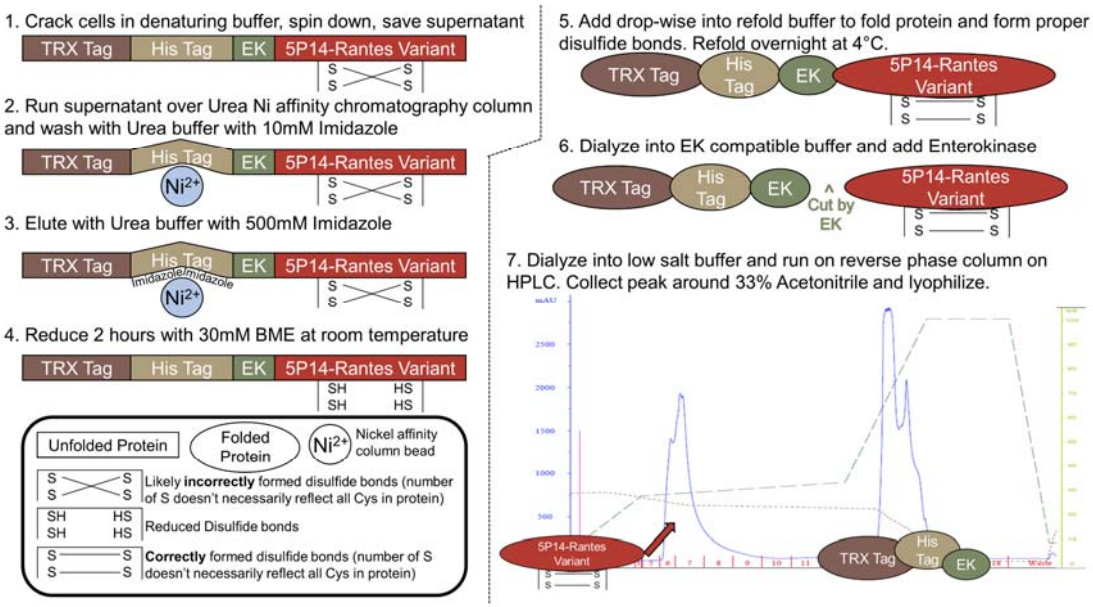


Figure 3.4 Purification scheme for 5P14-RANTES variants.
A general overview of the purification of 5P14-RANTES variants. Tan sections represent the His tag, used in Ni column purification, brown sections represent the Thioredoxin (TRX) tag used to assist in protein refolding, the green section indicates the location of the enterokinase (EK) recognition sequence (DDDK) and the red section represents the 5P14-RANTES protein. Steps with rectangular shapes are where the protein is not properly folded yet, meanwhile steps with oval shapes should have properly folded protein. A full legend is boxed in the lower left corner.

Plasmids were transformed into *Escherichia coli* BL21(DE3) (Novagen) competent cells and expressed in minimal media with $^{15}NH_4Cl$ as the sole nitrogen source; 0.5mM isopropyl β-d-1-thiogalactopyranoside (IPTG) was added to induce expression and then cells were incubated at 30°C for 16 hours, 220rpm. Cells were pelleted 5krpm for 10-

minutes in a Sorval RC6 Plus centrifuge with a F10S rotor. The cell pellet was resuspended in 8 M urea, 500 mM NaCl, 20 mM NaPi (pH 7.8), 5% glycerol and 10 mM imidazole and frozen to begin cracking cells. After thawing, the pellet was homogenized for two minutes, then centrifuged to remove the non-soluble fraction. The soluble portion was loaded onto a HiTrap IMAC nickel chelating column equilibrated with the resuspension buffer. Proteins were eluted from the column using 8 M urea buffer with 1M imidazole (8 M urea, 500 mM NaCl, 20 mM NaPi (pH 7.8), 5% glycerol and 1 M imidazole). The proteins were then refolded by dropwise addition to 9x volume of refolding buffer (550 mM L-Arginine Hydrochloride, 200 mM NaCl, 1 mM EDTA, 1 mM reduced glutathione (GSH), 0.1 mM oxidized glutathione (GSSG), 50 mM Tris, pH 8), and then allowed to stir overnight at 4°C. The solution was dialyzed twice into 4 L 200 mM NaCl, 2 mM $CaCl_2$, 20 m Tris, pH 8 buffer at 4°C. To cleave fusion tags from the purified protein, the samples were incubated for 12 hours with the protease enterokinase. The samples were then dialyzed twice into 4 L 80 mM NaCl, 2 mM $CaCl_2$, and 20 mM Tris, pH8 and then purified on a C_4 reversed-phase chromatography column (Vydac, Hesperia, CA), using an acetonitrile gradient. The proteins eluted between 33-38% acetonitrile. The fractions were analyzed on an SDS-PAGE gel to confirm purity (expected protein size is ~9 kDa) and then lyophilized in a Labconco freeze-dry system.

I collaborated with Arjan Bains to produce Tat. We were unable to induce protein expression using protocols derived from previous literature[861,862]. This may have been due to Tat's cytotoxicity, which may have resulted in killing cells that were successfully transfected. To improve expression, we tested Tat production in various cell lines and linked to common protein-fusion tags: Sumo, Thioredoxin, MBP, and GST. The GST-Tat construct was obtained from the NIH AIDS Reagent program; it was HIV strain HXB2[863–865]. I produced the other DNA constructs through PCR techniques. All constructs resulted in increased expression of Tat except Sumo-Tat (Figure 3.5).

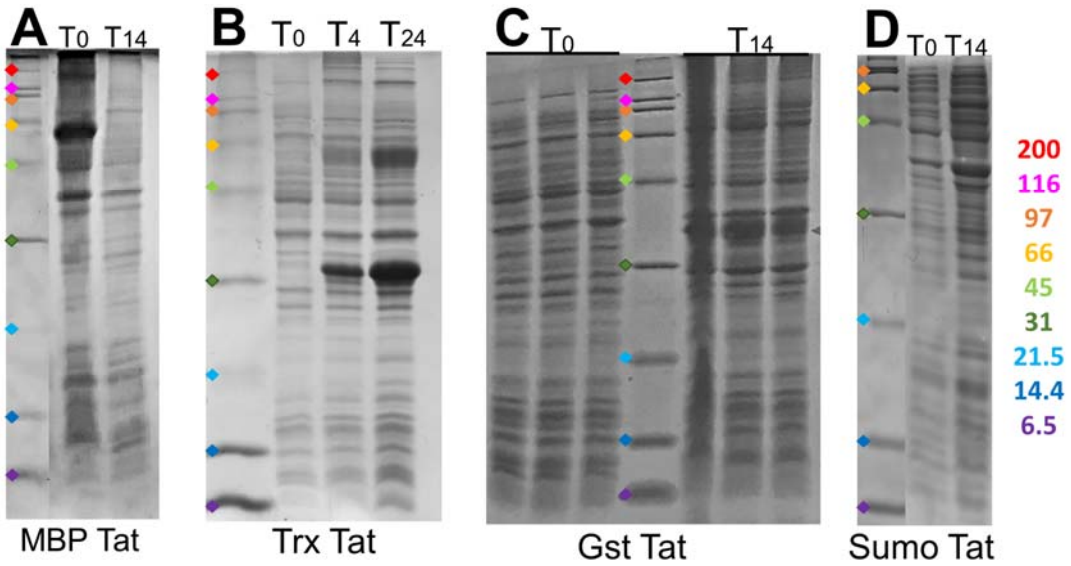


Figure 3.5 Expression of Tat with various fusion tags

A-D) Shows the SDS Page gels of Tat expressed with various fusion tags. The protein ladder standards, color coded, are shown to the right of the gel images. **A)** Shows the expression of MBP Tat at 30°C, the expected size is approximately 53kDa. **B)** Shows the expression of Trx Tat at 37°C(T4) and 20°C(T24), the expected size is approximately 26kDa. **C)** Shows the expression of Gst Tat at 30°C, the expected size is approximately 38kDa. **D)** Shows the lack of expression of Sumo Tat, this gel was from an expression test at 30°C, the expected size is 23kDa. Expression was also tested at 37°C and 20°C, no expression was seen (Data not shown).

Numerous complications were then discovered upon the attempts at purification: the thioredoxin tag co-eluted with the Tat protein even after enzymatic cleavage, GST required specialized and expensive purification techniques, MBP was extremely large and required smaller batch sizes to avoid overloading of the columns. The MBP tag was chosen as the best option, but MBP contamination in the final fractions was still a hindrance. Minimizing the amount of protein loaded onto the final C_4 reversed-phase chromatography column helped with Tat purity.

The full purification protocol for the MBP-Tat variant is described here. Plasmids were transformed into *Escherichia coli* BL21(DE3) (Novagen) competent cells and expressed in Luria broth; 0.5mM isopropyl β-d-1-thiogalactopyranoside (IPTG) was added to induce expression and then cells were incubated at 30°C for 14-16 hours, 220rpm. Cells were pelleted 5krpm for 10-minutes in a Sorval RC6 Plus centrifuge with a F10S rotor. The cell pellet was resuspended in 400 mM NaCl, 20 mM Tris (pH 7.5). The cells were homogenized for two minutes, then centrifuged to remove the non-soluble fraction. The soluble portion was loaded onto a HiTrap IMAC nickel chelating column equilibrated with 200 mM NaCl, 50 mM Tris (pH 7.5). Proteins were eluted from the column using 200 mM NaCl, 50 mM Tris (pH 7.5), and 750 mM imidazole). The solution was dialyzed twice into 4 L 200 mM NaCl, 2 mM $CaCl_2$, 20 m Tris, pH 7.5 buffer at 4°C. To cleave fusion tags from the purified protein, the samples were incubated for at least 12 hours with the

protease enterokinase. Completion of cutting was confirmed by running samples on SDS page gels. The samples were then dialyzed twice into 4 L 50 mM NaCl, and 10 mM NaPi, pH7.5-8. Protein solution was thoroughly reduced by stirring with 20-50 mM BME at room temperature for a minimum of 2 hours and then purified on a C_4 reversed-phase chromatography column (Vydac, Hesperia, CA), using an acetonitrile gradient. The proteins eluted between 28.5-30% acetonitrile. The fractions were analyzed on an SDS-PAGE gel to confirm purity (expected protein size is ~8.1 kDa, but Tat tends to run heavy, ~15 kDa) and then lyophilized in a Labconco freeze-dry system.

3.3.4 Testing of Tat Activity

The preliminary testing of the chimeric protein, performed by Dr. Nguyen, used free Tat as a positive control, but the protein had been obtained through the NIH AIDS Reagent program. TZM-bl colorimetric assays were conducted to compare activation induced by Tat made in our lab verses Tat obtained from the NIH AIDS Reagent program. The TZM colorimetric assay is essentially the same the TZM assay described in section 2.5.6 (Single-cycle Neutralization Assay (TZM-bl assay)). Briefly, TZM-bl cells were added to a 96-well plate and incubated 37°C for 18-20 hours. On day two, the media was replaced, then a series of Tat dilutions was added; Tat protein had been taken up in PBS. There was only enough Tat obtained from the NIH AIDS Reagent program for a single replicate. The wells on the edge of the plate were not used due to evaporation effects. On day 3, after 18-20 hours, the media was changed, and the cells allowed to incubate for 48 hours. On day five, the cells were lysed using 0.5% NP-40 in PBS and a substrate solution was added (20 mM KCl, 100 mM 2-Mercaptoethanol (βME) and 8.3 mM Chlorophenol red-β-D-galactopyranoside (CPRG) (CalBiochem) in PBS). The absorbance signal was measured at 570 nm 30 minutes later. Interestingly, the Tat activation followed an exponential curve rather than a linear one. At the same concentration, the Tat acquired from the NIH AIDS Reagent program showed approximately two-fold higher activation than Tat made in-house (Figure 3.6).

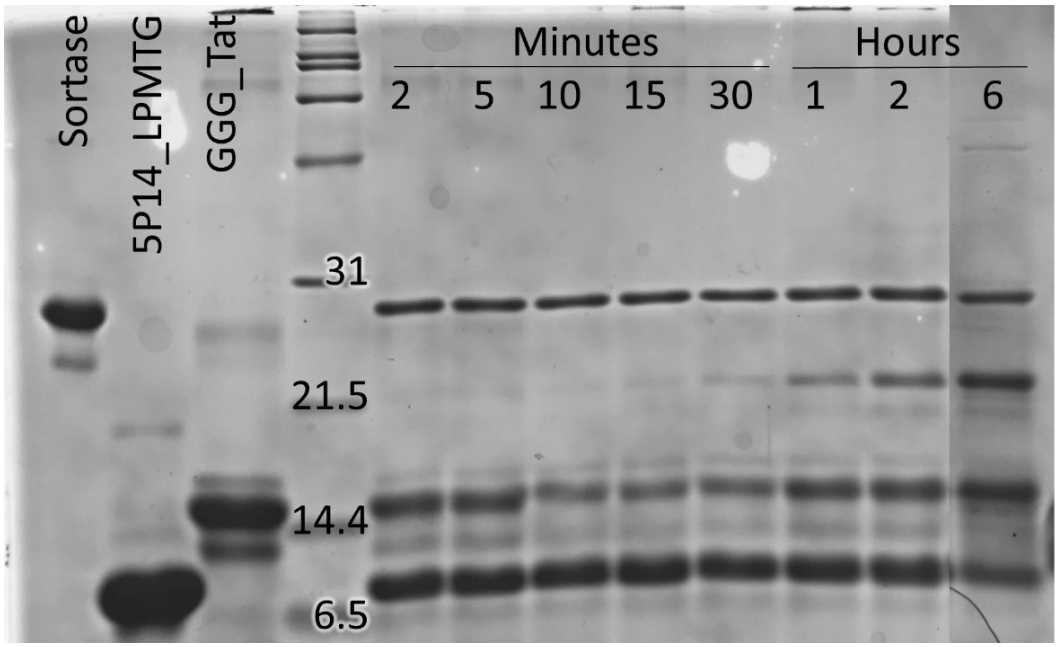


Figure 3.7 SDS Page gel showing timescale of sortase reaction
An SDS page gel shows the progression of a sortase reaction from 2 minutes to 6 hours. Controls with each individual protein are shown in the first three lanes. Expected sizes of all proteins: Sortase – 22 kDa (runs heavy at ~27 kDa), 5P14_LPMTG_His – 9.5 kDa, GGG_Tat – 8.1 kDa (runs heavy at ~15 kDa), 5P14-Linker-Tat – 16.8 kDa (runs heavy at ~22 kDa).

3.3.5 Sortase Assay to produce Chimeric Protein 5P14-Linker-Tat

The protocol for combining 5P14-RANTES and Tat using sortase was based on protocols from Theile et al[866] and Antos et al[867]. Briefly, 5P14-RANTES constructs were taken up in 20mM NaPi pH 2.5 buffer, while Tat constructs were taken up in 1X sortase reaction buffer (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4 pH 8, 1mM $CaCl_2$ and 0.5 mM $MgCl_2$). Proteins were added to a final concentration of 50uM to a 1X sortase reaction buffer with 10mM $CaCl_2$ and 5uM Sortase enzyme. Various times and temperatures were tested. A reaction left at room temperature for ~6 hours appeared to do the best based on time scale reactions (Figure 3.7). Small scale testing (200uL) showed the production of the chimeric protein, but the reaction would utilize less than half of the initial substrates and significant amounts of precipitation, composed of both the final chimeric protein and the individual components, formed over the course of the reaction.

It was previously known known that chemokines aggregate at high concentrations and physiologic pH. Another group had published a set of mutations, E26A and E66S, that they found to minimize chemokine precipitation[851,853]. These mutations were introduced into 5P14-RANTES through two step PCR to minimize precipitation during purification and the Sortase protocol. Neither the purification nor Sortase protocols were altered from what has already been outlined above. Mutants were produced and tested with the assistance of Julissa Garcia, Giselle San Roman and David Chen. The individual E66S mutation in 5P14-RANTES resulted in significantly reduced precipitate formation compared to 'wild type' 5P14-RANTES (Figure 3.8) and so medium and large-scale sortase assays were done to begin testing purification of the chimera away from the individual proteins and the sortase enzyme.

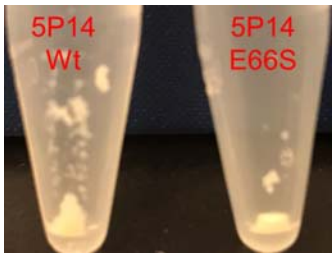


Figure 3.8 Comparison of precipitate produced during sortase reactions using 5P14-RANTES and 5P14 - RNATES E66S.

While the reaction was successful, purification of the chimera away from the individual components proved challenging. The chimera eluted over a wide range on both a C4 column and a heparin column. There was also MBP contamination present in all fractions, which highlighted a need for better production and purification of the individual Tat component. It was proposed to use a heparin column first to separate out free 5P14-RANTES and sortase enzyme, then use a C4 reverse phase column to separate free MBP and Tat from the full chimeric protein (5P14-Linker-Tat). The two-column method failed due to loss of the chimera; minimal initial yields from the enzymatic reaction limited amounts of chimeric protein and the use of two columns increased the rate of loss due to protein handling and exposure to equipment.

A large scale sortase reaction was prepared and purified using a C4 reverse phase column, resulting in fractions with large amount of the desired chimera but also significant presence of contaminating proteins (Figure 3.9). The cleanest of these fractions (Fractions 18-21), that contained minimal to none free 5P14-RANTES, was run on NMR HSQC (Figure 3.10B) and compared to the NMR HSQC that had been run of the $5P14_{E66S}$_LPMTG protein that had been used in the sortase reaction (Figure 3.10A) The sample appeared to have a folded 5P14-RANTES component but did contain evidence of unfolded protein as well (Figure 3.10C). It was hoped that the contaminating proteins were the unfolded elements see in the spectrum and it was decided to use those fractions to begin testing in the cellular assay. This batch of protein was used for all FACS based assays.

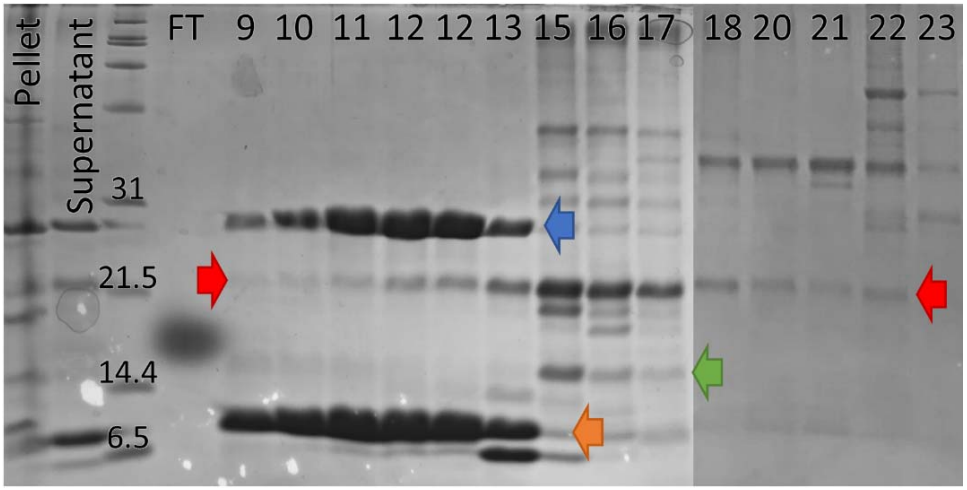
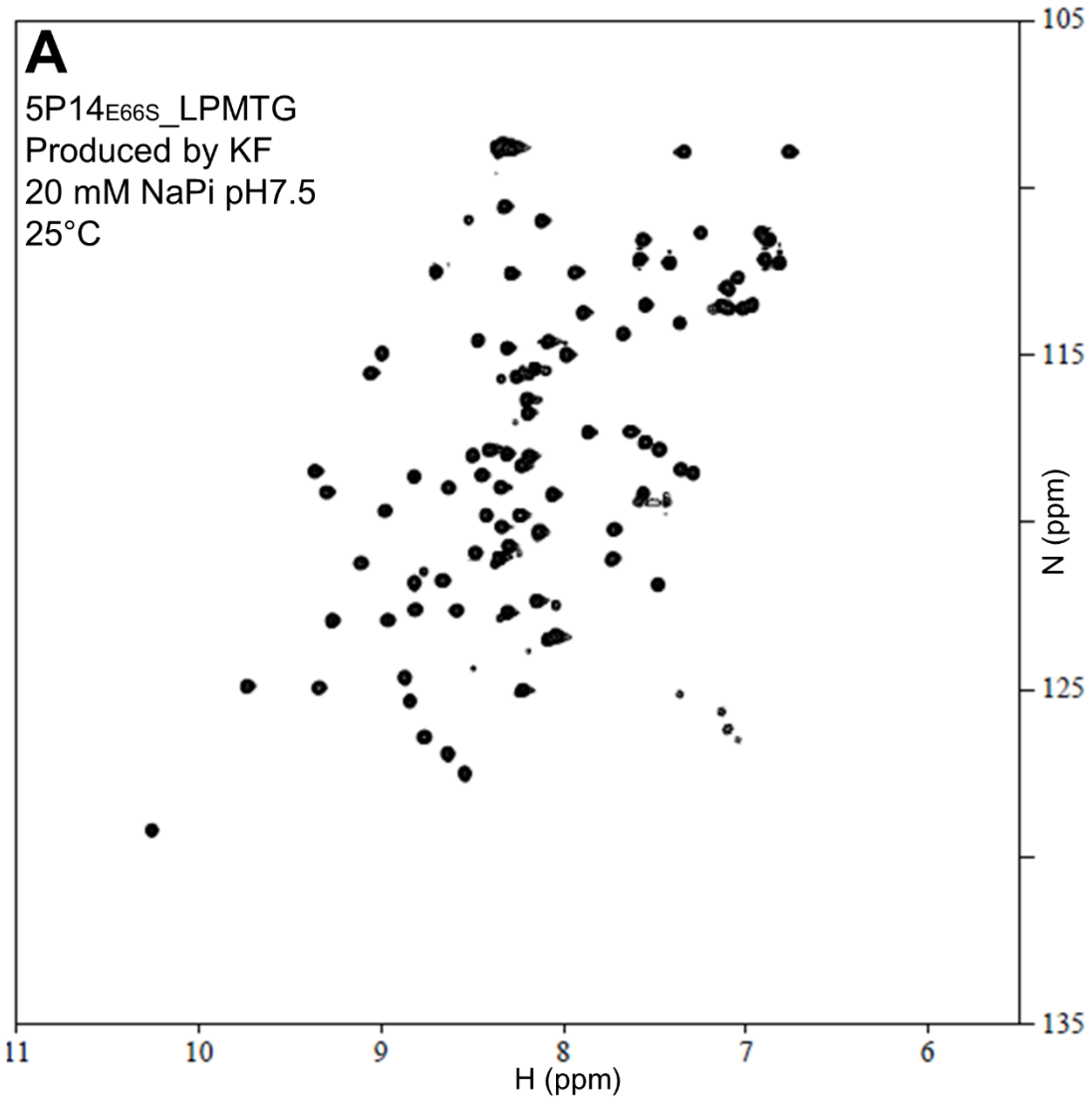


Figure 3.9 SDS Page gel of Sortased 5P14-L-Tat chimera C4 column fractions.
The gel shows the elution fractions of the sortase reaction to produce 5P14-Linker-Tat of off a reverse phase column. FT indicates flow through of the reaction, while the numbers at the top of each column indicate C4 column elution fractions. The blue arrow indicates the Sortase enzyme, the red arrows indicate the 5P14-Linker-Tat chimeric protein, the green arrow indicates free GGG_Tat 3B, the orange arrow indicates free $5P14_{E66S}$_LPMTG. The extended elution of the 5P14-Linker-Tat chimera and the absence of clean (single protein) fractions can be observed.


A
5P14E66S_LPMTG
Produced by KF
20 mM NaPi pH7.5
25°C
105
115
125
135
N (ppm)
11
10
9
8
7
6
H (ppm)

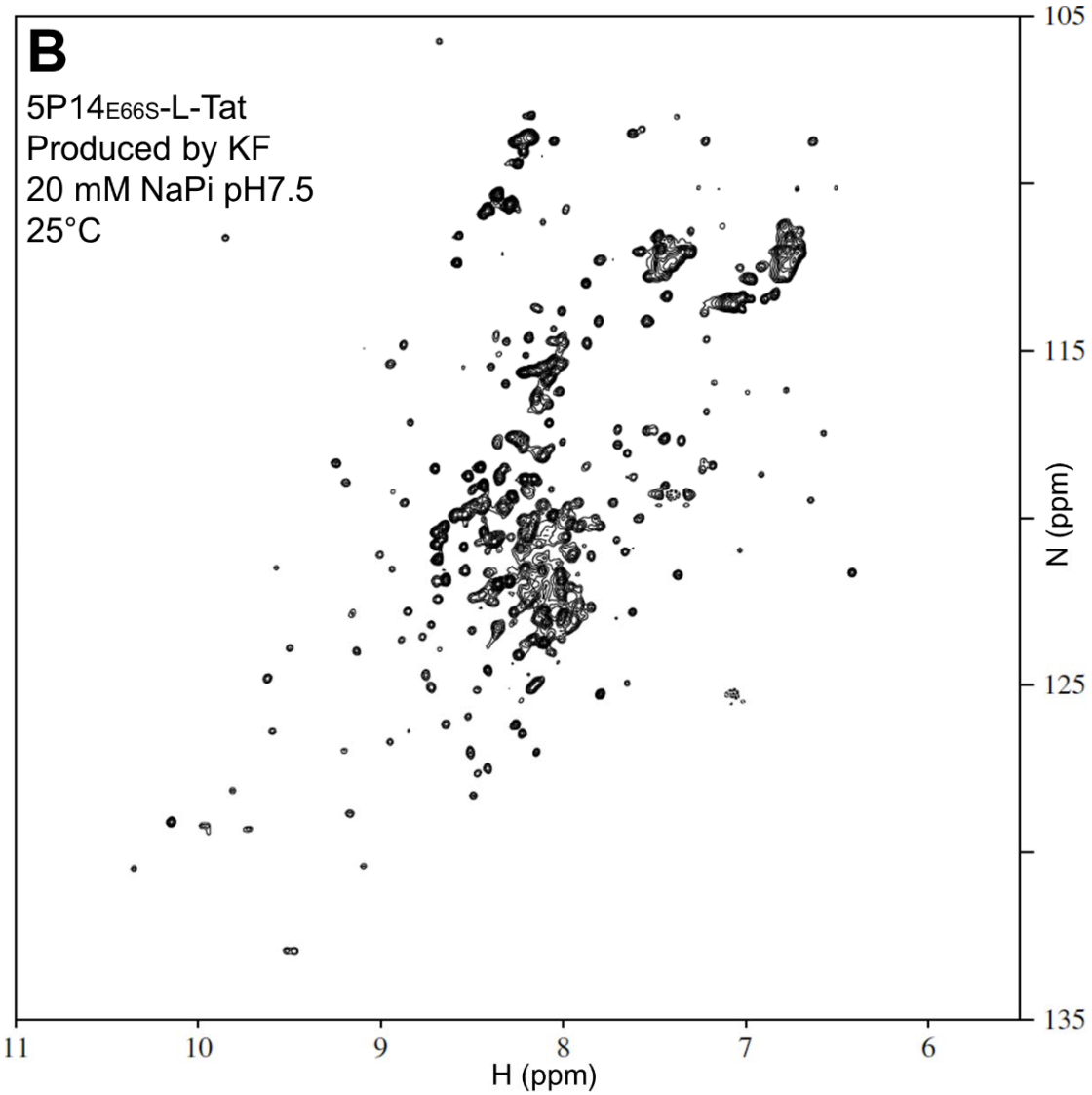


Figure 3.10 HSQC NMR of free 5P14-RANTES and sortased 5P14-Linker-Tat.
^{15}N-Labeled HSQC spectra of 5P14$_{E66S}$_LPMTG **(A)** and 5P14$_{E66S}$-Linker-Tat **(B)** run at 25°C in 20 mM NaPi Buffer pH 2.5. **A)** Shows that the 5P14-RANTES mutant (5P14 $_{E66S}$_LPMTG) was clearly folded, with sharp resolution of peaks. **B)** Shows that the sample contains unfolded protein, as evidenced by the smear of peaks around 8.2 ppm. The sample does appear to have some amount of folded 5P14-RANTES, shown by the few peaks above 9ppm in approximately the same position observed for the correctly folded protein shown in A. However, there are additional peaks that were not observed in the spectra for 5P14-RANTES alone (A).

3.3.6 Testing of the 5P14-Linker-Tat chimera for functionality

With semi-pure, folded protein, preliminary testing could begin. The goals were to ascertain that our 5P14-Linker-Tat chimeric construct 1) was able to activate cells containing an HIV LTR linked to a reporter gene, 2) demonstrated selectivity for CCR5 bearing cells, and 3) was able to reactivate latently infected cells (utilizing a cell line that had an LTR linked reporter, not live virus). This protocol required extensive troubleshooting, which is summarized in Table 3.1 and elaborated on in the section below. In our initial iteration of the protocol we aimed to test goals 1 and 2 (activation and selectivity) at the same time, using a suspended cell line containing a GFP reporter that could be tested using Fluorescence-activated cell sorting (FACS). FACS enabled the determination of the level of cellular activation by the Tat and the presence and density of CCR5 for each individual cell, therefore allowing us to determine whether there was a correlation between activation and the density of CCR5 on the surface rather than a simple, binary presence or absence of the receptor. The cell lines JLTRG and JLTRG+R5 were chosen as a pair since they came from the same progenitor cell line, thus having the same reporter and hopefully similar sensitivities to Tat activation. The only purported difference between the two cell lines was that JLTRG+R5 expressed low levels of CCR5. After determination that the JLTRG cell lines were entirely CCR5+, a pair of MAGI cell lines was procured, one with CCR5, the other without. These MAGI cell lines are adherent cell lines that contains a B-galactosidase reporter. Due to the nature of a MAGI colorimetric assay, the cells cannot be mixed to show selective activation of $CCR5^{+}$ cells, therefore, we were only able to directly test activation and had to infer selectivity by comparing activation levels of the two cell lines.

In our finalized protocol to test the functionality of both orientations of our chimeric proteins: 5P14-Linker-Tat and Tat-Linker-5P14, we tested the two proteins side by side in a MAGI colorimetric assay. These proteins were produced by Arjan Bains, both of which contained the E66S mutation to minimize precipitation at neutral pH. The 5P14-Linker-Tat construct was produced as a single polypeptide following Dr. Nguyen's protocol, while the Tat-Linker-5P14 was produced as two separate proteins and then linked together using sortase as described in the protocols above. Neither protein was able to be completely purified (Figure 3.11 A & B) and an HSQC NMR spectra was only run of the 5P14-Linker-Tat construct (Figure 3.11C). The protein did not appear to be completely folded, but we decided to proceed with preliminary testing.

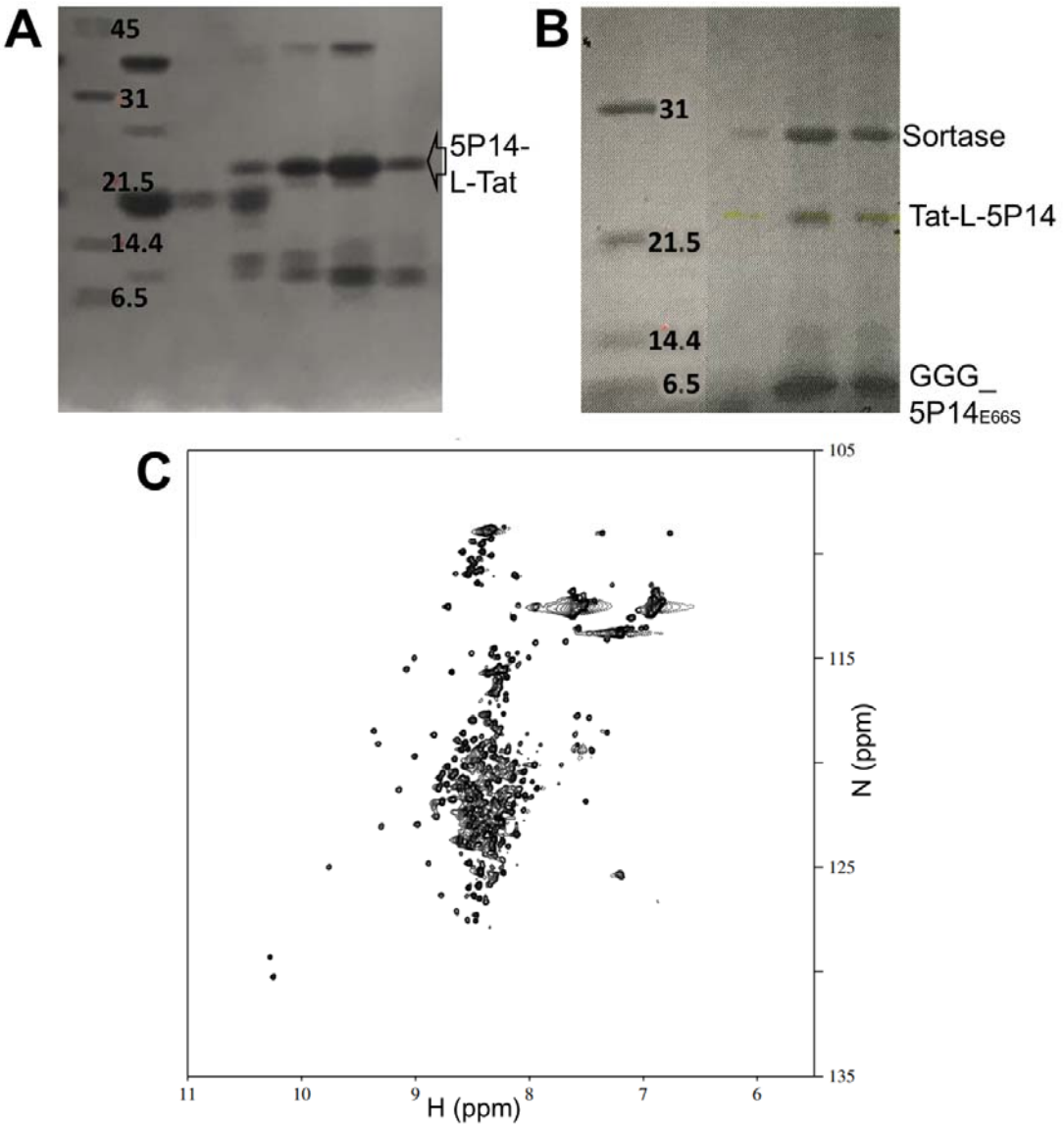


Figure 3.11 SDS Page gels and NMR of chimeric proteins produced by Arjan Bains
A) shows the SDS page gel of the C4 reversed phase column elution fractions of 5P14-Linker-Tat. The chimera is indicated with an arrow. Other proteins seen are unidentified bacterial proteins. **B)** shows the SDS page gel of the C4 column elution fractions of 5P14-Linker-Tat. The chimera and other proteins from the sortase reaction are labeled. C) shows the ^{15}N-Labeled HSQC spectrum of 5P14$_{E66S}$-Linker-Tat run at 25°C in 20 mM NaPi Buffer pH 2.5. The spectra are largely unfolded, as exhibited by the smear of peaks around 8.2 ppm. The sample appears to have a small amount of folded 5P14-RANTES, shown by the few peaks above 9ppm in approximately the same position observed for the correctly folded protein shown in
Figure 3.3C and Figure 3.10A.

The chimeric constructs were tested side by side, with GGG-Tat as a positive control and 5P14-RANTES as a negative control. The MAGI colorimetric assay is essentially the same the TZM assay described above. On day one, 10^4 cells were added to a 96-well plate and incubated at 37°C for 18-20 hours. On day two, the media was replaced, then proteins were added to a final concentration of 20uM. 5P14 and chimera constructs were taken up in 20mM NaPi pH 2.5. Tat was taken up in 1X sortase buffer (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4 pH 8, 1mM $CaCl_2$ and 0.5 mM $MgCl_2$). Buffer only controls were included to ensure that the low pH of the buffer used to take up the 5P14-RANTES constructs did not affect cell growth or baseline activation. The wells on the edge of the plate were not used due to evaporation effects. On day 3, after 18-20 hours, the media was changed, and the cells allowed to incubate for another 48 hours. On day five, the cells were lysed using 0.5% NP-40 in PBS and a substrate solution was added (20 mM KCl, 100 mM 2-Mercaptoethanol (βME) and 8.3 mM Chlorophenol red-β-D-galactopyranoside (CPRG) (CalBiochem) in PBS). The absorbance signal at 570nm was measured when absorbance was approximately 2.5.

Testing of the chimeric proteins, 5P14-Linker-Tat and Tat-Linker-5P14, did not show selective targeting for CCR5+ cells (Figure 3.12). In other words, the activation in both MAGI and MAGI +CCR5 cells was nearly identical. The 5P14-Linker-Tat chimera did not show activation. The Tat-Linker-5P14 orientation showed roughly 20% activation compared to 20uM Tat, which was only slightly lower than the 10uM tat control. The level of activation may be influenced by inaccuracies accrued during protein quantification since the protein was not completely pure and therefore the concentration of the chimera had to be estimated. The difference in activation between the two orientations may not actually be due to order of the proteins but due to the process of protein production. The 5P14-Linker-Tat was produced as a single chimeric construct(3.3.2 Protein Production), while the Tat-Linker-5P14 chimera was produced as two separate proteins that were then linked together with Sortase (3.3.5 Sortase Assay to produce Chimeric Protein 5P14-Linker-Tat). This means that since the prep of Tat-Linker-5P14 was not completely clean at the end of production (Figure 3.11B), there was potentially free Tat that could non-specifically be activating the cells.

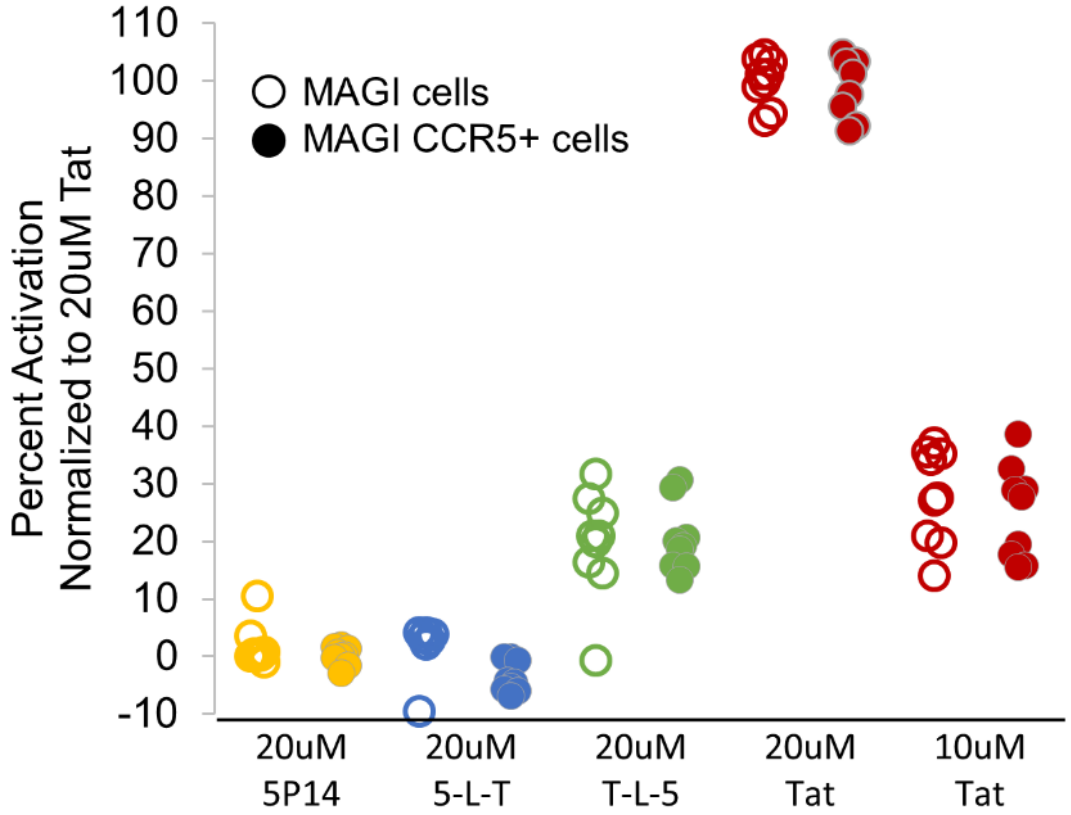


Figure 3.12 Testing of chimeric constructs in MAGI colorimetric assay.
Testing of 5P14-Linker-Tat (5-L-T) and Tat-Linker-5P14 (T-L-5) chimeric constructs in triplicate. Open circles indicate MAGI cells (CCR5-) and filled circles indicate MAGI CCR5+ cells. Levels of activation were linearly normalized with 20uM Tat set as 100% activation and 20uM 5P14 as 0% activation. It should be noted that Tat does not seem to follow a linear activation curve, as evidenced by the addition of 10uM Tat only yielding 15-35% activation. There is no discernible difference between the two cell lines in activation caused by either chimera. A figure showing raw data (direct 570nm absorbance) and a description of how data is processed can be found in Appendix C (pages 168-169).

Table 3.1 Troubleshooting of protocol to test chimera functionality

Problem	Solution
Testing of JLTRG & JLTRG+R5 cell lines for CCR5	Quantify surface expression of CCR5 using FACS and anti-CCR5 Ab 2D7
Balancing the cytotoxicity of Tat with activation	Test gradients of Tat to compare activation and cell death
Internalization of the CCR5 receptor	Intracellular staining protocol, results in lower GFP fluorescence
Competition between 5P14-RANTES and anti-CCR5 antibody 2D7	Purchased new anti-CCR5 antibody CTC5 that has been shown to not compete with chemokines
CTC5 antibody shows that 100% of JLTRG population is CCR5+, means assay does not have internal control to show selectivity for cells bearing the CCR5 receptor	CTC5 has been shown to have off target binding, use new anti-CCR5 antibody, J418F1, from Dr. Hoyer
J418F1 antibody shows that 100% of JLTRG cells are CCR5+.	Purchase new MAGI & MAGI+CCR5 cell lines
No direct testing of MAGI cell line to confirm absence of CCR5	Tested cells with live stain, with J418F1, under confocal microscopy, confirmed presence of CCR5 on MAGI+CCR5 and absence of CCR5 on the MAGI cells

FACS Assay Protocol

All the following FACS assays followed the same general protocol unless otherwise specified. On day one, cells were plated into a 96 well plate and proteins were added to their final concentrations; 5P14 and chimera constructs were all taken up in 20 mM NaPi buffer pH 2.5, while Tat was taken up in 1X sortase buffer (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4 pH 8, 1mM $CaCl_2$ and 0.5 mM $MgCl_2$). Cells were then allowed to grow for approximately 72 hours. For surface staining of the cells, cells were collected by centrifuging at 400g for 5 minutes, then stained with a 1:100 dilution of the desired antibody and ~0.5µg/ml of DAPI for 15 minutes. Cells were washed with 500uL FACS buffer (PBS with 2% FBS) two times then resuspended in 100uL FACS buffer and stored on ice until read. If intracellular staining also occurred, 506 viability dye was used instead of DAPI. For intracellular staining, cells were fixed with either Invitrogen Fox3/Transcription Fixation/Permeabilization buffer (completely negates GFP absorbance) or BD cytofix/cytoperm buffer for approximately 40 minutes to an hour, then washed with a permeabilization buffer (PBS with 0.5% Saponin) and permeabilized for approximately 10 minutes. Cells were stained with a 1:100 dilution of the desired antibody for 45 minutes to an hour, then rinsed and resuspended in permeabilization buffer.

Initial testing of the JLTRG & JLTRG+R5 cell lines

Initial testing of the JLTRG & JLTRG+R5 cell lines with anti-CCR5 antibodies 2D7 & 45531 indicated that while JLTRG+R5 did express CCR5 on >90% of the population, it was predominately only detectable with antibody 2D7 (Figure 3.13). This was unsurprising since the other antibody, 45531, recognizes epitopes produced by aggregation of CCR5. Results indicating that the JLTRG cell line had CCR5 expression on ~50% of the population, as identified by the 2D7 antibody, were surprising but not unwelcome (Figure 3.13). The original design of the assay had required eventually mixing the JLTRG and JLTRG+R5 cell lines to achieve a 50:50 mixture of $CCR5^{+}$ to $CCR5^{-}$ cells in order to demonstrate our chimera's selectivity for CCR5 bearing cells. Since the JLTRG cell line already had a 50:50 population, creating an internal control, it was decided to use the JLTRG cell line alone. This also ensured an exact match of Tat sensitivity within the assayed cells, which would have needed to be controlled for in prior iterations of the assay design.

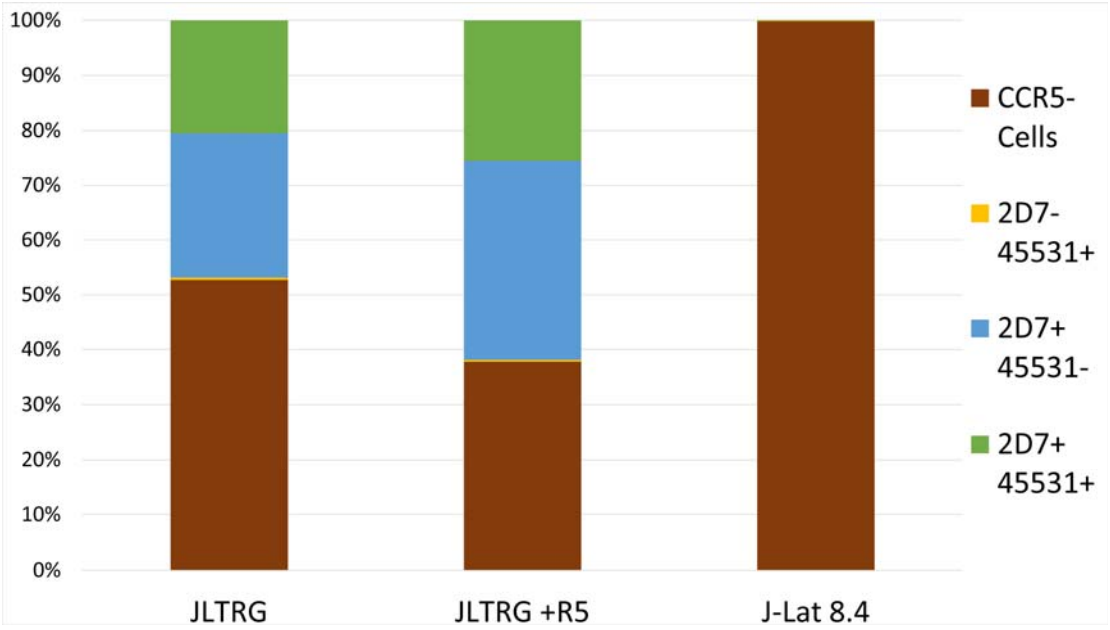


Figure 3.13 Expression of CCR5 on Jurkat cell lines (Trial 1)
Cell lines JLTRG, JLTRG+R5 and J-Lat 8.4, were tested by FACS using anti-CCR5 antibodies 2D7 and 45531. 2D7 targets monomeric CCR5, while 45531 targets aggregated/oligomeric CCR5. Brown indicates the percent of cells without any CCR5 present, yellow indicates the percent of cells recognized by only 45531, blue indicates the percent of cells recognized by only 2D7, green indicates the percent of cells recognized by both 2D7 and 45531.

Balancing Tat cytotoxicity and activation

A FACS assay was used to test a range of protein concentrations, from 10uM down to 2uM. The results are shown as dot-plots, with the intensity of GFP emission on the y axis and the intensity of anti-CCR5 2D7 antibody (APC channel) on the x axis. Gating was based on the unstained and untreated sample (Figure 3.14 unstained - untreated), with Q1 corresponding to cells that are GFP+/CCR5-, Q2 corresponding to cells that are CCR5+/GFP+, Q3 corresponding to cells that are CCR5+/GFP- and Q4 corresponding to cells that are CCR5-/GFP. The addition of Tat showed high levels of activation of the GFP reporter that was accompanied by increasing cell death at higher concentrations of Tat (Figure 3.14 2.5uM Tat, 5uM Tat and 10uM Tat). The activation can be seen in the increase of GFP positive cells (found in Q1 and Q2) from 32.6% in the untreated cells to greater than 90% activation in all cell treated with Tat. Increasing cell death is shown in the number of live cells decreasing from over 53 thousand in the untreated sample down to 10.6 thousand in the 10uM Tat sample. Ranges of 5P14-l-Tat protein showed that the chimera was approximately 25% as effective as Tat alone at similar concentrations, with corresponding levels of cell death (Figure 3.14 2.5uM 5P14-L-Tat through 10uM 5P14-L-Tat). This result must be taken with a grain of salt considering that the protein was not completely pure and therefore the concentration was an estimate to begin with. The addition of 5P14 alone resulted in neither activation nor cell death (Figure 3.14 2.5uM 5P14 through 10uM 5P14).

Internalization of the CCR5 receptor

An unexpected result from our results was the absence of CCR5 visible by 2D7 surface staining when the cells were exposed to any of our proteins. This is evidenced by the shift in cells out of Q2 and Q3 as compared to the stained but untreated sample (Figure 3.14). Samples exposed to Tat showed a dose-response reduction in CCR5 presence, with the 2.5uM Tat sample showing ~20% of cells as CCR5+, while the 10uM Tat sample showed only 7% CCR5+ (Figure 3.14). All samples containing a form of 5P14 were less than 2% CCR5+ (Figure 3.14). While this was partially expected for 5P14 and 5P14-L-Tat since 5P14 causes internalization of the CCR5 receptor, it was definitely not expected for Tat alone. It was not an initial concern for our longer scale experiments since the CCR5 receptor has been shown to be recycled back to the surface of the cell a few hours after removal of the chemokine[854].

According to Kraft 2001, receptors returned to the membrane after 120 minutes after being incubated with 50nM chemokine for 1 hour, but in their experiments the chemokine was removed by washing with fresh medium[854]. We decided to adjust our protocol accordingly, comparing the result of washing the cells 3 times with RPMI media either 12 hours or four hours before staining and reading. Washing of the cells at either time removed the internalization of the CCR5 receptor for cells exposed to Tat only. Apparently, even picomolar levels of 5P14 were still enough to result in nearly complete internalization of the receptor (Figure 3.15 & Figure 3.16). This is demonstrated by all Tat samples having cell populations that are ~50% CCR5+, while any sample containing 5P14 shows less than 2% of cells as CCR5+ (Figure 3.15 & Figure 3.16).

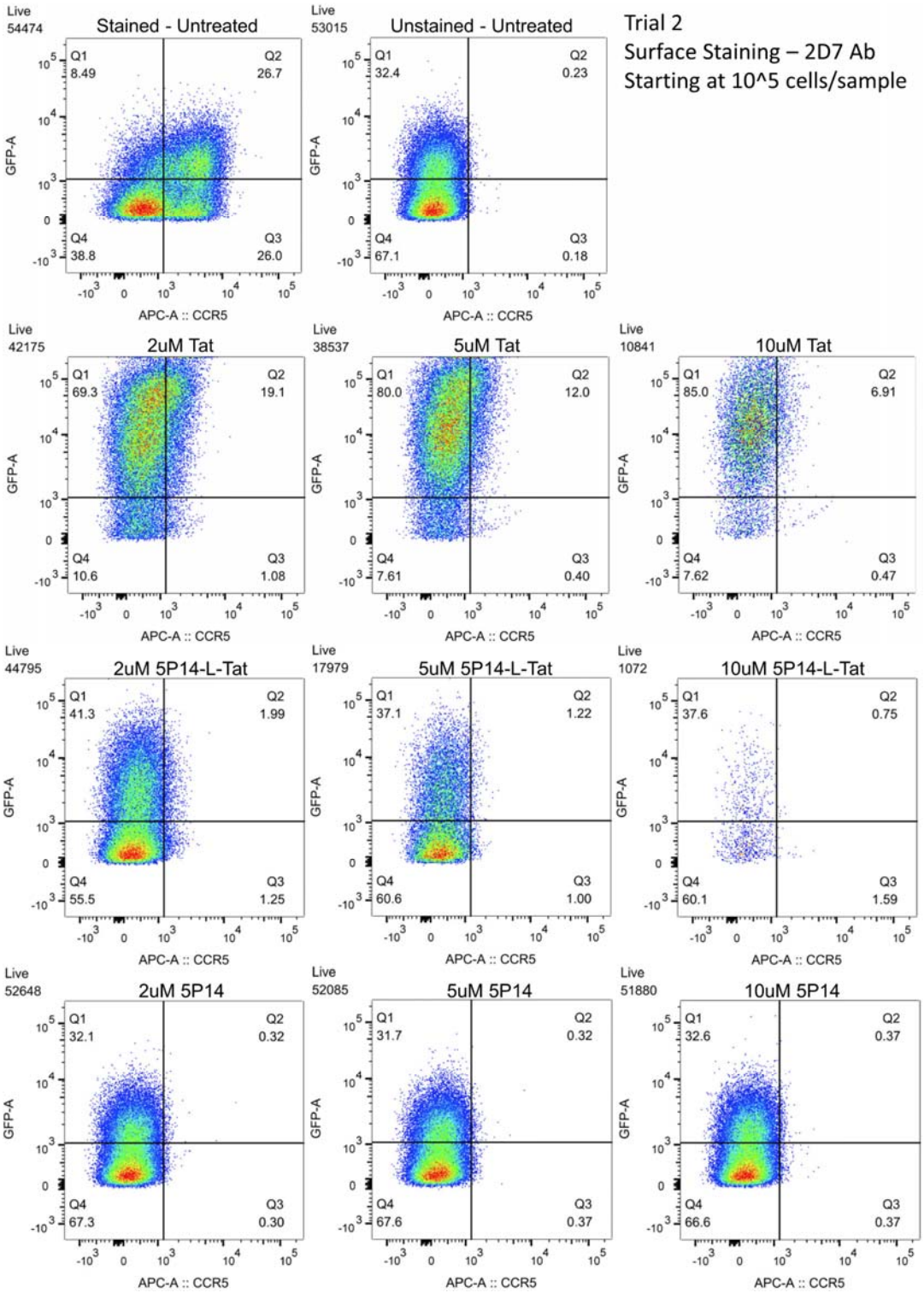


Figure 3.14 FACS Trial 2 - Surface Stain with 2D7 Ab
These dotplots show the intensity of GFP emission (y axis) versus the intensity of anti-CCR5 2D7 antibody (APC channel, x axis). The number of live cells remaining is found in the upper left corner of each plot. Frequencies of cells that fall within the gates are indicated. Full gating can be found in Appendix C.

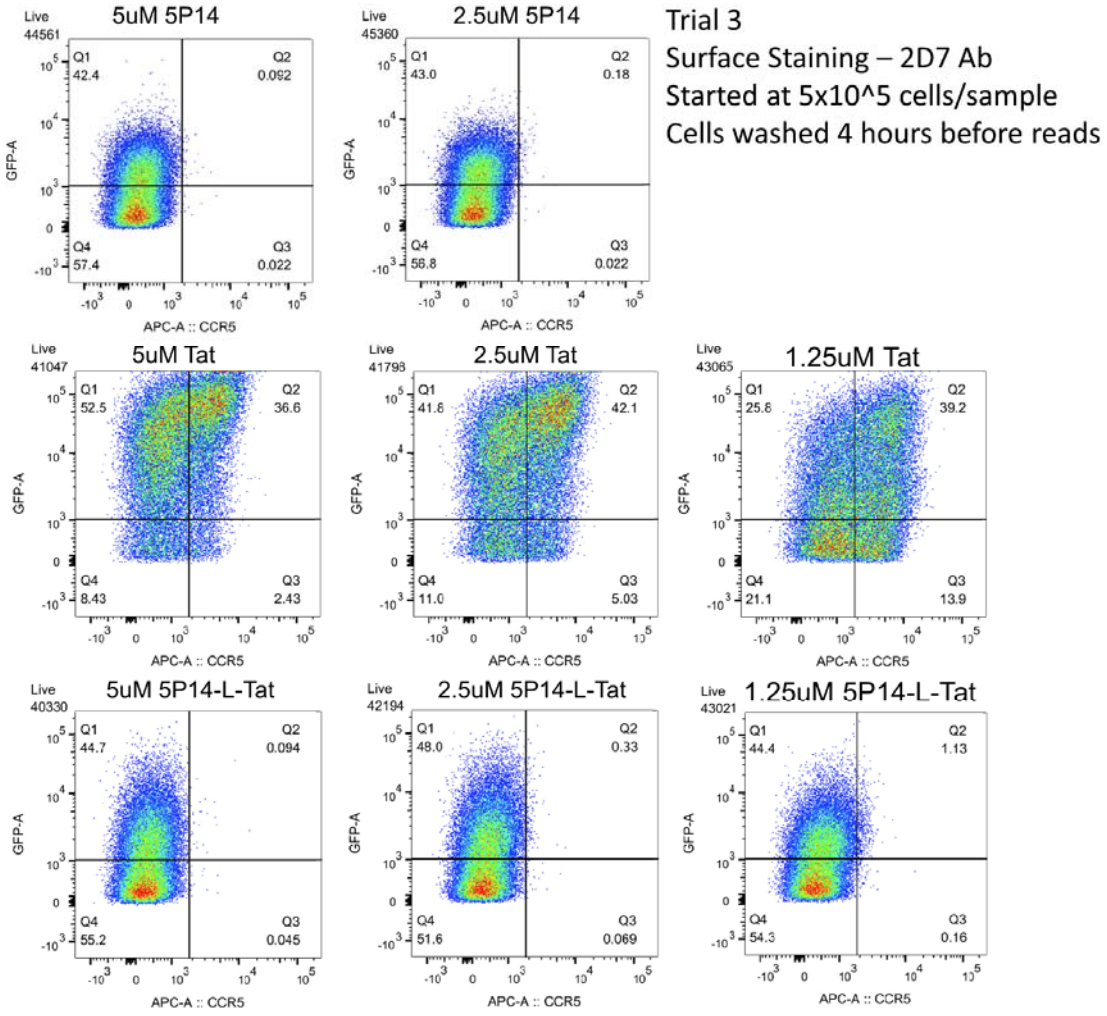


Figure 3.15 FACS Trial 3 - 4 hour wash - Surface stain with 2D7
These dotplots show the intensity of GFP emission (y axis) versus the intensity of anti-CCR5 2D7 antibody (APC channel, x axis). The number of live cells remaining is found in the upper left corner of each plot. Full gating can be found in Appendix C (pages 175-177).

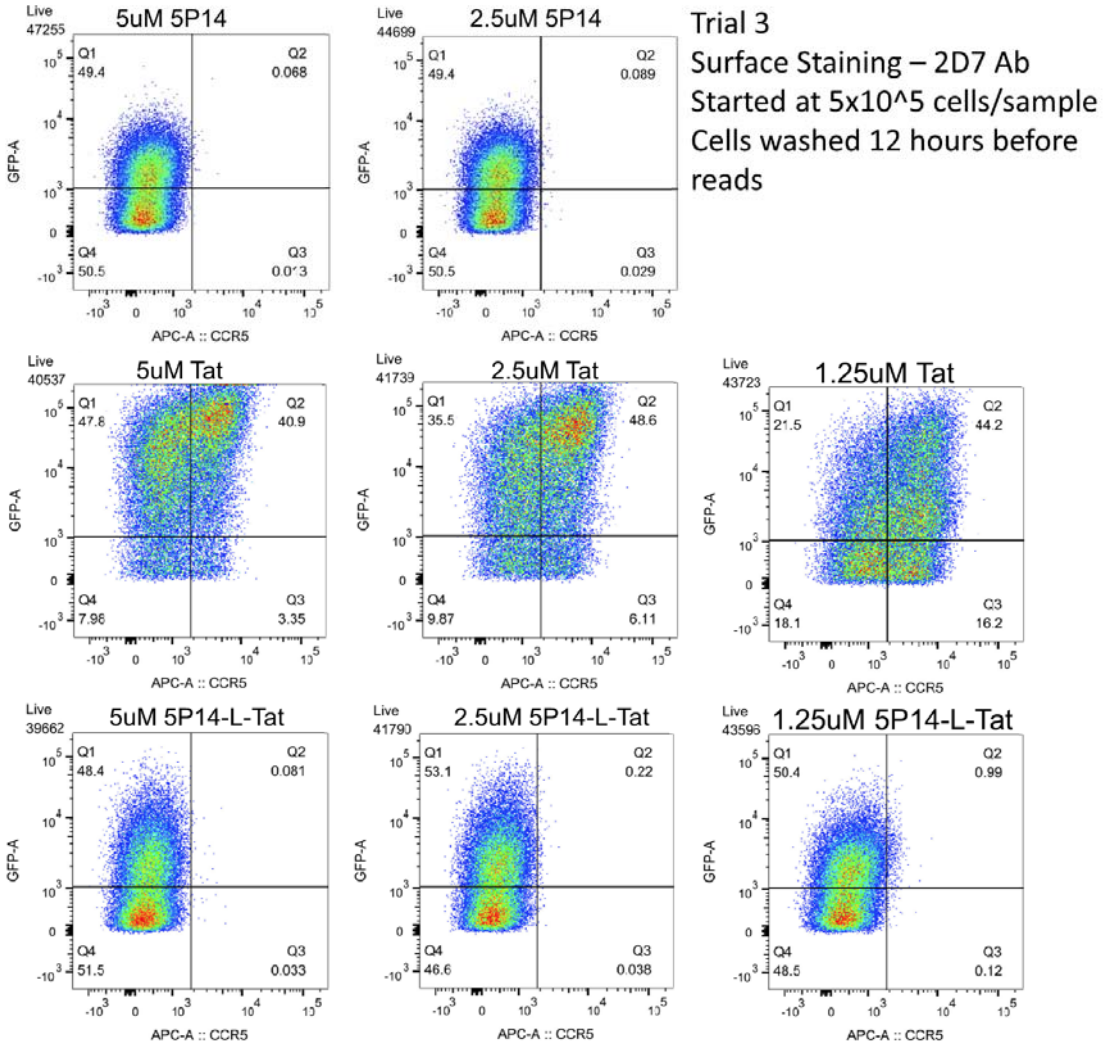


Figure 3.16 FACS Trial 3 - 12 hour wash - Surface stain with 2D7
These dotplots show the intensity of GFP emission (y axis) versus the intensity of anti-CCR5 2D7 antibody (APC channel, x axis). The number of live cells remaining is found in the upper left corner of each plot. Full gating can be found in Appendix C (pages 178-179).

Competition between 5P14 and 2D7 Ab for binding to CCR5

Our suspicion was that the 5P14 protein, both alone and in the chimera, was competing with the 2D7 antibody for binding to the CCR5 receptor and literature confirmed this theory[868,869]. The anti-CCR5 antibody, CTC5, was demonstrated to be compatible with chemokine binding assays and so we purchased this new antibody to replace the 2D7 in our assay. Shockingly, when the JLTRG cells were tested with the CTC5 antibody, nearly 100% of the cells were indicated to have CCR5 (Figure 3.17 E & F). A deeper literature review of the CTC5 antibody revealed that it has a propensity for off-target binding[870]. A third antibody, J418F1, was donated to our use by Dr. Hoyer, who had used it previously[871] with creditable specificity and robustness of signal; whether it competed with the chemokine binding site was unknown and untested based on a literature review.

JLTRG cells are >90% CCR5+

Trial of the J418F1 antibody confirmed the results found by the CTC5 antibody: nearly 100% of the JLTRG cells, despite distributor reports[856,857], expressed the CCR5 receptor(Figure 3.17 H & I). A FACS assay of JLTRG cells was used to compare the amount of CCR5 labeled with the antibodies J418F1, CTC5 and 2D7 in the presence and absence of 5P14 RANTES. The cells used to test binding of J418F1 and CTC5 were surface and intracellularly stained while the cells used to test 2D7 were only surface stained. Gating was determined based on the untreated/unstained cells for each grouping (Figure 3.17 A, D & E) with GFP absorbance on the y axis and absorbance of the fluorophore conjugated to each anti-CCR5 antibody on the x axis. Q1 corresponds to cells that are GFP+/CCR5-, Q2 corresponds to cells that are CCR5+/GFP+, Q3 corresponds to cells that are CCR5+/GFP- and Q4 corresponds to cells that are CCR5-/GFP.

Two intercellular staining protocols were also tested to ensure CCR5 would be visible, even if internalized by 5P14 or the chimera. The first resulted in complete quenching of the GFP fluorescence which was unacceptable (Figure 3.17D-F). The second decreased GFP fluorescent intensity but it was still within acceptable levels (Figure 3.17A-C).

The original scientific literature detailing the production of the JLTRG and JLTRG+R5 cell lines did not test the JLTRG cell line for CCR5 expression, and did not indicate which antibody clones were used on the cell lines that were tested so we had no way to try to repeat their findings[856,857]. When informed of these results, the distributor was unaware of the issue but offered to ship a new aliquot of cells for us to test. Further testing of this cell line may be considered at a future time (see Chapter 4).

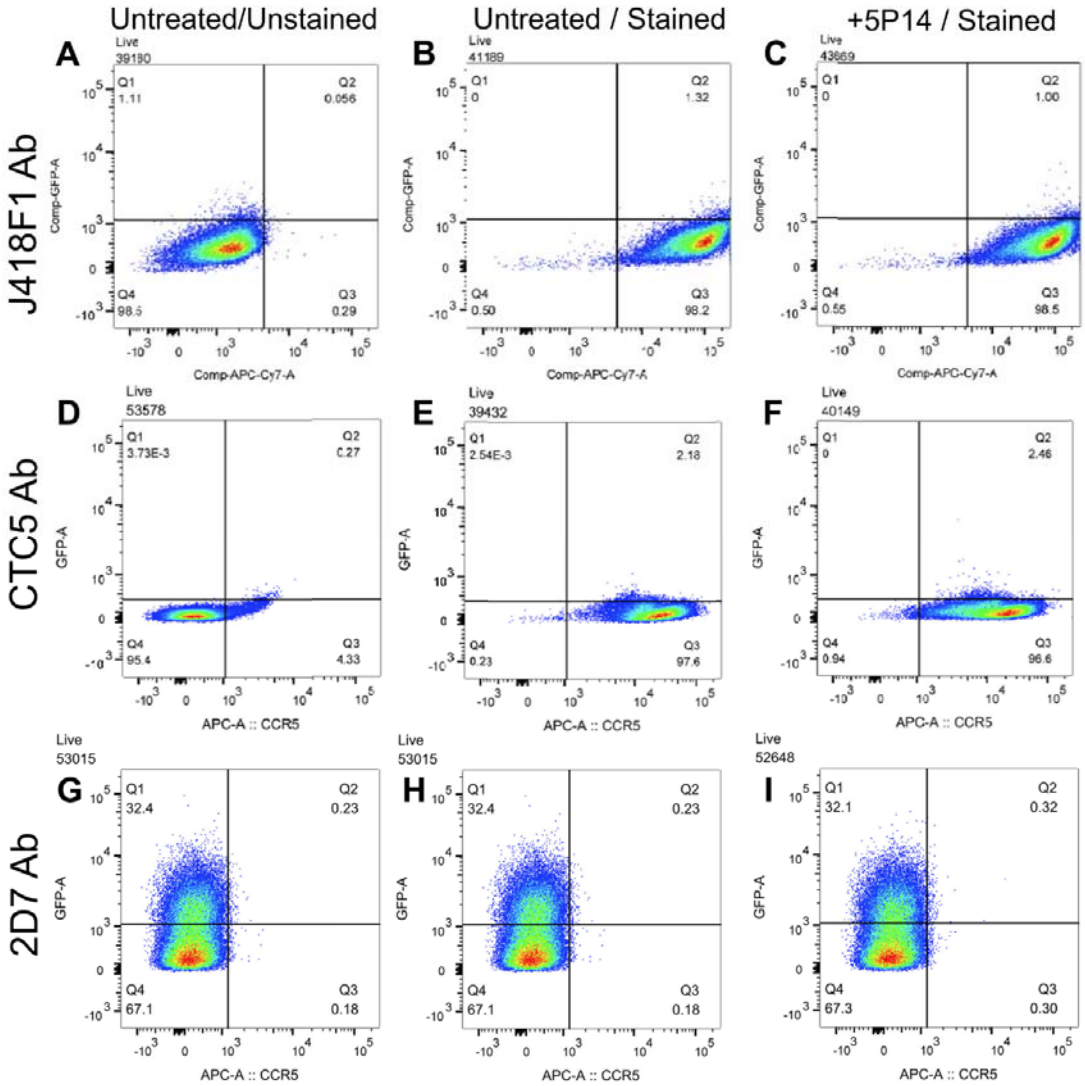


Figure 3.17 Comparison of anti-CCR5 antibodies J418F1, CTC5 and 2D7.
A FACS assay of JLTRG cells was used to compare the amount of CCR5 labeled with the antibodies J418F1, CTC5 and 2D7 in the presence and absence of 5P14 RANTES. The cells used to test binding of J418F1 and CTC5 were surface and intracellularly stained while the cells used to test 2D7 were only surface stained. Gating was determined based on the untreated/unstained cells for each grouping (A, D & E).

Initial testing of MAGI and MAGI +R5 cell lines

The MAGI cell line is an adherent cell line that contains a B-galactosidase reporter. This presented several disadvantages. Firstly, because we needed to measure a surface receptor, FACS would be difficult since trypsinization, the common method to encourage adherent cells to release from the plate, could potentially damage the receptor, the antibody or both. Secondly, the B-galactosidase reporter is measured using a colorimetric assay, which has a rather limited range of measurement which will make determination of significance even more difficult. The B-galactosidase assay also reports on the entire population of cells within a well, not on a single cellular basis, meaning that we could not mix the two populations of cells and be able to clearly show targeting to $CCR5^{+}$ cells. One potential workaround would be to use a fixed cell colorimetric assay in conjunction with fluorescent anti-CCR5 antibodies and measure the results manually using microscopy. The disadvantage to this method is that the level of activation would be a binary measurement, which could be highly problematic since preliminary data has shown that the chimeric protein yields significantly lower levels of activation that the Tat positive control at equal protein concentrations.

Because our previous cell line, JLTRG, expressed CCR5 in defiance of the literature, we decided to confirm the respective presence, and absence, of CCR5 in our MAGI cell lines. Using the J418F1 antibody, the cells were stained and then compared to the TZM-bl cell line, which has been shown to have high levels of CCR5 expression[872–874] and was previously our standard cell line for colorimetric assays. The cells were stained with a 1:100 dilution of the antibody for 25-30 minutes. Cells were then imaged at 40X resolution against white light and light at 650nm (since the APC conjugated to the J418F1 antibody is excited at 650nm and emits near 780nm). Successful microscopy results confirmed the presence of CCR5 on the MAGI+CCR5 cells(Figure 3.18E), and its absence on the MAGI cells(Figure 3.18C). The TZM-bl cells, our positive control, have clearly defined cell membranes when exposed to 650nm wavelength light, indicating high levels of CCR5 expression (Figure 3.18A). When the TZM-bl cells are unstained, the cell outline cannot be clearly visualized, but there is slight background cellular fluorescence at 650nm in the nucleus of each cell (Figure 3.18B). Regardless of whether the MAGI cells were stained or unstained, the cell outline cannot be clearly visualized at 650nm, indicating that this cell line is CCR5- (Figure 3.18 C & D). Stained MAGI +R5 cells have faint cell outlines at 650nm (Figure 3.18E), compared to how their unstained counterparts have only slight background fluorescence in the nuclei of the cells (Figure 3.18F), indicating low levels of CCR5 expression.

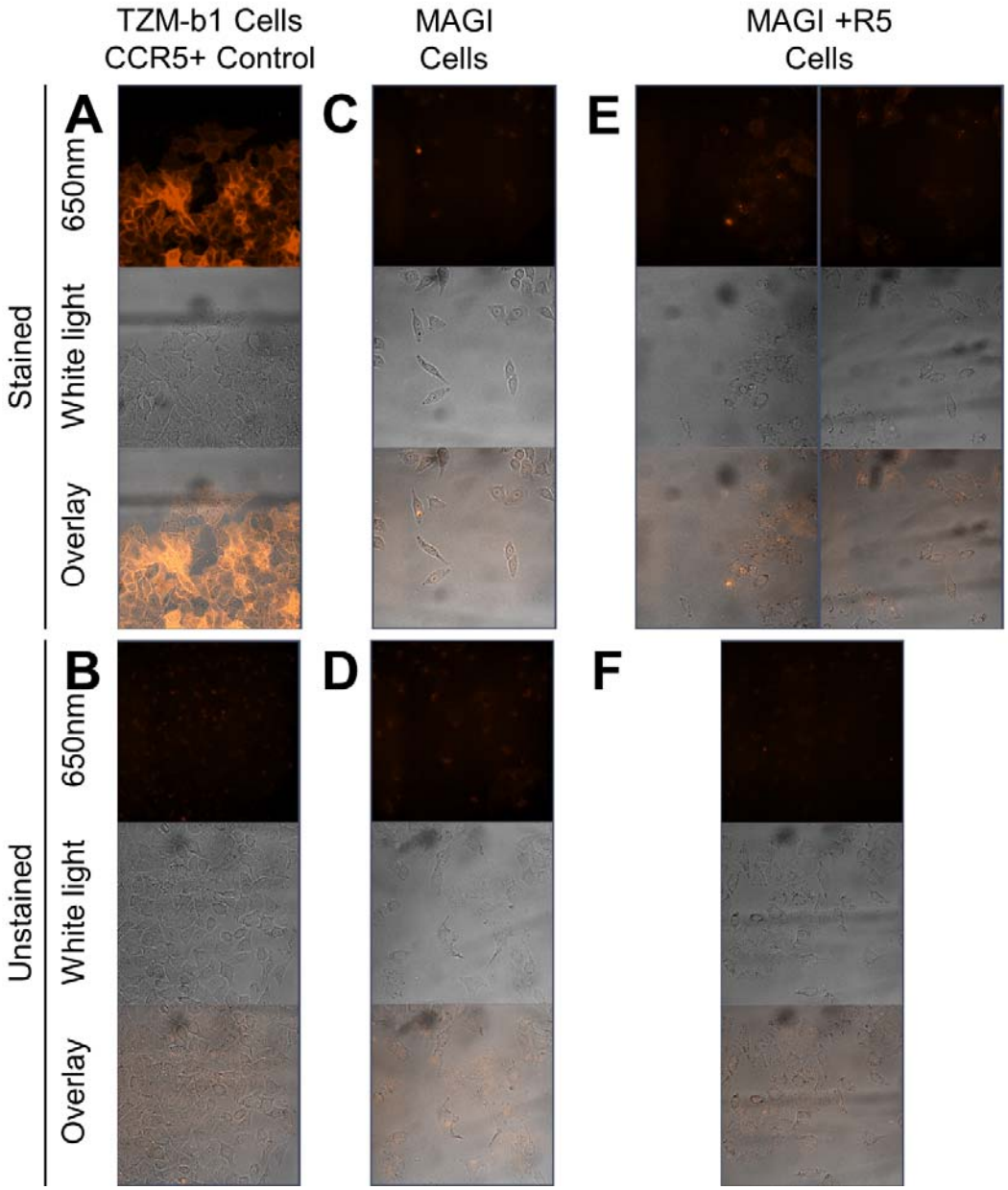


Figure 3.18 Microscopy images of cell lines to show absence of CCR5 in MAGI cells and presence of CCR5 in MAGI+R5 cells.

Images show 40X resolution on a confocal microscope when exposed to light at 650nm and white light. **A, C & F)** are stained with anti-CCR5 Ab J418F1. **A)** shows TZM-bl cells are clearly defined at 650nm, indicating high levels of CCR5 expression. **B)** shows when TZM-bl cells are unstained the cell outline cannot be clearly visualized, but there is slight background cellular fluorescence at 650nm in the nucleus of each cell. **C-D)** shows that when MAGI cells are stained or unstained, the cell outline cannot be clearly visualized at 650nm, indicating that this cell line is CCR5-. **E)** shows in stained MAGI +R5 cells the cell outlines can be faintly seen at 650nm, indicating low levels of CCR5 expression. **F)** shows slight background cellular fluorescence at 650nm in the nucleus of each MAGI +R5 cell.

3.3.7 Testing of J-Lat 8.4 Cell line for presence of CCR5 and activation by Tat

Since the third goal of the project was to confirm our chimera's ability to reactivate latently infected cells, the cell line J-Lat 8.4 was chosen as our latency model. We needed to confirm that this model would suit our purposes by testing its ability to be activated by Tat. Since another one of our Jurkat cell lines, JLTRG, surprised us by expressing CCR5, we wanted to determine whether these J-Lat 8.4 cells expressed CCR5, even though they shouldn't. We used a FACS assay, adding range of Tat concentrations (20uM, 10uM, 5uM and 2.5uM) to $4x10^5$ initial count of J-Lat cells and tested untreated cells with the anti-CCR5 antibodies J418F1, CTC5 and 2D7. The results were gated based on the untreated samples that had been stained only with DAPI (referred to as Untreated), with Q1/5 indicating cells that were CCR5-/GFP+, Q2/6 indicating cells that were CCR5+/GFP+, Q3/7 indicating cells that were CCR5+/GFP- and Q4/8 indicating cells that were CCR5-/GFP-. Results indicated that the J-Lat cells were CCR5- (cells were found in Q4/8), regardless of the antibody used (Figure 3.19 Untreated vs J418F1, CTC5 and 2D7). A concentration of 20uM Tat stimulated GFP activation in only 13% of live cells (cells extended up into Q1/5) compared to the untreated control with only 0.08% of live cells activated (Figure 3.19 20uM Tat). The remaining samples acted in a dose-based response: 10uM Tat activated 5.29% (Figure 3.19 10uM Tat), 5uM Tat activated 2.81% (Figure 3.19 5uM Tat), and 2.5uM Tat barely activated an cells at all, only 0.64% (Figure 3.19 2.5uM Tat). This indicates that the J-Lat 8.4 cells are significantly less responsive to Tat than our JLTRG and JLTRG +R5 cell lines. This is partially to be expected since the LTR-GFP in likely inserted into a heterochromatin region where transcription would be highly difficult, even with Tat there to act as a transcriptional activator.

3.4 Conclusions

The purpose of this project was to create a chimeric protein composed of 5P14-RANTES and HIV Tat, that could selectively target to, and activate CCR5+ cells. Based on a MAGI colorimetric assay, there was no discernable difference in activation between CCR5- and CCR5+ MAGI cells, indicating a lack of targeting. This result may be due to improper or incomplete folding of the 5P14-RANTES, which would negate binding to CCR5. The sample of 20uM 5P14-Linker-Tat did not induce activation noticeably above the negative control of 5P14-RANTES alone. The sample with 20uM Tat-Linker-5P14 protein showed equivalent levels of cellular activation as compared to 10uM Tat. Since the protein was not completely pure, and there was possibly trace amounts of free Tat in the sample, it is impossible to currently determine whether the level of activation induced by the Tat-Linker-5P14 was due to the orientation of the protein, or contaminating free Tat. I recommend that no further testing of function should be attempted until purification of the chimeric protein, in either orientation, is perfected to yield pure, fully folded protein.

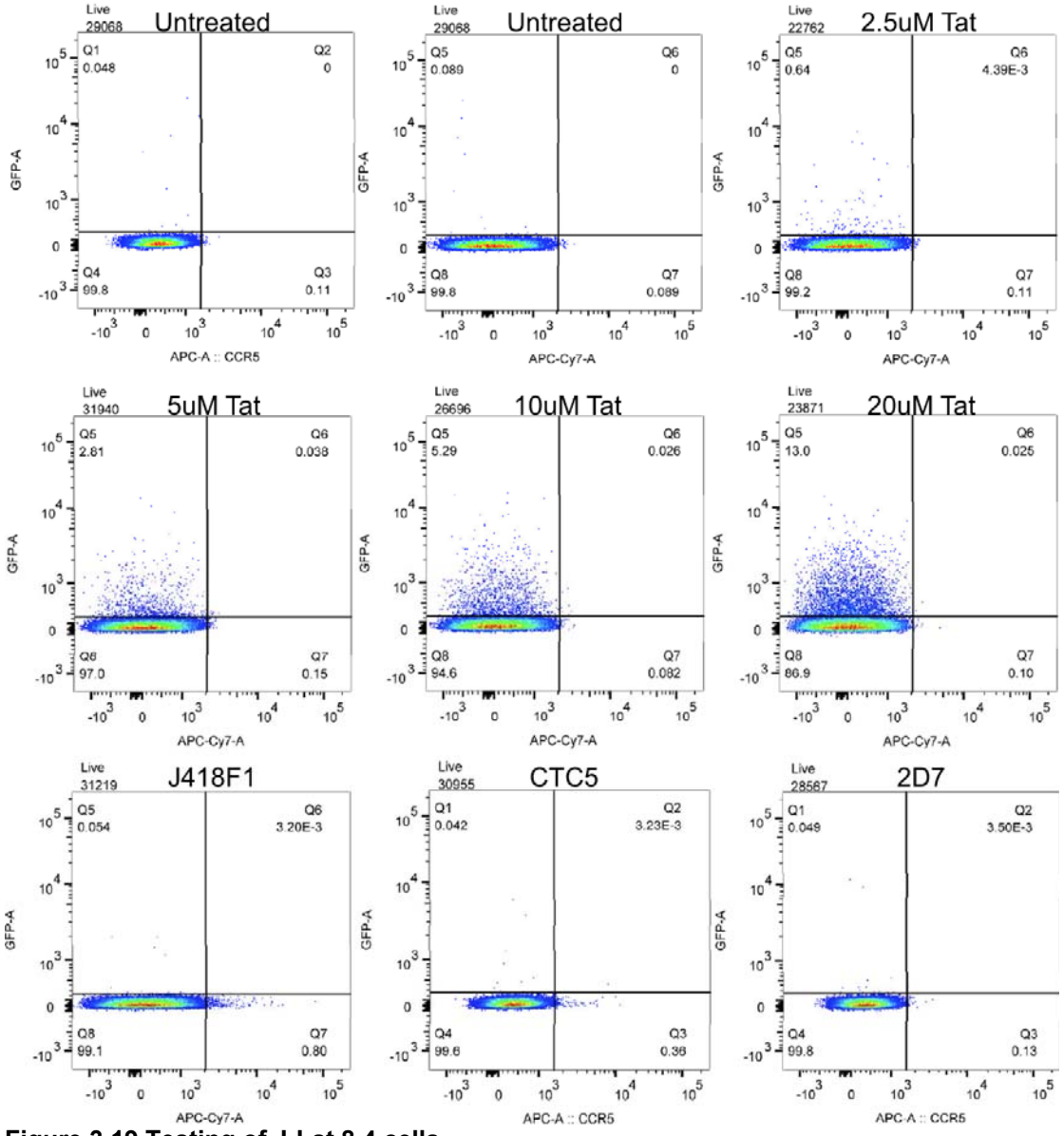


Figure 3.19 Testing of J-Lat 8.4 cells.

J-Lat 8.4 cells were tested for the presence of CCR5 on the surface and the ability of Tat to induce activation of the LTR-GFP located in a transcriptionally silenced region of the DNA. GFP absorbance (GFP-A) is on the y axis, and an anti-CCR5 antibody (with either APC or APC-Cy7 absorbance) is on the x axis. All Tat samples were stained with the J418F1 antibody. Untreated samples were stained with DAPI only. For full gating see Appendix C (pages 181-183).

Chapter 4

Conclusions and Future Directions

The purpose of this research was to develop therapeutic agents to help with the prevention of initial infection and reactivation of latent HIV infection. Our work on Grft clarified the significance of the individual glycan N295 and indicated that glycan density is also generally important for Grft potency. The glycosite N295 proved critical for Grft induced conformational change in B clade HIV, but not C clade. Finally, in synergy tests, we showed that Grft had additive effects when combined with bNAbs PGT121 and PGT126, despite the antibodies' epitope belonging to the same glycan cluster as N295. Unfortunately, our work to characterize the functionality of chimeric protein 5P14-L-Tat was not as successful. Production of pure, fully folded and functional protein remains a challenge, and cellular assays indicated that the chimera does not have selective targeting to CCR5+ cells.

4.1 Future directions of Griffithsin research

While our work has clarified the role of glycan location and density on Grft potency, new questions have come to light. In our work we proposed that the glycan density, which has been shown to affect the level of glycan processing[352,370,556], contributes to the potency of Grft. Previous research has shown that Grft has better binding to nonamannoside, which mimics the structure of minimally processed Man-9 glycans, than manno-pentose, which mimics the structure of the more highly processes Man-5 glycans[478]. There are several avenues which can be explored in relation to this theory. The easiest would be to explore Grft's potency against HIV strain CAP45, Table 2.1, which is surprising considering that this strain lacks glycosylation at N295 (Figure 2.3). My hypothesis is that the presence of an additional, non-conserved glycan at N335 increases the local glycan density in the region, minimizing glycan processing in that glycan cluster, which compensates for the lack of glycosylation at 295. Mutations to delete this glycan (N335), individually and in combination with surrounding glycans, would be an easy first step to elucidating this particular glycans effect. This theory could alternately be tested by adding additional, non-conserved, glycans in the region of cluster 3 to other strains that naturally lack glycosylation at 295 or have had the glycosite removed. A virus independent method could utilize mannose-nano-disk constructs, recently generated by Julia Stauber, to test Grft binding against a variety of mannose densities in NMR studies (unpublished data).

Another glycan that should be removed to test its effect on Grft potency is the glycan at N262. It was previously believed that removal of glycosite N262 resulted in complete loss of viral infectivity[372,373]. Rathore et al recently demonstrated that it was not the loss of the glycan that negated infectivity, but the mutation itself[364]. Considering these findings, an obvious next step in the research of Grft would be to test Grft against strains mutated to lack glycosite N262, since it is the final glycan in cluster 3 (and we have demonstrated that the other three glycans in this cluster are the most important glycans for Grft potency). It would also be advisable to test other evolutionary based mutants against Grft, particularly in viral capture assays where mutations that effect viral folding may unduly influence results. Table 4.1 below shows the mutants that I have already prepared that could be tested and compared to each other. My evolutionary based

mutations were generated in 2014 (far prior to Rathore's publication). They are based on a meta-analysis of 2960 strains, 1708 B clade and 1252 C clade, to determine what the most common mutation found in the absence of a glycosite at each site.

Table 4.1 HIV mutants already prepared to test evolutionary-based removal of glycosites

	Mutants Prepared			
Glycosite	N>Q	S/T>A	Evolutionary based	Mutations by Rathore et al[364]
234	✓	✓	✓ (S > K)	NA
295	✓	✓	✓ (N > T)	N > T
262				N > D
332	✓			N > T
448	✓	✓	✓ (N > K)	N > K
234, 295	✓		✓	
234, 448	✓		✓	
295, 332	✓			
295, 448	✓		✓	
332, 448	✓			
234, 259, 448	✓		✓	
295, 332, 448	✓			

Additional research could also be done to illuminate whether the presence of individual glycans affects Grft's ability to prevent HIV infection at various stages in the infection process. Grft can bind to HIV simultaneously with CD4, but not with CCR5 binding site specific Ab 17b, indicating that Grft prevents inhibition somewhere between CD4 binding and CCR5 binding, but it is unclear whether Grft simply occludes the CCR5 binding site or also prevents the conformational change in HIV to expose the CCR5 binding site or the subsequent conformational change to expose the gp41 helical bundle. This could be tested using CHO-K1 cells, which contain aquaporin (an enzyme that luminesces upon Ca^{2+} release induced by ligand binding to, and subsequent activation of CCR5)[875,876].

We demonstrated that Grft mediates conformational change (exposing the epitope of Ab b12) in B clade strains to a much greater degree than C clade strains, which seems to be partially due to the presence of glycosite N295 in B clade strains. This work could be expanded with additional glycosylation deletion mutants in viral capture ELISAs to test Grft's ability to induce this conformational change in the absence of glycosites other than N295, particularly other cluster 3 glycans and glycan N386. I hypothesize that other glycans in cluster 3 may be involved in Grft's ability to mediate conformational change. In companion to this work, deletion mutants should be tested for Grft's ability to compete with 17b, another antibody that binds to a CD4 induced epitope near the CCR5 binding site. Other groups have shown that Grft prevents binding of 17b Ab[480,877], but it has not been extensively confirmed, and never against any HIV-1 deglycosylation mutants. Based on our work with Ab b12, there may be differences in B vs C clades.

All of this information could more clearly indicate exactly how Grft prevents initial infection of HIV. Determination of which strains Grft will be most effective against facilitates tailoring of combinational therapies and prevention strategies. The amount of Grft necessary for complete coverage against the HIV strains predominate in the area the patient lives could be easily quantified, minimizing costs. Griffithsin is already well on its

way to becoming a microbicide available for public use; Dr. Palmer and Dr. Brand are currently in phase 1 testing of Grft in clinical trials with Grft delivered through an investigational enema[493,497,878]. Other delivery methods of Grft are also being tested, including carrageen gels[446–448,746,879,880], electrospun silk fibers[503–505], and of course our own silk disks and films[496,506]. Grft tandemers are also being developed, with trimeric and tetrameric forms of Grft showing potency in the absence of glycan N295 for cell free and cell-to-cell transmission[474,881,882].

4.2 Future directions of targeted reactivation of the latent HIV reservoir

Any future work with this project will need to begin with a thorough evaluation and improvement of all protein purification procedures to increase yield and purity. This may be done by troubleshooting the purification protocol of the full-length chimera or the sortase reaction and subsequent purification. It should be noted that new Tat expression vectors have become available through the NIH AIDS Reagent program that may have higher expression and simplified purification schemes, which would simplify production of Tat for any future work. It would also be beneficial to pursue the acquisition of a pair of Jurkat cell lines that are truly CCR5- and CCR5+, since that would significantly ease testing of any future chimeric proteins targeting to CCR5. The discovery that the JLRTG cell line was CCR5+ is highly unexpected and should be investigated further since Jurkats are generally presumed to be CCR5-[322,883–885]. It has been indicated that epigenetic factors control CCR5 expression in Jurkat cells[886,887], so it may be prudent for all labs to independently test their Jurkat cells with a panel of anti-CCR5 antibodies to confirm that lab conditions are not stimulating expression of CCR5. The anti-CCR5 antibody 2D7 has been shown to be conformationally specific[869,888,889], but J418F1 has been used extensively in other work and shows high specificity for CCR5 without being conformationally specific[890–896].

It may be beneficial to look more into the specific interactions between 5P14 and CCR5 to confirm and enhance specificity for targeting of not only HIV therapies but other epigenetic conditions as well. Once the ability of RANTES variants to selectively target to CCR5+ cells has been more thoroughly characterized, work could be done to include a protein transduction domain that would allow escape from the endosome after internalization of the receptor. This would allow the protein construct into the cytosol where the attached shock or kill agent could achieve its desired function[835,897–899]. A nuclear localization signal could also be attached, as necessary, for drug function.

VII. References

1. Growth and Maintenance of the 293FT Cell Line. (2010).
2. 293FT Cell Line - Thermo Fisher Scientific. Available at: https://www.thermofisher.com/order/catalog/product/R70007. (Accessed: 16th May 2018)
3. Reagent Datasheet Detail: Catalog 11051 - pSG3$^{\Delta env}$ - NIH AIDS Reagent Program. (2018). Available at: https://www.aidsreagent.org/reagentdetail.cfm?t=molecular_clones&id=860. (Accessed: 16th May 2018)
4. Schwetz, T. A. & Fauci, A. S. The Extended Impact of Human Immunodeficiency Virus/AIDS Research. *J. Infect. Dis.* (2018). doi:10.1093/infdis/jiy441
5. amfAR. *The Broad Benefits of AIDS Research*. (2019).
6. Moir, S. & Fauci, A. S. B-cell exhaustion in HIV infection: The role of immune activation. *Current Opinion in HIV and AIDS* **9**, 472–477 (2014).
7. Hoffmann, M. *et al.* Exhaustion of Activated CD8 T Cells Predicts Disease Progression in Primary HIV-1 Infection. *PLOS Pathog.* **12**, e1005661 (2016).
8. Deeks, S. G., Tracy, R. & Douek, D. C. Systemic Effects of Inflammation on Health during Chronic HIV Infection. *Immunity* **39**, 633–645 (2013).
9. Hope, T. J. Inflammation weakens HIV prevention. *Nat. Med.* **24**, 384–385 (2018).
10. Triant, V. A. Cardiovascular disease and HIV infection. *Curr. HIV/AIDS Rep.* **10**, 199–206 (2013).
11. Nou, E., Lo, J. & Grinspoon, S. K. Inflammation, immune activation, and cardiovascular disease in HIV. *AIDS* **30**, 1495–1509 (2016).
12. Boccara, F. *et al.* HIV and coronary heart disease: Time for a better understanding. *Journal of the American College of Cardiology* **61**, 511–523 (2013).
13. Anderson, K. W. *et al.* Mapping of the allosteric site in cholesterol hydroxylase CYP46A1 for efavirenz, a drug that stimulates enzyme activity. *J. Biol. Chem.* **291**, 11876–11886 (2016).
14. Mast, N., Verwilst, P., Wilkey, C. J., Guengerich, F. P. & Pikuleva, I. A. In Vitro Activation of Cytochrome P450 46A1 (CYP46A1) by Efavirenz-Related Compounds. *J. Med. Chem.* (2019). doi:10.1021/acs.jmedchem.9b01383
15. Serrano-Villar, S., Zhou, Y., Rodgers, A. J. & Moreno, S. Different impact of raltegravir versus efavirenz on CD4/CD8 ratio recovery in HIV-infected patients. *J. Antimicrob. Chemother.* dkw375 (2016). doi:10.1093/jac/dkw375
16. Durand, S. & Cimarelli, A. The inside out of lentiviral vectors. *Viruses* **3**, 132–59 (2011).
17. Berkhout, B. HIV, leukemia, and new horizons in molecular therapy. *Journal of the Formosan Medical Association* **112**, 441–444 (2013).
18. Bosticardo, M., Ferrua, F., Cavazzana, M. & Aiuti, A. Gene Therapy for Wiskott-Aldrich Syndrome. *Curr. Gene Ther.* **14**, 413–421 (2014).
19. Kwo, P. Y. & Vinayek, R. The therapeutic approaches for hepatitis C virus: Protease inhibitors and polymerase inhibitors. *Gut and Liver* **5**, 406–417 (2011).
20. Meuleman, P. *et al.* Griffithsin has antiviral activity against hepatitis C virus. *Antimicrob. Agents Chemother.* **55**, 5159–5167 (2011).
21. Jardim, A. C. G., Shimizu, J. F., Rahal, P. & Harris, M. Plant-derived antivirals against hepatitis c virus infection. *Virol. J.* **15**, 34 (2018).
22. Wyles, D. L. Antiviral resistance and the future landscape of hepatitis C virus

infection therapy. *Journal of Infectious Diseases* **207**, (2013).
23. Eckert, D. M. & Kim, P. S. Mechanisms of Viral Membrane Fusion and Its Inhibition. *Annu. Rev. Biochem.* **70**, 777–810 (2001).
24. Julien, J. P., Lee, P. S. & Wilson, I. A. Structural insights into key sites of vulnerability on HIV-1 Env and influenza HA. *Immunol. Rev.* **250**, 180–198 (2012).
25. Raymond, D. D. *et al.* Conserved epitope on influenza-virus hemagglutinin head defined by a vaccine-induced antibody. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 168–173 (2018).
26. Routy, J.-P. *et al.* Highlights from the 2016 International Symposium on HIV & Emerging Infectious Diseases (ISHEID): 25-27 May, Marseille, France. *J. virus Erad.* **2**, 187–92 (2016).
27. Jeanene Bengoa, A. *et al.* Zika Virus and HIV/AIDS. in *Global Virology II - HIV and NeuroAIDS* 731–750 (Springer New York, 2017). doi:10.1007/978-1-4939-7290-6_30
28. Kumar, R. & King, S. Advances in Developing Therapies to Combat Zika Virus : Current Knowledge and Future Perspectives Advances in Developing Therapies to Combat Zika Virus : Current Knowledge and Future Perspectives. *Front. Microbiol.* **8**, 1–19 (2017).
29. Sapparapu, G. *et al.* Neutralizing human antibodies prevent Zika virus replication and fetal disease in mice. *Nature* **540**, 443–447 (2016).
30. Röhn, T. A. & Bachmann, M. F. Vaccines against non-communicable diseases. *Current Opinion in Immunology* **22**, 391–396 (2010).
31. Tian, D. *et al.* Structural basis of respiratory syncytial virus subtype-dependent neutralization by an antibody targeting the fusion glycoprotein. *Nat. Commun.* **8**, (2017).
32. Lucas, S. & Nelson, A. M. HIV and the spectrum of human disease. *J. Pathol.* **235**, 229–241 (2015).
33. van Lier, J. E. New Therapeutic Targets for Brain Function and Disease. *J. Med. Chem.* acs.jmedchem.9b01947 (2019). doi:10.1021/acs.jmedchem.9b01947
34. Gottlieb, M. S., Schanker, H. M., Fan, P. T., Saxon, A. & Weisman, J. D. *Pneumocystis Pneumonia --- Los Angeles*. *Morbitity and Mortality Weekly Report* (1981).
35. Gottlieb, M. S. *et al.* Pneumocystis carinii pneumonia and mucosal candidiasis in previously healthy homosexual men: evidence of a new acquired cellular immunodeficiency. *N. Engl. J. Med.* **305**, 1425–31 (1981).
36. Hymes, K. *et al.* KAPOSI'S SARCOMA IN HOMOSEXUAL MEN—A REPORT OF EIGHT CASES. *Lancet* **318**, 598–600 (1981).
37. Centers for Disease Control (CDC). Kaposi's sarcoma and Pneumocystis pneumonia among homosexual men--New York City and California. *MMWR. Morb. Mortal. Wkly. Rep.* **30**, 305–8 (1981).
38. What is Kaposi sarcoma? (2016). Available at: http://www.cancer.org/cancer/kaposisarcoma/detailedguide/kaposi-sarcoma-what-is-kaposi-sarcoma.
39. Templeton, A. C. *Cancer of the skin : biology, diagnosis, management [WorldCat.org]*. (WB Saunders, 1976).
40. Centers for Disease Control (CDC). *Current Trends Update on Acquired Immune Deficiency Syndrome (AIDS) --United States*. (1982).
41. Center for Disease Control (CDC). *Unexplained Immunodeficiency and Opportunistic Infections in Infants -- New York, New Jersey, California*. (1982).

42. Center for Disease Control (CDC). *Epidemiologic Notes and Reports Possible Transfusion-Associated Acquired Immune Deficiency Syndrome (AIDS) -- California*. *MMWR* (1982).
43. Center for Disease Control (CDC). *Update on Acquired Immune Deficiency Syndrome (AIDS) among Patients with Hemophilia A*. (1982).
44. Ziegler, J. B., Johnson, R. O. & Cooper, D. A. POSTNATAL TRANSMISSION OF AIDS-ASSOCIATED RETROVIRUS FROM MOTHER TO INFANT. *Lancet* (1985).
45. Dowdle, W. R. The epidemiology of AIDS. *Public Health Rep.* **98**, 308–12 (1983).
46. Haris, C. *et al.* Immunodeficiency in Female Sexual Partners of Men with the Aquired Immunodeficiency Syndrome. *N. Engl. J. Med.* **308**, (1983).
47. Weiss, S. H. *et al.* HTLV-III Infection Among Health Care Workers. *JAMA* **254**, 2089 (1985).
48. Letvin, N. L. *et al.* EXPERIMENTAL TRANSMISSION OF MACAQUE AIDS BY MEANS OF INOCULATION OF MACAQUE LYMPHOMA TISSUE. *Ann Clin Res* **14**, 15–26 (1982).
49. Sande, M. A. The Transmission of AIDS. *New Engl. J. Med. Downloaded from nejm.org MERCED August* **314**, (1986).
50. Barré-Sinoussi, F. *et al.* Isolation of a T-Lymphotropic Retrovirus from a Patient at Risk for Acquired Immune Deficiency Syndrome (AIDS). *Mont. Source Sci. New Ser.* **220**, 868–871 (1983).
51. Gallo, R. C. *et al.* Isolation of Human T-Cell Leukemia Virus in Acquired Immune Deficiency Syndrome (AIDS). *Source Sci. New Ser.* **220**, 865–867 (1983).
52. Gallo, R. C. The discovery of the first human retrovirus: HTLV-1 and HTLV-2. *Retrovirology* **2**, 17 (2005).
53. Sarngadharan, M. G., Popovic, M., Bruch, L., Schüpbach, J. & Gallo, R. C. Antibodies Reactive with Human T-Lymphotropic Retroviruses (HTLV-III) in the Serum of Patients with AIDS. *Source Sci. New Ser.* **224**, 506–508 (1984).
54. Popovic, M., Sarngadharan, M. G., Read, E. & Gallo, R. C. Detection, Isolation, and Continuous Production of Cytopathic Retroviruses (HTLV-III) from Patients with AIDS and Pre-AIDS. *Science* **224**, 497–500 (1984).
55. Gallo, R. C. *et al.* Frequent Detection and Isolation of Cytopathic Retroviruses (HTLV-III) from Patients with AIDS and at Risk for AIDS. *Science* **224**, 500–503 (1984).
56. Klatzmann, D. *et al.* Selective Tropism of Lymphadenopathy Associated Virus (LAV) for Helper-Inducer T. *Source Sci. New Ser.* **225**, 59–63 (1984).
57. Levy, J. A. *et al.* Isolation of Lymphocytopathic Retroviruses from San Francisco Patients with AIDS. *Source Sci. New Ser.* **225**, 840–842 (1984).
58. Wain-Hobson, S., Sonigo, P., Danos, O., Cole, S. & Alizon, M. Nucleotide Sequence of the AIDS Virus, LAV. *Cell* **40**, 9–17 (1985).
59. Alizon, M. *et al.* Molecular cloning of lymphadenopathy-associated virus. *Nature* **312**, 757–60 (1984).
60. Marx, J. L. Strong New Candidate for AIDS Agent. *Source Sci. New Ser.* **224**, 475–477 (1984).
61. Marx, J. L. AIDS Virus Has New Name--Perhaps. *Source Sci. New Ser.* **2322369**, 699–700 (1986).
62. Nájera, R., Herrera, M. I. & de Andrés, R. Human immunodeficiency virus and related retroviruses. *West. J. Med.* **147**, 702–8 (1987).
63. Weiss, R. A. Retrovirus classification and cell interactions. *J. Antimicrob. Chemother.* **37**, 1–11 (1996).

64. Yasunaga, T., Sagata, N. & Ikawa, Y. Protease gene structure and env gene variability of the AIDS virus. *FEBS Lett.* **199**, 145–150 (1986).
65. Stowring, L., Haase, A. T. & Charman2, H. P. Serological Definition of the Lentivirus Group of Retroviruses. *J. Virol.* **29**, 523–528 (1979).
66. Sigurdsson, B., Grfmsson, H. & Palsson, P. A. MAEDI, A CHRONIC, PROGRESSIVE INFECTION OF SHEEP'S LUNGS. *J. Infect. Dis.* **90**, 233–241 (1952).
67. Sigurdsson, B. Transmission Experiments with Maedi. *J. Infect. Dis.* **93**, (1953).
68. Lentivirus | definition of lentivirus by Medical dictionary.
69. Curran, J. W. & Jaffe, H. W. AIDS: the Early Years and CDC's Response. *Morbitity Mortal. Wkly. Rep. Suppl.* **60**, 64–69 (2011).
70. Nobel Media AB. Françoise Barré-Sinoussi - Facts. *Nobelprize.org* (2014). Available at: https://www.nobelprize.org/nobel_prizes/medicine/laureates/2008/barre-sinoussi-facts.html.
71. Robertson, D. L. *et al.* HIV-1 Nomenclature Proposal. in *Human Retroviruses and AIDS 1999* (eds. Kuiken C.L. et al.) 492–505 (Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, 1999).
72. De Leys, R. *et al.* Isolation and Partial Characterization of an Unusual Human Immunodeficiency Retrovirus from Two Persons of West-Central African Origin. *J. Virol.* 1207–1216 (1990).
73. Gurtler, L. G. *et al.* A New Subtype of Human Immunodeficiency Virus Type 1 (MVP-5180) from Cameroon. *J. Virol.* **68**, 1581–1585 (1994).
74. Charneau, P. *et al.* Isolation and Envelope Sequence of a Highly Divergent HIV-1 Isolate: Definition of a New HIV-1 Group. *Virology* **205**, 247–253 (1994).
75. Vanden Haesevelde, M. *et al.* Genomic Cloning and Complete Sequence Analysis of a Highly Divergent African Human Immunodeficiency Virus Isolate. *J. Virol.* 1586–1596 (1994).
76. Simon, F. *et al.* Identification of a new human immunodeficiency virus type 1 distinct from group M and group O. *Nat. Med.* **4**, 1032–1037 (1998).
77. Plantier, J.-C. *et al.* A new human immunodeficiency virus derived from gorillas. *Nat. Publ. Gr.* (2009). doi:10.1038/nm.2016
78. Myers, G., Macinnes, K. & Korber, B. The Emergence of Simian/Human Immunodeficiency Viruses. *AIDS Res. Hum. Retroviruses* **8**, (1992).
79. Menndez-Arias, L. Molecular basis of human immunodeficiency virus type 1 drug resistance: Overview and recent developments. *Antiviral Res.* **98**, 93–120 (2013).
80. Taylor, B. S., Sobieszczyk, M. E., Mccutchan, F. E. & Hammer, S. M. The Challenge of HIV-1 Subtype Diversity Origin of HIV a nd Mechanisms of HIV Diversity. **15**, (2008).
81. Jabara, C. B., Jones, C. D., Roach, J., Anderson, J. A. & Swanstrom, R. Accurate sampling and deep sequencing of the HIV-1 protease gene using a Primer ID. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 20166–71 (2011).
82. Zhou, T. *et al.* Multidonor Analysis Reveals Structural Elements, Genetic Determinants, and Maturation Pathway for HIV-1 Neutralization by VRC01-Class Antibodies. *Immunity* **39**, 245–258 (2013).
83. Li, M. *et al.* Human immunodeficiency virus type 1 env clones from acute and early subtype B infections for standardized assessments of vaccine-elicited neutralizing antibodies. *J. Virol.* **79**, 10108–25 (2005).
84. Menéndez-Arias, L. Mutation Rates and Intrinsic Fidelity of Retroviral Reverse

Transcriptases. *Viruses* **1**, 1137–1165 (2009).
85. Sosin, D. V. & Tchurikov, N. A. Molecular mechanisms of HIV-1 genetic diversity. *Mol. Biol.* **51**, 483–495 (2017).
86. Santoro, M. M. & Perno, C. F. HIV-1 Genetic Variability and Clinical Implications. *ISRN Microbiol.* **2013**, 1–20 (2013).
87. Roberts, J. D., Bebenek, K. & Kunkel, T. A. The accuracy of reverse transcriptase from HIV-1. *Science* **242**, 1171–3 (1988).
88. Preston, B. D., Poiesz, B. J. & Loeb, L. A. Fidelity of HIV-1 reverse transcriptase. *Science* **242**, 1168–71 (1988).
89. Louwagie, J. *et al.* Genetic diversity of the envelope glycoprotein from human immunodeficiency virus type 1 isolates of African origin. *J. Virol.* **69**, 263–71 (1995).
90. Ashokkumar, M., Aralaguppe, S. G., Tripathy, S. P., Hanna, L. E. & Neogi, U. Unique Phenotypic Characteristics of Recently Transmitted HIV-1 Subtype C Envelope Glycoprotein gp120: Use of CXCR6 Coreceptor by Transmitted Founder Viruses. *J. Virol.* **92**, JVI.00063-18 (2018).
91. Nedellec, R. *et al.* High Sequence Diversity and Rapid Virus Turnover Contribute to Higher Rates of Coreceptor Switching in Treatment-Experienced Subjects with HIV-1 Viremia. *AIDS Res. Hum. Retroviruses* AID.2016.0153 (2016). doi:10.1089/AID.2016.0153
92. Modrow, S. *et al.* Computer-assisted analysis of envelope protein sequences of seven human immunodeficiency virus isolates: prediction of antigenic epitopes in conserved and variable regions. *J. Virol.* **61**, 570–578 (1987).
93. Araújo, L. A. L. & Almeida, S. E. M. HIV-1 diversity in the envelope glycoproteins: Implications for viral entry inhibition. *Viruses* **5**, 595–604 (2013).
94. Nkengasong, J. N. *et al.* Antigenic evidence of the presence of the aberrant HIV-1ant70 virus in Cameroon and Gabon. *AIDS* **7**, 1536–8 (1993).
95. Loussert-Ajaka, I. *et al.* Variability of human immunodeficiency virus type 1 group O strains isolated from Cameroonian patients living in France. *J. Virol.* **69**, 5640–9 (1995).
96. Mourez, T., Simon, F. & Plantier, J.-C. Non-M variants of human immunodeficiency virus type 1. *Clin. Microbiol. Rev.* **26**, 448–61 (2013).
97. Peeters, M. *et al.* Geographical distribution of HIV-1 group O viruses in Africa. *Aids* **11**, 493–498 (1997).
98. Mauclère, P. *et al.* Serological and virological characterization of HIV-1 group O infection in Cameroon. *AIDS* **11**, 445–453 (1997).
99. Vallari, A. *et al.* Four New HIV-1 Group N Isolates from Cameroon: Prevalence Continues to Be Low. *AIDS Res. Hum. Retroviruses* **26**, (2010).
100. Ayouba, A. *et al.* HIV-1 group N among HIV-1-seropositive individuals in Cameroon. *AIDS* **14**, 2623–5 (2000).
101. Takehisa, J. *et al.* Origin and Biology of Simian Immunodeficiency Virus in Wild-Living Western Gorillas. *J. Virol.* **83**, 1635–1648 (2009).
102. Vallari, A. *et al.* Confirmation of putative HIV-1 group P in Cameroon. *J. Virol.* **85**, 1403–7 (2011).
103. Faria, N. R. *et al.* HIV epidemiology. The early spread and epidemic ignition of HIV-1 in human populations. *Science* **346**, 56–61 (2014).
104. Lynch, R. M., Shen, T., Gnanakaran, S. & Derdeyn, C. a. Appreciating HIV type 1 diversity: subtype differences in Env. *AIDS Res. Hum. Retroviruses* **25**, 237–248 (2009).

105. Louwagie, J. *et al.* Phylogenetic analysis of gag genes from 70 international HIV-1 isolates provides evidence for multiple genotypes. *AIDS (London, England)* **7**, 769–780 (1993).
106. Janssens, W. *et al.* Genetic and Phylogenetic Analysis of env Subtypes G and H in Central Africa. *ADDS Res. Hum. RETROVIRUSES* **10**, (1994).
107. Kostrikis, L. G. *et al.* Genetic Analysis of Human Immunodeficiency Virus Type 1 Strains from Patients in Cyprus: Identification of a New Subtype Designated Subtype I. *J. Virol.* **69**, 6122–6130 (1995).
108. LEITNER, T. *et al.* Yet Another Subtype of HIV Type 1? *AIDS Res. Hum. Retroviruses* **11**, 995–997 (1995).
109. Hemelaar, J., Gouws, E., Ghys, P. D. & Osmanov, S. Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. *AIDS* **20**, W13–W23 (2006).
110. McCutchan, F. E. Global epidemiology of HIV. *J. Med. Virol.* **78**, S7–S12 (2006).
111. Hemelaar, J., Gouws, E., Ghys, P. D., Osmanov, S. & WHO-UNAIDS Network for HIV Isolation and Characterisation. Global trends in molecular epidemiology of HIV-1 during 2000-2007. *AIDS* **25**, 679–89 (2011).
112. Buonaguro, L., Tornesello, M. L. & Buonaguro, F. M. Human Immunodeficiency Virus Type 1 Subtype Distribution in the Worldwide Epidemic: Pathogenetic and Therapeutic Implications. *J. Virol.* **81**, 10209–10219 (2007).
113. Borm, K. *et al.* Frequency and Env determinants of HIV-1 subtype C strains from antiretroviral therapy-naive subjects that display incomplete inhibition by maraviroc. *Retrovirology* **13**, 74 (2016).
114. Beyrer, C. & Abdool Karim, Q. The changing epidemiology of HIV in 2013. *Curr Opin HIV AIDS* **8**, 306–10 (2013).
115. Maartens, G., Celum, C. & Lewin, S. R. HIV infection: Epidemiology, pathogenesis, treatment, and prevention. *Lancet* **384**, 258–271 (2014).
116. Letvin, N. L. *et al.* Acquired immunodeficiency syndrome in a colony of macaque monkeys. *Proc. Natl. Acad. Sci. U. S. A.* **80**, 2718–22 (1983).
117. King, N. W., Hunt, R. D. & Letvin, N. L. Histopathologic changes in macaques with an acquired immunodeficiency syndrome (AIDS). *Am. J. Pathol.* **113**, 382–8 (1983).
118. Daniel, M. D. *et al.* Isolation of T-Cell Tropic HTLV-III-Like Retrovirus from Macaques. *Source Sci. New Ser.* **228**, 1201–1204 (1985).
119. Barin, F. *et al.* SEROLOGICAL EVIDENCE FOR VIRUS RELATED TO SIMIAN T-LYMPHOTROPIC RETROVIRUS III IN RESIDENTS OF WEST AFRICA. *Lancet* **326**, 1387–1389 (1985).
120. Clavel, F. *et al.* Isolation of a New Human Retrovirus from West African Patients with AIDS. *Source Sci. New Ser.* **233**, 343–346 (1986).
121. Hirsch, V. M., Olmsted, R. A., Murphey-Corb, M., Purcell, R. H. & Johnson, P. R. An African primate lentivirus (SIVsm) closely related to HIV-2. *Nature* **339**, 389–392 (1989).
122. Gao, F. *et al.* Human infection by genetically diverse SIVSM-related HIV-2 in West Africa. *Nature* **358**, 495–499 (1992).
123. 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 t. *J. Virol.* **70**, 3617–27 (1996).
124. Peeters, M. *et al.* Isolation and partial characterization of an HIV-related virs

occuring naturally in chimpanzees in Gabon. *Aids* **3**, 625–630 (1989).
125. Huet, T., Cheynier, R., Meyerhans, A., Roelants, G. & Wain-Hobson, S. Genetic organization of a chimpanzee lentivirus related to HIV-1. *Nature* **345**, 356–359 (1990).
126. Gao, F. *et al.* Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes. *Nature* **397**, 436–441 (1999).
127. Hahn, B. H. *et al.* AIDS as a zoonosis: scientific and public health implications. *Science* **287**, 607–14 (2000).
128. Ho, D. D. *et al.* An African HIV-1 sequence from 1959 and implications for the origin of the epidemic. *Nature* **391**, 594–597 (1998).
129. Worobey, M. *et al.* Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. *Nature* **455**, 661–4 (2008).
130. Mann, J. M. AIDS: A worldwide pandemic. in *Current topics in AIDS* (ed. Gottlieb, M. S. et al.) **2**, (John Wiley & Sons, 1989).
131. Korber, B. *et al.* Timing the ancestor of the HIV-1 pandemic strains. *Science* **288**, 1789–96 (2000).
132. Salemi, M. *et al.* Dating the common ancestor of SIVcpz and HIV-1 group M and the origin of HIV-1 subtypes by using a new method to uncover clock-like molecular evolution. *FASEB J.* **15**, 276–278 (2000).
133. Keele, B. F. *et al.* Chimpanzee reservoirs of pandemic and nonpandemic HIV-1. *Science* **313**, 523–6 (2006).
134. Heuverswyn, F. Van *et al.* Genetic diversity and phylogeographic clustering of SIVcpzPtt in wild chimpanzees in Cameroon. *Virology* **368**, 155–171 (2007).
135. Van Heuverswyn, F. *et al.* Human immunodeficiency viruses: SIV infection in wild gorillas. *Nature* **444**, 164–164 (2006).
136. Unaids. *GLOBAL HIV STATISTICS*. (2019).
137. Kariuki, S. M., Selhorst, P., Ariën, K. K. & Dorfman, J. R. The HIV-1 transmission bottleneck. *Retrovirology* **14**, 22 (2017).
138. Joseph, S. B., Swanstrom, R., Kashuba, A. D. M. & Cohen, M. S. Bottlenecks in HIV-1 transmission: Insights from the study of founder viruses. *Nature Reviews Microbiology* **13**, 414–425 (2015).
139. UNAIDS. *MILES TO GO CLOSING GAPS BREAKING BARRIERS RIGHTING INJUSTICES*. (2018).
140. Unaids. *UNAIDS data 2019*.
141. Patel, P. *et al.* Estimating per-act HIV transmission risk: A systematic review. *AIDS* **28**, 1509–1519 (2014).
142. Statistics Overview | Statistics Center | HIV/AIDS | CDC. Available at: https://www.cdc.gov/hiv/statistics/overview/index.html. (Accessed: 21st January 2020)
143. U.S. Statistics | HIV.gov. Available at: https://www.hiv.gov/hiv-basics/overview/data-and-trends/statistics. (Accessed: 21st January 2020)
144. Gallo, R. C. History of the discoveries of the first human retroviruses: HTLV-1 and HTLV-2. *Oncogene* **24**, 5926–5930 (2005).
145. Popovic, M. *et al.* Isolation and Transmission of Human Retrovirus (Human T-Cell Leukemia Virus). *Source Sci. New Ser.* 856–859 (1983).
146. Popovic, M. T4 positive human neoplastic cell lines susceptible to and permissive for HTLV-III. *Lancet (London, England)* **2**, (1984).
147. Gallo, R. C., de-Thé, G. B. & Ito, Y. Kyoto Workshop on some specific recent advances in human tumor virology. *Cancer Res.* **41**, 4738–9 (1981).

148. Kalyanaraman, V. S. *et al.* A New Subtype of Human T-Cell Leukemia Virus (HTLV-II) Associated with a T-Cell Variant of Hairy Cell Leukemia. *Source Sci. New Ser.* **218**, 571–573 (1982).
149. Miyoshi, I. *et al.* Type C virus particles in a cord T-cell line derived by co-cultivating normal human cord leukocytes and human leukaemic T cells. *Nature* **294**, 770–771 (1981).
150. Gelmann, E. P. *et al.* Proviral DNA of a Retrovirus, Human T-Cell Leukemia Virus, in Two Patients with AIDS. *Source Sci. New Ser.* **220**, 862–865 (1983).
151. FAUCI, A. S. *et al.* Acquired Immunodeficiency Syndrome: Epidemiologic, Clinical, Immunologic, and Therapeutic Considerations. *Ann. Intern. Med.* **100**, 92 (1984).
152. Klatzmann, D. *et al.* T-lymphocyte T4 molecule behaves as the receptor for human retrovirus LAV. *Nature* **312**, 767–768 (1984).
153. Dalgleish, A. G. *et al.* The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. *Nature* **312**, 763–7 (1984).
154. McDougal, J. *et al.* Immunoassay for the Detection and Quantitation of Infectious Human Retrovirus, Lymphadenopathy-Associated Virus (LAV). *J. oflmmunologicalMethods* **76**, 171–183 (1985).
155. Chess, L., MacDermott, R. P. & Schlossman, S. F. Immunologic functions of isolated human lymphocyte subpopulations. I. Quantitative isolation of human T and B cells and response to mitogens. *J. Immunol.* **113**, 1113–21 (1974).
156. Shiku, H. *et al.* Expression of T cell differentiation antigens on effector cells in cell mediated cytotoxicity in vitro. Evidence for functional heterogeneity related to the surface phenotype of T cells. *J. Exp. Med.* **141**, 227–241 (1975).
157. Cantor, H. & Boyse, E. A. Functional subclasses of T lymphocytes bearing different Ly antigens. II. Cooperation between subclasses of Ly+ cells in the generation of killer activity. *J. Exp. Med.* **141**, 1390–1399 (1975).
158. Contor, H. & Boyse, E. A. Functional subclasses of T lymphocytes bearing different Ly antigens. I. The generation of functionally distinct T cell subclasses is a differentiative process independent of antigen. *J. Exp. Med.* **141**, 1376–1389 (1975).
159. Jandinski, J., Cantor, H., Tadakuma, T., Peavy, D. L. & Pierce, C. W. Separation of helper T cells from suppressor T cells expressing different ly components: I. polyclonal activation: Suppressor and helper activities are inherent properties of distinct T-cell subclasses*. *J. Exp. Med.* **143**, 1382–1390 (1976).
160. Reinherz, E. L. & Schlossman, S. F. The differentiation and function of human T lymphocytes. *Cell* **19**, 821–827 (1980).
161. Bernard, A., Boumsell, L. & Hill, C. Joint Report of the First International Workshop on Human Leucocyte Differentiation Antigens by the Investigators of the Participating Laboratories. in *Leucocyte Typing* 9–142 (Springer Berlin Heidelberg, 1984). doi:10.1007/978-3-642-68857-7_3
162. Wang, J. *et al.* Atomic structure of a fragment of human CD4 containing two immunoglobulin-like domains. *Nature* **348**, (1990).
163. Garrett, T. P. J., Wang, J., Yan, Y., Liu, J. & Harrison, S. C. Refinement and Analysis of the Structure of the First Two Domains of Human CD4. *J. Mol. Biol.* **234**, 763–778 (1993).
164. Brady, R. L. *et al.* Crystal structure of domains 3 and 4 of rat CD4: relation to the NH2-terminal domains. *Science* **260**, 979–83 (1993).
165. Willbold, D. & Rösch, P. Solution structure of the human CD4 (403-419) receptor

peptide. *J. Biomed. Sci.* **3**, 435–441 (1996).
166. Wittlich, M., Thiagarajan, P., Koenig, B. W., Hartmann, R. & Willbold, D. NMR structure of the transmembrane and cytoplasmic domains of human CD4 in micelles. *Biochim. Biophys. Acta* **1798**, 122–7 (2010).
167. Robey, W. G. *et al.* Characterization of Envelope and Core Structural Gene Products of HTLV-III with Sera from AIDS Patients. *Source Sci. New Ser.* **228**, 593–595 (1985).
168. Henderson, L. E., Copeland, T. D., Smythers, G. W., Marquardt, H. & Oroszlan ', S. Amino-Terminal Amino Acid Sequence and Carboxyl-Terminal Analysis of Rauscher Murine Leukemia Virus Glycoproteins. *Virology* **85**, 319–322 (1978).
169. Navia, M. A. *et al.* Three-dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1. *Nature* **337**, 615–620 (1989).
170. Allan, J. S. *et al.* Major Glycoprotein Antigens that Induce Antibodies in AIDS Patients are Encoded by HTVL-III. *Source Sci. New Ser.* **228**, 1091–1094 (1985).
171. McDougal, J. S. *et al.* Cellular tropism of the human retrovirus. *J. Immunol.* **135**, 3151–3162 (1985).
172. Mcdougal, J. S., Kennedy, J. M., Sligh, S. P., Cort, A. & Mawle, J. K. A. Binding of HTLV-III/LAV to T4 T Cells by a Complex of the 110K Viral Protein and the T4 Molecule. *Nicholson Source Sci. New Ser.* **231**, 382–385 (1986).
173. Jay Maddon, P. *et al.* The T4 Gene Encodes the AIDS Virus Recepto ressed in the Immune System and the Brain. *Cell* **47**, 333–348 (1986).
174. Koenig, S. *et al.* Detection of AIDS virus in macrophages in brain tissue from AIDS patients with encephalopathy. *Science* **233**, 1089–93 (1986).
175. Gartner, S. *et al.* The Role of Mononuclear Phagocytes in HTLV-III/LAV Infection. *Source Sci. New Ser.* **233**, 215–219 (1986).
176. Matthews, T. J. *et al.* Interaction between the human T-cell lymphotropic virus type IIIB envelope glycoprotein gpl20 and the surface antigen CD4: Role of carbohydrate in binding and cell fusion. *Med. Sci.* **84**, 5424–5428 (1987).
177. Myers, G. *et al. Human Retroviruses and AIDS 1990: A Compilation and Analysis of Nucleic Acid and Amino Acid Sequences*. (Los Alamos National Laboratory, 1990).
178. Guo, H. G. *et al.* Characterization of an HIV-1 point mutant blocked in envelope glycoprotein cleavage. *Virology* **174**, 217–24 (1990).
179. Weissenhorn, W., Dessen, A., Harrison, S. C., Skehel, J. J. & Wiley, D. C. Atomic structure of the ectodomain from HIV-1 gp41. *Nature* **387**, 426–430 (1997).
180. Leonard, C. K. *et al.* Assignment of Intrachain Disulfide Bonds and Characterization of Potential Glycosylation Sites of the Type 1 Recombinant Human Immunodeficiency Virus Envelope Glycoprotein (gp120) Expressed in Chinese Hamster Ovary Cells. *J. Biol. Chem.* **265**, 10373–10382 (1990).
181. Fenouillet, E. *et al.* Role of N-linked glycans in the interaction between the envelope glycoprotein of human immunodeficiency virus and its CD4 cellular receptor. Structural enzymatic analysis. *J. Exp. Med.* **169**, 807–822 (1989).
182. Montefiori, D. C., Robinson, W. E. & Mitchell, W. M. Medical Sciences Role of protein N-glycosylation in pathogenesis of human immunodeficiency virus type 1. **85**, 9248–9252 (1988).
183. Gruters, R. Interference with HIV-induced syncytium formation and viral infectivity by inhibitors of trimming glucosidase. *Nature* **330**, (1987).
184. Fennie, C. & Lasky, L. A. Model for intracellular folding of the human immunodeficiency virus type 1 gp120. *J. Virol.* **63**, 639–46 (1989).

185. Lasky, L. A. *et al.* Delineation of a Region of the Human Immunodeficiency Virus Type 1 gp120 Glycoprotein Critical for Interaction with the CD4 Receptor. *Cell* **50**, 975–985 (1987).
186. Kowalski, M. *et al.* Functional Regions of the Envelope Glycoprotein of Human Immunodeficiency Virus Type. *Source Sci. New Ser.* **237**, 1351–1355 (1987).
187. Olshevsky, U. *et al.* Identification of individual human immunodeficiency virus type 1 gp120 amino acids important for CD4 receptor binding. *J.Virol.* **64**, 5701–5707 (1990).
188. Ryu, S.-E., Kwong, P. D., Truneh, A., Porter, T. G. & Arthos, J. Crystal structure of an HIV-binding recombinant fragment of human CD4. *Nature* **348**, (1990).
189. Richardson, N. E. *et al.* Binding site for human immunodeficiency virus coat protein gp120 is located in the NH2-terminal region of T4 (CD4) and requires the intact variable-region-like domain. *Proc. Natl. Acad. Sci. U. S. A.* **85**, 6102–6 (1988).
190. Maddon, P. J. *et al.* The Isolation and Nucleotide Sequence of a cDNA Encoding the T Cell Surface Protein 1"4: A New Member of the Immunoglobulin Gene Family. *Cell* **42**, 93–104 (1985).
191. Ashkenazi, a *et al.* Mapping the CD4 binding site for human immunodeficiency virus by alanine-scanning mutagenesis. *Proc. Natl. Acad. Sci. U. S. A.* **87**, 7150–7154 (1990).
192. Moebius, U., Clayton, L. K., Abraham, S., Harrison, S. C. & Reinherz, E. L. The Human Immunodeficiency Virus gp120 Binding Site on CD4: Deliniation by Quantitative Equilibrium and Kinetic Binding Studies of Mutants in Conjunction with a High-Resolution CD4 atomic structure. *J. Exp. Med.* **176**, 507–517 (1992).
193. Wu, H. *et al.* Kinetic and structural analysis of mutant CD4 receptors that are defective in HIV gp120 binding. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 15030–15035 (1996).
194. Kwong, P. D. *et al.* Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* **393**, 648–659 (1998).
195. Myszka, D. G. *et al.* Energetics of the HIV gp120-CD4 binding reaction. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 9026–31 (2000).
196. Capon, D. J. & Ward, R. H. R. THE CD4-gp120 INTERACTION AND AIDS PATHOGENSIS. *Annu. Rev. Immunol.* **9**, 649–678 (1991).
197. Stein, B. S. *et al.* pH-independent HIV entry into CD4-positive T cells via virus envelope fusion to the plasma membrane. *Cell* **49**, 659–668 (1987).
198. Bedinger, P., Moriarty, A., von Borstel, R. C., Donovan, N. J. & Steimer, K. S. Internalization of the human immunodeficiency virus does not require the cytoplasmic domain of CD4. *Nature* **334**, (1988).
199. Maddon, P. J. *et al.* HIV infection does not require endocytosis of its receptor, CD4. *Cell* **54**, 865–874 (1988).
200. Gallaher, W. R. Detection of a Fusion Peptide Sequence in the Transmembrane Protein of Human Immunodeficiency Virus. *Cell* **50**, 327–328 (1987).
201. Moore, J. Dissociation of gp120 from HIV-1 virions induced by soluble CD4. *Science* **250**, (1990).
202. Wyatt, R. *et al.* Involvement of the V1/V2 Variable Loop Structure in the Exposure of Human Immunodeficiency Virus Type 1 gp120 Epitopes Induced by Receptor Binding. *J. Virol.* **69**, 5723–5733 (1995).
203. Sweet, R. W., Truneh, A. & Hendrickson, W. A. CD4 : its structure , role in

immune function and AIDS pathogenesis , and potential as a pharmacological target. *Curr. Opin. Biotechnol.* **2**, 622–633 (1991).
204. Westervelt, P., Gendelmant, H. E. & Ratner, L. Identification of a determinant within the human immunodeficiency virus 1 surface envelope glycoprotein critical for productive infection of primary monocytes (macrophage/tropism/third variable domain of gpl20/CD4). *Med. Sci.* **88**, 3097–3101 (1991).
205. Hwang, S. S. *et al.* Identification of the Envelope V3 Loop as the Primary Determinant of Cell Tropism in. *Source Sci. New Ser. Knowl. Zool. R. I. Bowman, Ed. (Univ. Calif. Press G. J. Bakus, Atoll Res. Bull. Environ. Conserv* **253**, 71–74 (1991).
206. Schuitemaker, H. *et al.* Biological phenotype of human immunodeficiency virus type 1 clones at different stages of infection: progression of disease is associated with a shift from monocytotropic to T-cell-tropic virus population. *J. Virol.* **66**, 1354–60 (1992).
207. Rusche, J. R. *et al.* SC WMJ-2 LAV-MAL SF-2 Antibodies that inhibit fusion of human immunodeficiency virus-infected cells bind a 24-amino acid sequence of the viral envelope, gpl20. **85**, 3198–3202 (1989).
208. Javaherian, K. *et al.* Principal neutralizing domain of the human immunodeficiency virus type 1 envelope protein. *Med. Sci.* **86**, 6768–6772 (1989).
209. Skinner, M. A. *et al.* Neutralizing Antibodies to an Immunodominant Envelope Sequence Do Not Prevent gpl20 Binding to CD4. *J. Virol.* **62**, 4195–4200 (1988).
210. Palker, T. J. *et al.* Type-specific neutralization of the human immunodeficiency virus with antibodies to env-encoded synthetic peptides. *Immunology* **85**, 1932–1936 (1988).
211. Matsushita, S. *et al.* Characterization of a human immunodeficiency virus neutralizing monoclonal antibody and mapping of the neutralizing epitope. *J. Virol.* **62**, 2107–14 (1988).
212. Goudsmit, J. *et al.* Human immunodeficiency virus type 1 neutralization epitope with conserved architecture elicits early type-specific antibodies in experimentally infected chimpanzees. *Proc. Natl. Acad. Sci. U. S. A.* **85**, 4478–82 (1988).
213. Freed, E. 0, Myers, D. J. & Rissert, R. Identification of the Principal Neutralizing Determinant of Human Immunodeficiency Virus Type 1 as a Fusion Domain. *J. Virol.* 190–194 (1991).
214. Grimaila, R. J. *et al.* Mutations in the Principal Neutralization Determinant of Human Immunodeficiency Virus Type 1 Affect Syncytium Formation, Virus Infectivity, Growth Kinetics, and Neutralization. *J. Virol.* **66**, 1875–1883 (1992).
215. Weiner, D. Human Genes Other than CD4 Facilitate HIV-1 Infection of Murine Cells. *Pathobiology* **59**, (1991).
216. Weiner, D. B., Hubner, K., Williams, W. V & Greene, M. I. Species tropism of HIV-1 infectivity of interspecific cell hybridomas implies non-CD4 structures are required for cell entry. *Cancer Detect. Prev.* **14**, 317–20 (1990).
217. Clapham, P. R., Blanc, D. & Weiss ', R. A. Specific Cell Surface Requirements for the Infection of CD4-Positive Cells by Human Immunodeficiency Virus Types 1 and 2 and by Simian Immunodeficiency Virus. *Virology* **181**, 703–715 (1991).
218. Dragic, T. & Alizon, M. Different Requirements for Membrane Fusion Mediated by the Envelopes of Human Immunodeficiency Virus Types 1 and 2. *J. Virol.* 2355–2359 (1993).
219. Cocchi, F. *et al.* Identification of RANTES, MIP-1 alpha, and MIP-1 beta as the Major HIV-Suppressive Factors Produced by CD8+ T Cells. *Science (80-.).* **270**,

1811–1815 (1995).
220. Nussbaum, O., Broder, C. C. & Berger, E. A. Fusogenic Mechanisms of Enveloped-Virus Glycoproteins Analyzed by a Novel Recombinant Vaccinia Virus-Based Assay Quantitating Cell Fusion-Dependent Reporter Gene Activation. *J. Virol.* 5411–5422 (1994).
221. Feng, Y., Broder, C. C., Kennedy, P. E. & Berger, E. A. HIV-1 Entry Cofactor: Functional cDNA Cloning of a Seven-Transmembrane, G Protein. *Source Sci. New Ser.* **272**, 872–877 (1996).
222. D'Souza, M. P. & Harden, V. Chemokines and HIV–1 second receptors. *Nat. Med.* **2**, 1293–1300 (1996).
223. Choe, H. *et al.* The Beta-Chemokine Receptors CCR3 and CCR5 Facilitate Infection by Primary HIV-1 Isolates. *Cell* **85**, 1135–1148 (1996).
224. Deng, H. *et al.* Identification of a major co-receptor for primary isolates of HIV-1. *Nature* **381**, 661–666 (1996).
225. Dragic, T. *et al.* HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5. *Nature* **381**, 667–673 (1996).
226. Alkhatib, G. *et al.* CC CKR5: a RANTES, MIP-1alpha, MIP-1beta receptor as a fusion cofactor for macrophage-tropic HIV-1. *Science* **272**, 1955–8 (1996).
227. Berson, J. F. *et al.* A Seven-Transmembrane Domain Receptor Involved in Fusion and Entry of T-Cell-Tropic Human Immunodeficiency Virus Type 1 Strains. *J. Virol.* **70**, 6288–6295 (1996).
228. Doranz, B. J. *et al.* A dual-tropic primary HIV-1 isolate that uses fusin and the β-chemokine receptors CKR-5, CKR-3, and CKR-2b as fusion cofactors. *Cell* (1996). doi:10.1016/S0092-8674(00)81314-8
229. Weiss, R. A. *et al.* Thirty years on: HIV receptor gymnastics and the prevention of infection. *BMC Biol.* **11**, 57 (2013).
230. Forsman, A. & Weiss, R. A. Why is HIV a pathogen? *Trends Microbiol.* **16**, 555–560 (2008).
231. Doms, R. W. Beyond Receptor Expression: The Influence of Receptor Conformation, Density, and Affinity in HIV-1 Infection. *Virology* **276**, 229–237 (2000).
232. Rizzuto, C. D. *et al.* A conserved HIV gp120 glycoprotein structure involved in chemokine receptor binding. *Science* **280**, 1949–1953 (1998).
233. Moore, J. P., Trkolat, A. & Dragic, T. Co-receptors for HIV-1 entry. *Curr. Opin. Immunol.* **9**, 551–562 (1997).
234. Wu, L. *et al.* CD4-induced interaction of primary HIV-1 gp120 glycoproteins with the chemokine receptor CCR-5. *Nature* **384**, (1996).
235. Cocchi, F. *et al.* The V3 domain of the HIV-1 gp120 envelope glycoprotein is critical for chemokine-mediated blockade of infection. *Nat. Med.* **2**, (1996).
236. Rucker, J. *et al.* Regions in Beta-Chemokine Receptors CCR5 and CCR2b That Determine HIV-1 Cofactor Specificity. *Cell* **87**, 437–446 (1996).
237. Atchison, R. E. *et al.* Response to Multiple Extracellular Elements of CCR5 and HIV-1 Entry: Dissociation from Response to Chemokines. *Science (80-.).* **274**, 1924–1926 (1996).
238. Bieniasz, P. D. *et al.* HIV-1-induced cell fusion is mediated by multiple regions within both the viral envelope and the CCR-5 co-receptor. *EMBO J.* **16**, 2599–2609 (1997).
239. Edinger, A. L. *et al.* Differential utilization of CCR5 by macrophage and T cell tropic simian immunodeficiency virus strains. *Med. Sci.* **94**, 4005–4010 (1997).

240. Tamamis, P. & Floudas, C. A. Molecular recognition of CCR5 by an HIV-1 gp120 V3 loop. *PLoS One* **9**, (2014).
241. Wu, B. *et al.* Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science* **330**, 1066–1071 (2010).
242. Scholten, D. J. *et al.* Pharmacological modulation of chemokine receptor function. *British Journal of Pharmacology* **165**, 1617–1643 (2012).
243. Tan, Q. *et al.* Structure of the CCR5 chemokine receptor-HIV entry inhibitor maraviroc complex. *Science* **341**, 1387–90 (2013).
244. Zheng, Y. *et al.* Structure of CC Chemokine Receptor 5 with a Potent Chemokine Antagonist Reveals Mechanisms of Chemokine Recognition and Molecular Mimicry by HIV. *Immunity* **46**, 1005-1017.e5 (2017).
245. Huang, C.-C. *et al.* Structures of the CCR5 N terminus and of a tyrosine-sulfated antibody with HIV-1 gp120 and CD4. *Science* **317**, 1930–4 (2007).
246. Peng, P. *et al.* Structure-Based Design of 1-Heteroaryl-1,3-propanediamine Derivatives as a Novel Series of CC-Chemokine Receptor 5 Antagonists. *J. Med. Chem.* **61**, 9621–9636 (2018).
247. Shaik, M. M. *et al.* Structural basis of coreceptor recognition by HIV-1 envelope spike. *Nature* **565**, 318–323 (2019).
248. Wyatt, R. *et al.* The antigenic structure of the HIV gp120 envelope glycoprotein. *Nature* **393**, 705–11 (1998).
249. Ward, A. B. & Wilson, I. A. Insights into the trimeric HIV-1 envelope glycoprotein structure. *Trends Biochem. Sci.* **40**, 101–107 (2015).
250. Kwon, Y. Do *et al.* Crystal structure, conformational fixation, and entry-related interactions of mature ligand-free HIV-1 Env. *Nat. Struct. Mol. Biol.* **22**, 522–531 (2015).
251. Binley, J. M. *et al.* A Recombinant Human Immunodeficiency Virus Type 1 Envelope Glycoprotein Complex Stabilized by an Intermolecular Disulfide Bond between the gp120 and gp41 Subunits Is an Antigenic Mimic of the Trimeric Virion-Associated Structure. *J. Virol.* **74**, 627–643 (2000).
252. Sanders, R. W. *et al.* Stabilization of the Soluble, Cleaved, Trimeric Form of the Envelope Glycoprotein Complex of Human Immunodeficiency Virus Type 1. *J. Virol.* **76**, 8875–8889 (2002).
253. Khayat, R. *et al.* Structural Characterization of Cleaved, Soluble HIV-1 Envelope Glycoprotein Trimers. *J. Virol.* **87**, 9865–9872 (2013).
254. Lai, Y. T. *et al.* Lattice engineering enables definition of molecular features allowing for potent small-molecule inhibition of HIV-1 entry. *Nat. Commun.* **10**, 1–11 (2019).
255. Yamada, H. *et al.* 'Crystal lattice engineering,' an approach to engineer protein crystal contacts by creating intermolecular symmetry: Crystallization and structure determination of a mutant human RNase 1 with a hydrophobic interface of leucines. *Protein Sci.* **16**, 1389–1397 (2007).
256. Kalyoncu, S. *et al.* Effects of protein engineering and rational mutagenesis on crystal lattice of single chain antibody fragments. *Proteins Struct. Funct. Bioinforma.* **82**, 1884–1895 (2014).
257. Koide, S. Engineering of recombinant crystallization chaperones. *Current Opinion in Structural Biology* **19**, 449–457 (2009).
258. Doranz, B. J., Baik, S. S. & Doms, R. W. Use of a gp120 binding assay to dissect the requirements and kinetics of human immunodeficiency virus fusion events. *J. Virol.* **73**, 10346–58 (1999).

259. Anglister, J., Tugarinov, V., Zvi, A. & Levy, R. A cis proline turn linking two |[beta]|-hairpin strands in the solution structure of an antibody-bound HIV-1IIIB V3 peptide. *Nat. Struct. Biol.* **6**, 331–335 (1999).
260. Tugarinov, V. *et al.* NMR structure of an anti-gp120 antibody complex with a V3 peptide reveals a surface important for co-receptor binding. *Structure* **8**, 385–395 (2000).
261. Hu, Q. *et al.* Identification of ENV determinants in V3 that influence the molecular anatomy of CCR5 utilization. *J. Mol. Biol.* **302**, 359–375 (2000).
262. Kassa, A. *et al.* Transitions to and from the CD4-Bound Conformation Are Modulated by a Single-Residue Change in the Human Immunodeficiency Virus Type 1 gp120 Inner Domain. *J. Virol.* **83**, 8364–8378 (2009).
263. Kwon, Y. Do *et al.* Unliganded HIV-1 gp120 core structures assume the CD4-bound conformation with regulation by quaternary interactions and variable loops. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 5663–8 (2012).
264. Liu, Q. *et al.* Quaternary contact in the initial interaction of CD4 with the HIV-1 envelope trimer. *Nat. Struct. Mol. Biol.* 1–12 (2017). doi:10.1038/nsmb.3382
265. Duenas-Decamp, M. *et al.* Saturation Mutagenesis of the HIV-1 Envelope CD4 Binding Loop Reveals Residues Controlling Distinct Trimer Conformations. *PLOS Pathog.* **12**, e1005988 (2016).
266. Wang, H. *et al.* Cryo-EM structure of a CD4-bound open HIV-1 envelope trimer reveals structural rearrangements of the gp120 V1V2 loop. *Proc. Natl. Acad. Sci. U. S. A.* 201615939 (2016). doi:10.1073/pnas.1615939113
267. Ozorowski, G. *et al.* Open and closed structures reveal allostery and pliability in the HIV-1 envelope spike. *Nature* **547**, 360–363 (2017).
268. Chien, M.-P., Jiang, S. & Chang, D.-K. The function of coreceptor as a basis for the kinetic dissection of HIV type 1 envelope protein-mediated cell fusion. *FASEB J.* **22**, 1179–1192 (2008).
269. Deng, J. *et al.* HIV envelope gp120 alters T-cell receptor mobilization in the immunological synapse of uninfected CD4 T cells and augments the T cell activation. *J. Virol.* JVI.01532-16 (2016). doi:10.1128/JVI.01532-16
270. Guttman, M. *et al.* CD4-Induced Activation in a Soluble HIV-1 Env Trimer. *Struct. Des.* **22**, 974–984 (2014).
271. Pan, Y., Ma, B. & Nussinov, R. CD4 binding partially locks the bridging sheet in gp120 but leaves the β2/3 strands flexible. *J. Mol. Biol.* **350**, 514–527 (2005).
272. Kaplan, G., Roitburd-Berman, A., Lewis, G. K. & Gershoni, J. M. Range of CD4-Bound Conformations of HIV-1 gp120, as Defined Using Conditional CD4-Induced Antibodies. *J. Virol.* **90**, 4481–4493 (2016).
273. Garrett, T. P. J., Wang, J., Yan, Y., Liu, J. & Harrison, S. C. Refinement and analysis of the structure of the first two domains of human CD4. *J. Mol. Biol.* **234**, 763–778 (1993).
274. Ryu, S.-E. *et al. Crystal structure of an HIV-binding recombinant fragment of human CD4*.
275. Zerbib, A. C., Reske-Kunz, A. B., Lock, P. & Sékaly, R. P. Cd4-mediated enhancement or inhibition of t cell activation does not require the cd4:p56kk association. *J. Exp. Med.* **179**, 1973–1983 (1994).
276. Chu, K. & Littman, D. R. Requirement for kinase activity of CD4-associated p56lck in antibody-triggered T cell signal transduction. *J. Biol. Chem.* **269**, 24095–101 (1994).
277. Wilen, C. B., Tilton, J. C. & Doms, R. W. Molecular Mechanisms of HIV Entry. in

Advances in Experimental Medicine and Biology **726**, 223–242 (Springer US, 2012).
278. Shaik, M. M. *et al.* Structural basis of coreceptor recognition by HIV-1 envelope spike. *Nature* **565**, 318–323 (2019).
279. Chang, M. I., Panorchan, P., Dobrowsky, T. M., Tseng, Y. & Wirtz, D. Single-Molecule Analysis of Human Immunodeficiency Virus Type 1 gp120-Receptor Interactions in Living Cells. *J. Virol.* **79**, 14748–14755 (2005).
280. Dobrowsky, T. M., Zhou, Y., Sun, S. X., Siliciano, R. F. & Wirtz, D. Monitoring Early Fusion Dynamics of Human Immunodeficiency Virus Type 1 at Single-Molecule Resolution. *J. Virol.* **82**, 7022–7033 (2008).
281. Kumar, S. *et al.* Capturing the inherent structural dynamics of the HIV-1 envelope glycoprotein fusion peptide. *Nat. Commun.* **10**, 1–10 (2019).
282. Liu, J., Bartesaghi, A., Borgnia, M. J., Sapiro, G. & Subramaniam, S. Molecular architecture of native HIV-1 gp120 trimers. *Nature* **455**, 109–113 (2008).
283. Pancera, M. *et al.* Structure and immune recognition of trimeric pre-fusion HIV-1 Env. *Nature* **514**, (2014).
284. Wang, H. *et al.* Cryo-EM structure of a CD4-bound open HIV-1 envelope trimer reveals structural rearrangements of the gp120 V1V2 loop. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E7151–E7158 (2016).
285. Zhang, W., Cao, S., Martin, J. L., Mueller, J. D. & Mansky, L. M. Morphology and ultrastructure of retrovirus particles. *AIMS Biophysics* **2**, 343–369 (2015).
286. Klug, Y. A., Rotem, E., Schwarzer, R. & Shai, Y. Mapping out the intricate relationship of the HIV envelope protein and the membrane environment. *Biochim. Biophys. Acta - Biomembr.* **1859**, 550–560 (2017).
287. White, J. M., Delos, S. E., Brecher, M. & Schornberg, K. Structures and Mechanisms of Viral Membrane Fusion Proteins: Multiple Variations on a Common Theme. *Crit. Rev. Biochem. Mol. Biol.* **43**, 189–219 (2008).
288. Schibli, D. J. & Weissenhorn, W. Class I and class II viral fusion protein structures reveal similar principles in membrane fusion. *Molecular Membrane Biology* **21**, 361–371 (2004).
289. Durell, S. R., Martin, I., Ruysschaert, J. M., Shai, Y. & Blumenthal, R. What studies of fusion peptides tell us about viral envelope glycoprotein-mediated membrane fusion. *Molecular Membrane Biology* **14**, 97–112 (1997).
290. Khasnis, M. D. *et al.* Receptor Activation of HIV-1 Env Leads to Asymmetric Exposure of the gp41 Trimer. *PLOS Pathog.* **12**, e1006098 (2016).
291. Gallo, R. C. & Montagnier, L. The Discovery of HIV as the Cause of AIDS. *N. Engl. J. Med.* **349**, 2283–2285 (2003).
292. Moore, J. P. & Doms, R. W. The entry of entry inhibitors: a fusion of science and medicine. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 10598–602 (2003).
293. White, T. A. *et al.* Molecular architectures of trimeric SIV and HIV-1 envelope glycoproteins on intact viruses: Strain- dependent variation in quaternary structure. *PLoS Pathog.* **6**, (2010).
294. Chan, D. C. & Kim, P. S. HIV Entry and Its Inhibition. *Cell* **93**, 681–684 (1998).
295. Colman, P. M. & Lawrence, M. C. The structural biology of type I viral membrane fusion. *Nat. Rev. Mol. Cell Biol.* **4**, 309–319 (2003).
296. DeSantis, M. C., Kim, J. H., Song, H., Klasse, P. J. & Cheng, W. Quantitative correlation between infectivity and Gp120 density on HIV-1 virions revealed by optical trapping virometry. *J. Biol. Chem.* **292**, 3061 (2017).
297. Yuste, E., Johnson, W., Pavlakis, G. N. & Desrosiers, R. C. Virion Envelope

Content, Infectivity, and Neutralization Sensitivity of Simian Immunodeficiency Virus. *J. Virol.* **79**, 12455–12463 (2005).
298. Earl, P. L., Doms, R. W. & Moss, B. Multimeric CD4 binding exhibited by human and simian immunodeficiency virus envelope protein dimers. *J. Virol.* **66**, 5610–4 (1992).
299. Kim, M. *et al.* The Stoichiometry of Trimeric SIV Glycoprotein Interaction with CD4 Differs from That of Anti-envelope Antibody Fab Fragments. *J. Biol. Chem.* **276**, 42667–42676 (2001).
300. Salzwedel, K. & Berger, E. A. Cooperative subunit interactions within the oligomeric envelope glycoprotein of HIV-1: Functional complementation of specific defects in gp120 and gp41. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 12794–12799 (2000).
301. Fischer, P. B., Karlsson, G. B., Dwek, R. A. & Platt, F. M. N-butyldeoxynojirimycin-mediated inhibition of human immunodeficiency virus entry correlates with impaired gp120 shedding and gp41 exposure. *J. Virol.* **70**, 7153–7160 (1996).
302. Fu, Y. K., Hart, T. K., Jonak, Z. L. & Bugelski, P. J. Physicochemical dissociation of CD4-mediated syncytium formation and shedding of human immunodeficiency virus type 1 gp120. *J. Virol.* **67**, 3818–25 (1993).
303. Thali, M., Furman, C., Helseth, E., Repke, H. & Sodroski, J. Lack of correlation between soluble CD4-induced shedding of the human immunodeficiency virus type 1 exterior envelope glycoprotein and subsequent membrane fusion events. *J. Virol.* **66**, 5516–24 (1992).
304. Willey, R. L., Martin, M. A. & Peden, K. W. Increase in soluble CD4 binding to and CD4-induced dissociation of gp120 from virions correlates with infectivity of human immunodeficiency virus type 1. *J. Virol.* **68**, 1029–39 (1994).
305. Frankel, A. D. & Young, J. A. T. HIV-1: Fifteen Proteins and an RNA. *Annu. Rev. Biochem* **67**, 1–25 (1998).
306. Checkley, M. A., Luttge, B. G. & Freed, E. O. HIV-1 envelope glycoprotein biosynthesis, trafficking, and incorporation. *J. Mol. Biol.* **410**, 582–608 (2011).
307. Anderson, J. L. & Hope, T. J. HIV accessory proteins and surviving the host cell. *Curr. HIV/AIDS Rep.* **1**, 47–53 (2004).
308. Bhargava, A., Lahaye, X. & Manel, N. Let me in: Control of HIV nuclear entry at the nuclear envelope. *Cytokine Growth Factor Rev.* (2018). doi:10.1016/J.CYTOGFR.2018.02.006
309. Knyazhanskaya, E. S., Shadrina, O. A., Anisenko, A. N. & Gottikh, M. B. Role of DNA-Dependent Protein Kinase in the HIV-1 Replication Cycle. *Mol. Biol. Orig. Russ. Text* © **50**, 26–8933 (2016).
310. Singh, A. Acquired Immunodeficiency Syndrome (AIDS): A Review. **4**,
311. Li, G. & De Clercq, E. HIV Genome-Wide Protein Associations: a Review of 30 Years of Research. *Microbiol. Mol. Biol. Rev.* **80**, 679–731 (2016).
312. Malim, M. H. & Emerman, M. HIV-1 Accessory Proteins-Ensuring Viral Survival in a Hostile Environment. *Cell Host and Microbe* **3**, 388–398 (2008).
313. Fassati, A. & Goff, S. P. Characterization of intracellular reverse transcription complexes of human immunodeficiency virus type 1. *J. Virol.* **75**, 3626–35 (2001).
314. Warrilow, D., Tachedjian, G. & Harrich, D. Maturation of the HIV reverse transcription complex: Putting the jigsaw together. *Reviews in Medical Virology* **19**, 324–337 (2009).
315. Warrilow, D., Warren, K. & Harrich, D. Strand Transfer and Elongation of HIV-1 Reverse Transcription Is Facilitated by Cell Factors In Vitro. *PLoS One* **5**, e13229

(2010).
316. Iordanskiy, S. N. & Bukrinsky, M. I. Reverse transcription complex: The key player of the early phase of HIV replication. *Future Virology* **2**, 49–64 (2007).
317. Larsen, K. P. *et al.* Architecture of an HIV-1 reverse transcriptase initiation complex. *Nature* **557**, 118–122 (2018).
318. Gonçalves, J. *et al.* Integration of HIV in the Human Genome: Which Sites Are Preferential? A Genetic and Statistical Assessment. *Int. J. Genomics* **2016**, (2016).
319. Lusic, M. & Siliciano, R. F. Nuclear landscape of HIV-1 infection and integration. *Nature Reviews Microbiology* **15**, 69–82 (2017).
320. Debyser, Z., Vansant, G., Bruggemans, A., Janssens, J. & Christ, F. Insight in HIV Integration Site Selection Provides a Block-and-Lock Strategy for a Functional Cure of HIV Infection. *Viruses* **11**, 12 (2018).
321. Battivelli, E. *et al.* Chromatin Functional States Correlate with HIV Latency Reversal in Infected Primary CD4+ T Cells. *bioRxiv* 242958 (2018). doi:10.1101/242958
322. Jordan, A., Bisgrove, D. & Verdin, E. HIV reproducibly establishes a latent infection after acute infection of T cells in vitro. *EMBO J.* **22**, 1868–77 (2003).
323. Zerbato, J. M. *et al.* Establishment and Reversal of HIV-1 Latency in Naïve and Central Memory CD4+ T Cells *in Vitro*. *J. Virol.* JVI.00553-16 (2016). doi:10.1128/JVI.00553-16
324. Lesbats, P., Engelman, A. N. & Cherepanov, P. Retroviral DNA Integration. *Chemical Reviews* **116**, 12730–12757 (2016).
325. Lewinski, M. K. & Bushman, F. D. Retroviral DNA Integration-Mechanism and Consequences. *Advances in Genetics* **55**, 147–181 (2005).
326. Michieletto, D., Lusic, M., Marenduzzo, D. & Orlandini, E. Physical principles of retroviral integration in the human genome. *Nat. Commun.* **10**, 1–11 (2019).
327. Keane, S. & Summers, M. NMR Studies of the Structure and Function of the HIV-1 5′-Leader. *Viruses* **8**, 338 (2016).
328. Nomaguchi, M., Doi, N., Koma, T. & Adachi, A. HIV-1 mutates to adapt in fluxing environments. *Microbes Infect.* 1–5 (2017). doi:10.1016/j.micinf.2017.08.003
329. Structure, Expression, and Regulation of the HIV Genome. Available at: http://hivinsite.ucsf.edu/InSite?page=kb-02-01-02. (Accessed: 13th February 2020)
330. Quinones-Kochs, M. I., Buonocore, L. & Rose, J. K. Role of N-Linked Glycans in a Human Immunodeficiency Virus Envelope Glycoprotein: Effects on Protein Function and the Neutralizing Antibody Response. *J. Virol.* **76**, 4199–4211 (2002).
331. Klug, Y. A., Rotem, E., Schwarzer, R. & Shai, Y. Mapping out the intricate relationship of the HIV envelope protein and the membrane environment. *Biochim. Biophys. Acta - Biomembr.* (2016). doi:10.1016/j.bbamem.2016.10.012
332. Pritchard, L. K., Harvey, D. J., Bonomelli, C., Crispin, M. & Doores, K. J. Cell- and Protein-Directed Glycosylation of Native Cleaved HIV-1 Envelope. *J. Virol.* **89**, 8932–44 (2015).
333. Sundquist, W. I. & Kräusslich, H. G. HIV-1 assembly, budding, and maturation. *Cold Spring Harbor Perspectives in Medicine* **2**, (2012).
334. Liu, Y. *et al.* HIV-1 Sequence Necessary and Sufficient to Package Non-viral RNAs into HIV-1 Particles. *J. Mol. Biol.* (2017). doi:10.1016/j.jmb.2017.06.018
335. Bell, N. M. & Lever, A. M. L. HIV Gag polyprotein: Processing and early viral particle assembly. *Trends Microbiol.* **21**, 136–144 (2013).

336. Henne, W. M., Buchkovich, N. J. & Emr, S. D. The ESCRT Pathway. *Developmental Cell* **21**, 77–91 (2011).
337. Crispin, M., Ward, A. B. & Wilson, I. A. Structure and Immune Recognition of the HIV Glycan Shield. *Annu. Rev. Biophys.* **47**, 499–523 (2018).
338. Ferreira, R. C. *et al.* Structural Rearrangements Maintain the Glycan Shield of an HIV-1 Envelope Trimer After the Loss of a Glycan. *Sci. Rep.* **8**, (2018).
339. Wei, X. *et al.* Antibody neutralization and escape by HIV-1. *Nature* **422**, 307–312 (2003).
340. Behrens, A.-J. J. & Crispin, M. Structural principles controlling HIV envelope glycosylation. *Curr. Opin. Struct. Biol.* **44**, 125–133 (2017).
341. Watanabe, Y., Bowden, T. A., Wilson, I. A. & Crispin, M. Exploitation of glycosylation in enveloped virus pathobiology. *Biochimica et Biophysica Acta - General Subjects* **1863**, 1480–1497 (2019).
342. Chakraborty, S. *et al.* A Network-based approach for Quantifying the Resilience and Vulnerability of HIV-1 Native Glycan Shield. *bioRxiv Biophys.* doi:10.1101/846071
343. Taylor, M. E., Drickamer, K., Schnaar, R. L., Etzler, M. E. & Varki, A. Discovery and Classification of Glycan-Binding Proteins. (2017). doi:10.1101/GLYCOBIOLOGY.3E.028
344. Varki, A. C. R. D. . E. J. D. . F. H. H. . S. P. . B. C. R. . H. G. W. . E. M. & E. *Essentials of Glycobiology, 3rd edition*. *Cold Spring Harbor (NY)* (Cold Spring Harbor Laboratory Press, 2015).
345. Stansell, E. *et al.* Gp120 on HIV-1 virions lacks O-linked carbohydrate. *PLoS One* **10**, (2015).
346. Lin, T.-Y. & Lai, J. R. Interrogation of Side Chain Biases for Oligomannose Recognition by Antibody 2G12 via Structure-Guided Phage Display Libraries. *Bioorg. Med. Chem.* (2017). doi:10.1016/j.bmc.2017.09.013
347. Polonoff, E., Machida, C. A. & Kabat, D. Glycosylation and Intracellular Transport of Membrane Glycoproteins Encoded by Murine Leukemia Viruses INHIBITION BY AMINO ACID ANALOGUES AND BY TUNICAMYCIN*. **57**, 34023–14028 (1982).
348. Shakin-Eshleman, S. H., Spitalnik, S. L. & Kasturi, L. The Amino Acid at the X Position of an Asn-X-Ser Sequon Is an Important Determinant of N-Linked Core-glycosylation Efficiency. *J. Biol. Chem.* **271**, 6363–6366 (1996).
349. Gavel, Y. & Heijne, G. von. Sequence differences between glycosylated and non-glycosylated Asn-X-Thr/Ser acceptor sites: implications for protein engineering. *"Protein Eng. Des. Sel.* **3**, 433–442 (1990).
350. Bieberich, E. Synthesis, Processing, and Function of N-glycans in N-glycoproteins. in 47–70 (2014). doi:10.1007/978-1-4939-1154-7_3
351. Aebi, M. N-linked protein glycosylation in the ER. *Biochimica et Biophysica Acta - Molecular Cell Research* **1833**, 2430–2437 (2013).
352. Coss, K. P. *et al.* HIV-1 Glycan Density Drives the Persistence of the Mannose Patch within an Infected Individual. *J. Virol.* **90**, 11132–11144 (2016).
353. Breitling, J. & Aebi, M. N-Linked Protein Glycosylation in the Endoplasmic Reticulum. *Cold Spring Harb. Perspect. Biol.* **5**, a013359–a013359 (2013).
354. Malaby, H. L. H. & Kobertz, W. R. Molecular determinants of co-and post-translational N-glycosylation of type I transmembrane peptides. *Biochem. J.* **453**, 427–434 (2013).
355. Thaysen-Andersen, M. & Packer, N. H. Site-specific glycoproteomics confirms

that protein structure dictates formation of N-glycan type, core fucosylation and branching. *Glycobiology* **22**, 1440–52 (2012).
356. Alqudah, M. A. Y., Yaseen, M. M. M. S. & Yaseen, M. M. M. S. HIV-1 strategies to overcome the immune system by evading and invading innate immune system. *HIV AIDS Rev.* **15**, 1–12 (2016).
357. Klein, J. S. & Bjorkman, P. J. Few and Far Between: How HIV May Be Evading Antibody Avidity. *PLoS Pathog.* **6**, e1000908 (2010).
358. Go, E. P. *et al.* A glycosylation benchmark profile for HIV-1 envelope glycoprotein production based on eleven Env trimers. *J. Virol.* JVI.02428-16 (2017). doi:10.1128/JVI.02428-16
359. Go, E. P. *et al.* Characterization of Host-Cell Line Specific Glycosylation Profiles of Early Transmitted/Founder HIV-1 gp120 Envelope Proteins. (2013).
360. Zhou, T. *et al.* Structural basis for broad and potent neutralization of HIV-1 by antibody VRC01. *Science* **329**, 811–7 (2010).
361. Zhou, T. *et al.* A Neutralizing Antibody Recognizing Primarily N-Linked Glycan Targets the Silent Face of the HIV Envelope. *Immunity* **48**, 500-513.e6 (2018).
362. Moore, P. L. *et al.* Evolution of an HIV glycan–dependent broadly neutralizing antibody epitope through immune escape. *Nat. Med.* **18**, 1688–1692 (2012).
363. Sok, D. *et al.* Promiscuous glycan site recognition by antibodies to the high-mannose patch of gp120 broadens neutralization of HIV. *Sci. Transl. Med.* **6**, 236ra63 (doi: 10.1126/scitranslmed.3008104) (2014).
364. Rathore, U. *et al.* Glycosylation of the core of the HIV-1 envelope subunit protein gp120 is not required for native trimer formation or viral infectivity. *J. Biol. Chem.* **292**, 10197–10219 (2017).
365. Gristick, H. B., Wang, H., Bjorkman, P. J. & IUCr. X-ray and EM structures of a natively glycosylated HIV-1 envelope trimer. *Acta Crystallogr. Sect. D Struct. Biol.* **73**, 822–828 (2017).
366. Panico, M. *et al.* Mapping the complete glycoproteome of virion-derived HIV-1 gp120 provides insights into broadly neutralizing antibody binding. *Sci. Rep.* **6**, 32956 (2016).
367. Stewart-Jones, G. B. E. *et al.* Trimeric HIV-1-Env Structures Define Glycan Shields from Clades A, B, and G. *Cell* **165**, 813–826 (2016).
368. Bonomelli, C. *et al.* The Glycan Shield of HIV Is Predominantly Oligomannose Independently of Production System or Viral Clade. *PLoS One* **6**, e23521 (2011).
369. Lemmin, T., Soto, C., Stuckey, J. & Kwong, P. D. Microsecond Dynamics and Network Analysis of the HIV-1 SOSIP Env Trimer Reveal Collective Behavior and Conserved Microdomains of the Glycan Shield. *Structure* **25**, 1–9 (2017).
370. Pritchard, L. K. *et al.* Glycan clustering stabilizes the mannose patch of HIV-1 and preserves vulnerability to broadly neutralizing antibodies. *Nat. Commun.* **6**, 7479 (2015).
371. Behrens, A.-J. *et al.* Molecular architecture of the cleavage-dependent mannose patch on a soluble HIV-1 envelope glycoprotein trimer. *J. Virol.* JVI.01894-16 (2016). doi:10.1128/JVI.01894-16
372. Kong, L., Wilson, I. A. & Kwong, P. D. Crystal structure of a fully glycosylated HIV-1 gp 120 core reveals a stabilizing role for the glycan at Asn262. *Proteins Struct. Funct. Bioinforma.* **83**, 590–596 (2015).
373. Wang, W. *et al.* A systematic study of the N-glycosylation sites of HIV-1 envelope protein on infectivity and antibody-mediated neutralization. *Retrovirology* **10**, (2013).

374. Ward, A. B. & Wilson, I. A. The HIV-1 envelope glycoprotein structure: nailing down a moving target. *Immunol. Rev.* **275**, 21–32 (2017).
375. Travers, S. A. Conservation, Compensation, and Evolution of N-Linked Glycans in the HIV-1 Group M Subtypes and Circulating Recombinant Forms. *Isrn Aids* **2012**, 823605 (2012).
376. Behrens, A.-J. J. *et al.* Composition and Antigenic Effects of Individual Glycan Sites of a Trimeric HIV-1 Envelope Glycoprotein. *Cell Rep.* **14**, 2695–2706 (2016).
377. Behrens, A. Fine structure of the HIV-1 glycan shield. (2017).
378. Sanders, R. W. *et al.* The carbohydrate at asparagine 386 on HIV-1 gp120 is not essential for protein folding and function but is involved in immune evasion. *Retrovirology* **5**, 10 (2008).
379. AIDSinfo. Approval of AZT. *HIV/AIDS News* (1987). Available at: https://aidsinfo.nih.gov/news/274/approval-of-azt.
380. Montefiori, D. C., Robinson, W. E., Schuffman, S. S. & Mitchell, W. M. Evaluation of Antiviral Drugs and Neutralizing Antibodies to Human Immunodeficiency Virus by a Rapid and Sensitive Microtiter Infection Assay. *J. Clin. Microbiol.* 231–235 (1988).
381. Drugs@FDA: FDA-Approved Drugs.
382. Qian, K., Morris-Natschke, S. L. & Lee, K. H. HIV entry inhibitors and their potential in HIV therapy. *Medicinal Research Reviews* **29**, 369–393 (2009).
383. Curreli, F. *et al.* Synthesis, antiviral potency, in vitro ADMET and X-ray structure of potent CD4-mimics as entry inhibitors that target the Phe43 cavity of HIV-1 gp120. *J. Med. Chem.* acs.jmedchem.7b00179 (2017). doi:10.1021/acs.jmedchem.7b00179
384. Zhao, B., Mankowski, M. K., Snyder, B. A., Ptak, R. G. & Liwang, P. J. Highly potent chimeric inhibitors targeting two steps of HIV cell entry. *J. Biol. Chem.* **286**, 28370–81 (2011).
385. Xue, J., Kuo, N.-W., Schill, M. S. & LiWang, P. J. A Comparison of 5P12-vMIP-II and vMIP-II as HIV-1 Entry Inhibitors. *Biochem. Physiol. Open Access* **s2**, (2013).
386. Nguyen, A. F., Schill, M. S., Jian, M. & LiWang, P. J. The Effect of N-Terminal Cyclization on the Function of the HIV Entry Inhibitor 5P12-RANTES. *Int. J. Mol. Sci.* **18**, 1575 (2017).
387. Haqqani, A. A. & Tilton, J. C. Entry inhibitors and their use in the treatment of HIV-1 infection. *Antiviral Res.* **98**, 158–170 (2013).
388. Ahn, K. W. & Root, M. J. Complex interplay of kinetic factors governs the synergistic properties of HIV-1 entry inhibitors. *J. Biol. Chem.* **292**, 16498–16510 (2017).
389. De Clercq, E. Anti-HIV drugs: 25 compounds approved within 25 years after the discovery of HIV. *Int. J. Antimicrob. Agents* **33**, 307–320 (2009).
390. Gaertner, H. *et al.* Highly potent, fully recombinant anti-HIV chemokines: reengineering a low-cost microbicide. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 17706–11 (2008).
391. Hu, Q. *et al.* Blockade of attachment and fusion receptors inhibits HIV-1 infection of human cervical tissue. *J. Exp. Med.* **199**, 1065–1075 (2004).
392. Madani, N. *et al.* Activation and Inactivation of Primary Human Immunodeficiency Virus (HIV-1) Envelope Glycoprotein Trimers by CD4-Mimetic Compounds. *J. Virol.* JVI.01880-16 (2016). doi:10.1128/JVI.01880-16
393. Kagiampakis, I. *et al.* Potent strategy to inhibit HIV-1 by binding both gp120 and gp41. *Antimicrob. Agents Chemother.* **55**, 264–275 (2011).

394. Curreli, F. *et al.* Design, synthesis and evaluation of small molecule CD4-mimics as entry inhibitors possessing broad spectrum anti-HIV-1 activity. *Bioorg. Med. Chem.* (2016). doi:10.1016/j.bmc.2016.09.057
395. Tomkowicz, B. & Collman, R. G. HIV-1 entry inhibitors: closing the front door. *http://dx.doi.org/10.1517/14728222.8.2.65* (2005).
396. Guilloux, L. & Cedex, R. Efficient HIV-1 trans infection of CD4+ T cells occurs in the presence of antiretroviral therapy. *Infect. Dis. Soc. Am.* (2019).
397. Wilkin, T. J. CROI 2018 : Advances in Antiretroviral Therapy. **26**, (2018).
398. Smith, S. J., Zhao, X. Z., Burke, T. R. & Hughes, S. H. HIV-1 integrase inhibitors that are broadly effective against drug-resistant mutants. *Antimicrob. Agents Chemother.* **62**, (2018).
399. Wei, X. *et al.* Emergence of Resistant Human Immunodeficiency Virus Type 1 in Patients Receiving Fusion Inhibitor (T-20) Monotherapy Emergence of Resistant Human Immunodeficiency Virus Type 1 in Patients Receiving Fusion Inhibitor (T-20) Monotherapy. *Antimicrob. Agents Chemother.* **46**, 1896–1905 (2002).
400. Ferreira, R.-C. *et al.* Structural Rearrangements Maintain the Glycan Shield of an HIV-1 Envelope Trimer After the Loss of a Glycan. *Sci. Rep.* **8**, 15031 (2018).
401. Diskin, R. *et al.* Restricting HIV-1 pathways for escape using rationally designed anti-HIV-1 antibodies. *J. Exp. Med.* **210**, 1235–49 (2013).
402. Li, H., Zony, C., Chen, P. & Chen, B. K. Reduced potency and incomplete neutralization of broadly neutralizing antibodies against cell-to-cell transmission of HIV-1 with transmitted founder Envs. *J. Virol.* JVI.02425-16 (2017). doi:10.1128/JVI.02425-16
403. Balzarini, J. Inhibition of HIV entry by carbohydrate-binding proteins. *Antiviral Res.* **71**, 237–247 (2006).
404. Raman, R., Tharakaraman, K., Sasisekharan, V. & Sasisekharan, R. Glycan–protein interactions in viral pathogenesis. *Curr. Opin. Struct. Biol.* **40**, 153–162 (2016).
405. Huskens, D. & Schols, D. Algal lectins as potential HIV microbicide candidates. *Mar. Drugs* **10**, 1476–1497 (2012).
406. Balzarini, J. Pointer Carbohydrate-binding agents: a potential future cornerstone for the chemotherapy of enveloped viruses? *Antivir. Chem. Chemother.* **18**, 1–11 (2007).
407. Ziółkowska, N. E. & Wlodawer, A. Structural studies of algal lectins with anti-HIV activity. (2006).
408. Takebe, Y. *et al.* Antiviral Lectins from Red and Blue-Green Algae Show Potent In Vitro and In Vivo Activity against Hepatitis C Virus. *PLoS One* **8**, 1–10 (2013).
409. Mitchell, C. A., Ramessar, K. & O'Keefe, B. R. Antiviral lectins: Selective inhibitors of viral entry. *Antiviral Res.* **142**, 37–54 (2017).
410. van der Vlist, M. & Geijtenbeek, T. B. H. Langerin functions as an antiviral receptor on Langerhans cells. *Immunol. Cell Biol.* **88**, 410–415 (2010).
411. De Witte, L. *et al.* Langerin is a natural barrier to HIV-1 transmission by Langerhans cells. *Nat. Med.* **13**, 367–371 (2007).
412. Rodriguez-Garcia, M. *et al.* Dendritic cells from the human female reproductive tract rapidly capture and respond to HIV. *Nat. Publ. Gr.* (2016). doi:10.1038/mi.2016.72
413. Dutartre, H., Clavière, M., Journo, C. & Mahieux, R. Cell-Free versus Cell-to-Cell Infection by Human Immunodeficiency Virus Type 1 and Human T-Lymphotropic Virus Type 1: Exploring the Link among Viral Source, Viral Trafficking, and Viral

Replication. *J. Virol.* **90**, 7607–7617 (2016).
414. Geijtenbeek, T. B. H. *et al.* DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells. *Cell* **100**, 587–597 (2000).
415. Trybala, E. *et al.* Glycosaminoglycan-binding ability is a feature of wild-type strains of herpes simplex virus type 1. *Virology* **302**, 413–419 (2002).
416. Patel, M. *et al.* Cell-Surface Heparan Sulfate Proteoglycan Mediates HIV-1 Infection of T-Cell Lines. *AIDS Res. Hum. Retroviruses* **9**, 167–174 (1993).
417. Huskens, D., Vermeire, K., Vandemeulebroucke, E., Balzarini, J. & Schols, D. Safety concerns for the potential use of cyanovirin-N as a microbicidal anti-HIV agent. *Int. J. Biochem. Cell Biol.* **40**, 2802–2814 (2008).
418. Akkouh, O. *et al.* Lectins with Anti-HIV Activity: A Review. *Molecules* **20**, 648–668 (2015).
419. O'Keefe, B. R. *et al.* Engineering soya bean seeds as a scalable platform to produce cyanovirin-N, a non-ARV microbicide against HIV. *Plant Biotechnol. J.* **13**, 884–892 (2015).
420. Matei, E. *et al.* Structure and glycan binding of a new Cyanovirin-N homolog. (2016). doi:10.1074/jbc.M116.740415
421. Pusch, O. *et al.* Bioengineering lactic acid bacteria to secrete the HIV-1 virucide cyanovirin. *J. Acquir. Immune Defic. Syndr.* **40**, 512–520 (2005).
422. Dennis, J. W. & Brewer, C. F. Density-dependent lectin-glycan interactions as a paradigm for conditional regulation by posttranslational modifications. *Molecular and Cellular Proteomics* **12**, 913–920 (2013).
423. Chandra, N. Common Scaffolds, Diverse Recognition Profiles. *Structure* **14**, 1093–1094 (2006).
424. Barre, A., Bourne, Y., Van Damme, E. & Rougé, P. Overview of the Structure–Function Relationships of Mannose-Specific Lectins from Plants, Algae and Fungi. *Int. J. Mol. Sci.* **20**, 254 (2019).
425. Mazalovska, M. & Kouokam, J. C. Lectins as Promising Therapeutics for the Prevention and Treatment of HIV and Other Potential Coinfections. **2018**, (2018).
426. Hoorelbeke, B. *et al.* Differences in the mannose oligomer specificities of the closely related lectins from Galanthus nivalis and Zea mays strongly determine their eventual anti-HIV activity. *Retrovirology* **8**, 10 (2011).
427. Nakamura-Tsuruta, S. *et al.* Analysis of the sugar-binding specificity of mannose-binding-type Jacalin-related lectins by frontal affinity chromatography - An approach to functional classification. *FEBS J.* **275**, 1227–1239 (2008).
428. Balzarini, J. *et al.* Profile of Resistance of Human Immunodeficiency Virus to Mannose-Specific Plant Lectins. *J. Virol.* **78**, 10617–10627 (2004).
429. O'Keefe, B. R. *et al.* Broad-spectrum in vitro activity and in vivo efficacy of the antiviral protein griffithsin against emerging viruses of the family Coronaviridae. *J. Virol.* **84**, 2511–21 (2010).
430. Balzarini, J. *et al.* Mutational pathways, resistance profile, and side effects of cyanovirin relative to human immunodeficiency virus type 1 strains with N-glycan deletions in their gp120 envelopes. *J. Virol.* **80**, 8411–21 (2006).
431. Alexandre, K. B. *et al.* Mechanisms of HIV-1 subtype C resistance to GRFT, CV-N and SVN. *Virology* **446**, 66–76 (2013).
432. Alexandre, K. B. *et al.* The lectins griffithsin, cyanovirin-N and scytovirin inhibit HIV-1 binding to the DC-SIGN receptor and transfer to CD4+ cells. *Virology* **423**, 175–186 (2012).
433. Lusvarghi, S. *et al.* Binding Site Geometry and Subdomain Valency Control

Effects of Neutralizing Lectins on HIV-1 Viral Particles. *ACS Infect. Dis.* **2**, 882–891 (2016).
434. Balzarini, J. *et al.* Mannose-Specific Plant Lectins from the Amaryllidaceae Family Qualify as Efficient Microbicides for Prevention of Human Immunodeficiency Virus Infection. *Antimicrob. Agents Chemother.* **48**, 3858–3870 (2004).
435. Jan, M., Upadhyay, C., Sharma, A., Hioe, C. E. & Arora, S. K. Manα1-2Man binding anti-HIV lectins enhance the exposure of V2i and V3 crown neutralization epitopes on the V1V2 and V3 hypervariable loops of HIV-1 Envelope. *AIDS Res. Hum. Retroviruses* AID.2016.0262 (2017). doi:10.1089/AID.2016.0262
436. Raval, S., Gowda, S. B., Singh, D. D. & Chandra, N. R. A database analysis of jacalin-like lectins: Sequence-structure-function relationships. *Glycobiology* **14**, 1247–1263 (2004).
437. Johannssen, T. & Lepenies, B. Glycan-Based Cell Targeting To Modulate Immune Responses. *Trends Biotechnol.* (2016). doi:10.1016/j.tibtech.2016.10.002
438. Botos, I. & Wlodawer, A. *Proteins that bind high-mannose sugars of the HIV envelope*. *Progress in Biophysics and Molecular Biology* **88**, (2005).
439. Hopper, J. T. S. *et al.* The Tetrameric Plant Lectin BanLec Neutralizes HIV through Bidentate Binding to Specific Viral Glycans. *Structure* **46**, 233–239 (2017).
440. Mahomoodally, M. F. *et al.* Marine Algae : A Potential Resource of Anti-HSV Molecules. *Process Biochem.* **1**, 1–18 (2019).
441. Gondim, A. C. S. *et al.* Potent antiviral activity of carbohydrate-specific algal and leguminous lectins from the Brazilian biodiversity. *Medchemcomm* **10**, 390–398 (2019).
442. Smee, D. F. *et al.* Treatment of influenza A (H1N1) virus infections in mice and ferrets with cyanovirin-N. *Antiviral Res.* **80**, 266–271 (2008).
443. Millet, J. K. *et al.* Middle East respiratory syndrome coronavirus infection is inhibited by griffithsin. *Antiviral Res.* **133**, 1–8 (2016).
444. Keyaerts, E. *et al.* Plant lectins are potent inhibitors of coronaviruses by interfering with two targets in the viral replication cycle. *Antiviral Res.* **75**, 179–187 (2007).
445. Nixon, B. *et al.* Griffithsin Protects Mice from Genital Herpes by Preventing Cell-to-Cell Spread. *J. Virol.* **87**, 6257–6269 (2013).
446. Lal, M. *et al.* Development of a Vaginal Fast-Dissolving Insert Combining Griffithsin and Carrageenan for Potential Use Against Sexually Transmitted Infections. *J. Pharm. Sci.* **107**, 2601–2610 (2018).
447. Levendosky, K. *et al.* Griffithsin and carrageenan combination to target HSV-2 and HPV. *Antimicrob Agents Chemother* **59**, 7290–7298 (2015).
448. Derby, N. *et al.* Griffithsin carrageenan fast dissolving inserts prevent SHIV HSV-2 and HPV infections in vivo. *Nat. Commun.* **9**, 3881 (doi: 10.1038/s41467-018-06349–0) (2018).
449. Guo, Y. *et al.* Lectin microarray and mass spectrometric analysis of hepatitis C proteins reveals N-linked glycosylation. *Medicine (Baltimore)*. **97**, e0208 (2018).
450. Helle, F. *et al.* Cyanovirin-N inhibits hepatitis C virus entry by binding to envelope protein glycans. *J. Biol. Chem.* **281**, 25177–25183 (2006).
451. Bertaux, C. *et al.* Entry of hepatitis C virus and human immunodeficiency virus is selectively inhibited by carbohydrate-binding agents but not by polyanions. *Virology* **366**, 40–50 (2007).
452. Cao, X. *et al.* Inhibition on hepatitis B virus in vitro of lectin from Musca domestica pupa via the activation of NF-κB. *Virus Res.* **170**, 53–58 (2012).

453. He, M., Su, D., Liu, Q., Gao, W. & Kang, Y. Mushroom lectin overcomes hepatitis B virus tolerance via TLR6 signaling. *Sci. Rep.* **7**, (2017).
454. Martin, B., Hoenen, T., Canard, B. & Decroly, E. Filovirus proteins for antiviral drug discovery: A structure/function analysis of surface glycoproteins and virus entry. *Antiviral Res.* (2016). doi:10.1016/j.antiviral.2016.09.001
455. Barrientos, L. G., Lasala, F., Otero, J. R., Sanchez, A. & Delgado, R. In Vitro Evaluation of Cyanovirin-N Antiviral Activity, by Use of Lentiviral Vectors Pseudotyped with Filovirus Envelope Glycoproteins. *J. Infect. Dis.* **189**, 1440–1443 (2004).
456. SCOPe 2.07: Fold b.78: beta-Prism II. Available at: http://scop.berkeley.edu/sunid=51109. (Accessed: 23rd January 2020)
457. SCOPe 2.07: Fold b.77: beta-Prism I. Available at: http://scop.berkeley.edu/sunid=51091. (Accessed: 23rd January 2020)
458. Agardh, C. A. *Synopsis algarum Scandinaviae*. (Ex officina Berlingiana, 1817).
459. M.D. Guiry in Guiry, M.D. & Guiry, G. M. Algaebase. *World-wide electronic publication, National University of Ireland* (2019). Available at: https://www.algaebase.org/search/genus/detail/?genus_id=32770&-session=abv4:AC1F0365164532ECFAhOF0565F75. (Accessed: 12th September 2019)
460. O' Carra, P. Purification and N-terminal Analysis of Algal Biliproteins. *Biochem. J.* **94**, 171–4 (1965).
461. Kaixian, Q., Franklin, M. & Borowitzka, M. A. The study for isolation and purification of R-phycoerythrin from a red alga. *Appl. Biochem. Biotechnol.* **43**, 133–139 (1993).
462. O'Keefe*, B. R. Biologically Active Proteins from Natural Product Extracts1. (2001). doi:10.1021/NP0103362
463. Boyd, M. R. *et al.* Discovery of cyanovirin-N, a novel human immunodeficiency virus-inactivating protein that binds viral surface envelope glycoprotein gp120: potential applications to microbicide development. *Antimicrob. Agents Chemother.* **41**, 1521–30 (1997).
464. Bokesch, H. R. *et al.* A Potent Novel Anti-HIV Protein from the Cultured Cyanobacterium Scytonema varium. *Biochemistry* **42**, 2578–2584 (2003).
465. Ziółkowska, N. E., Shenoy, S. R., O'Keefe, B. R. & Wlodawer, A. Crystallographic studies of the complexes of antiviral protein griffithsin with glucose and N-acetylglucosamine. *Protein Sci.* **16**, 1485–1489 (2007).
466. Mori, T. *et al.* Isolation and characterization of Griffithsin, a novel HIV-inactivating protein, from the red alga Griffithsia sp. *J. Biol. Chem.* **280**, 9345–9353 (2005).
467. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
468. Morgulis, A. *et al.* Database indexing for production MegaBLAST searches. in *Bioinformatics* **24**, 1757–1764 (2008).
469. McGinnis, S. & Madden, T. L. BLAST: At the core of a powerful and diverse set of sequence analysis tools. *Nucleic Acids Res.* **32**, (2004).
470. Johnson, M. *et al.* NCBI BLAST: a better web interface. *Nucleic Acids Res.* **36**, (2008).
471. Mouquet, H. & Nussenzweig, M. C. Polyreactive antibodies in adaptive immune responses to viruses. *Cell. Mol. Life Sci.* **69**, 1435–1445 (2012).
472. Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Research* **25**, 3389–3402 (1997).

473. Barre, A. *et al.* Artocarpin is a polyspecific jacalin-related lectin with a monosaccharide preference for mannose. *Biochimie* **86**, 685–691 (2004).
474. Moulaei, T. *et al.* Griffithsin tandemers: flexible and potent lectin inhibitors of the human immunodeficiency virus. *Retrovirology* **12**, 6 (doi: 10.1186/s12977-014-0127–3) (2015).
475. Hoorelbeke, B., Xue, J., LiWang, P. J. & Balzarini, J. Role of the Carbohydrate-Binding Sites of Griffithsin in the Prevention of DC-SIGN-Mediated Capture and Transmission of HIV-1. *PLoS One* **8**, 1–10 (2013).
476. Xue, J. *et al.* The role of individual carbohydrate-binding sites in the function of the potent anti-HIV lectin griffithsin. *Mol. Pharm.* **9**, 2613–2625 (2012).
477. Xue, J. *et al.* The griffithsin dimer is required for high-potency inhibition of HIV-1: evidence for manipulation of the structure of gp120 as part of the griffithsin dimer mechanism. *Antimicrob. Agents Chemother.* **57**, 3976–3989 (2013).
478. Moulaei, T. *et al.* Monomerization of viral entry inhibitor griffithsin elucidates the relationship between multivalent binding to carbohydrates and anti-HIV activity. *Structure* **18**, 1104–1115 (2010).
479. Ziółkowska, N. E. *et al.* Crystallographic, thermodynamic, and molecular modeling studies of the mode of binding of oligosaccharides to the potent antiviral protein griffithsin. *Proteins Struct. Funct. Bioinforma.* **67**, 661–670 (2007).
480. Alexandre, K. B. *et al.* Binding of the Mannose-Specific Lectin, Griffithsin, to HIV-1 gp120 Exposes the CD4-Binding Site. *J. Virol.* **85**, 9039–9050 (2011).
481. Hahn, S., Giritch, A., Bartels, D., Bortesi, L. & Gleba, Y. A novel and fully scalable *Agrobacterium* spray-based process for manufacturing cellulases and other cost-sensitive proteins in plants. *Plant Biotechnol. J.* **13**, 708–716 (2015).
482. Huang, X., Jin, W., Griffin, G. E., Shattock, R. J. & Hu, Q. Removal of two high-mannose N-linked glycans on gp120 renders human immunodeficiency virus 1 largely resistant to the carbohydrate-binding agent griffithsin. *J. Gen. Virol.* **92**, 2367–2373 (2011).
483. Hu, B. *et al.* Sensitivity of transmitted and founder HIV-1 envelopes to carbohydrate binding agents griffithsin, cyanovirin-N and Galanthus nivalis agglutinin. *J. Gen. Virol.* 3660–3666 (2015). doi:10.1099/jgv.0.000299
484. Barton, C. *et al.* Activity of and effect of subcutaneous treatment with the broad-Spectrum antiviral lectin griffithsin in two laboratory rodent models. *Antimicrob. Agents Chemother.* **58**, 120–127 (2014).
485. Giomarelli, B. *et al.* Recombinant production of anti-HIV protein, griffithsin, by auto-induction in a fermentor culture. *Protein Expr. Purif.* **47**, 194–202 (2006).
486. Vamvaka, E. *et al.* Rice endosperm is cost-effective for the production of recombinant griffithsin with potent activity against HIV. *Plant Biotechnol. J.* **14**, 1427–37 (2016).
487. Fuqua, J. L., Hamorsky, K., Khalsa, G., Matoba, N. & Palmer, K. E. Bulk production of the antiviral lectin griffithsin. *Plant Biotechnol. J.* **13**, 1160–1168 (2015).
488. Vafaee, Y., Staniek, A., Mancheno-Solano, M. & Warzecha, H. A modular cloning toolbox for the generation of chloroplast transformation vectors. *PLoS One* **9**, e110222 (2014).
489. O'Keefe, B. R. *et al.* Scaleable manufacture of HIV-1 entry inhibitor griffithsin and validation of its safety and efficacy as a topical microbicide component. *Proc. Natl. Acad. Sci.* **106**, 6099–6104 (2009).
490. Hoelscher, M. *et al.* High-level expression of the HIV entry inhibitor griffithsin from

the plastid genome and retention of biological activity in dried tobacco leaves. *Plant Mol. Biol.* **0**, 0 (2018).
491. Kouokam, J. C., Lasnik, A. B. & Palmer, K. E. Studies in a murine model confirm the safety of griffithsin and advocate its further development as a microbicide targeting HIV-1 and other enveloped viruses. *Viruses* **8**, 4–5 (2016).
492. Barton, C., Kouokam, J. C., Hurst, H. & Palmer, K. E. Pharmacokinetics of the Antiviral Lectin Griffithsin Administered by Different Routes Indicates Multiple Potential Uses. *Viruses* **8**, v8120331 (2016).
493. Girard, L. *et al.* Impact of the griffithsin anti-HIV microbicide and placebo gels on the rectal mucosal proteome and microbiome in non-human primates. *Sci. Rep.* **8**, 8059 (doi: 10.1038/s41598-018-26313–8) (2018).
494. Férir, G., Palmer, K. E. & Schols, D. Synergistic activity profile of griffithsin in combination with tenofovir, maraviroc and enfuvirtide against HIV-1 clade C. *Virology* **417**, 253–258 (2011).
495. Alexandre, K. B. *et al.* Mannose-rich glycosylation patterns on HIV-1 subtype C gp120 and sensitivity to the lectins, Griffithsin, Cyanovirin-N and Scytovirin. *Virology* **402**, 187–196 (2010).
496. Zhang, L. *et al.* Stabilization and sustained release of HIV inhibitors by encapsulation in silk fibroin disks. *ACS Biomater. Sci. Eng.* **3**, 1654–1665 (2017).
497. Palmer, K. *Griffithsin-based Rectal Microbicides for PREvention of Viral ENTry (PREVENT)*. *Available at: http://grantome.com/grant/NIH/U19-AI113182-04* (2014).
498. Ziółkowska, N. E. *et al.* Domain-Swapped Structure of the Potent Antiviral Protein Griffithsin and Its Mode of Carbohydrate Binding. *Structure* **14**, 1127–1135 (2006).
499. Nixon, B. *et al.* Griffithsin protects mice from genital herpes by preventing cell-to-cell spread. *J. Virol.* **87**, 6257–69 (2013).
500. Ishag, H. Z. A. *et al.* Griffithsin inhibits Japanese encephalitis virus infection in vitro and in vivo. *Arch. Virol.* **158**, 349–358 (2013).
501. Lo, M. K. *et al.* Griffithsin Inhibits Nipah Virus Entry and Fusion and Can Protect Syrian Golden Hamsters From Lethal Nipah Virus Challenge. *J. Infect. Dis.* 1–13 (2020). doi:10.1093/infdis/jiz630
502. Klatt, N. R. *et al.* Vaginal bacteria modify HIV tenofovir microbicide efficacy in African women. **945**, 938–945 (2017).
503. Grooms, T. N. *et al.* Griffithsin-Modified Electrospun Fibers as a Delivery Scaffold to Prevent HIV Infection. *Antimicrob. Agents Chemother.* **60**, 6518–6531 (2016).
504. Vuong, H. R., Tyo, K. M. & Steinbach-Rankins, J. M. Fabrication and Characterization of Griffithsin-modified Fiber Scaffolds for Prevention of Sexually Transmitted Infections. *J. Vis. Exp.* **31**, e56492 (2017).
505. Tyo, K. M. *et al.* Sustained-release Griffithsin nanoparticle-fiber composites against HIV-1 and HSV-2 infections. *J. Control. Release* (2020). doi:10.1016/j.jconrel.2020.02.006
506. Yavuz, B. *et al.* Sustained release silk fibroin discs: Antibody and protein delivery for HIV prevention. *J. Control. Release* **301**, 1–12 (2019).
507. Yavuz, B., Chambre, L. & Kaplan, D. L. Extended release formulations using silk proteins for controlled delivery of therapeutics. *Expert Opin. Drug Deliv.* **16**, 741–756 (2019).
508. Poku, N. K. HIV Prevention: The Key to Ending AIDS by 2030. *Open AIDS J.* **10**, 65–77 (2016).
509. Clutter, D. S., Jordan, M. R., Bertagnolio, S. & Shafer, R. W. HIV-1 drug

resistance and resistance testing. *Infect. Genet. Evol.* (2016). doi:10.1016/j.meegid.2016.08.031
510. Health - United Nations Sustainable Development. Available at: https://www.un.org/sustainabledevelopment/health/. (Accessed: 20th February 2020)
511. UNAIDS. *Global AIDS Update 2016*. (2016). doi:10.1073/pnas.86.15.5781
512. Bulm-Mhisoun, ~, Heiroann, N., Na, C. & Aontllitho, N. *PERMANENT MISSION OF THE REPUBLIC OF KENYA TO TIIE UNITED NATIONS*. (2015).
513. Bekker, L. G. *et al.* The complex challenges of HIV vaccine development require renewed and expanded global commitment. *The Lancet* **395**, 384–388 (2020).
514. Vekemans, J. *et al.* HIV immunoprophylaxis: preparing the pathway from proof of concept to policy decision and use. *The Lancet HIV* **7**, e141–e148 (2020).
515. Moodley, K. *et al.* The psychology of 'cure' - Unique challenges to consent processes in HIV cure research in South Africa. *BMC Med. Ethics* **20**, 9 (2019).
516. Adamson, B., Dimitrov, D., Devine, B. & Barnabas, R. The Potential Cost-Effectiveness of HIV Vaccines: A Systematic Review. *PharmacoEconomics - Open* **1**, 1–12 (2017).
517. Harmon, T. M. *et al.* Exploring the Potential Health Impact and Cost-Effectiveness of AIDS Vaccine within a Comprehensive HIV/AIDS Response in Low- and Middle-Income Countries. *PLoS One* **11**, e0146387 (2016).
518. Vax Report - Vax Report. Available at: https://www.vaxreport.org/index.php?option=com_content&view=article&id=695&Itemid=884. (Accessed: 20th February 2020)
519. Louie, R. H. Y., Kaczorowski, K. J., Barton, J. P., Chakraborty, A. K. & McKay, M. R. Fitness landscape of the human immunodeficiency virus envelope protein that is targeted by antibodies. *Proc. Natl. Acad. Sci. U. S. A.* 201717765 (2018). doi:10.1073/pnas.1717765115
520. Hemelaar, J. *et al.* Global and regional molecular epidemiology of HIV-1, 1990-2015: a systematic review, global survey, and trend analysis. *Lancet. Infect. Dis.* **19**, 143–155 (2019).
521. Robinson, H. L. HIV/AIDS Vaccines: 2018. *Clinical Pharmacology and Therapeutics* **104**, 1062–1073 (2018).
522. Safrit, J. T. *et al.* Status of vaccine research and development of vaccines for HIV-1. *Vaccine* **34**, 2921–2925 (2016).
523. Leal, L. *et al.* New challenges in therapeutic vaccines against HIV infection. *Expert Rev. Vaccines* **16**, 587–600 (2017).
524. Burton, D. R. Advancing an HIV vaccine; advancing vaccinology. *Nat. Rev. Immunol.* **19**, 77–78 (2019).
525. Torrents de la Peña, A. *et al.* Similarities and differences between native HIV-1 envelope glycoprotein trimers and stabilized soluble trimer mimetics. *PLOS Pathog.* **15**, e1007920 (2019).
526. Ringel, O. *et al.* The Hard Way towards an Antibody-Based HIV-1 Env Vaccine: Lessons from Other Viruses. *Viruses* **10**, 197 (2018).
527. Pancera, M., Changela, A. & Kwong, P. D. How HIV-1 entry mechanism and broadly neutralizing antibodies guide structure-based vaccine design. *Curr. Opin. HIV AIDS* 1 (2017). doi:10.1097/COH.0000000000000360
528. The Rise of Broadly Neutralizing Antibodies | AVAC. Available at: https://www.avac.org/blog/rise-broadly-neutralizing-antibodies. (Accessed: 20th February 2020)

529. Schoofs, T. *et al.* Broad and Potent Neutralizing Antibodies Recognize the Silent Face of the HIV Envelope. *Immunity* **50**, 1513–1529 (2019).
530. Liu, J. *et al.* Antibody-mediated protection against SHIV challenge includes systemic clearance of distal virus. *Science (80-.).* **353**, (2016).
531. Montefiori, D. C., Roederer, M., Morris, L. & Seaman, M. S. Neutralization tiers of HIV-1. *Current Opinion in HIV and AIDS* **13**, (2018).
532. Wei, X. *et al.* Antibody neutralization and escape by HIV-1. *Nature* **422**, 307–312 (2003).
533. Guttman, M. *et al.* Antibody potency relates to the ability to recognize the closed, pre-fusion form of HIV Env. *Nature* **6**, (2015).
534. Guenaga, J. *et al.* Structure-Guided Redesign Increases the Propensity of HIV Env To Generate Highly Stable Soluble Trimers. *J. Virol.* **90**, 2806–17 (2015).
535. van Gils, M. J. & Sanders, R. W. Hitting HIV's Harpoon. *Immunity* **49**, (2018).
536. Lewis, G. K., Pazgier, M. & DeVico, A. L. Survivors Remorse: antibody-mediated protection against HIV-1. *Immunological Reviews* **275**, (2017).
537. Kong, L. *et al.* Supersite of immune vulnerability on the glycosylated face of HIV-1 envelope glycoprotein gp120. *Nat. Struct. Mol. Biol.* **20**, 796–803 (2013).
538. Burton, D. R. & Mascola, J. R. Antibody responses to envelope glycoproteins in HIV-1 infection. *Nat. Immunol.* **16**, 571–576 (2015).
539. Flyak, A. I. *et al.* Broadly neutralizing antibodies from human survivors target a conserved site in the ebola virus glycoprotein hr2-mper region. *Nat. Microbiol.* **3**, 670–677 (2018).
540. Crispin, M., Ward, A. B. & Wilson, I. A. Structure and Immune Recognition of the HIV Glycan Shield. *Annu. Rev. Biophys.* **47**, (2018).
541. Kouyos, R. D. *et al.* Tracing HIV-1 strains that imprint broadly neutralizing antibody responses. *Nature* 1 (2018). doi:10.1038/s41586-018-0517-0
542. Doria-Rose, N. A. *et al.* Frequency and Phenotype of Human Immunodeficiency Virus Envelope-Specific B Cells from Patients with Broadly Cross-Neutralizing Antibodies. *J. Virol.* **83**, 188–199 (2009).
543. Sather, D. N. *et al.* Factors Associated with the Development of Cross-Reactive Neutralizing Antibodies during Human Immunodeficiency Virus Type 1 Infection. *J. Virol.* **83**, 757–769 (2009).
544. Fauci, A. Host factors and the pathogenesis of HIV-induced disease. *Nature* **384**, (1996).
545. Sather, D. N. *et al.* Emergence of Broadly Neutralizing Antibodies and Viral Coevolution in Two Subjects during the Early Stages of Infection with Human Immunodeficiency Virus Type 1. *J. Virol.* **88**, 12968–12981 (2014).
546. Verkoczy, L., Alt, F. W. & Tian, M. Human Ig knockin mice to study the development and regulation of HIV-1 broadly neutralizing antibodies. *Immunol. Rev.* **275**, 89–107 (2017).
547. Steichen, J. M. *et al.* HIV Vaccine Design to Target Germline Precursors of Glycan-Dependent Broadly Neutralizing Antibodies. *Immunity* **45**, 483–496 (2016).
548. Briney, B. *et al.* Tailored Immunogens Direct Affinity Maturation toward HIV Neutralizing Antibodies. *Cell* **166**, 1459-1470.e11 (2016).
549. 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.* **12**, (2016).
550. Subbaraman, H., Schanz, M. & Trkola, A. Broadly neutralizing antibodies: What is

needed to move from a rare event in HIV-1 infection to vaccine efficacy? *Retrovirology* **15**, 1–14 (2018).
551. CDC & Ncird. *Understanding How Vaccines Work The Immune System-The Body's Defense Against Infection*.
552. How Vaccines Work | PublicHealth.org. Available at: https://www.publichealth.org/public-awareness/understanding-vaccines/vaccines-work/. (Accessed: 20th February 2020)
553. Vaccine Basics - How Vaccines Work. Available at: https://vaccineinformation.org/how-vaccines-work/. (Accessed: 20th February 2020)
554. How Vaccines Work | History of Vaccines. Available at: https://www.historyofvaccines.org/content/how-vaccines-work. (Accessed: 20th February 2020)
555. Basics of Vaccines | CDC. Available at: https://www.cdc.gov/vaccines/vpd/vpd-vac-basics.html. (Accessed: 20th February 2020)
556. Pritchard, L. K. *et al.* Structural Constraints Determine the Glycosylation of HIV-1 Envelope Trimers. *Cell Rep.* **11**, 1604–1613 (2015).
557. Pugach, P. *et al.* A native-like SOSIP.664 trimer based on an HIV-1 subtype B env gene. *J. Virol.* **89**, 3380–95 (2015).
558. Julien, J.-P. *et al.* Design and structure of two HIV-1 clade C SOSIP.664 trimers that increase the arsenal of native-like Env immunogens. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 11947–52 (2015).
559. Georgiev, I. S. *et al.* Single-Chain Soluble BG505.SOSIP gp140 Trimers as Structural and Antigenic Mimics of Mature Closed HIV-1 Env. *J. Virol.* **89**, 5318–29 (2015).
560. Joyce, M. G. *et al.* Soluble Prefusion Closed DS-SOSIP.664-Env Trimers of Diverse HIV-1 Strains. *Cell Rep.* **21**, 2992–3002 (2017).
561. Ringe, R. P. *et al.* Improving the expression and purification of soluble, recombinant native-like HIV-1 envelope glycoprotein trimers by targeted sequence changes. *J. Virol.* JVI.00264-17 (2017). doi:10.1128/JVI.00264-17
562. Sliepen, K. *et al.* Structure and immunogenicity of a stabilized HIV-1 envelope trimer based on a group-M consensus sequence. *Nat. Commun.* **10**, (2019).
563. Ingale, J. *et al.* High-Density Array of Well-Ordered HIV-1 Spikes on Synthetic Liposomal Nanoparticles Efficiently Activate B Cells. *Cell Rep.* **15**, 1986–1999 (2016).
564. Guenaga, J. *et al.* Well-Ordered Trimeric HIV-1 Subtype B and C Soluble Spike Mimetics Generated by Negative Selection Display Native-like Properties. *PLoS Pathog.* **11**, 1–16 (2015).
565. Liang, Y. *et al.* Changes in structure and antigenicity of HIV-1 Env trimers resulting from removal of a conserved CD4 binding site-proximal glycan. *J. Virol.* JVI.01116-16 (2016). doi:10.1128/JVI.01116-16
566. Go, E. P. *et al.* Glycosylation site-specific analysis of clade C HIV-1 envelope proteins. *J. Proteome Res.* **8**, 4231–4242 (2009).
567. Zhou, T. *et al.* Structural definition of a conserved neutralization epitope on HIV-1 gp120. **445**, (2007).
568. Qualls, Z. M. *et al.* Identification of novel structural determinants in MW965 Env that regulate the neutralization phenotype and conformational masking potential of primary HIV-1 isolates. *J. Virol.* JVI.01779-17 (2017). doi:10.1128/JVI.01779-17
569. Duan, H. *et al.* Glycan Masking Focuses Immune Responses to the HIV-1 CD4-

Binding Site and Enhances Elicitation of VRC01-Class Precursor Antibodies. *Immunity* **49**, 301-311.e5 (2018).
570. 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.* **100**, 15812–7 (2003).
571. Sliepen, K. *et al.* Binding of inferred germline precursors of broadly neutralizing HIV-1 antibodies to native-like envelope trimers. *Virology* **486**, 116–120 (2015).
572. Kesavardhana, S. *et al.* Structure based design of cyclically permuted HIV-1 gp120 trimers that elicit neutralizing antibodies. *J. Biol. Chem.* jbc.M116.725614 (2016). doi:10.1074/jbc.M116.725614
573. Koch, M. *et al.* Structure-based, targeted deglycosylation of HIV-1 gp120 and effects on neutralization sensitivity and antibody recognition. *Virology* **313**, 387–400 (2003).
574. Kulp, D. W. *et al.* Structure-based design of native-like HIV-1 envelope trimers to silence non-neutralizing epitopes and eliminate CD4 binding. *Nat. Commun.* **8**, 1655 (2017).
575. Zhang, P. *et al.* Interdomain Stabilization Impairs CD4 Binding and Improves Immunogenicity of the HIV-1 Envelope Trimer. *Cell Host Microbe* **23**, 832-844.e6 (2018).
576. Gautam, R. *et al.* A single injection of anti-HIV-1 antibodies protects against repeated SHIV challenges. *Nature* **533**, 105–109 (2016).
577. Moldt, B. *et al.* Neutralizing antibody affords comparable protection against vaginal and rectal simian/human immunodeficiency virus challenge in macaques. *AIDS* **30**, 1543–51 (2016).
578. Todd, C. A. *et al.* Development and implementation of an international proficiency testing program for a neutralizing antibody assay for HIV-1 in TZM-bl cells. *J. Immunol. Methods* **375**, 57–67 (2012).
579. Escolano, A., Dosenovic, P. & Nussenzweig, M. C. Progress toward active or passive HIV-1 vaccination. *J. Exp. Med* **214**, 3–16 (2017).
580. Huang, Y. *et al.* Selection of HIV vaccine candidates for concurrent testing in an efficacy trial. *Curr. Opin. Virol.* **17**, 57–65 (2016).
581. Wagh, K. *et al.* Optimal Combinations of Broadly Neutralizing Antibodies for Prevention and Treatment of HIV-1 Clade C Infection. *PLoS Pathog.* **12**, (2016).
582. Halper-Stromberg, A. & Nussenzweig, M. C. Towards HIV-1 remission: Potential roles for broadly neutralizing antibodies. *Journal of Clinical Investigation* **126**, (2016).
583. Bar-On, Y. *et al.* Safety and antiviral activity of combination HIV-1 broadly neutralizing antibodies in viremic individuals. *Nature Medicine* **24**, (2018).
584. Sun, Z. *et al.* Brief introduction of current technologies in isolation of broadly neutralizing HIV-1 antibodies. *Virus Res.* **243**, 75–82 (2018).
585. Horwitz, J. A. *et al.* HIV-1 suppression and durable control by combining single broadly neutralizing antibodies and antiretroviral drugs in humanized mice. *Proc. Natl. Acad. Sci.* **110**, (2013).
586. Kumar, R., Qureshi, H., Deshpande, S. & Bhattacharya, J. Broadly neutralizing antibodies in HIV-1 treatment and prevention. *Ther. Adv. Vaccines Immunother.* **6**, 61–68 (2018).
587. Bhiman, J. N. & Lynch, R. M. Broadly Neutralizing Antibodies as Treatment: Effects on Virus and Immune System. *Curr. HIV/AIDS Rep.* **14**, 54–62 (2017).
588. Cohen, Y. Z. *et al.* Neutralizing activity of broadly neutralizing anti-HIV-1

antibodies against clade B clinical isolates produced in peripheral blood mononuclear cells. *J. Virol.* JVI.01883-17 (2017). doi:10.1128/JVI.01883-17
589. Caskey, M., Klein, F. & Nussenzweig, M. C. Broadly neutralizing anti-HIV-1 monoclonal antibodies in the clinic. *Nat. Med.* **25**, 547–553 (2019).
590. Stephenson, K. E. & Barouch, D. H. Broadly Neutralizing Antibodies for HIV Eradication. *Current HIV/AIDS Reports* **13**, (2016).
591. Pejchal, R. *et al.* Structure and function of broadly reactive antibody PG16 reveal an H3 subdomain that mediates potent neutralization of HIV-1. *Proc. Natl. Acad. Sci.* **107**, 11483–11488 (2010).
592. Andrabi, R. *et al.* Glycans Function as Anchors for Antibodies and Help Drive HIV Broadly Neutralizing Antibody Development. *Immunity* **47**, 524-537.e3 (2017).
593. Lynch, R. M. *et al.* HIV-1 fitness cost associated with escape from the VRC01 class of CD4 binding site neutralizing antibodies. *J. Virol.* **89**, 4201–13 (2015).
594. Scheid, J. F. *et al.* Sequence and structural convergence of broad and potent HIV antibodies that mimic CD4 binding. *Science* **333**, 1633–7 (2011).
595. Scott, Y. & Dezzutti, C. S. Non-Antiretroviral Microbicides for HIV Prevention. *AIDS Rev.* **18**, (2016).
596. McGuire, A. T. *et al.* Engineering HIV envelope protein to activate germline B cell receptors of broadly neutralizing anti-CD4 binding site antibodies. *J. Exp. Med.* **210**, (2013).
597. Freund, N. T. *et al.* A New Glycan-Dependent CD4-Binding Site Neutralizing Antibody Exerts Pressure on HIV-1 In Vivo. *PLoS Pathog.* **11**, (2015).
598. Alam, M. *et al.* Enhanced antibody-mediated neutralization of HIV-1 variants that are resistant to fusion inhibitors. *Retrovirology* **13**, 70 (2016).
599. Riddler, S. A. *et al.* Randomized clinical trial to assess the impact of the broadly neutralizing HIV-1 monoclonal antibody VRC01 on HIV-1 persistence in individuals on effective ART. *Open Forum Infect. Dis.* **5**, ofy242 (2018).
600. Wu, X. *et al.* Focused evolution of HIV-1 neutralizing antibodies revealed by structures and deep sequencing. *Science* **333**, 1593–602 (2011).
601. Wu, X. *et al.* Maturation and Diversity of the VRC01-Antibody Lineage over 15 Years of Chronic HIV-1 Infection. *Cell* **161**, 470–485 (2015).
602. Chun, T. W. *et al.* Early establishment of a pool of latently infected, resting CD4+ T cells during primary HIV-1 infection. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 8869–8873 (1998).
603. Chun, T.-W. *et al.* Presence of an inducible HIV-1 latent reservoir during highly active antiretroviral therapy. *Proc. Natl. Acad. Sci.* **94**, 13193–13197 (1997).
604. Chun, T. W. *et al.* In vivo fate of HIV-1-infected T cells: Quantitative analysis of the transition to stable latency. *Nat. Med.* **1**, 1284–1290 (1995).
605. Yavuz, B. *et al.* Pharmaceutical Approaches to HIV Treatment and Prevention. *Adv. Ther.* **1**, 1800054 (2018).
606. Saison, J., Cotte, L., Chidiac, C. & Ferry, T. Fatal cumulative toxicities of HAART in a stable, AIDS-free, HIV-infected patient. *BMJ Case Rep.* **2012**, (2012).
607. Flash, C., Krakower, D. & Mayer, K. H. The Promise of Antiretrovirals for HIV Prevention. *Curr. Infect. Dis. Rep.* **14**, 185–93 (2012).
608. Schwartz, C. *et al.* On the way to find a cure: Purging latent HIV-1 reservoirs. *Biochemical Pharmacology* (2017). doi:10.1016/j.bcp.2017.07.001
609. Ryom, L. *et al.* Use of antiretroviral therapy and risk of end-stage liver disease and hepatocellular carcinoma in HIV-positive persons. *AIDS* **30**, 1731–1743 (2016).

610. Barton, K., Winckelmann, A. & Palmer, S. HIV-1 Reservoirs During Suppressive Therapy. *Trends in Microbiology* **24**, 345–355 (2016).
611. Hamlyn, E. *et al.* Plasma HIV Viral Rebound following Protocol-Indicated Cessation of ART Commenced in Primary and Chronic HIV Infection. *PLoS One* **7**, e43754 (2012).
612. Zhang, L. *et al.* Quantifying residual HIV-1 replication in patients receiving combination antiretroviral therapy. *N. Engl. J. Med.* **340**, 1605–1613 (1999).
613. Finzi, D. *et al.* Latent infection of CD4+ T cells provides a mechanism for lifelong persistence of HIV-1, even in patients on effective combination therapy. *Nat. Med.* **5**, 512–517 (1999).
614. Blankson, J. N., Persaud, D. & Siliciano, R. F. The Challenge of Viral Reservoirs in HIV-1 Infection. *Annu. Rev. Med.* **53**, 557–593 (2002).
615. Ahmad, A. & Rinaldo, C. R. A novel anti-HIV immunotherapy to cure HIV. *Aids* **31**, 447–449 (2017).
616. Rezaei, S. D. *et al.* The Pathway To Establishing HIV Latency Is Critical to How Latency Is Maintained and Reversed. *J. Virol.* **92**, e02225-17 (2018).
617. Hu, W.-S. & Hughes, S. H. HIV-1 reverse transcription. *Cold Spring Harb. Perspect. Med.* **2**, a006882 (2012).
618. Ballandras-Colas, A. *et al.* A supramolecular assembly mediates lentiviral DNA integration. *Science (80-.).* **355**, 7002–6 (2017).
619. Craigie, R. & Bushman, F. D. HIV DNA integration. *Cold Spring Harb. Perspect. Med.* **2**, a006890 (2012).
620. Burnett, J. C., Miller-Jensen, K., Shah, P. S., Arkin, A. P. & Schaffer, D. V. Control of stochastic gene expression by host factors at the HIV promoter. *PLoS Pathog.* **5**, (2009).
621. Burnett, J. C. *et al.* Combinatorial latency reactivation for HIV-1 subtypes and variants. *J. Virol.* **84**, 5958–74 (2010).
622. Hataye, J. M. *et al.* A Scalable RNA RT-PCR Based Assay to Quantify Latency Disruption in HIV-1 Infected Cells | CROI Conference. in *Conference on Retroviruses and Opportunistic Infections* 399LB (2014).
623. Ho, Y. C. *et al.* Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure. *Cell* **155**, 540 (2013).
624. Yu, Q. *et al.* Single-strand specificity of APOBEC3G accounts for minus-strand deamination of the HIV genome. *Nat. Struct. Mol. Biol.* **11**, 435–442 (2004).
625. Archin, N. M., Sung, J. M., Garrido, C., Soriano-Sarabia, N. & Margolis, D. M. Eradicating HIV-1 infection: Seeking to clear a persistent pathogen. *Nature Reviews Microbiology* **12**, 750–764 (2014).
626. Méndez, C., Ledger, S., Petoumenos, K., Ahlenstiel, C. & Kelleher, A. D. RNA-induced epigenetic silencing inhibits HIV-1 reactivation from latency. *Retrovirology* **15**, (2018).
627. Sadowski, I. & Hashemi, F. B. Strategies to eradicate HIV from infected patients: elimination of latent provirus reservoirs. *Cellular and Molecular Life Sciences* **76**, 3583–3600 (2019).
628. Kimpton, J. & Emerman, M. Detection of replication-competent and pseudotyped human immunodeficiency virus with a sensitive cell line on the basis of activation of an integrated beta-galactosidase gene. *J. Virol.* **66**, 2232–2239 (1992).
629. Weinberger, L. S., Dar, R. D. & Simpson, M. L. Transient-mediated fate determination in a transcriptional circuit of HIV. *Nat. Genet.* **40**, 466–470 (2008).
630. Singh, A., Razooky, B., Cox, C. D., Simpson, M. L. & Weinberger, L. S.

Transcriptional bursting from the HIV-1 promoter is a significant source of stochastic noise in HIV-1 gene expression. *Biophys. J.* **98**, L32-4 (2010).
631. Hosmane, N. N. *et al.* Proliferation of latently infected CD4+ T cells carrying replication-competent HIV-1: Potential role in latent reservoir dynamics. *J. Exp. Med.* **214**, 959–972 (2017).
632. Vanhamel, J., Bruggemans, A. & Debyser, Z. Establishment of latent HIV-1 reservoirs: what do we really know? *J. virus Erad.* **5**, 3–9 (2019).
633. Chomont, N. *et al.* HIV reservoir size and persistence are driven by T cell survival and homeostatic proliferation. *Nat. Med.* **15**, 893–900 (2009).
634. Musick, A. *et al.* HIV Infected T Cells Can Proliferate in vivo Without Inducing Expression of the Integrated Provirus. *Front. Microbiol.* **10**, 2204 (2019).
635. Vanhamel, J., Bruggemans, A. & Debyser, Z. Establishment of latent HIV-1 reservoirs: what do we really know? *J. virus Erad.* **5**, 3–9 (2019).
636. Kulpa, D. A. & Chomont, N. HIV persistence in the setting of antiretroviral therapy: when, where and how does HIV hide? *J. virus Erad.* **1**, 59–66 (2015).
637. Chomont, N. *et al.* HIV reservoir size and persistence are driven by T cell survival and homeostatic proliferation. *Nat. Med.* **15**, 893–900 (2009).
638. Coffin, J. & Swanstrom, R. HIV pathogenesis: Dynamics and genetics of viral populations and infected cells. *Cold Spring Harb. Perspect. Med.* **3**, a012526 (2013).
639. Khochenkov, D. A. Biology of dendritic cells. *Biologicheskie Membrany* **25**, 403–419 (2008).
640. McDonald, D. *et al.* Recruitment of HIV and its receptors to dendritic cell-T cell junctions. *Science (80-.).* **300**, 1295–1297 (2003).
641. Coulon, P.-G. *et al.* HIV-Infected Dendritic Cells Present Endogenous MHC Class II-Restricted Antigens to HIV-Specific CD4+ T Cells. *J. Immunol.* (2016). doi:10.4049/jimmunol.1600286
642. Weissman, D. *et al.* Three populations of cells with dendritic morphology exist in peripheral blood, only one of which is infectable with human immunodeficiency virus type 1 (infection/dendritic cell/precursor). *Immunology* **92**, 826–830 (1995).
643. Pope, M. *et al.* Conjugates of Dendritic Cells and Memory T Lymphocytes from Skin Facilitate Productive Infection with HIV-1. *Cell* **78**, 389–398 (1994).
644. Rowland-Jones, S. L. HIV: The deadly passenger in dendritic cells. *Current Biology* **9**, R248–R250 (1999).
645. Mitsuki, Y. -y., Tuen, M. & Hioe, C. E. Differential effects of HIV transmission from monocyte-derived dendritic cells vs. monocytes to IL-17+CD4+ T cells. *J. Leukoc. Biol.* jlb.4A0516-216R (2016). doi:10.1189/jlb.4A0516-216R
646. Nath, A. Eradication of human immunodeficiency virus from brain reservoirs. *J. Neurovirol.* **21**, 227–234 (2015).
647. Stevenson, M. Role of myeloid cells in HIV-1-host interplay. *J. Neurovirol.* **21**, 242–248 (2015).
648. Haldar, M. & Murphy, K. M. Origin, development, and homeostasis of tissue-resident macrophages. *Immunol. Rev.* **262**, 25–35 (2014).
649. Stevenson, M. CROI 2018 : Advances in Basic Science Understanding of HIV. **26**, (2018).
650. Castro-Gonzalez, S., Colomer-Lluch, M. & Serra-Moreno, R. Barriers for a HIV cure: the latent reservoir. *AIDS Res. Hum. Retroviruses* AID.2018.0118 (2018). doi:10.1089/AID.2018.0118
651. Rouzioux, C. & Richman, D. How to best measure HIV reservoirs? *Current*

Opinion in HIV and AIDS **8**, 170–175 (2013).
652. Cary, D. C., Fujinaga, K. & Peterlin, B. M. Molecular mechanisms of HIV latency. *Journal of Clinical Investigation* **126**, 448–454 (2016).
653. Spina, C. A. *et al.* An In-Depth Comparison of Latent HIV-1 Reactivation in Multiple Cell Model Systems and Resting CD4+ T Cells from Aviremic Patients. *PLoS Pathog.* **9**, 1–15 (2013).
654. van Nierop, K. & de Groot, C. Human follicular dendritic cells: function, origin and development. *Semin. Immunol.* **14**, 251–7 (2002).
655. King, D. F. L. *et al.* Mucosal Tissue Tropism and Dissemination of HIV-1 Subtype B Acute Envelope-Expressing Chimeric Virus. *J. Virol.* **87**, 890–899 (2013).
656. DiNapoli, S. R., Hirsch, V. M. & Brenchley, J. M. Macrophages in progressive HIV/SIV infections. *J. Virol.* JVI.00672-16 (2016). doi:10.1128/JVI.00672-16
657. Woodham, A. W. *et al.* Human Immunodeficiency Virus Immune Cell Receptors, Coreceptors, and Cofactors: Implications for Prevention and Treatment. *http://www.liebertpub.com/apc* (2016). doi:10.1089/apc.2016.0100
658. Mzingwane, M. L. & Tiemessen, C. T. Mechanisms of HIV persistence in HIV reservoirs. *Reviews in Medical Virology* **27**, (2017).
659. Wong, J. K. & Yukl, S. A. Tissue reservoirs of HIV. *Current Opinion in HIV and AIDS* **11**, 362–370 (2016).
660. Henderson, L. J., Reoma, L. B., Kovacs, J. A. & Nath, A. Advances toward Curing HIV-1 Infection in Tissue Reservoirs. *J. Virol.* **94**, (2019).
661. Ait-Ammar, A. *et al.* Current Status of Latency Reversing Agents Facing the Heterogeneity of HIV-1 Cellular and Tissue Reservoirs. *Frontiers in Microbiology* **10**, (2020).
662. Rose, R. *et al.* Eradication of HIV from Tissue Reservoirs: Challenges for the Cure. *AIDS Res. Hum. Retroviruses* **34**, 3–8 (2018).
663. Deleage, C. *et al.* Defining HIV and SIV Reservoirs in Lymphoid Tissues. *Pathog. Immun.* **1**, 68 (2016).
664. Garcia-Tellez, T. *et al.* Non-human primates in HIV research: Achievements, limits and alternatives. *Infect. Genet. Evol.* (2016). doi:10.1016/j.meegid.2016.07.012
665. Darcis, G., Van Driessche, B. & Van Lint, C. HIV Latency: Should We Shock or Lock? *Trends Immunol.* **38**, 217–228 (2017).
666. Mousseau, G. *et al.* Resistance to the Tat Inhibitor Didehydro-Cortistatin A Is Mediated by Heightened Basal HIV-1 Transcription. *MBio* **10**, (2019).
667. Li, C., Mousseau, G. & Valente, S. T. Tat inhibition by didehydro-Cortistatin A promotes heterochromatin formation at the HIV-1 long terminal repeat. *Epigenetics and Chromatin* **12**, 23 (2019).
668. Kessing, C. F. *et al.* In Vivo Suppression of HIV Rebound by Didehydro-Cortistatin A, a "Block-and-Lock" Strategy for HIV-1 Treatment. *Cell Rep.* **21**, 600–611 (2017).
669. Sarabia, I. & Bosque, A. HIV-1 Latency and Latency Reversal : Does Subtype Matter ? *Viruses* **11**, 1–29 (2019).
670. Shirakawa, K., Chavez, L., Hakre, S., Calvanese, V. & Verdin, E. Reactivation of latent HIV by histone deacetylase inhibitors. *Trends Microbiol.* **21**, 277–285 (2013).
671. Hashemi, P. *et al.* Compounds producing an effective combinatorial regimen for disruption of HIV-1 latency. *EMBO Mol. Med.* e201708193 (2017). doi:10.15252/emmm.201708193
672. Perreau, M., Banga, R. & Pantaleo, G. Targeted Immune Interventions for an HIV-

1 Cure. *Trends Mol. Med.* **23**, 945–961 (2017).
673. Deeks, S. G. HIV: Shock and kill. *Nature* **487**, 439–440 (2012).
674. Marsden, M. D. & Zack, J. A. HIV cure strategies: A complex approach for a complicated viral reservoir? *Future Virology* **14**, 5–8 (2019).
675. Elsheikh, M. M., Tang, Y., Li, D. & Jiang, G. Deep latency: A new insight into a functional HIV cure. *EBioMedicine* **45**, 624–629 (2019).
676. Delannoy, A., Poirier, M. & Bell, B. Cat and mouse: HIV transcription in latency, immune evasion and cure/remission strategies. *Viruses* **11**, (2019).
677. Yin, C. *et al.* In Vivo Excision of HIV-1 Provirus by saCas9 and Multiplex Single-Guide RNAs in Animal Models. *Mol. Ther.* **25**, 1168–1186 (2017).
678. Limsirichai, P., Gaj, T. & Schaffer, D. V. CRISPR-mediated Activation of Latent HIV-1 Expression. *Mol. Ther.* **24**, 499–507 (2016).
679. Silas, S. *et al.* Direct CRISPR spacer acquisition from RNA by a natural reverse transcriptase-Cas1 fusion protein. *Science (80-.).* **351**, aad4234–aad4234 (2016).
680. Amsterdam, D. Unique natural and adaptive response mechanisms to control and eradicate HIV infection. *AIMS Allergy Immunol.* **2**, 113–125 (2018).
681. Wang, H., La Russa, M. & Qi, L. S. CRISPR/Cas9 in Genome Editing and Beyond. *Annu. Rev. Biochem.* **85**, 227–264 (2016).
682. Pegu, A. *et al.* Activation and lysis of human CD4 cells latently infected with HIV-1. *Nat. Commun.* **6**, 8447 (2015).
683. Kumar, N., Chahroudi, A. & Silvestri, G. Animal models to achieve an HIV cure. *Current Opinion in HIV and AIDS* **11**, 432–441 (2016).
684. Takata, H. *et al.* Modeling HIV-1 Latency Using Primary CD4 + T Cells from Virally Suppressed HIV-1-Infected Individuals on Antiretroviral Therapy . *J. Virol.* **93**, (2019).
685. Flemming, A. Shocking HIV out of hiding. *Nature Reviews Immunology* **20**, 138–139 (2020).
686. Llewellyn, G. N. *et al.* Humanized Mouse Model of HIV-1 Latency with Enrichment of Latent Virus in PD-1 + and TIGIT + CD4 T Cells . *J. Virol.* **93**, (2019).
687. Nixon, C. C. *et al.* Systemic HIV and SIV latency reversal via non-canonical NF-κB signalling in vivo. *Nature* **578**, 160–165 (2020).
688. McBrien, J. B. *et al.* Robust and persistent reactivation of SIV and HIV by N-803 and depletion of CD8+ cells. *Nature* **578**, 154–159 (2020).
689. Van Praag, R. M. E. *et al.* OKT3 and IL-2 treatment for purging of the latent HIV-1 reservoir in vivo results in selective long-lasting CD4+ T cell depletion. *J. Clin. Immunol.* **21**, 218–226 (2001).
690. Gama, L. *et al.* Reactivation of simian immunodeficiency virus reservoirs in the brain of virally suppressed macaques. *AIDS* **31**, 5–14 (2017).
691. Knights, H. D. J. A Critical Review of the Evidence Concerning the HIV Latency Reversing Effect of Disulfiram, the Possible Explanations for Its Inability to Reduce the Size of the Latent Reservoir In Vivo, and the Caveats Associated with Its Use in Practice. *AIDS Res. Treat.* **2017**, 8239428 (2017).
692. Kim, Y., Anderson, J. L. & Lewin, S. R. Getting the “Kill” into “Shock and Kill”: Strategies to Eliminate Latent HIV. *Cell Host and Microbe* **23**, 14–26 (2018).
693. Bouchat, S. *et al.* Histone methyltransferase inhibitors induce HIV-1 recovery in resting CD4+ T cells from HIV-1-infected HAART-treated patients. *AIDS* **26**, 1473–1482 (2012).
694. Macedo, A. B., Novis, C. L. & Bosque, A. Targeting Cellular and Tissue HIV

Reservoirs With Toll-Like Receptor Agonists. *Frontiers in Immunology* **10**, 2450 (2019).
695. Tsai, A. *et al.* Toll-Like Receptor 7 Agonist GS-9620 Induces HIV Expression and HIV-Specific Immunity in Cells from HIV-Infected Individuals on Suppressive Antiretroviral Therapy. *J. Virol.* **91**, (2017).
696. Appleman, J. R. & Webber, S. E. Discovery of a novel Toll-like Receptor 7 agonist for systemic immunotherapy of cancer. *J. Clin. Oncol.* **37**, e14246–e14246 (2019).
697. Chi, H. *et al.* Anti-tumor activity of toll-like receptor 7 agonists. *Frontiers in Pharmacology* **8**, (2017).
698. Macedo, A. B. *et al.* Dual TLR2 and TLR7 agonists as HIV latency-reversing agents. *JCI insight* **3**, (2018).
699. Archin, N. M. *et al.* Interval dosing with the HDAC inhibitor vorinostat effectively reverses HIV latency. *J. Clin. Invest.* **127**, 3126–3135 (2017).
700. Margolis, D. M. Histone deacetylase inhibitors and HIV latency. *Curr. Opin. HIV AIDS* **6**, 25–29 (2011).
701. Beliakova-Bethell, N. *et al.* Histone deacetylase inhibitors induce complex host responses that contribute to differential potencies of these compounds in HIV reactivation. *J. Biol. Chem.* **294**, 5576–5589 (2019).
702. Zaikos, T. D., Painter, M. M., Sebastian Kettinger, N. T., Terry, V. H. & Collins, K. L. Class 1-Selective Histone Deacetylase (HDAC) Inhibitors Enhance HIV Latency Reversal while Preserving the Activity of HDAC Isoforms Necessary for Maximal HIV Gene Expression. *J. Virol.* **92**, (2018).
703. Chan, J. K. L. & Greene, W. C. NF-κB/Rel: Agonist and antagonist roles in HIV-1 latency. *Curr. Opin. HIV AIDS* **6**, 12–18 (2011).
704. An, W. Histone acetylation and methylation: combinatorial players for transcriptional regulation. *Subcell. Biochem.* **41**, 351–69 (2007).
705. Gilmore, T. D. Introduction to NF-κB: players, pathways, perspectives. *Oncogene* **25**, 6680–6684 (2006).
706. Ott, M., Geyer, M. & Zhou, Q. The control of HIV transcription: Keeping RNA polymerase II on track. *Cell Host Microbe* **10**, 426–435 (2011).
707. Bayer, P. *et al.* Structural studies of HIV-1 Tat protein. *J. Mol. Biol.* **247**, 529–35 (1995).
708. Bigalke, J. M. *et al.* Formation of Tatâ€“TAR containing ribonucleoprotein complexes for biochemical and structural analyses. (2011). doi:10.1016/j.ymeth.2010.04.001
709. Gu, J. *et al.* Crystal structure of HIV-1 Tat complexed with human P-TEFb and AFF4. *Cell Cycle* **13**, 1788–1797 (2014).
710. Gatignol, A. Transcription of HIV: Tat and Cellular Chromatin. *Advances in Pharmacology* **55**, 137–159 (2007).
711. Easley, R. *et al.* Chromatin dynamics associated with HIV-1 Tat-activated transcription. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms* **1799**, 275–285 (2010).
712. Siddappa, N. B. *et al.* Transactivation and signaling functions of Tat are not correlated: biological and immunological characterization of HIV-1 subtype-C Tat protein. *Retrovirology* **3**, 53 (2006).
713. Jeang, K. T., Xiao, H. & Rich, E. A. Multifaceted activities of the HIV-1 transactivator of transcription, Tat. *J. Biol. Chem.* **274**, 28837–28840 (1999).
714. Rice, A. P. The HIV-1 Tat Protein: Mechanism of Action and Target for HIV-1 Cure Strategies. *Curr. Pharm. Des.* **23**, 4098 (2017).

715. Romani, B., Engelbrecht, S. & Glashoff, R. H. Functions of Tat: The versatile protein of human immunodeficiency virus type 1. *Journal of General Virology* **91**, 1–12 (2010).
716. Das, A. T., Harwig, A. & Berkhout, B. The HIV-1 Tat Protein Has a Versatile Role in Activating Viral Transcription. *J. Virol.* **85**, 9506–9516 (2011).
717. Donahue, D. A., Kuhl, B. D., Sloan, R. D. & Wainberg, M. A. The Viral Protein Tat Can Inhibit the Establishment of HIV-1 Latency. *J. Virol.* **86**, 3253–3263 (2012).
718. Geng, G. *et al.* Development of an attenuated tat protein as a highly-effective agent to specifically activate HIV-1 latency. *Mol. Ther.* **24**, 1528–1537 (2016).
719. Kamori, D. & Ueno, T. HIV-1 tat and viral latency: What we can learn from naturally occurring sequence variations. *Frontiers in Microbiology* **8**, (2017).
720. Khoury, G. *et al.* HIV latency reversing agents act through Tat post translational modifications. *Retrovirology* **15**, (2018).
721. Mousseau, G. *et al.* The tat inhibitor didehydro-cortistatin a prevents HIV-1 reactivation from latency. *MBio* **6**, (2015).
722. Donahue, D. A., Kuhl, B. D., Sloan, R. D. & Wainberg, M. A. The Viral Protein Tat Can Inhibit the Establishment of HIV-1 Latency. *J. Virol.* **86**, 3253–3263 (2012).
723. Tang, X. *et al.* Exosomal Tat protein activates latent HIV-1 in primary, resting CD4+ T lymphocytes. *JCI insight* **3**, (2018).
724. Donahue, D. A., Bastarache, S. M., Sloan, R. D. & Wainberg, M. A. Latent HIV-1 can be reactivated by cellular superinfection in a Tat-dependent manner, which can lead to the emergence of multidrug-resistant recombinant viruses. *J. Virol.* **87**, 9620–32 (2013).
725. Razooky, B. S., Pai, A., Aull, K., Rouzine, I. M. & Weinberger, L. S. A hardwired HIV latency program. *Cell* **160**, 990–1001 (2015).
726. Bagashev, A. & Sawaya, B. E. Roles and functions of HIV-1 Tat protein in the CNS: an overview. *Virol. J.* **10**, 358 (2013).
727. King, J. E., Eugenin, E. A., Buckner, C. M. & Berman, J. W. HIV tat and neurotoxicity. *Microbes and Infection* **8**, 1347–1357 (2006).
728. WHO. Progress report on HIV, viral hepatitis and sexually transmitted infections, 2019. 2016–2021 (2019).
729. Marrazzo, J. M. *et al.* Tenofovir-Based Preexposure Prophylaxis for HIV Infection among African Women. *N. Engl. J. Med.* **372**, 509–518 (2015).
730. Smith, J. M. *et al.* Intravaginal ring eluting tenofovir disoproxil fumarate completely protects macaques from multiple vaginal simian-HIV challenges. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 16145–50 (2013).
731. Nel, A. *et al.* Safety and Efficacy of a Dapivirine Vaginal Ring for HIV Prevention in Women. *N. Engl. J. Med.* **375**, 2133–2143 (2016).
732. Baeten, J. M. *et al.* Use of a Vaginal Ring Containing Dapivirine for HIV-1 Prevention in Women. *N. Engl. J. Med.* **375**, 2121–2132 (2016).
733. Gunawardana, M., Baum, M. M., Smith, T. J. & Moss, J. A. An intravaginal ring for the sustained delivery of antibodies. *J. Pharm. Sci.* **103**, 3611–20 (2014).
734. Wibmer, C. K., Moore, P. L. & Morris, L. HIV broadly neutralizing antibody targets. *Curr. Opin. HIV AIDS* **10**, 135–43 (2015).
735. Huang, J. *et al.* Identification of a CD4-Binding-Site Antibody to HIV that Evolved Near-Pan Neutralization Breadth. *Immunity* **45**, 1108–1121 (2016).
736. Jin, H., Kagiampakis, I., Li, P. & Liwang, P. J. Structural and functional studies of the potent anti-HIV chemokine variant P2-RANTES. *Proteins* **78**, 295–308 (2010).
737. Boyd, M. R. *et al.* Discovery of Cyanovirin-N, a Novel Human Immunodeficiency

Virus-Inactivating Protein That Binds Viral Surface Envelope Glycoprotein gp120: Potential Applications to Microbicide Development. **41**, 1521–1530 (1997).
738. Lasnik, A. Preclinical safety assessment of Griffithsin-based vaginal microbicides. (Masters Thesis. University of Louisville, 2013).
739. Tsai, C.-C. *et al.* Cyanovirin-N Inhibits AIDS Virus Infections in Vaginal Transmission Models. *AIDS Res. Hum. Retroviruses* **20**, 11–18 (2004).
740. Kish-Catalone, T. M., Lu, W., Gallo, R. C. & DeVico, A. L. Preclinical evaluation of synthetic -2 RANTES as a candidate vaginal microbicide to target CCR5. *Antimicrob. Agents Chemother.* **50**, 1497–509 (2006).
741. Fuqua, J. L., Wanga, V. & Palmer, K. E. Improving the large scale purification of the HIV microbicide, griffithsin. *BMC Biotechnol.* **15**, 12 (doi: 10.1186/s12896-015-0120–5) (2015).
742. Kouokam, J. C. *et al.* Investigation of Griffithsin's Interactions with Human Cells Confirms Its Outstanding Safety and Efficacy Profile as a Microbicide Candidate. *PLoS One* **6**, e22635 (2011).
743. Balzarini, J. *et al.* Targeting the glycans of gp120: a novel approach aimed at the Achilles heel of HIV. *Lancet. Infect. Dis.* **5**, 726–31 (2005).
744. Moncla, B. J., Pryke, K., Rohan, L. C. & Graebing, P. W. Degradation of naturally occurring and engineered antimicrobial peptides by proteases. *Adv. Biosci. Biotechnol.* **2**, 404–408 (2011).
745. Emau, P. *et al.* Griffithsin, a potent HIV entry inhibitor, is an excellent candidate for anti-HIV microbicide. *J. Med. Primatol.* **36**, 244–253 (2007).
746. Lal, M. *et al.* Self-administered Griffithsin and Carrageenan Containing Microbicide Fast-dissolving Insert as Pre-exposure Prophylaxis Against HIV and HPV Infections. in *Abstracts of the HIV Research for Prevention Meeting, HIVR4P, 21-25 October, 2018, Madrid* **34**, 1–407 (Mary Ann Liebert, Inc., publishers 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA, 2018).
747. Hamorsky, K. T. *et al.* Efficient single tobamoviral vector-based bioproduction of broadly neutralizing anti-HIV-1 monoclonal antibody VRC01 in Nicotiana benthamiana plants and utility of VRC01 in combination microbicides. *Antimicrob. Agents Chemother.* **57**, 2076–86 (2013).
748. Lusvarghi, S. & Bewley, C. Griffithsin: An Antiviral Lectin with Outstanding Therapeutic Potential. *Viruses* **8**, 296 (doi: 10.3390/v8100296) (2016).
749. Center, R. J. *et al.* Oligomeric structure of virion-associated and soluble forms of the simian immunodeficiency virus envelope protein in the prefusion activated conformation. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 14877–82 (2001).
750. Go, E. P. *et al.* Glycosylation site-specific analysis of HIV envelope proteins (JR-FL and CON-S) reveals major differences in glycosylation site occupancy, glycoform profiles, and antigenic epitopes' accessibility. *J. Proteome Res.* **7**, 1660–1674 (2008).
751. Stanfield, R. *et al.* Dual conformations for the HIV-1 gp120 V3 loop in complexes with different neutralizing Fabs. *Structure* **7**, 131–142 (1999).
752. Ma, X. *et al.* HIV-1 Env trimer opens through an asymmetric intermediate in which individual protomers adopt distinct conformations. *Elife* **7**, e34271 (2018).
753. Julg, B. *et al.* Protective Efficacy of Broadly Neutralizing Antibodies with Incomplete Neutralization Activity Against SHIV in Rhesus Monkeys. *J. Virol.* **91**, e01187-17 (2017).
754. Badamchi-Zadeh, A. *et al.* Therapeutic efficacy of vectored PGT121 gene delivery in HIV-1-infected humanized mice. *J. Virol.* **92**, e01925-17 (2018).

755. Barouch, D. H. *et al.* Therapeutic Efficacy of Potent Neutralizing HIV-1-Specific Monoclonal Antibodies in SHIV-Infected Rhesus Monkeys. *Nature* **503**, 224–228 (2013).
756. Stephenson, K. E. *et al.* Therapeutic Activity of PGT121 Monoclonal Antibody in HIV-Infected Adults. in *Conference on Retroviruses and Opportunistic Infections* (CROI Foundation/IAS-USA, 2019).
757. Moore, P. L. *et al.* Evolution of an HIV glycan – dependent broadly neutralizing antibody epitope through immune escape. *Nat Med* **18**, 1688–1692 (2012).
758. Rademeyer, C. *et al.* Features of Recently Transmitted HIV-1 Clade C Viruses that Impact Antibody Recognition: Implications for Active and Passive Immunization. *PLOS Pathog.* **12**, e1005742 (2016).
759. Bouvin-Pley, M. *et al.* Evidence for a Continuous Drift of the HIV-1 Species towards Higher Resistance to Neutralizing Antibodies over the Course of the Epidemic. *PLoS Pathog.* **9**, e1003477 (2013).
760. Garces, F. *et al.* Structural Evolution of Glycan Recognition by a Family of Potent HIV Antibodies. *Cell* **159**, 69–79 (2014).
761. Garces, F. *et al.* Affinity Maturation of a Potent Family of HIV Antibodies Is Primarily Focused on Accommodating or Avoiding Glycans. *Immunity* **43**, 1053–1063 (2015).
762. Huang, X. *et al.* Highly conserved HIV-1 gp120 glycans proximal to CD4-binding region affect viral infectivity and neutralizing antibody induction. *Virology* **423**, 97–106 (2012).
763. Pritchard, L. K. *et al.* Glycan clustering stabilizes the mannose patch of HIV-1 and preserves vulnerability to broadly neutralizing antibodies. *Nat. Commun.* **6**, 7479 (2015).
764. Coss, K. P. *et al.* HIV-1 glycan density drives the persistence of the mannose patch within an infected 1 individual 2 3. *Serv.* doi:10.1128/JVI.01542-16
765. Dam, T. K. & Brewer, C. F. Current Topics Effects of Clustered Epitopes in Multivalent Ligand-Receptor Interactions †. *Biochemistry* **47**, 41 (2008).
766. Zhang, M. *et al.* Tracking global patterns of N-linked glycosylation site variation in highly variable viral glycoproteins: HIV, SIV, and HCV envelopes and influenza hemagglutinin. *Glycobiology* **14**, 1229–1246 (2004).
767. Bolton, D. L. *et al.* Human Immunodeficiency Virus Type 1 Monoclonal Antibodies Suppress Acute Simian-Human Immunodeficiency Virus Viremia and Limit Seeding of Cell-Associated Viral Reservoirs. *J. Virol.* **90**, (2016).
768. Morrow, R. J. *et al.* Sustained release of proteins from a modified vaginal ring device. *Eur. J. Pharm. Biopharm.* **77**, 3–10 (2011).
769. Friend, D. R. Intravaginal rings: controlled release systems for contraception and prevention of transmission of sexually transmitted infections. *Drug Deliv. Transl. Res.* **1**, 185–193 (2011).
770. Férir, G. *et al.* Combinations of Griffithsin with Other Carbohydrate-Binding Agents Demonstrate Superior Activity Against HIV Type 1, HIV Type 2, and Selected Carbohydrate-Binding Agent-Resistant HIV Type 1 Strains. *AIDS Res. Hum. Retroviruses* **28**, 1513–23 (2012).
771. Balzarini, J. *et al.* Marked Depletion of Glycosylation Sites in HIV-1 gp120 under Selection Pressure by the Mannose-Specific Plant Lectins of Hippeastrum Hybrid and Galanthus nivalis. *Mol. Pharmacol.* **67**, 1556–1565 (2005).
772. Balzarini, J. *et al.* Carbohydrate-binding agents cause deletions of highly conserved glycosylation sites in HIV GP120: A new therapeutic concept to hit the

Achilles heel of HIV. *J. Biol. Chem.* **280**, 41005–41014 (2005).
773. Garces, F. *et al.* Structural evolution of HIV-1 gp120 glycan recognition by a family of potent HIV antibodies. *Cell* **159**, 69–79 (2014).
774. Moore, P. L. *et al.* Evolution of an HIV glycan-dependent broadly neutralizing antibody epitope through immune escape. *Nat. Med.* **18**, 1688–92 (2012).
775. Li, M. *et al.* Human Immunodeficiency Virus Type 1 env Clones from Acute and Early Subtype B Infections for Standardized Assessments of Vaccine-Elicited Neutralizing Antibodies. *J. Virol.* **79**, 10108–10125 (2005).
776. Derdeyn, C. A. *et al.* Envelope-constrained neutralization-sensitive HIV-1 after heterosexual transmission. *Science* **303**, 2019–22 (2004).
777. Li, M. *et al.* Genetic and Neutralization Properties of Subtype C Human Immunodeficiency Virus Type 1 Molecular env Clones from Acute and Early Heterosexually Acquired Infections in Southern Africa. *J. Virol.* **80**, 11776–11790 (2006).
778. Li, Y. *et al.* Characterization of antibody responses elicited by human immunodeficiency virus type 1 primary isolate trimeric and monomeric envelope glycoproteins in selected adjuvants. *J. Virol.* **80**, 1414–26 (2006).
779. Platt, E. J., Bilska, M., Kozak, S. L., Kabat, D. & Montefiori, D. C. Evidence that ecotropic murine leukemia virus contamination in TZM-bl cells does not affect the outcome of neutralizing antibody assays with human immunodeficiency virus type 1. *J. Virol.* **83**, 8289–92 (2009).
780. Takeuchi, Y., McClure, M. O. & Pizzato, M. Identification of gammaretroviruses constitutively released from cell lines used for human immunodeficiency virus research. *J. Virol.* **82**, 12585–8 (2008).
781. 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.* **74**, 8358–67 (2000).
782. Platt, E. J., Wehrly, K., Kuhmann, S. E., Chesebro, B. & Kabat, D. Effects of CCR5 and CD4 Cell Surface Concentrations on Infections by Macrophagetropic Isolates of Human Immunodeficiency Virus Type 1 Effects of CCR5 and CD4 Cell Surface Concentrations on Infections by Macrophagetropic Isolates of Human Immunodeficiency. *J. Virol.* **72**, 2855–2864 (1998).
783. Graham, F. L., Smiley, J., Russell, W. C. & Nairn, R. Characteristics of a Human Cell Line Transformed by DNA from Human Adenovirus Type 5. *J. Gen. Virol.* **36**, 59–72 (1977).
784. Harrison, T., Graham, F. & Williams, J. Host-range mutants of adenovirus type 5 defective for growth in HeLa cells. *Virology* **77**, 319–329 (1977).
785. Naldini, L., Blömer, U., Gage, F. H., Trono, D. & Verma, I. M. Efficient transfer, integration, and sustained long-term expression of the transgene in adult rat brains injected with a lentiviral vector. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 11382–8 (1996).
786. Burton, D. R. *et al.* A large array of human monoclonal antibodies to type 1 human immunodeficiency virus from combinatorial libraries of asymptomatic seropositive individuals. *Proc. Natl. Acad. Sci. USA* **88**, 10134–7 (1991).
787. Barbas, C. F. *et al.* Recombinant human Fab fragments neutralize human type 1 immunodeficiency virus in vitro. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 9339–43 (1992).
788. Roben, P. *et al.* Recognition properties of a panel of human recombinant Fab fragments to the CD4 binding site of gp120 that show differing abilities to

neutralize human immunodeficiency virus type 1. *J. Virol.* **68**, 4821–8 (1994).
789. Burton, D. *et al.* Efficient neutralization of primary isolates of HIV-1 by a recombinant human monoclonal antibody. *Science (80-.).* **266**, 1024–1027 (1994).
790. R Core Team. R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria* (2013). Available at: http://www.r-project.org/. (Accessed: 23rd September 2019)
791. Chou, T.-C. & Talalay, P. Quantitative analysis of dose-effect relationships: the combined effects of multiple drugs or enzyme inhibitors. *Adv. Enzyme Regul.* **22**, 27–55 (1984).
792. Chou, T. C. Drug combination studies and their synergy quantification using the chou-talalay method. *Cancer Research* **70**, 440–446 (2010).
793. Malik, T., Chauhan, G., Rath, G., Murthy, R. S. R. & Goyal, A. K. "Fusion and binding inhibition" key target for HIV-1 treatment and pre-exposure prophylaxis: targets, drug delivery and nanotechnology approaches. *Drug Deliv.* **24**, 608–621 (2017).
794. García, M., Buzón, M. J., Benito, J. M. & Rallón, N. Peering into the HIV reservoir. *Rev. Med. Virol.* e1981 (2018). doi:10.1002/rmv.1981
795. Agosto, L. M., Herring, M. B., Mothes, W. & Henderson, A. J. HIV-1-Infected CD4+ T Cells Facilitate Latent Infection of Resting CD4+ T Cells through Cell-Cell Contact. *Cell Rep.* **24**, 2088–2100 (2018).
796. Agosto, L. M. *et al.* HIV-1 integrates into resting CD4+ T cells even at low inoculums as demonstrated with an improved assay for HIV-1 integration. *Virology* **368**, 60–72 (2007).
797. Jordan, A., Defechereux, P. & Verdin, E. The site of HIV-1 integration in the human genome determines basal transcriptional activity and response to Tat transactivation. *EMBO J.* **20**, 1726–1738 (2001).
798. König, R. *et al.* Global Analysis of Host-Pathogen Interactions that Regulate Early-Stage HIV-1 Replication. *Cell* **135**, 49–60 (2008).
799. Wagner, T. A. *et al.* Proliferation of cells with HIV integrated into cancer genes contributes to persistent infection. *Science (80-.).* **345**, 570–573 (2014).
800. Maldarelli, F. *et al.* Specific HIV integration sites are linked to clonal expansion and persistence of infected cells. *Science (80-.).* **345**, 179–183 (2014).
801. Bosque, A. & Planelles, V. Induction of HIV-1 latency and reactivation in primary memory CD4+ T cells. *Blood* **113**, 58–65 (2009).
802. Khan, S., Telwatte, S., Trapecar, M., Yukl, S. & Sanjabi, S. Differentiating Immune Cell Targets in Gut-Associated Lymphoid Tissue for HIV Cure. *AIDS Res. Hum. Retroviruses* AID.2017.0153 (2017). doi:10.1089/AID.2017.0153
803. Dahabieh, M. S., Battivelli, E. & Verdin, E. Understanding HIV Latency: The Road to an HIV Cure. *Annu. Rev. Med.* **66**, 407–421 (2015).
804. Soriano-Sarabia, N. *et al.* Quantitation of replication-competent HIV-1 in populations of resting CD4+ T cells. *J. Virol.* **88**, 14070–7 (2014).
805. von Stockenstrom, S. *et al.* Longitudinal Genetic Characterization Reveals That Cell Proliferation Maintains a Persistent HIV Type 1 DNA Pool During Effective HIV Therapy. *J. Infect. Dis.* **212**, 596–607 (2015).
806. Sengupta, S. & Siliciano, R. F. Targeting the Latent Reservoir for HIV-1. *Immunity* **48**, 872–895 (2018).
807. Eriksson, S. *et al.* Comparative Analysis of Measures of Viral Reservoirs in HIV-1 Eradication Studies. *PLoS Pathog.* **9**, (2013).

808. Zhang, M. *et al.* Regulatory CD4 T cells inhibit HIV-1 expression of other CD4 T cell subsets via interactions with cell surface regulatory proteins. *Virology* **516**, 21–29 (2018).
809. Siliciano, J. D. *et al.* Long-term follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting CD4+ T cells. *Nat. Med.* **9**, 727–728 (2003).
810. Rong, L. & Perelson, A. S. Modeling latently infected cell activation: Viral and latent reservoir persistence, and viral blips in HIV-infected patients on potent therapy. *PLoS Comput. Biol.* **5**, (2009).
811. Hogan, L. E. *et al.* Increased HIV-1 transcriptional activity and infectious burden in peripheral blood and gut-associated CD4+ T cells expressing CD30. *PLOS Pathog.* **14**, e1006856 (2018).
812. Weinberger, L. S., Burnett, J. C., Toettcher, J. E., Arkin, A. P. & Schaffer, D. V. Stochastic Gene Expression in a Lentiviral Positive-Feedback Loop: HIV-1 Tat Fluctuations Drive Phenotypic Diversity. *Cell* **122**, 169–182 (2005).
813. Sloan, D. D. *et al.* Targeting HIV Reservoir in Infected CD4 T Cells by Dual-Affinity Re-targeting Molecules (DARTs) that Bind HIV Envelope and Recruit Cytotoxic T Cells. *PLoS Pathog.* **11**, (2015).
814. Abbas, A. K., Lichtman, A. H. & Pillai, S. Basic immunology: functions and disorders of the immune system. in *Basic Immunology: Functions and Disorders of the Immune System* 303–305 (2012). doi:10.1007/978-1-62703-589-7_1
815. Iglesias-Ussel, M., Vandergeeten, C., Marchionni, L., Chomont, N. & Romerio, F. High Levels of CD2 Expression Identify HIV-1 Latently Infected Resting Memory CD4+ T Cells in Virally Suppressed Subjects. *J. Virol.* **87**, 9148–9158 (2013).
816. Descours, B. *et al.* CD32a is a marker of a CD4 T-cell HIV reservoir harbouring replication-competent proviruses. *Nature* (2017). doi:10.1038/nature21710
817. Dhummakupt, A. *et al.* The Latent Human Immunodeficiency Virus (HIV) Reservoir Resides Primarily in CD32−CD4+ T Cells in Perinatally HIV-Infected Adolescents With Long-Term Virologic Suppression. *J. Infect. Dis.* 1–9 (2018). doi:10.1093/infdis/jiy461
818. Martin, G. E. Enrichment of the HIV reservoir in CD32+ CD4 T cells occurs early and is closely associated with immune checkpoint receptor expression. *IAS HIV Cure Cancer Forum* (2017).
819. García, M. *et al.* CD32 Expression is not Associated to HIV-DNA content in CD4 cell subsets of individuals with Different Levels of HIV Control. *Sci. Rep.* **8**, 1–8 (2018).
820. Bertagnolli, L. N. *et al.* The role of CD32 during HIV-1 infection. *Nature* **561**, E17–E19 (2018).
821. Abdel-Mohsen, M. *et al.* CD32 is expressed on cells with transcriptionally active HIV but does not enrich for HIV DNA in resting T cells. *Sci. Transl. Med.* **10**, (2018).
822. Griffith, J. W., Sokol, C. L. & Luster, A. D. Chemokines and Chemokine Receptors: Positioning Cells for Host Defense and Immunity. *Annu. Rev. Immunol.* **32**, 659–702 (2014).
823. Bönsch, C., Munteanu, M., Rossitto-Borlat, I., Fürstenberg, A. & Hartley, O. Potent anti-HIV chemokine analogs direct post-endocytic sorting of CCR5. *PLoS One* **10**, e0125396 (2015).
824. Gaertner, H. *et al.* Highly potent, fully recombinant anti-HIV chemokines: reengineering a low-cost microbicide. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 17706–11 (2008).

825. Hu, C., Chen, X., Huang, Y. & Chen, Y. Synergistic effect of the pro-apoptosis peptide kla-TAT and the cationic anticancer peptide HPRP-A1. *Apoptosis* **23**, 132–142 (2018).
826. Hyun, S. *et al.* Apoptosis inducing, conformationally constrained, dimeric peptide analogs of KLA with submicromolar cell penetrating abilities. *Biomacromolecules* **15**, 3746–3752 (2014).
827. Chen, W.-H. *et al.* Dual-Targeting Pro-apoptotic Peptide for Programmed Cancer Cell Death via Specific Mitochondria Damage. *Sci. Rep.* **3**, 3468 (2013).
828. Schott, K. & König, R. Picking the Survivor! CRISPR Reveals HIV Dependency Factors. *Trends Microbiol.* (2017). doi:10.1016/j.tim.2017.02.004
829. Kovalevich, J. & Langford, D. Neuronal toxicity in HIV CNS disease. *Future Virol.* **7**, 687–698 (2012).
830. Crakes, K. R. *et al.* PPARα-targeted mitochondrial bioenergetics mediate repair of intestinal barriers at the host–microbe intersection during SIV infection. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 24819–24829 (2019).
831. Wadia, J. S., Stan, R. V & Dowdy, S. F. Transducible TAT-HA fusogenic peptide enhances escape of TAT-fusion proteins after lipid raft macropinocytosis. *Nat. Med.* **10**, 310–315 (2004).
832. Li, Q. *et al.* Oligohistidine and targeting peptide functionalized TAT-NLS for enhancing cellular uptake and promoting angiogenesis in vivo. *J. Nanobiotechnology* **16**, 29 (2018).
833. Brock, D. J. *et al.* Efficient cell delivery mediated by lipid-specific endosomal escape of supercharged branched peptides. *Traffic* **19**, 421–435 (2018).
834. Pellois, J.-P. Mechanisms and optimization of endosomal escape for cell delivery applications.
835. Salerno, J. C. *et al.* Novel cell-penetrating peptide-adaptors effect intracellular delivery and endosomal escape of protein cargos. *J. Cell Sci.* **129**, 893–897 (2016).
836. Bukrinsky, M. I. *et al.* A nuclear localization signal within HIV-1 matrix protein that governs infection of non-dividing cells. *Nature* **365**, 666–669 (1993).
837. Truant, R. & Cullen, B. R. The Arginine-Rich Domains Present in Human Immunodeficiency Virus Type 1 Tat and Rev Function as Direct Importin β-Dependent Nuclear Localization Signals. *Mol. Cell. Biol.* **19**, 1210–1217 (1999).
838. Cardarelli, F., Serresi, M., Bizzarri, R., Giacca, M. & Beltram, F. In vivo study of HIV-1 tat arginine-rich motif unveils its transport properties. *Mol. Ther.* **15**, 1313–1322 (2007).
839. Ragin, A. D., Morgan, R. A. & Chmielewski, J. Cellular import mediated by nuclear localization signal peptide sequences. *Chem. Biol.* **9**, 943–948 (2002).
840. Efthymiadis, A., Briggs, L. J. & Jans, D. A. The HIV-1 tat nuclear localization sequence confers novel nuclear import properties. *J. Biol. Chem.* **273**, 1623–1628 (1998).
841. Smith, K. M., Himiari, Z., Tsimbalyuk, S. & Forwood, J. K. Structural Basis for Importin-α Binding of the Human Immunodeficiency Virus Tat. *Sci. Rep.* **7**, 1650 (2017).
842. Opi, S. *et al.* 1H-13C nuclear magnetic resonance assignment and structural characterization of HIV-1 Tat protein. *Comptes Rendus l'* **323**, 883–94 (2000).
843. Tahirov, T. H. *et al.* Crystal structure of HIV-1 Tat complexed with human P-TEFb. *Nature* **465**, 747–51 (2010).
844. Björndal, A. *et al.* Coreceptor usage of primary human immunodeficiency virus

type 1 isolates varies according to biological phenotype. *J. Virol.* **71**, 7478–87 (1997).
845. Duma, L., Häussinger, D., Rogowski, M., Lusso, P. & Grzesiek, S. Recognition of RANTES by Extracellular Parts of the CCR5 Receptor. *J. Mol. Biol.* **365**, 1063–1075 (2007).
846. Doranz, B. J. *et al.* Two distinct CCR5 domains can mediate coreceptor usage by human immunodeficiency virus type 1. *J. Virol.* **71**, 6305–14 (1997).
847. Wu, Y. & Yoder, A. Chemokine coreceptor signaling in HIV-1 infection and pathogenesis. *PLoS Pathog.* **5**, 1–8 (2009).
848. Mo, H. *et al.* Patterns of the HIV Type 1 Coreceptors Expression HIV-1. *New York* **14**, 607–617 (1998).
849. Hartley, O. *et al.* Human Immunodeficiency Virus Type 1 Entry Inhibitors Selected on Living Cells from a Library of Phage Chemokines. *J. Virol.* **77**, 6637–6644 (2003).
850. Kuo, N.-W. *et al.* Structural Insights into the Interaction between a Potent Anti-inflammatory Protein, Viral CC Chemokine Inhibitor (vCCI), and the Human CC Chemokine, Eotaxin-1. *J. Biol. Chem.* **289**, 6592–6603 (2014).
851. Appay, V., Brown, A., Cribbes, S., Randle, E. & Czaplewski, L. G. Aggregation of RANTES Is Responsible for Its Inflammatory Properties: Characterization of Nonaggregating, Noninflammatory RANTES Mutants. *J. Biol. Chem.* **274**, 27505–27512 (1999).
852. Nedellec, R. *et al.* Resistance to the CCR5 inhibitor 5P12-RANTES requires a difficult evolution from CCR5 to CXCR4 Coreceptor use. *PLoS One* **6**, e22020 (2011).
853. Czaplewski, L. G. *et al.* Identification of amino acid residues critical for aggregation of human CC chemokines macrophage inflammatory protein (MIP)-1alpha, MIP-1beta, and RANTES. Characterization of active disaggregated chemokine variants. *J. Biol. Chem.* **274**, 16077–84 (1999).
854. Kraft, K. *et al.* Characterization of sequence determinants within the carboxyl-terminal domain of chemokine receptor CCR5 that regulate signaling and receptor internalization. *J. Biol. Chem.* **276**, 34408–18 (2001).
855. Mueller, A. *et al.* Pathways for internalization and recycling of the chemokine receptor CCR5. *Blood* **99**, 785–91 (2002).
856. Ochsenbauer-Jambor, C., Jones, J., Heil, M., Zammit, K. P. & Kutsch, O. T-cell line for HIV drug screening using EGFP as a quantitative marker of HIV-1 replication. *Biotechniques* **40**, 91–100 (2006).
857. Kutsch, O. *et al.* Bis-anthracycline antibiotics inhibit human immunodeficiency virus type 1 transcription. *Antimicrob. Agents Chemother.* **48**, 1652–63 (2004).
858. Chackerian, B., Long, E. M., Luciw, P. A. & Overbaugh, J. Human immunodeficiency virus type 1 coreceptors participate in postentry stages in the virus replication cycle and function in simian immunodeficiency virus infection. *J. Virol.* **71**, 3932–3939 (1997).
859. Nguyen, A. F. Purification and Study of CC Chemokine-Based Strategies to Combat Chronic Inflammation and HIV. (University of California, Merced, 2017).
860. Clancy, K. W., Melvin, J. A. & McCafferty, D. G. Sortase transpeptidases: insights into mechanism, substrate specificity, and inhibition. *Biopolymers* **94**, 385–396 (2010).
861. Kirsch, T. *et al.* Cloning, high-yield expression in Escherichia coli, and purification of biologically active HIV-1 Tat protein. *Protein Expr. Purif.* **8**, 75–84 (1996).

862. Patki, A. H. & Lederman, M. M. HIV-1 Tat Protein and Its Inhibitor Ro 24-7429 Inhibit Lymphocyte Proliferation and Induce Apoptosis in Peripheral Blood Mononuclear Cells from Healthy Donors. *Cell. Immunol.* **169**, 40–46 (1996).
863. Rhim, H., Echetebu, C. O., Herrmann, C. H. & Rice, A. P. Wild-type and mutant HIV-1 and HIV-2 Tat proteins expressed in Escherichia coli as fusions with glutathione S-transferase. *J. Acquir. Immune Defic. Syndr.* **7**, 1116–21 (1994).
864. Herrmann, C. H. & Rice, A. P. Lentivirus Tat proteins specifically associate with a cellular protein kinase, TAK, that hyperphosphorylates the carboxyl-terminal domain of the large subunit of RNA polymerase II: candidate for a Tat cofactor. *J. Virol.* **69**, 1612–1620 (1995).
865. Herrmann, C. H. & Rice, A. P. Specific interaction of the human immunodeficiency virus Tat proteins with a cellular protein kinase. *Virology* **197**, 601–608 (1993).
866. Theile, C. S. *et al.* Site-specific N-terminal labeling of proteins using sortase-mediated reactions. *Nat. Protoc.* **8**, 1800–1807 (2013).
867. Popp, M. W. L., Antos, J. M. & Ploegh, H. L. Site-Specific protein labeling via sortase-mediated transpeptidation. *Current Protocols in Protein Science* **CHAPTER 15**, Unit (2009).
868. 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.* **186**, 1373–1381 (1997).
869. Lee, B. *et al.* Epitope mapping of CCR5 reveals multiple conformational states and distinct but overlapping structures involved in chemokine and coreceptor function. *J. Biol. Chem.* **274**, 9617–26 (1999).
870. Bernstone, L., Van Wilgenburg, B. & James, W. Several Commercially Available Anti-CCR5 Monoclonal Antibodies Lack Specificity and Should Be Used with Caution. doi:10.1089/hyb.2010.0092
871. De Bruyne, M. *et al.* Elevated regulatory T cells at diagnosis of Coccidioides infection associates with chronicity in pediatric patients. *J. Allergy Clin. Immunol.* **142**, 1971-1974.e7 (2018).
872. Platt, E. J., Wehrly, K., Kuhmann, S. E., Chesebro, B. & Kabat, D. Effects of CCR5 and CD4 Cell Surface Concentrations on Infections by Macrophagetropic Isolates of Human Immunodeficiency Virus Type 1. *J. Virol.* **72**, 2855–2864 (1998).
873. Sarzotti-Kelsoe, M. *et al.* Optimization and validation of the TZM-bl assay for standardized assessments of neutralizing antibodies against HIV-1. *J. Immunol. Methods* **409**, 131–46 (2014).
874. Ozaki, D. A. *et al.* International technology transfer of a GCLP-compliant HIV-1 neutralizing antibody assay for human clinical trials. *PLoS One* **7**, e30963 (2012).
875. Jin, H., Shen, X., Baggett, B. R., Kong, X. & LiWang, P. J. The human CC chemokine MIP-1beta dimer is not competent to bind to the CCR5 receptor. *J. Biol. Chem.* **282**, 27976–83 (2007).
876. Laurence, J. S., Blanpain, C., Burgner, J. W., Parmentier, M. & LiWang, P. J. CC chemokine MIP-1β can function as a monomer and depends on Phe13 for receptor binding. *Biochemistry* **39**, 3401–3409 (2000).
877. Banerjee, K. *et al.* Occluding the Mannose Moieties on Human Immunodeficiency Virus Type 1 gp120 with Griffithsin Improves the Antibody Responses to Both Proteins in Mice. *AIDS Res. Hum. Retroviruses* **28**, 206–214 (2012).
878. McGowan, I. Project 3: PREVENT: A first-in-human clinical trial of Griffithsin, an HIV entry.

879. Population Council. A Phase 1 Trial to Evaluate the Safety, Pharmacokinetics (PK) and Pharmacodynamics (PD) of PC-6500 (Griffithsin [GRFT] in a Carrageenan Gel) in Healthy Women. *ClinicalTrials.gov* NCT02875119 (2017). Available at: https://www.popcouncil.org/news/19417. (Accessed: 5th February 2019)
880. Study to Evaluate the Safety of Griffithsin in a Carrageenan Gel in Healthy Women - Full Text View - ClinicalTrials.gov. Available at: https://clinicaltrials.gov/ct2/show/NCT02875119. (Accessed: 6th July 2017)
881. Alexandre, B., Buckheit, W., Jr, R. W. B. & Keefe, B. R. O. Investigation of tandemers activities against cell-associated HIV-1 and against viruses with decreased sensitivity to griffithsin. **106**, 9353 (2010).
882. Alexandre, K., Malatji, K. & Mulaudzi, T. Comparison of the antiviral activity of the microbicide candidate griffithsin and its tandemers derivatives against different modes of HIV-1 transmission. *Virology* (2020). doi:10.1016/j.virol.2020.01.017
883. Ochsenbauer-Jambor, C., Jones, J., Heil, M., Zammit, K. P. & Kutsch, O. T-cell line for HIV drug screening using EGFP as a quantitative marker of HIV-1 replication. *Biotechniques* **40**, 91–100 (2006).
884. Feng, T. *et al.* HIV-1 downregulates the expression and phosphorylation of receptor tyrosine kinase by targeting the NF-κB pathway. *Mol. Med. Rep.* **14**, 1947 (2016).
885. Lee, B., Sharron, M., Montaner, L. J., Weissman, D. & Doms, R. W. Quantification of CD4, CCR5, and CXCR4 levels on lymphocyte subsets, dendritic cells, and differentially conditioned monocyte-derived macrophages. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 5215–5220 (1999).
886. Mummidi, S., Ahuja, S. S., McDaniel, B. L. & Ahuja, S. K. The human CC chemokine receptor 5 (CCR5) gene. Multiple transcripts with 5'-end heterogeneity, dual promoter usage, and evidence for polymorphisms within the regulatory regions and noncoding exons. *J. Biol. Chem.* **272**, 30662–30671 (1997).
887. Wierda, R. J. *et al.* Epigenetic control of CCR5 transcript levels in immune cells and modulation by small molecules inhibitors. *J. Cell. Mol. Med.* **16**, 1866–1877 (2012).
888. Khurana, S., Kennedy, M., King, L. R. & Golding, H. Identification of a Linear Peptide Recognized by Monoclonal Antibody 2D7 Capable of Generating CCR5-Specific Antibodies with Human Immunodeficiency Virus-Neutralizing Activity. *J. Virol.* **79**, 6791–6800 (2005).
889. Flegler, A. J., Cianci, G. C. & Hope, T. J. CCR5 Conformations Are Dynamic and Modulated by Localization, Trafficking and G Protein Association. *PLoS One* **9**, e89056 (2014).
890. Konuma, T. *et al.* Circulating monocyte subsets in human chronic graft-versus-host disease. *Bone Marrow Transplant.* **53**, 1532–1540 (2018).
891. Stoeckius, M. *et al.* Simultaneous epitope and transcriptome measurement in single cells. *Nat. Methods* **14**, 865–868 (2017).
892. Juno, J. A. *et al.* Modulation of the CCR5 Receptor/Ligand Axis by Seminal Plasma and the Utility of In Vitro versus In Vivo Models . *J. Virol.* **93**, (2019).
893. Kulkarni, S. *et al.* CCR5AS lncRNA variation differentially regulates CCR5, influencing HIV disease outcome. *Nat. Immunol.* **20**, 824–834 (2019).
894. Liu, J. *et al.* Local production of the chemokines CCL5 and CXCL10 attracts CD8++ T lymphocytes into esophageal squamous cell carcinoma. *Oncotarget* **6**,

24978–24989 (2015).
895. Ghulam-Smith, M. TWO SIDES TO THE SAME ANTIBODY: ASSESSING THE ROLE OF NEUTRALIZING AND NON-NEUTRALIZING ANTIBODIES IN MOTHER-TO-CHILD TRANSMISSION OF HIV-1 by MELISSA GHULAM-SMITH. (Boston University, 2021).
896. Villar-Vesga, J. *et al.* Platelet-derived microparticles generated in vitro resemble circulating vesicles of patients with rheumatoid arthritis and activate monocytes. *Cell. Immunol.* **336**, 1–11 (2019).
897. Lönn, P. *et al.* Enhancing Endosomal Escape for Intracellular Delivery of Macromolecular Biologic Therapeutics. *Sci. Rep.* **6**, 1–9 (2016).
898. Erazo-Oliveras, A., Muthukrishnan, N., Baker, R., Wang, T. Y. & Pellois, J. P. Improving the endosomal escape of cell-penetrating peptides and their cargos: Strategies and challenges. *Pharmaceuticals* **5**, 1177–1209 (2012).
899. LeCher, J. C., Nowak, S. J. & McMurry, J. L. Breaking in and busting out: Cell-penetrating peptides and the endosomal escape problem. *Biomolecular Concepts* **8**, 131–141 (2017).
900. Ryu, S.-E., Truneh, A., Sweet, R. W. & Hendrickson, W. A. Structures of an HIV and MHC binding fragment from human CD4 as refined in two crystal lattices. *Structure* **2**, 59–74 (1994).
901. Wu, H., Kwong, P. D. & Hendrickson, W. A. Dimeric association and segmental variability in the structure of human CD4. *Nature* **387**, 527–30 (1997).
902. Wang, J. H. *et al.* Crystal structure of the human CD4 N-terminal two-domain fragment complexed to a class II MHC molecule. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 10799–804 (2001).
903. Owen, G. R., Le, D., Stoychev, S., Cerutti, N. M. & Papathanasopoulos, M. Redox exchange of the disulfides of human two-domain CD4 regulates the conformational dynamics of each domain, providing insight into its mechanisms of control. *Biochem. Biophys. Res. Commun.* (2018). doi:10.1016/J.BBRC.2018.02.161
904. Klasse, P. J. *et al.* Influences on Trimerization and Aggregation of Soluble, Cleaved HIV-1 SOSIP Envelope Glycoprotein. *J. Virol.* **87**, 9873–9885 (2013).
905. Sanders, R. W. *et al.* A Next-Generation Cleaved, Soluble HIV-1 Env Trimer, BG505 SOSIP.664 gp140, Expresses Multiple Epitopes for Broadly Neutralizing but Not Non-Neutralizing Antibodies. *PLoS Pathog.* **9**, e1003618 (2013).
906. De Taeye, S. W. *et al.* Immunogenicity of Stabilized HIV-1 Envelope Trimers with Reduced Exposure of Non-neutralizing Epitopes. *Cell* **163**, 1702–1715 (2015).
907. TZM-bl Assay for Neutralizing Antibodies Protocol for Measuring Neutralizing Antibodies Against HIV-1, SIV and SHIV Using a Luciferase Reporter Gene Assay in TZM-BL Cells. (2006).
908. Protocol for Neutralizing Antibody Assay for HIV-1 in TZM-bl Cells. (2017).
909. Stuart, A. TZM-bl Neutralization Protocol Stamatatos Lab.
910. Protocol for the Preparation and Titration of HIV-1 Env-pseudotyped Viruses. (2017).
911. Sarzotti-Kelsoe, M. *et al.* Optimization and validation of the TZM-bl assay for standardized assessments of neutralizing antibodies against HIV-1. *J. Immunol. Methods* **409**, 131–146 (2014).
912. Rasheed, M., Bettadapura, R. & Bajaj, C. Computational Refinement and Validation Protocol for Proteins with Large Variable Regions Applied to Model HIV Env Spike in CD4 and 17b Bound State. *STFODE* **23**, 1138–1149 (2015).

913. Melikyan, G. B. Common principles and intermediates of viral protein-mediated fusion: the HIV-1 paradigm. *Retrovirology* **5**, 111 (2008).
914. Hu, X. *et al.* Cell surface assembly of HIV gp41 six-helix bundles for facile, quantitative measurements of hetero-oligomeric interactions. *J. Am. Chem. Soc.* **134**, 14642–14645 (2012).
915. Doms, R. W. & Moore, J. P. HIV-1 membrane fusion: Targets of opportunity. *Journal of Cell Biology* **151**, f9 (2000).
916. Liu, D. *et al.* Six-helix bundle completion in the distal C-terminal heptad repeat region of gp41 is required for efficient human immunodeficiency virus type 1 infection. *Retrovirology* **15**, 27 (2018).
917. Goddard, T. D. *et al.* UCSF ChimeraX: Meeting Modern Challenges in Visualization and Analysis. *Protein Sci.* **27**, 14–25 (2017).
918. Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–12 (2004).
919. Immunoglobulin M (IgM): Structure, Properties, Functions and Clinical Significance - Learn Microbiology Online. Available at: https://microbeonline.com/igm-antibody-structure-properties-functions-clinical-significance/. (Accessed: 20th February 2020)
920. Liu, J. *et al.* Role of the IgM Fc Receptor in Immunity and Tolerance. *Front. Immunol.* **10**, 529 (2019).
921. Schroeder, H. W. & Cavacini, L. Structure and function of immunoglobulins. *J. Allergy Clin. Immunol.* **125**, S41 (2010).
922. Astronomo, R. D. *et al.* Neutralization Takes Precedence Over IgG or IgA Isotype-related Functions in Mucosal HIV-1 Antibody-mediated Protection. *EBioMedicine* (2016). doi:10.1016/j.ebiom.2016.11.024
923. Woof, J. M. & Russell, M. W. Structure and function relationships in IgA. *Mucosal Immunology* **4**, 590–597 (2011).
924. Pabst, O. & Slack, E. IgA and the intestinal microbiota: the importance of being specific. *Mucosal Immunology* **13**, 12–21 (2020).
925. Woof, J. M. & Kerr, M. A. IgA function - Variations on a theme. *Immunology* **113**, 175–177 (2004).
926. Sutton, B., Davies, A., Bax, H. & Karagiannis, S. IgE Antibodies: From Structure to Function and Clinical Translation. *Antibodies* **8**, 19 (2019).
927. Scheepers, C. *et al.* Serum glycan-binding IgG antibodies in HIV-1 infection and during the development of broadly neutralizing responses. *AIDS* **31**, 2199–2209 (2017).
928. Cheeseman, H. M. *et al.* Broadly Neutralizing Antibodies Display Potential for Prevention of HIV-1 Infection of Mucosal Tissue Superior to That of Nonneutralizing Antibodies. *J. Virol.* **91**, (2017).
929. McCoy, L. E. The expanding array of HIV broadly neutralizing antibodies 11 Medical and Health Sciences 1108 Medical Microbiology Marit van Gils, m.j.vangils@amc.uva.nl; Rogier W Sanders, rws2002@med.cornell.edu. *Retrovirology* **15**, 1–12 (2018).
930. Landais, E. & Moore, P. L. Development of broadly neutralizing antibodies in HIV-1 infected elite neutralizers. *Retrovirology* **15**, 1–14 (2018).

Appendix A
Supplemental Materials for Chapter 1

Table Appendix A.1 Structures of CD4

Year	PDB#	AA covered	Bound	Advancement	Method	Res.
1990	-[162]	1-183	-	First crystal structures, determined interactions between domains and theorized interactions with HIV gp120	X-Ray	2.4 Å
1990	-[188]	1-183	-		X-Ray	2.3 Å
1993	1CID[164]	210-385	-	First structure of domains 3-4	X-Ray	2.8 Å
1993	3CD4[273]	1-200	-	First structure of CD4 D1-2 on RCSB PDB	X-Ray	2.2 Å
1994	1CDH[900]	1-178	-		X-Ray	2.3 Å
1996	1WBR[165]	428-444	-	First structure of cytoplasmic domain, showed sites of interaction with HIV Vpu & Nef	Sol NMR	-
1996	1CDJ[193] 1CDY 1CDU	26-203	-	Mutational studies showing defective HIV binding to CD4	X-Ray	2.5 Å
1997	1WIO[901] 1WIP	1-363	-	Full extracellular structure, D1-4, showed how CD4 could dimerize	X-Ray	3.9 Å
1998	1GC1[194]	1-183	gp120	First complex of CD4 with HIV gp120	X-Ray	2.5 Å
2001	1JL4[902]	26-203	Class II MHC	First complex of CD4 with class II MHC, a natural ligand of CD4	X-Ray	4.3 Å
2009	2KLU[166]	364-433	-	First structure of transmembrane domain and full cytoplasmic domain	Sol NMR	-

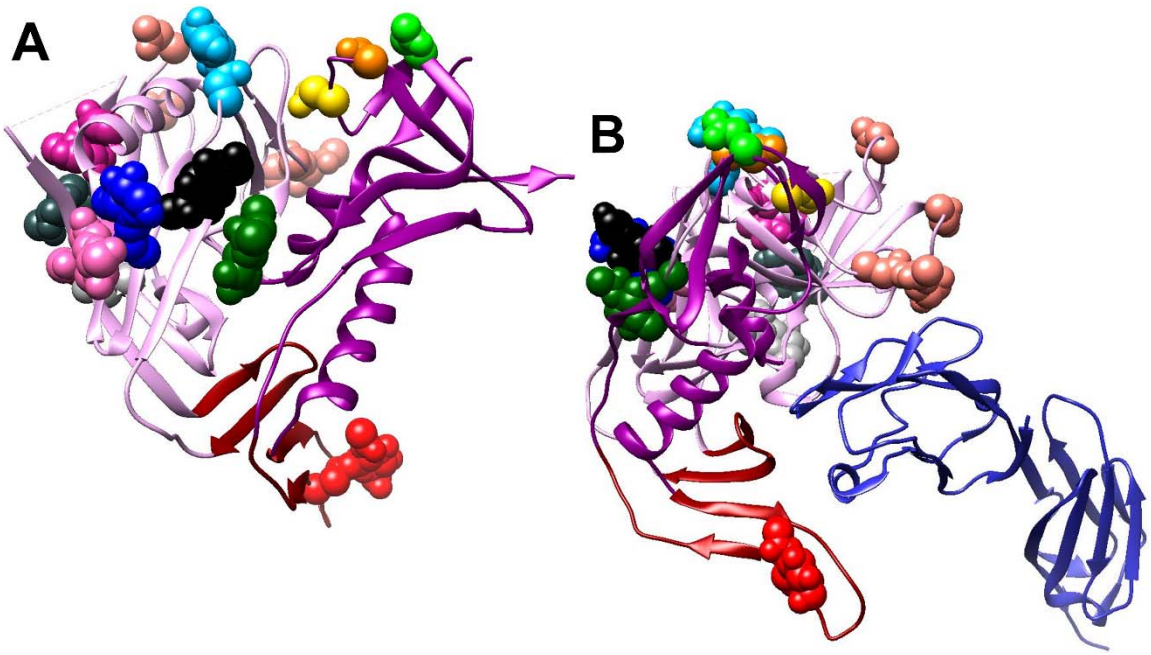


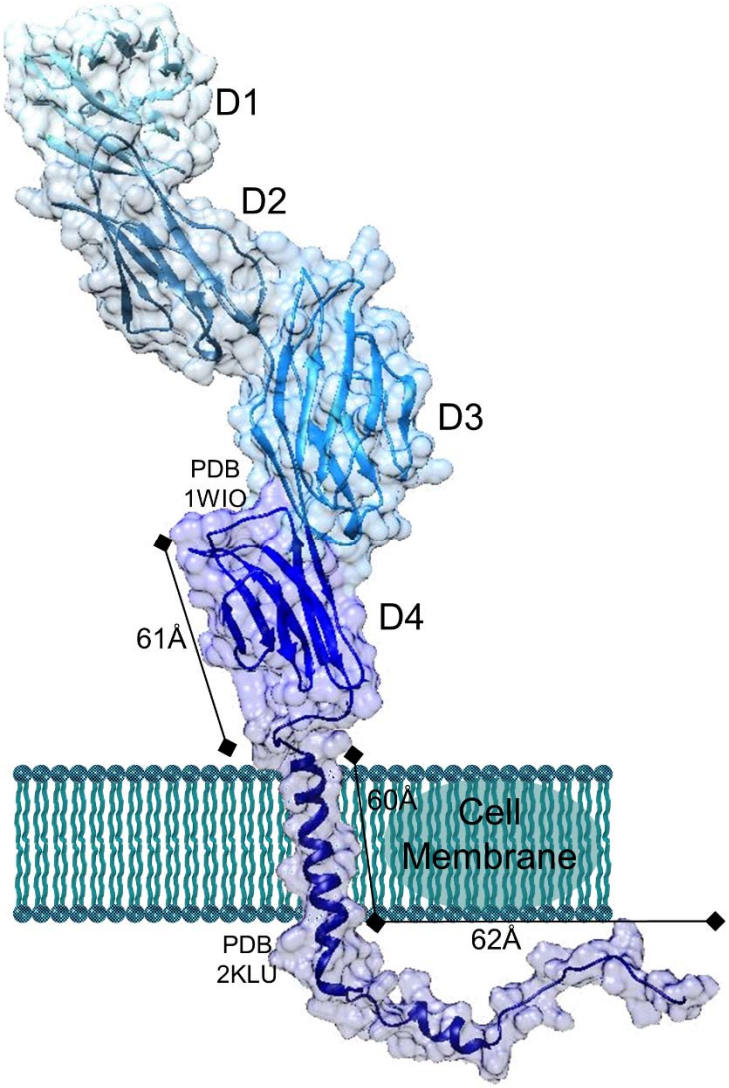
Figure Appendix A. 1 Basic domains of gp120 with high mannose glycans color coded.

These figures are extensions of Figure_ PDB 1GC1[194], showing the domains of gp120 (the outer domain in light purple, the inner domain medium purple and the bridging sheet in dark red) that has been bound to both CD4 and 17b, over-layed with the sites of the high mannose glycans as they are color coded in Chapter 2. A key for the glycan colors is given below. **A)** shows gp120 alone with the high mannose glycans color coded as indicated in the list below, with non-high mannose glycans shown in salmon. **B)** shows the same structure with CD4 visible in dusty blue.

197 – Dark Red
230 – Orange
234 – Goldenrod
241 – Green
262 – Dark Green
289 – Deep Sky Blue
295 – Blue
301 – Medium Purple
332 – Hot Pink
339 – Dark Magenta
386 – Light Grey
392 – Dark Slate Grey
448 – Black

Structure of CD4

The first CD4 structure published was rat CD4 in 1993 by the Barclay group in Science. CD4 from a Human T cell was imaged by X-ray crystallography at a resolution of 2.5 A in 1996 by the Hendrickson group. CD4 is composed of 4 extracellular domains (D1-4, ~370 residues), an ~25 aa transmembrane stretch, then a 38aa cytoplasmic domain Each monomeric extracellular domain is approximately 61A long by 26A wide. The transmembrane domain is ~60Å long and the cytosolic domain is ~62Å long. Human CD4 has two glycosylation sites, one on D3, & one on D4[203,274,903].



HIV SOSIP Trimers

SOSIP gp140 trimers are soluble HIV trimers designed to mimic native HIV envelope spikes. SOSIP stands for the mutations used to create a more stable soluble HIV trimer: SOS is for the disulfide bond introduced between A501 and T605 in gp120 and gp41 respectively[251] (Figure), IP is a point substitution, I559P, in the gp41 ectodomain that stabilizes intra-trimer interactions[252]. SOSIP.664 trimers have the membrane-proximal external region (MPER) of gp41 deleted to improve solubility; 644 indicates the total number of residues in the gp140 monomer[904]. SOSIP trimers have been created from most major clades of HIV, see Table for the most used SOSIPs and which clade they are derived from. SOSIP trimers differ from native trimers in that SOSIP trimers' glycosylation tends to be less extensively processed and they do not bind to non-neutralizing antibodies[525].

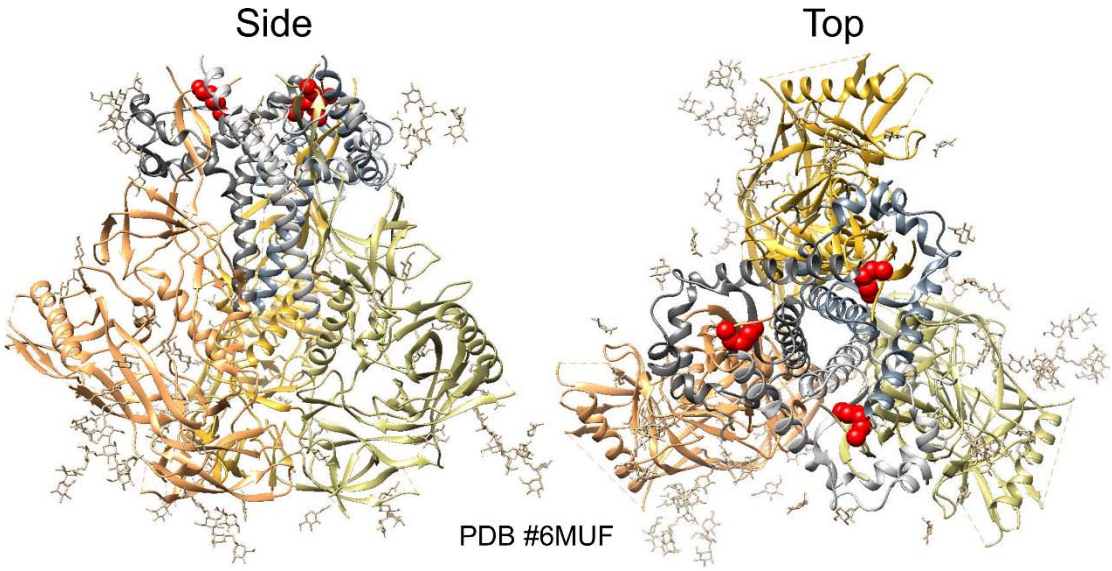


Figure Appendix A.2 SOS mutations in SOSIP trimer.
This figure shows B41 SOSIP.664 trimer, with each monomer of gp120 colored in a different shade of yellow, each monomer of gp41 colored in a different shade of grey, glycans colored in tan and the disulfide bond at 501 and 605 (the SOS) has been highlighted as red spheres.

Table Appendix A.2 Commonly used HIV SOSIP trimers

Clade	Title	Made
A	BG505 KNH1144	2013[905]
B	B41 AMC008 AMC011 JRFL	2014[557] 2015[906] 2015[564]
C	16055 DU422 ZM197M	2015[564]

Table Appendix A.3 Summary of HIV drugs currently available

Class	Drug		Approved[381]	Brand names
NRTIs	Zidovudine	AZT	1987	Zidovudine, Retrovir
	Abacavir	ABC	1998	Ziagen
	Didanosine	DDI	2000	Didanosine, Videx
	Stavudine	D4T	2002	Stavudine, Zerit
	Emtricitabine	FTC	2003	Emtricitabine, Emtriva
	Lamivudine	3TC	2001	Lamuvidine, Epivir
	Tenofovir GS 7340	TDF	2001	Viread
	ACH-126,433		ICT*	Elvucitabine
	SPD-754		ICT*	Apricitabine
NNRTIs	Delavirdine	DLV	1997	Rescriptor
	Efavirenz	EFV	1998	Sustiva, Stocrin
	Nevirapine	NVP	1996	Viramune
	TMC-125	ETR	2018	Etravirine
PIs	Atazanavir	ATV	2004	Reyataz
	Fosamprenavir	FPV	2003	Agenerase, Lexiva
	Indinavir	IDV	1996	Crixivan
	Lopinavir	LPV	2000	Kaletra
	Nelfinavir	NFV	2003	Viracept
	Ritonavir	RTV	1999	Norvir
	Saquinavir	SQV	1995	Invirase, Fortovase
	Tipranavir	TPV	2005	Aptivus
	TMC-114		2006	darunivir, Prezista
	Brecanavir		2006	
Entry Inhibitors	Enfuviride	T-20	2003	Enfurvitide, Fuzeon
	Maraviroc		2007	Maraviroc
	Ibalizumab		2018	Trogarzo, TNX-355
	BMS-4588043		2018	OPDIVO
	Leronlimab		ICT*	PRO-140
	Temsavir		ICT*	
	Fostemsavir		ICT*	
Integrase Inhibitors	MK-0518	RAL	2007	Raltegravir
	Dolutegravir	DGT	2013	Dovato
	Elvitegravir	EVG	2012	Vitekta

*ICT – In Clinical Trials

Appendix B
Supplemental Materials for Chapter 2

R code for ANOVA with a Bonferroni correction

```
d=read.csv(file.choose(),
        header = FALSE,
        stringsAsFactors=FALSE,
        sep = ",")
colnames(d)=c("Mutant","IC50")
anova(lm(IC50~Mutant, d))
options(width=10000)
pairwise.t.test(d$IC50, d$Mutant, p.adj = "bonferroni")
```

Note that files must be saved as a CSV file with only two columns, the first being the name/title of each mutant and the second being the calculated IC_{50} or fold difference value.

Matlab code for sigmoidal curve fit

```
clear %clears workspace variables
clc %clears command window

% Dilution:
Dil = [1
1
1
0.5
0.5
0.5
0.25
0.25
0.25
0.125
0.125
0.125
0.0625
0.0625
0.0625
0.03125
0.03125
0.03125
0.015625
0.015625
0.015625
0.0078125
0.0078125
0.0078125
];
Dil = log10(Dil);
```

```
% Percent Infection:
PI = [4.207167653
5.651521553
6.774351579
13.22902094
7.341243626
9.31315018
17.42860685
26.87057559
27.90005454
47.61951293
42.61087326
56.64754564
78.55750403
79.87274176
86.95331049
79.01237559
100.5609221
88.77263852
100.6850534
98.44685908
105.5629261
115.8801556
107.6391163
114.2106872
];
PI = (PI)./100;

[FR,GOF] = sigmoidal_fit(Dil,PI)

% figure(2)
% clf
% plot(Dil,PI,'or')

function [fitresult, gof] = sigmoidal_fit(xp, yp)
%CREATEFIT(XP,YP)
%  Create a fit.
%
%  Data for 'sigmoidal' fit:
%      X Input : xp
%      Y Output: yp
%  Output:
%      fitresult : a fit object representing the fit.
%      gof : structure with goodness-of fit info.
%
%  See also FIT, CFIT, SFIT.

%  Auto-generated by MATLAB on 12-Mar-2019 18:10:16
```

```
%% Fit: 'sigmoidal'.
[xData, yData] = prepareCurveData( xp, yp );

% Set up fittype and options.
ft = fittype( '(a/(b+exp(+c*x)))', 'independent', 'x', 'dependent', 'y' );
opts = fitoptions( 'Method', 'NonlinearLeastSquares' );
opts.Display = 'Off';
opts.StartPoint = [1 1 1];

% Fit model to data.
[fitresult, gof] = fit( xData, yData, ft, opts );

% Plot fit with data.
figure(1)%'Name', 'Sigmoidal Curve Fit' );
clf
h = plot( fitresult, xData, yData );
legend( h, 'Data', 'Sigmoidal Fit', 'Location', 'SouthEast' );
% Label axes
xlabel('Log Dilution')
ylabel('Percent Infection')
grid on
end
```

Notes on TZM neutralization Assay

Preparation of Virus

- Recommended to store frozen virus in 20% FBS, we only store in old media[907]
- "greater sensitivity is achieved in the neutralization assay at lower input TCID50 doses but overall this difference is less than 3-fold between 50 and 1,000 input TCID50 in most cases"
- Recommended to thaw virus in ambient temperature water[908]
- Recommended to dilute to 150,000 RLU but should have at least 15,000 RLU, virus only control should be greater than 10X the background[908]
- There is a viral database to consult for the optimal virus dilution for viruses that have been thawed one time, virus should not be thawed and refrozen more than once[908]

Protocol of Assay

- Plating of cells into preincubated virus, is this only for luminescence or B-galactosidase as well?
  - Samatatos lab preincubates then transfers to pre-plated cells @ 30-50% confluency[909], they then incubate for 72 hours but no media change. Samatatos lab is at U of Washington, Department of Global Health. They use polybrene instead of DEAE
- We do not add DEAE-Dextran, is this only for luminescence or B-galactosidase as well? Stated that DEAE is optional – is a transfection reagent, used to enhance infectivity[907]. Why would you want to enhance infectivity of a virus you are trying to inhibit? Why exactly is it being added to assays?
- In luciferase assay, they incubate 48-72 hours but do not change media. Is this because of luciferase? Is B-galactosidase assay different?
- Should be looking at presence of syncytia in virus control wells. Pictures? What should I have been looking for? Is there a way to tell after the fact?
  - "the presence of syncytia as too many syncytia indicate cell killing and thus the validity of the assay is compromised. If cell killing is present, the assays should be repeated using a lower dose of the virus. Also check the bottom row of the plate for the presence of toxicity. Cell toxicity could be erroneously interpreted as neutralization"[910]
- Percent difference between triplicate [911]wells should be <30% for sample dilutions that yield at least 40% neutralization & for all virus control wells
- Luciferase assay has been shown to have luminescence that is directly proportional to the number of infectious virus particles present[911]

Conditions to Check

- Read time of virus
- Confluency of cells at time of infection
- Times between steps
- Try using alternate calculation method (see pg Cell Calculations)
- Presence of syncytia? Able to tell after the fact?

Additions to existing procedure

- Record the generation number of the cells being used in each round of plates
- Length of each stage (will be done retroactively for all experiments)
- Calculate percent variation in positive controls and all triplicate wells

Cell Calculations[908]

To calculate the cell concentration, multiple the average number of cells per quadrant, the dilution factor, and 10,000 to yield the cell concentration, “C1”, in cells/ml. To calculate the total cell mixture volume, “V2”, that you need, multiply the number of plates by the total volume of cell mixture needed per plate. The concentration of cells desired is 100,000 cells/ml, “C2”. Thus, using the equation C1V1=C2V2, one can solve for “V1”, the volume of cells needed[908].

For example:

Total number of viable cells counted = 60

Number of quadrants counted = 4

Dilution factor = 10

Number of plates = 1

Cell mixture needed per plate = 10 ml

60 cells ÷ 4 quadrants = 15 cells/quadrant

15 x 10 x 10,000 cells/ml = 1,500,000 cells/ml = C1

1 plate x 10 ml / plate = 10 ml = V2

Optimum final concentration of cells = 100,000/ml = C2

Therefore: C1V1=C2V2

V1 = (100,000 x 10) ÷ 1,500,000 = 0.67 ml of cells

B.4 Example Calculations for TZM Assay

All plates were laid out with the mutant and wild type on the same plate to provide an internal control. The negative control was PBS only. The positive control only had the appropriate dilution of the virus. All the other wells had a 1:3 dilution series of Grft that centered the IC_{50} to ensure that there was at least one point each at both 0% and 100% infection.

Plate Map

A	*	*	*	*	*	*	*	*	*	*	*
B	**Neg**	**Pos**	**Griffithsin Dilution Series**								Wild Type
C			**1**	**1/3**	**1/9**	**1/27**	**1/81**	**1/243**	**1/729**	**1/2k**	
D											
E	**Neg**	**Pos**	**Griffithsin Dilution Series**								Mutant
F			**1**	**1/3**	**1/9**	**1/27**	**1/81**	**1/243**	**1/729**	**1/2k**	
G											
H	*	*	*	*	*	*	*	*	*	*	*

After lysis of the cells, releasing Beta-galactosidase, CPRG was added and the plates were read on a spectrophotometer with two wavelengths: 570nm and 630nm. The

cleavage product of CPRG absorbs at 570nm, in the yellow color range, and the optimal read was when the absorbance of the positive control was approximately 2 AU based on the linear range of the spectrophotometer (**Error! Reference source not found.**A). Absorbance at 630nm was read Development time of the plate determines how accurate the read variation will be.

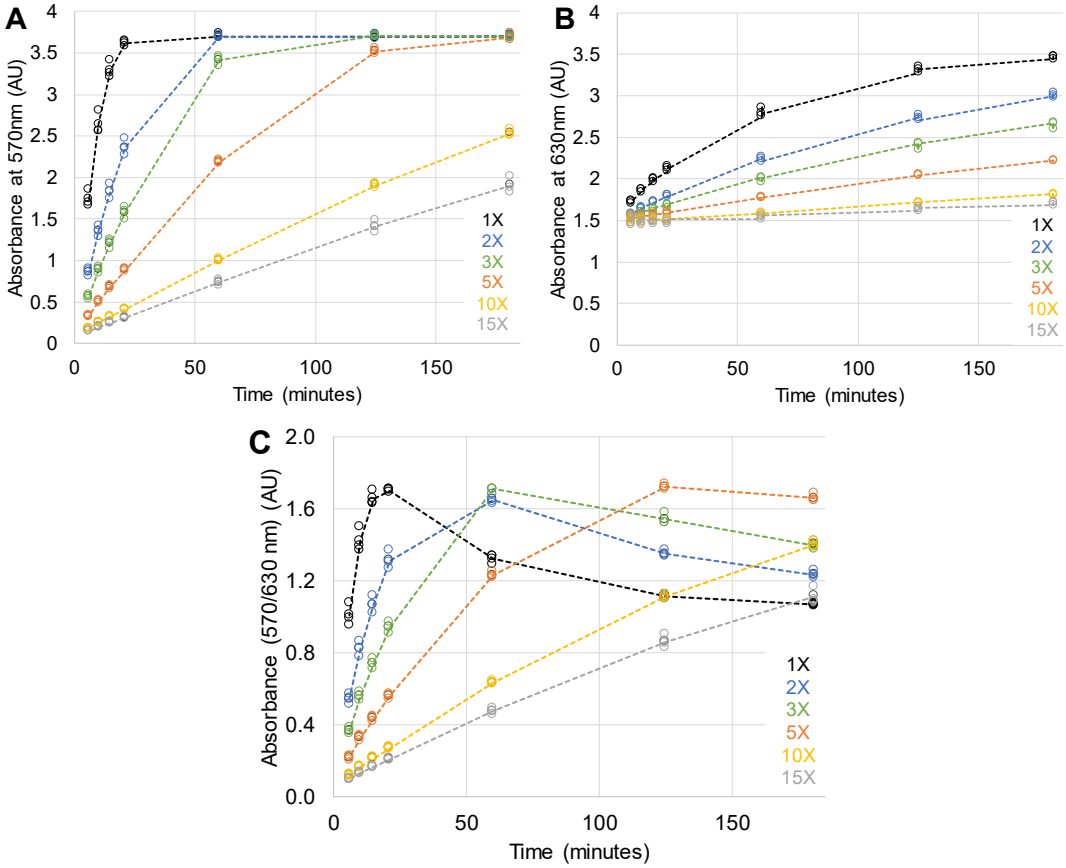


Figure Appendix B. 1 Absorbance over time of various PVO.4 Wt HIV-1 virus dilutions to show range of spectrophotometer.
A and **B** show the increasing absorbance over time at 570nm and 630nm respectively. **C)** shows how the absorbance at 570nm divided by the absorbance at 630nm over time peaks and then begins to decline

ΔAbs/sec = (2-0.1) / (60sec * Min to reach 2 AU @ 570nm)

2 is the desired absorbance threshold.
0.1 is the absorbance of the negative control, the baseline read in all wells.

ΔAbs across = ΔAbs/sec * time to read

Time to read given in parenthesis for each parameter

Min to reach 2 AU @ 570nm	ΔAbs/ sec	ΔAbs across triplicate (24 sec)	ΔAbs across plate (75 sec)	ΔAbs across total time (120 sec)
5	0.006	0.152	0.475	0.760
10	0.003	0.076	0.238	0.380
15	0.002	0.051	0.158	0.253
20	0.002	0.038	0.119	0.190
25	0.001	0.030	0.095	0.152
30	0.001	0.025	0.079	0.127
45	0.001	0.017	0.053	0.084
60	0.001	0.013	0.040	0.063
75	0.0004	0.010	0.032	0.051
90	0.0004	0.008	0.026	0.042
120	0.0003	0.006	0.020	0.032
150	0.0002	0.005	0.016	0.025
180	0.0002	0.004	0.013	0.021

Plate Reads

570	1	2	3	4	5	6	7	8	9	10	11
A	0.052	0.053	0.053	0.052	0.052	0.051	0.052	0.051	0.054	0.053	0.052
B	0.103	2.159	1.981	0.117	0.167	0.399	0.632	1.126	1.558	1.588	1.798
C	0.102	1.994	1.995	0.111	0.17	0.308	0.75	1.131	1.779	1.89	1.825
D	0.1	1.676	1.645	0.104	0.135	0.267	0.599	1.013	1.522	1.49	1.701
E	0.102	3.693	3.677	0.132	0.339	0.984	1.668	2.791	3.557	3.693	3.72
F	0.1	3.674	3.684	0.135	0.467	0.987	1.897	3.302	3.628	3.665	3.69
G	0.104	3.658	3.67	0.13	0.341	0.904	1.662	2.734	3.475	3.601	3.63
H	0.049	0.049	0.052	0.05	0.052	0.05	0.053	0.05	0.052	0.051	0.05
630	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
A	0.054	0.056	0.056	0.055	0.054	0.054	0.054	0.053	0.057	0.056	0.055
B	1.602	1.896	1.888	1.532	1.511	1.66	1.708	1.732	1.811	1.67	1.83
C	1.583	1.893	1.864	1.662	1.588	1.698	1.727	1.77	1.857	1.867	1.848
D	1.604	1.826	1.772	1.523	1.638	1.633	1.711	1.742	1.825	1.823	1.82
E	1.603	2.354	2.292	1.619	1.65	1.744	1.829	1.987	2.15	2.257	2.326
F	1.554	2.346	2.384	1.595	1.641	1.716	1.855	2.065	2.155	2.221	2.239
G	1.592	2.267	2.213	1.611	1.621	1.72	1.781	1.926	2.093	2.168	2.168
H	0.052	0.052	0.055	0.052	0.054	0.052	0.055	0.053	0.055	0.053	0.052

=570/630

	1	2	3	4	5	6	7	8	9	10	11
A	0.963	0.946	0.946	0.945	0.963	0.944	0.963	0.962	0.947	0.946	0.945
B	0.064	1.139	1.049	0.076	0.111	0.240	0.370	0.650	0.860	0.951	0.983
C	0.064	1.053	1.070	0.067	0.107	0.181	0.434	0.639	0.958	1.012	0.988
D	0.062	0.918	0.928	0.068	0.082	0.164	0.350	0.582	0.834	0.817	0.935
E	0.064	1.569	1.604	0.082	0.205	0.564	0.912	1.405	1.654	1.636	1.599
F	0.064	1.566	1.545	0.085	0.285	0.575	1.023	1.599	1.684	1.650	1.648
G	0.065	1.614	1.658	0.081	0.210	0.526	0.933	1.420	1.660	1.661	1.674
H	0.942	0.942	0.945	0.962	0.963	0.962	0.964	0.943	0.945	0.962	0.962

Negative control: Average (1B-G) = 0.064
Positive control: Average (2E-G + 3E-G) = 1.026
Positive – negative = 0.962

Percent Infection:

(570/630)-Neg / (Pos-Neg)*100								
	1.28	4.83	18.32	31.80	60.91	82.75	92.16	95.45
	0.28	4.47	12.19	38.47	59.75	92.90	98.55	95.97
	0.44	1.91	10.33	29.72	53.78	80.01	78.28	90.47
Conc. (nM)	10	3.3	1.1	0.370	0.123	0.041	0.014	0.005

IC_{50} Calculations:

M = slope
B = %Inf - (M*Con)
IC50 = (50 - B)/M

M	B	IC_{50}
-117.9	75.46	0.22
-86.16	70.39	0.24
-97.41	65.80	0.16

Average IC_{50} (nM): 0.20
SD: 0.04

Slope is determined by the points closest to 50% inhibition, which are outlined in red.

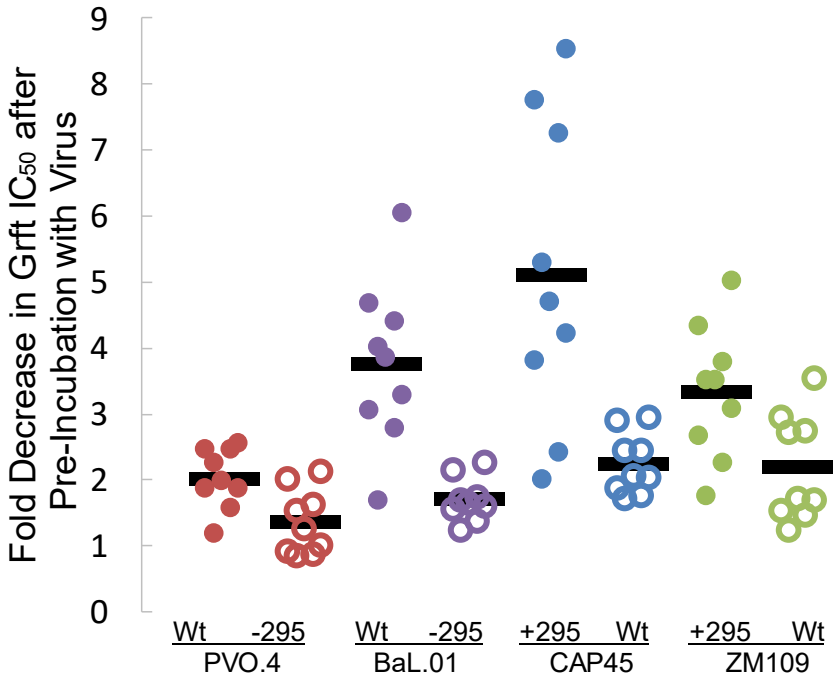


Figure Appendix B. 2 Methods comparison of the single cycle neutralization assay. Comparison of the effect of pre-incubation of Grft and virus for an hour before addition to the cells (compared to no pre-incubation of Grft with the virus) on the IC_{50} of Grft based on the presence of glycosite N295. When a glycan is present at N295, pre-incubation of the virus with Grft can cause a sizable decrease in Grft's IC_{50}.

Troubleshooting a Shift in Pseudo-Viral IC_{50}

Midway through collection of data for Grft potency against various HIV mutants, a sudden drop in the IC_{50} occurred, shown in **Error! Reference source not found.**, from a steady average of 0.2nM to 0.1nM. This 2-fold drop was consistent, which resulted in a stable fold difference between wild type and mutants. To my best guess, the media composition was altered in a way that affected the sugar level or type. Since Grft binds

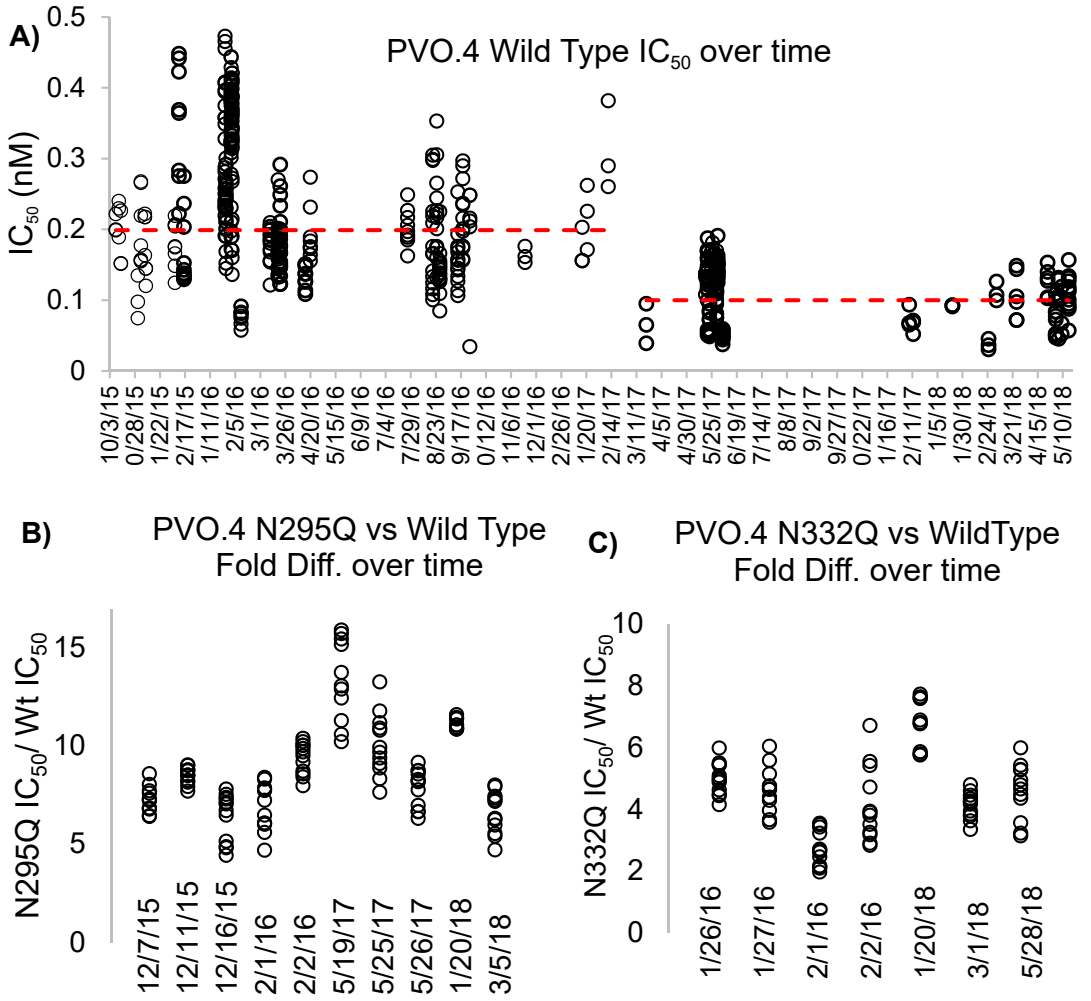


Figure Appendix B. 3 Analysis of Grft Inhibition (IC50 values) over time.
Despite using identical techniques, there was an abrupt change in measured IC_{50} of Grft. **A)** shows the progression of IC_{50} against PVO.4 over time from October 9th 2015 to May 16th 2018. A distinct drop in IC_{50} occurred between February 11th, 2017 and March 21st, 2017. Median IC_{50} before and after the drop is indicated by a dotted trendline. Prior to February 11th, the IC_{50} was 0.2nM ± 0.06, after the drop the IC_{50} was 0.1nM ± 0.03. This is likely due to a change in media components. **B and C**. Despite specific IC_{50} values apparently decreasing, the fold difference between the Grft IC_{50} against wild type virus compared to mutant virus is largely the same. That is, the effect of each mutation of gp120 on Grft inhibition is largely the same when considered in the context of its comparison to Grft inhibition of the wild type virus (done at the same time). Therefore, Table 2.1 reports the IC_{50} of Grft against the wild type of each virus, then reports the fold difference for each variant in comparison with the wild type virus.

to sugars, they could have potentially been inhibiting Grft from binding to HIV, resulting in the higher IC_{50}. The subsequent removal or alteration of the sugar composition would therefore lower the IC_{50} of Grft since it can now be more effective against HIV. This conclusion was reached after testing numerous variables, listed in **Error! Reference source not found.** below.

Table Appendix B 1 TZM assay parameters tested for influence on Grft IC50.

Virus	Production	Confluency at time of transfection Passage number Length of virus production before collection
	Use	Freeze thaw cycles Dilution used
Assay	TZM cells	Confluency at time of infection Batch Passage number
	Time	Growth time after plating cells Time of virus exposure Growth time before read

Appendix C
Supplemental Materials for Chapter 3

Calculations for MAGI Assay

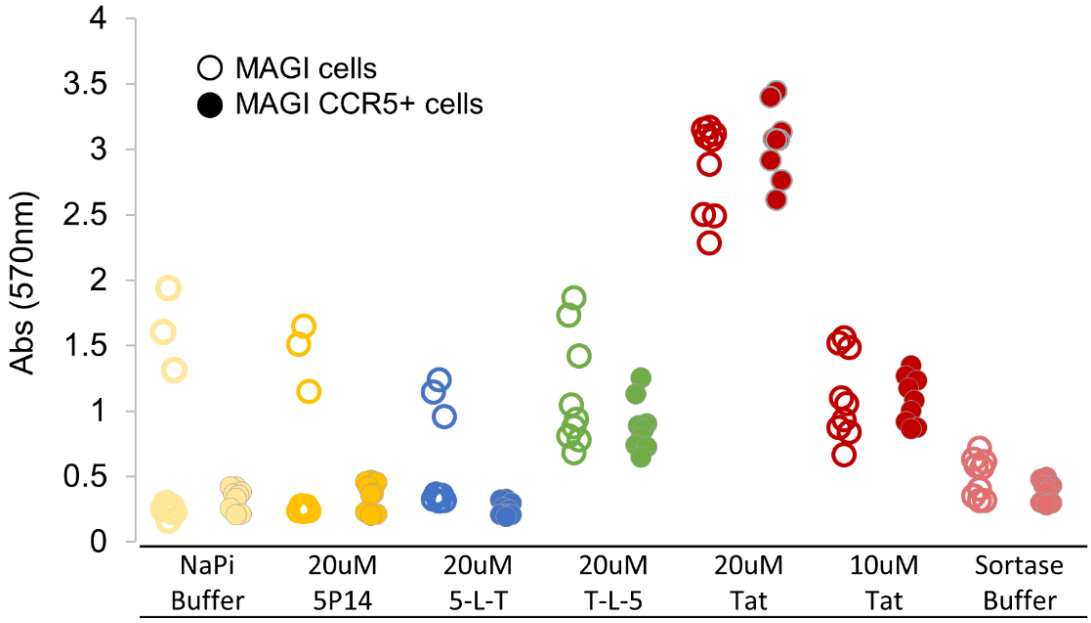


Figure Appendix C.1 Raw Data of MAGI Assay
Testing of 5P14-Linker-Tat (5-L-T) and Tat-Linker-5P14 (T-L-5) chimeric constructs in triplicate. Open circles indicate MAGI cells (CCR5-) and filled circles indicate MAGI CCR5+ cells. It should be noted that Tat does not seem to follow a linear activation curve, as evidenced by the addition of 10uM Tat only giving an absorbance of ~1 at 570nm compared to 20uM Tat giving an absorbance of ~3 at 570nm.

The average baseline absorbance caused by addition of the buffers that the proteins were taken up in was subtracted from all samples' absorbance (Table Appendix C.2). Tat samples were taken up in 1X sortase buffer, all other samples were taken up in 20mM NaPi buffer pH 2.5. The adjusted reads of the 20uM Tat and 20uM 5P14-RANTES were then each averaged and set as the 100% and 0% activation points respectively. To convert reads of the chimeric proteins, 5P14-Linker-Tat (5-L-T) and Tat-Linker-5P14 (T-L-5), and 10uM Tat to percent activation, the average 20uM 5P14-RANTES (0%) adjusted absorbance was subtracted from the sample adjusted absorbance, then that value was divided by the average adjusted absorbance of 20uM Tat (Table Appendix C.3).

Table Appendix C.1 Absorbance at 570nm - Raw Data

2.5 NaPi Buffer	20uM 5P14	20uM 5P14-L-Tat	20uM Tat-L-5P14	20uM Tat	10uM Tat	1X Sortase Buffer
0.2961	0.2763	0.362	1.0468	3.0974	1.099	0.5811
0.2685	0.2699	0.354	0.9383	3.0746	1.0572	0.567
0.2603	0.238	0.3472	0.8782	2.887	0.931	0.403

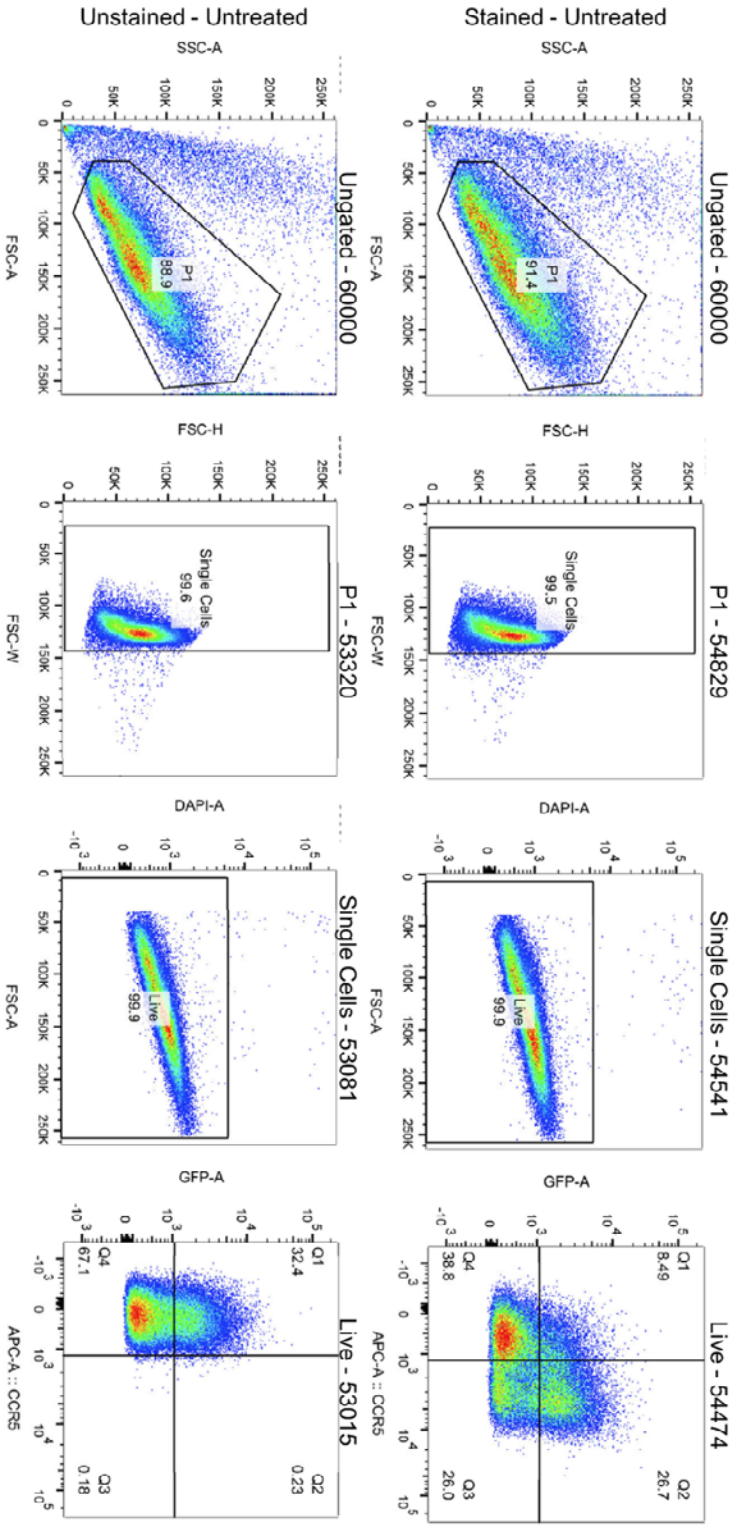
Table Appendix C.2 Average of Buffer controls subtracted from samples

20uM 5P14	20uM 5P14-L-Tat	20uM Tat-L-5P14	20uM Tat	10uM Tat
-0.06	4.08	19.81	94.30	20.87
-0.04	2.84	20.93	101.14	19.68
0.10	3.25	16.29	104.55	27.59

Table Appendix C.3 Samples Normalized as Percent Activation with 20uM 5P14 as 0% and 20uM Tat as 100%

20uM 5P14	20uM 5P14-L-Tat	20uM Tat-L-5P14	20uM Tat	10uM Tat
-1.65	-7.02	29.27	92.08	19.40
0.11	-6.10	13.21	104.73	27.59
1.54	-5.43	15.50	103.19	28.81

Figure Appendix C.2 FACS Assay in JLTRG cells Trial 2 - Unstained and untreated



These figures show the full gating for FACS assay trial 2, shown in Figure 3.14. Cells were surface stained with 2D7.

For all FACS assays, the gating was determined as follows: The P1 gate was based on the expected cellular shape of Jurkat cells with side scatter area (SSC-A) on the y axis, and forward scatter area (FSC-A) on the x axis. The single cell gate was based on width of cells with forward scatter height (FSC-H) on the x axis, and forward scatter width (FSC-W) on the x axis. The live cell gate was based on DAPI absorbance in surface stained samples with DAPI absorbance (DAPI-A) on the y axis and forward scatter area (FSC-A) on the x axis. Final quadrant gates were placed so that the majority of cells in the unstained and untreated sample were located in the $CCR5^{-}/GFP^{-}$ quadrant (Q4). There were about 30% of cells with baseline GFP fluorescent intensity above 10^{3}. All gates were placed under the guidance of Dr. Gravano.

Figure Appendix C.3 FACS Assay in JLTRG cells Trial 2 – Tat (Positive Control)

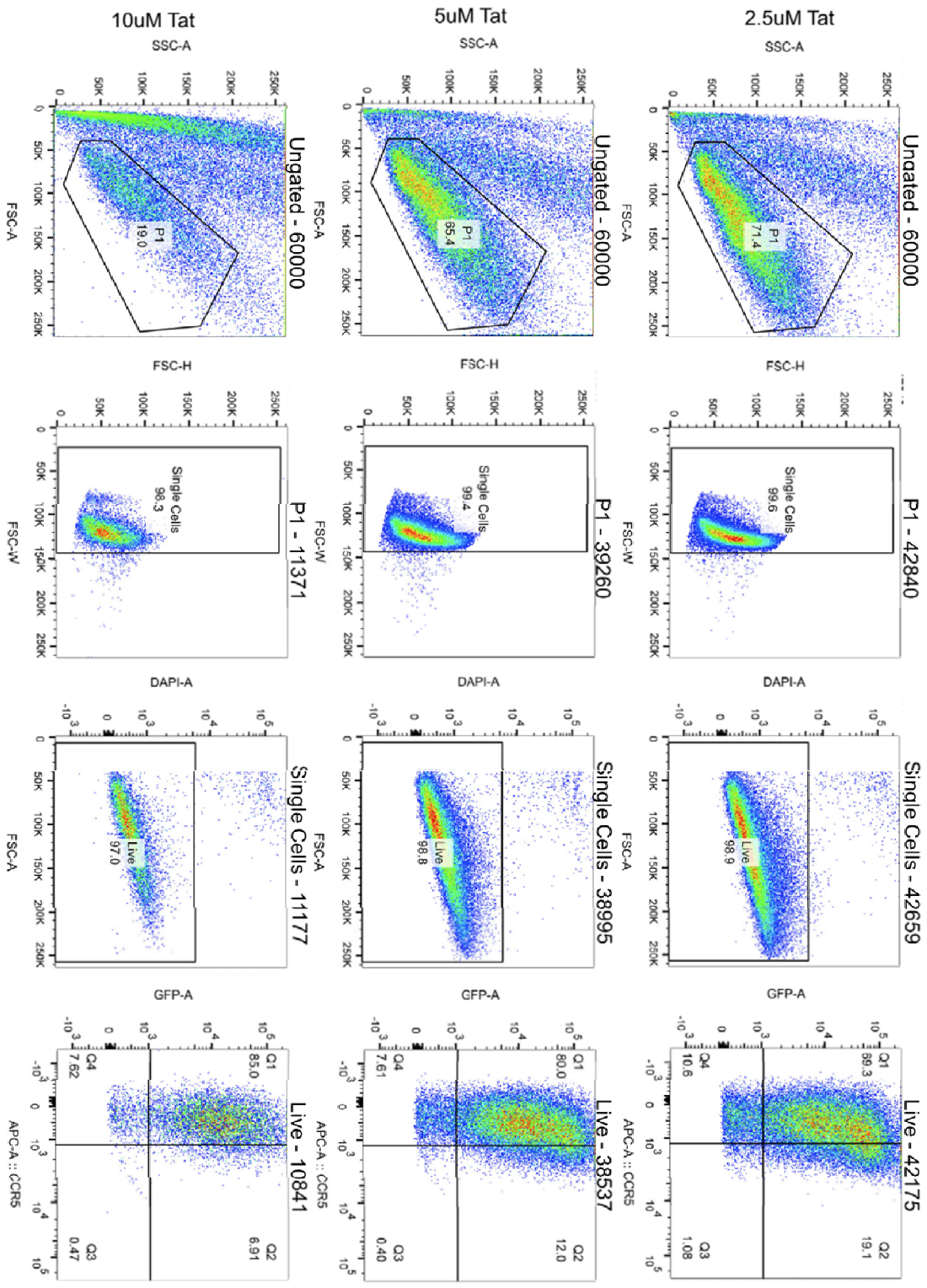


Figure Appendix C.4 FACS Assay in JLTRG cells Trial 2 – 5P14 (Negative Control)

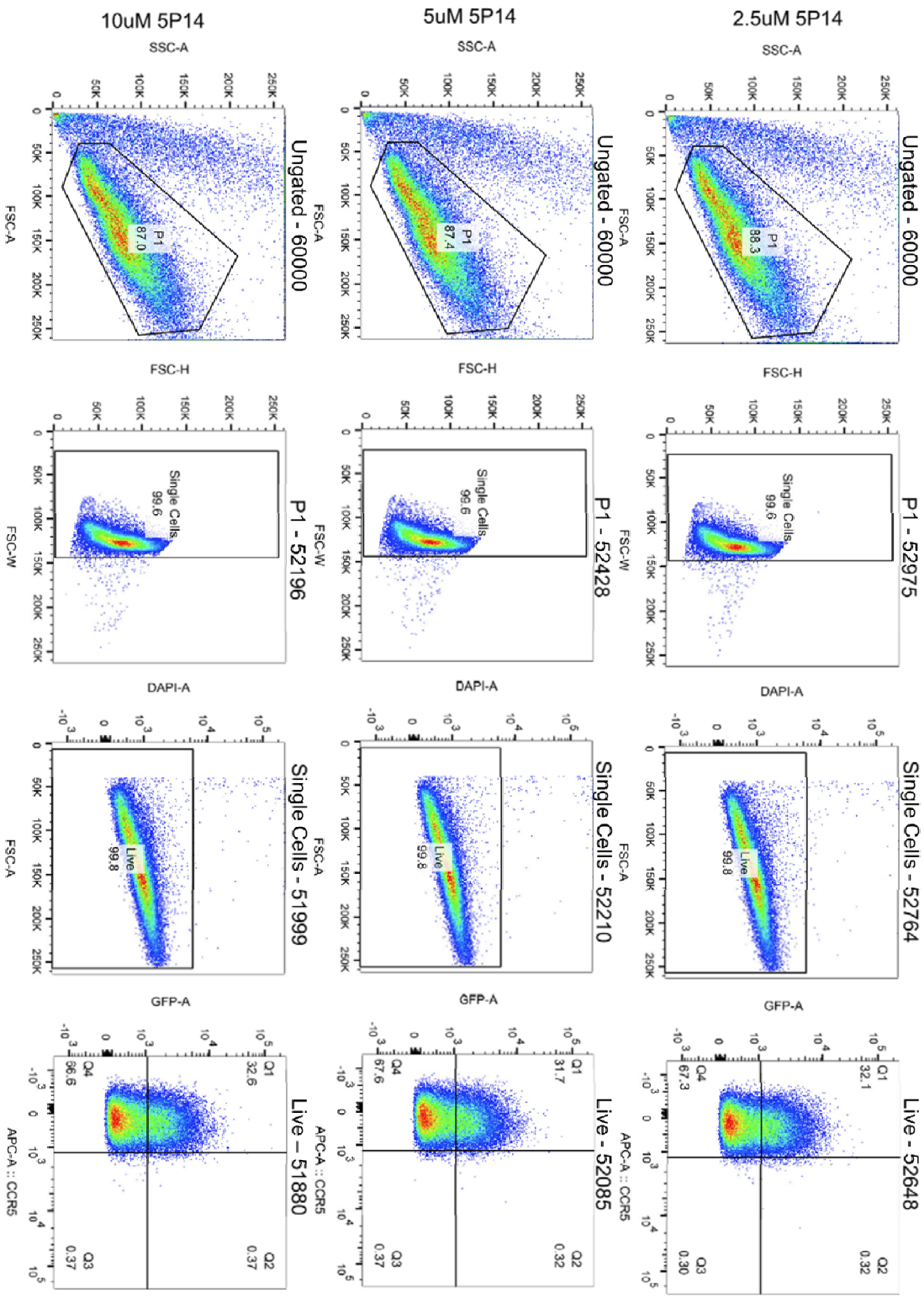


Figure Appendix C.5 FACS Assay in JLTRG cells Trial 2 – 5P14-L-Tat (Sample)

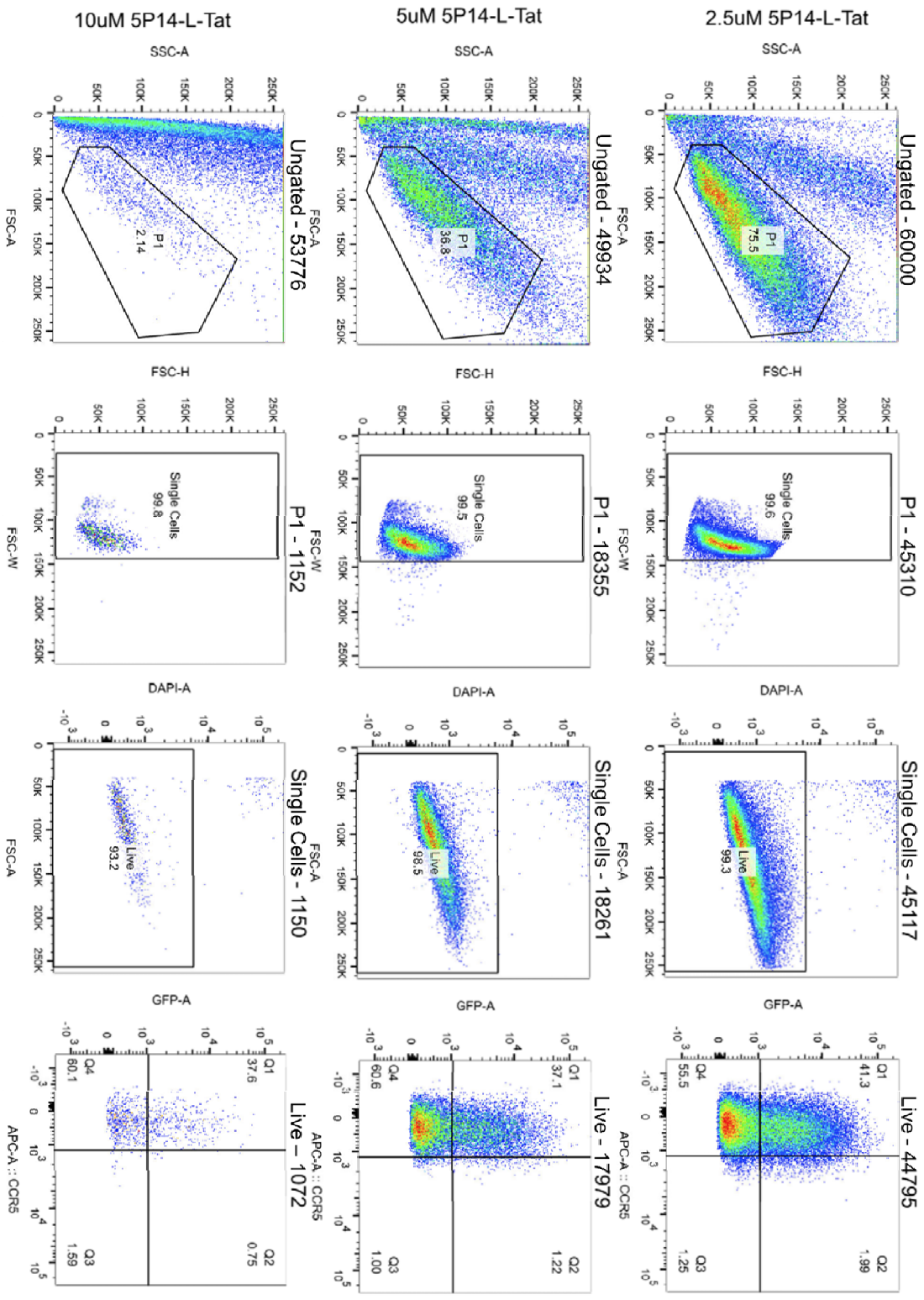


Figure Appendix C.6 FACS Assay in JLTRG cells Trial 3 – 5P14 (Negative Control) 4 hour wash

These figures show the full gating for FACS assay trial 3, shown in Figure 3.16. All samples are surface stained with 2D7 after washing 4 hours prior to staining.

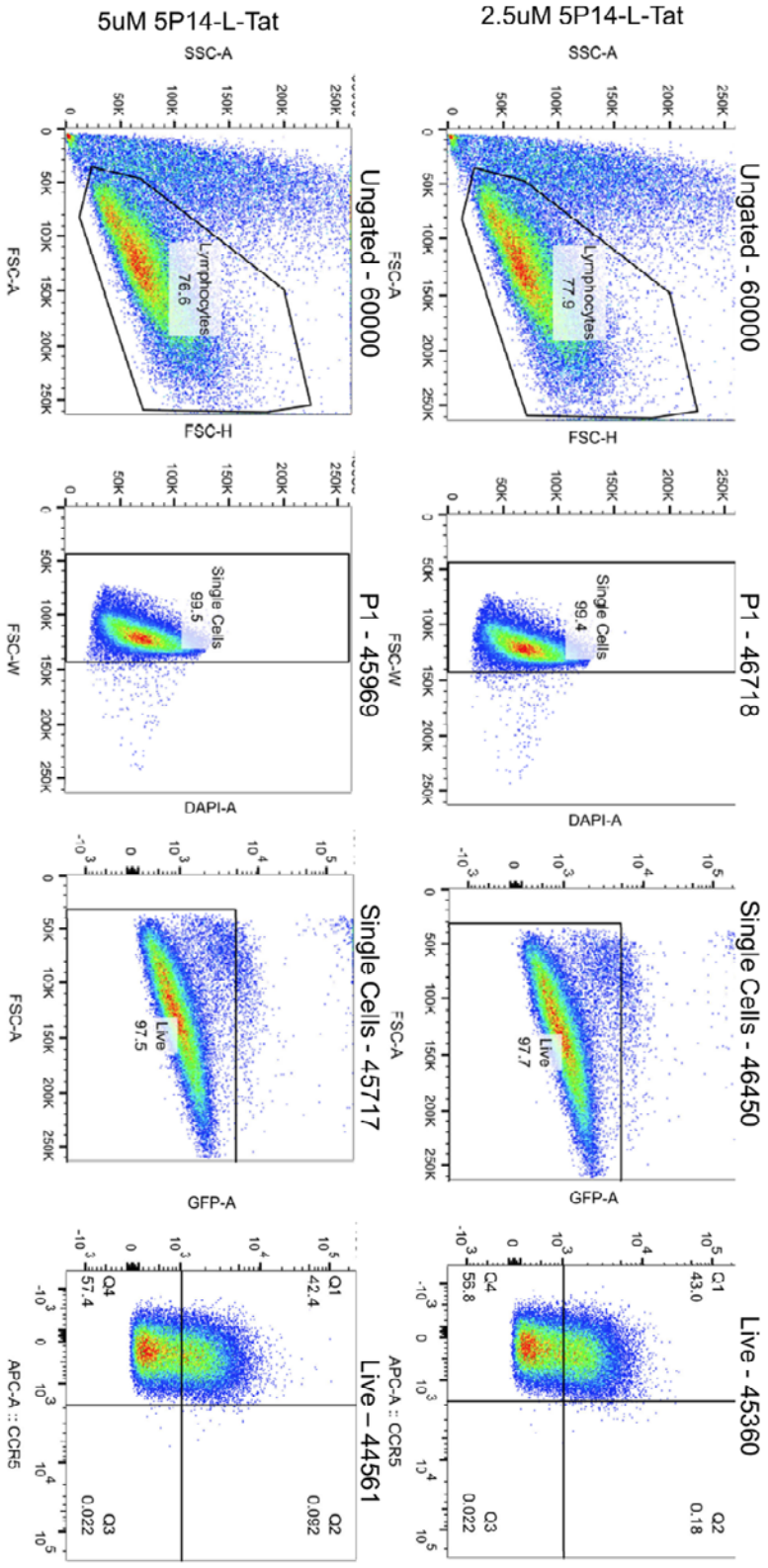


Figure Appendix C.7 FACS Assay in JLTRG cells Trial 3 – Tat (Positive Control) 4 hour wash

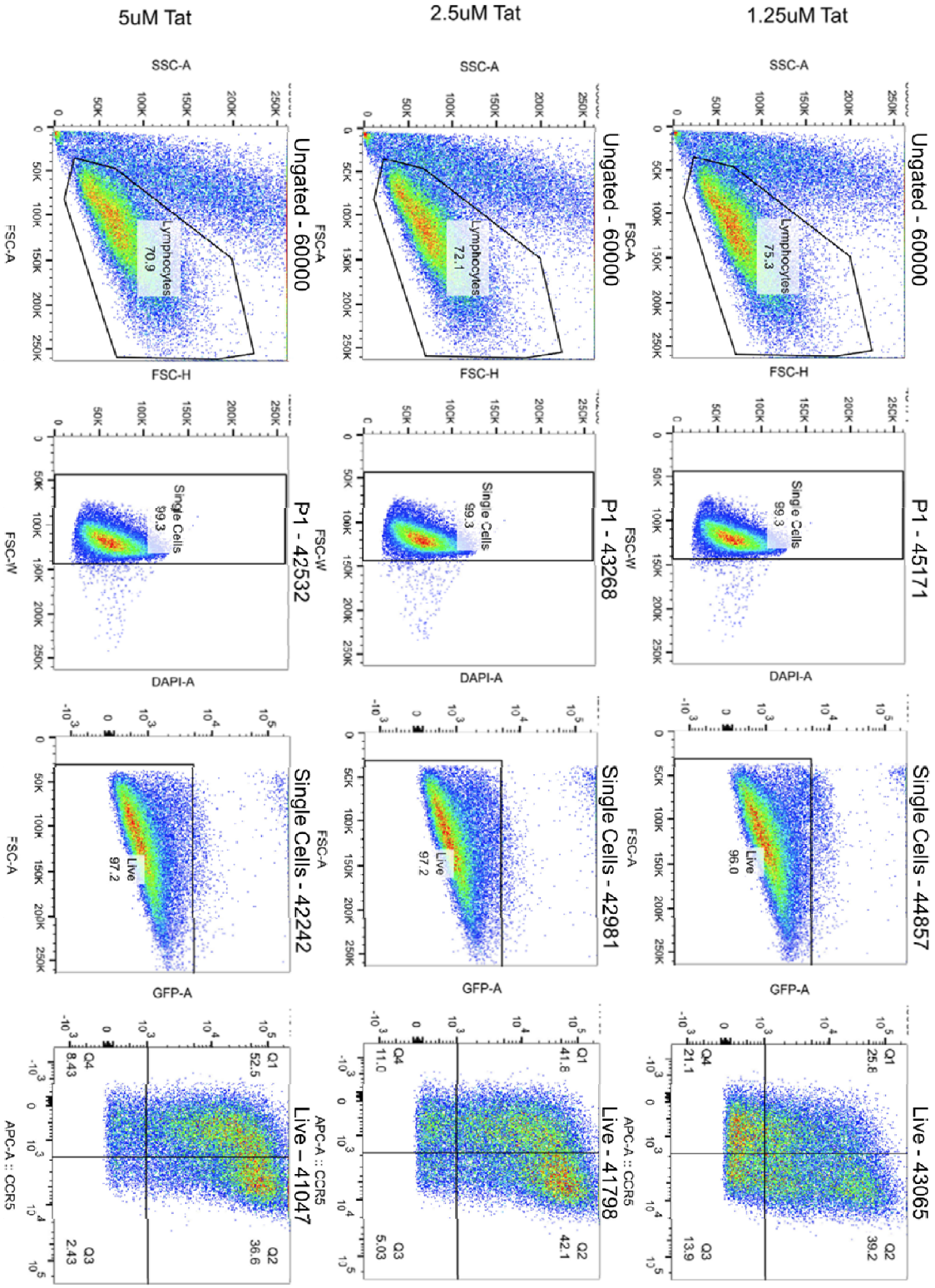


Figure Appendix C.8 FACS Assay in JLTRG cells Trial 3 – 5P14-L-Tat (Sample) 4 hour wash

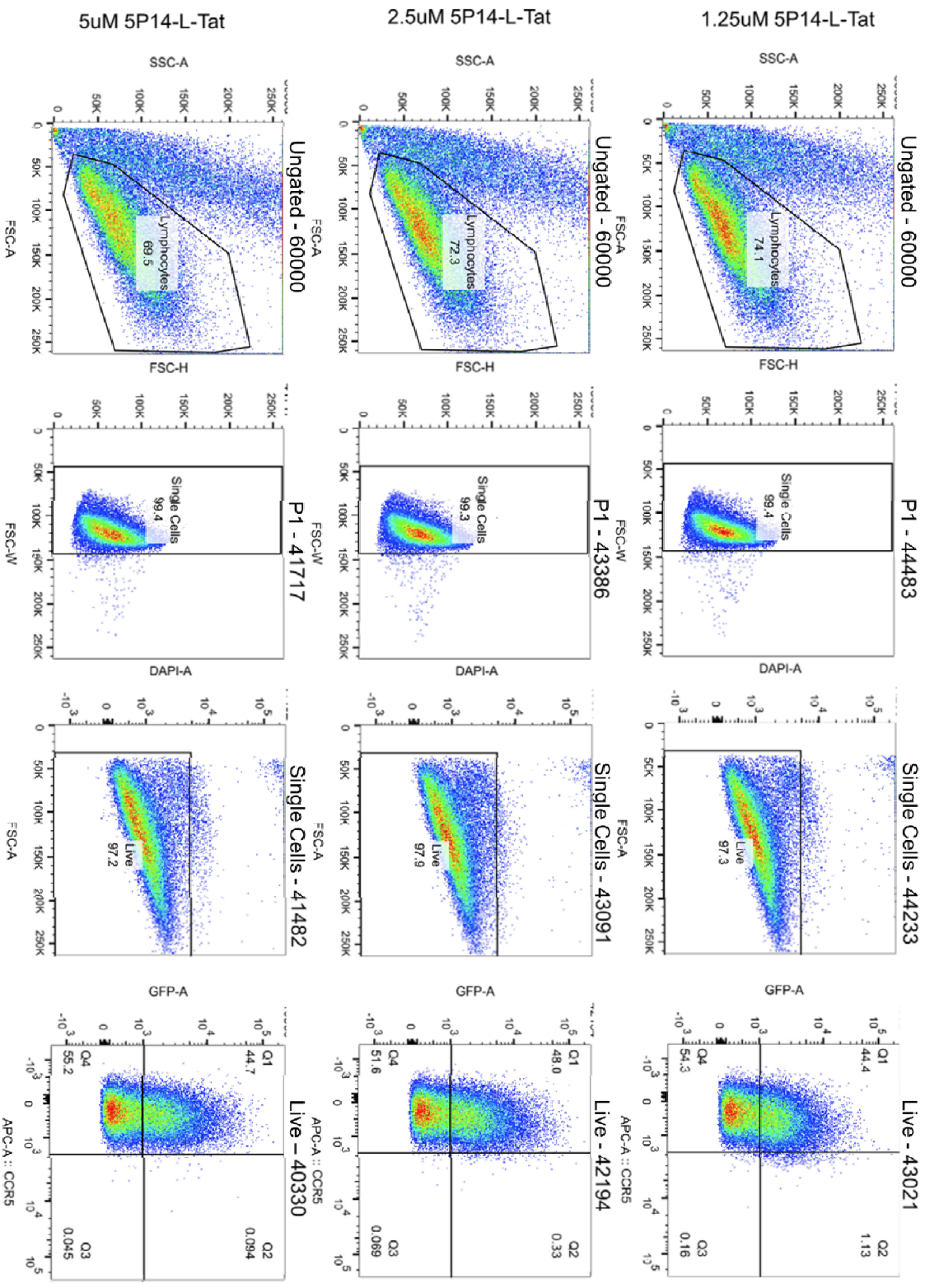


Figure Appendix C.9 FACS Assay in JLTRG cells Trial 3 – 5P14 (Negative Control) 12 hour wash

These figures show the full gating for FACS assay trial 3, shown in Figure 3.16. All samples are surface stained with 2D7 after washing 12 hours prior to staining.

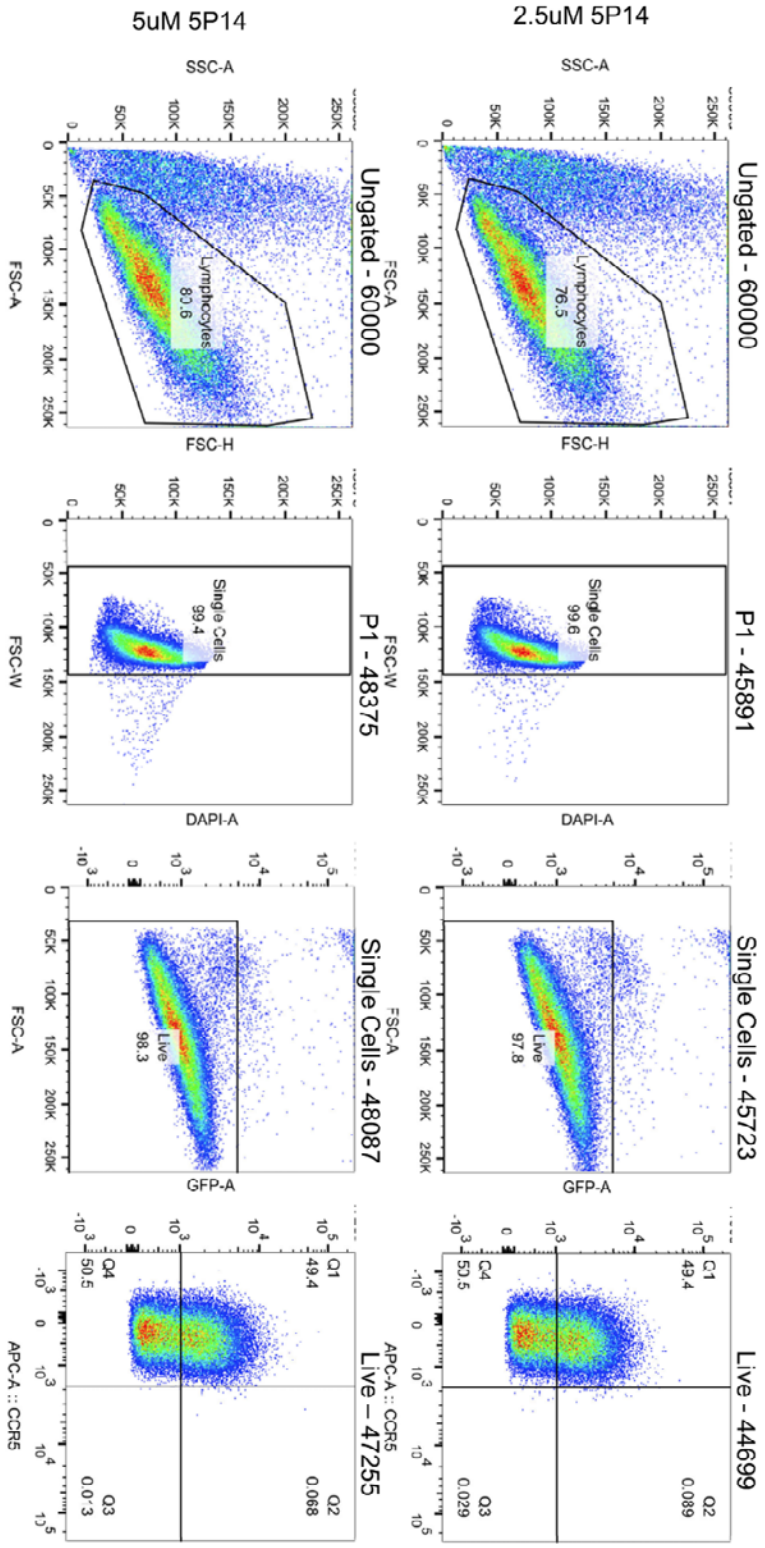


Figure Appendix C.10 FACS Assay in JLTRG cells Trial 3 – Tat (Positive Control) 12 hour wash

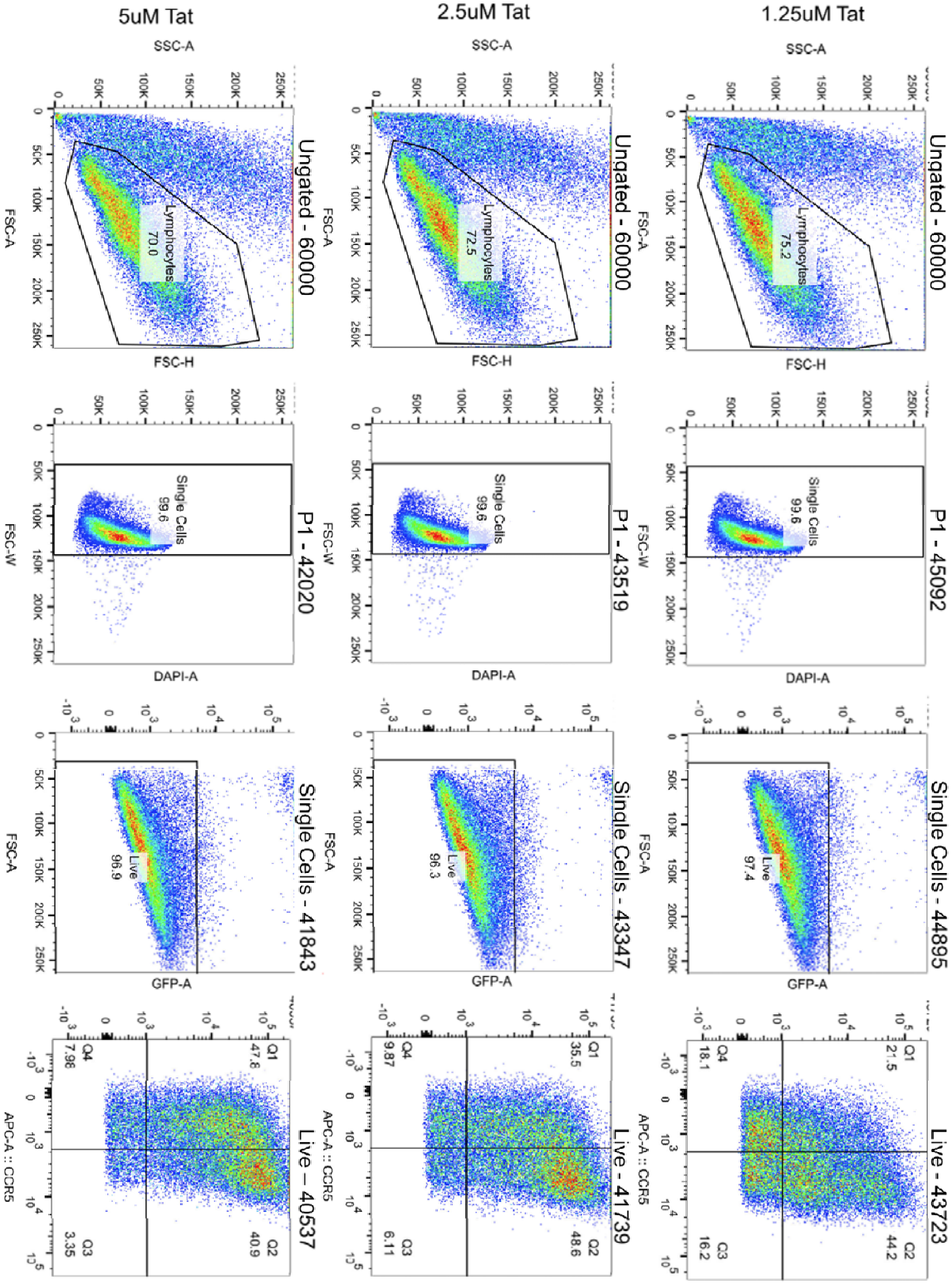


Figure Appendix C.11 FACS Assay in JLTRG cells Trial 3 – 5P14-L-Tat (Sample) 12 hour wash

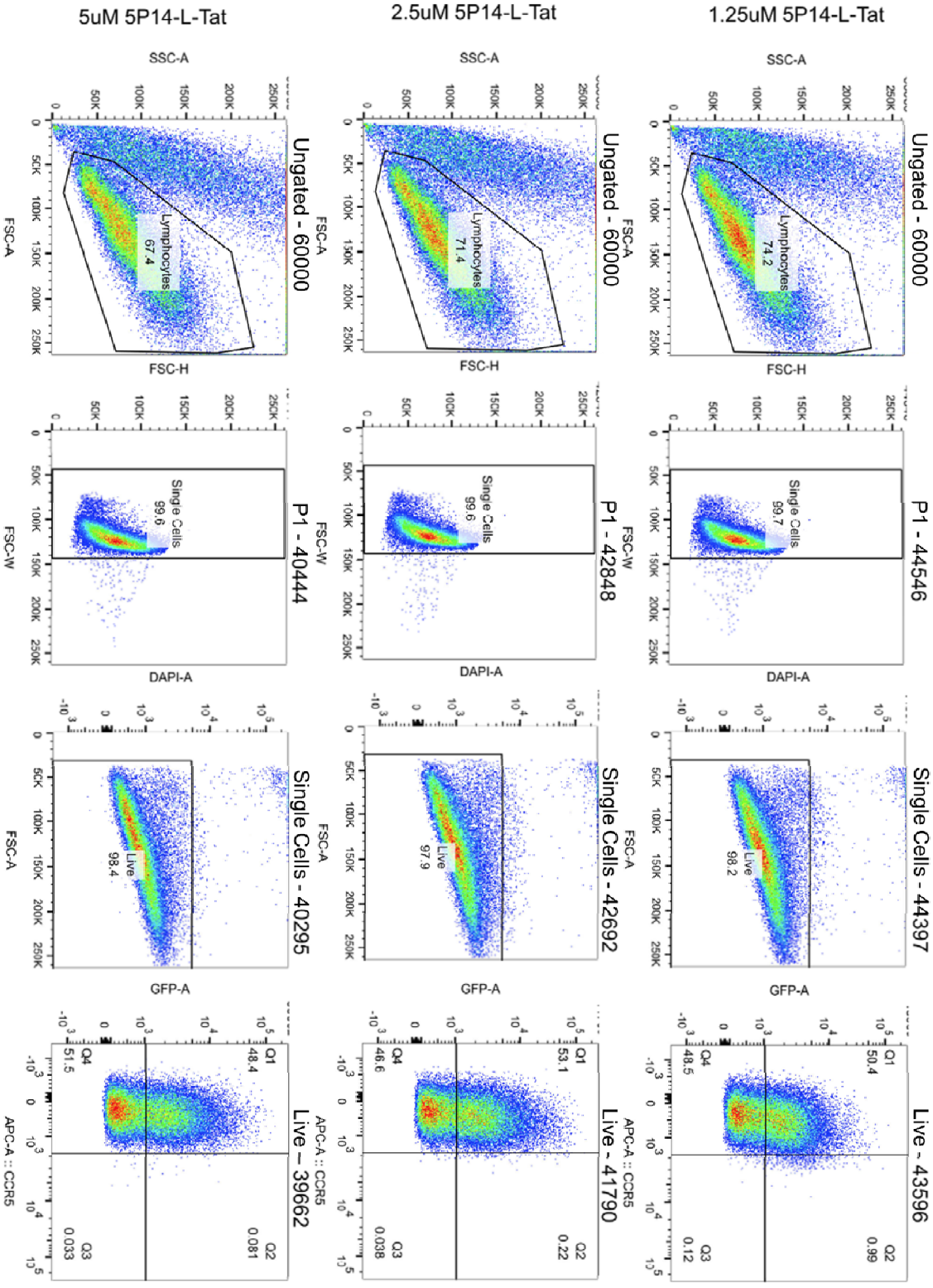


Figure Appendix C.12 FACS assay to test J-Lat 8.4 cells – unstained and untreated controls

These figures show the full gating for FACS assay to test J-Lat 8.4 cells for the presence of CCR5 and the ability of Tat to activate the latent LTR-HIV, shown in Figure 3.19. All samples are surface stained with the indicated antibody. Tat samples are stained with J418F1.

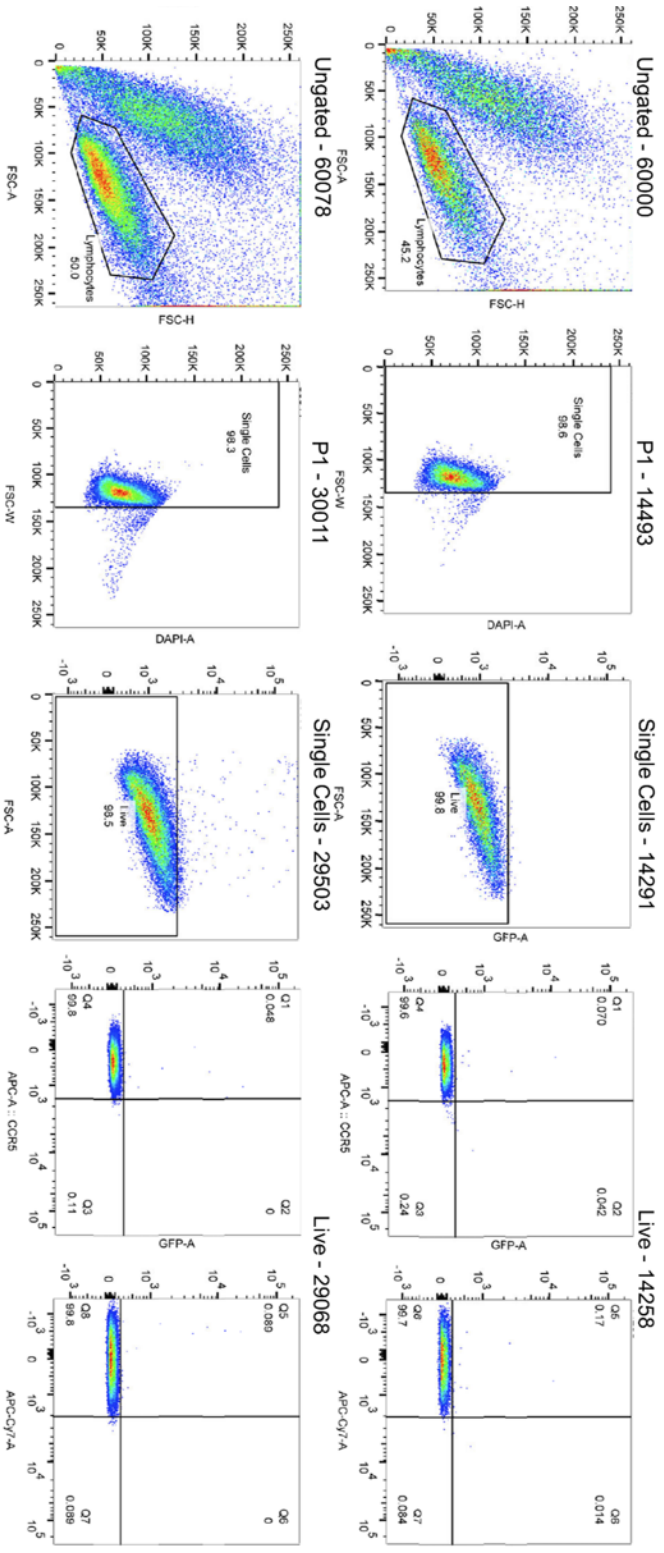


Figure Appendix C.13 FACS Assay to Test J-Lat 8.4 Cells – Staining with anti-CCR5 antibodies

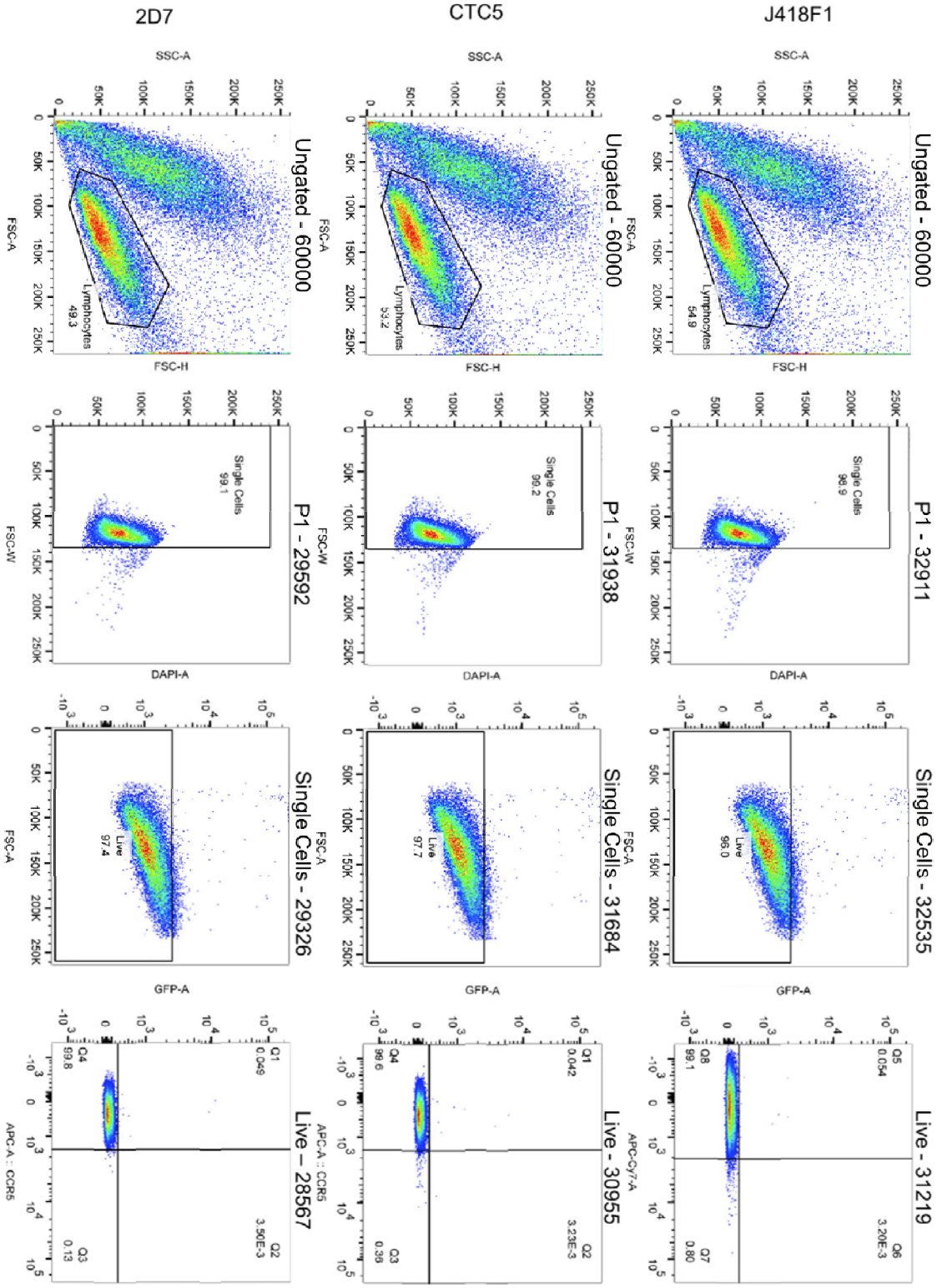


Figure Appendix C.14 FACS Assay to Test J-Lat 8.4 Cells – Treated with Tat (5uM & 2.5uM)

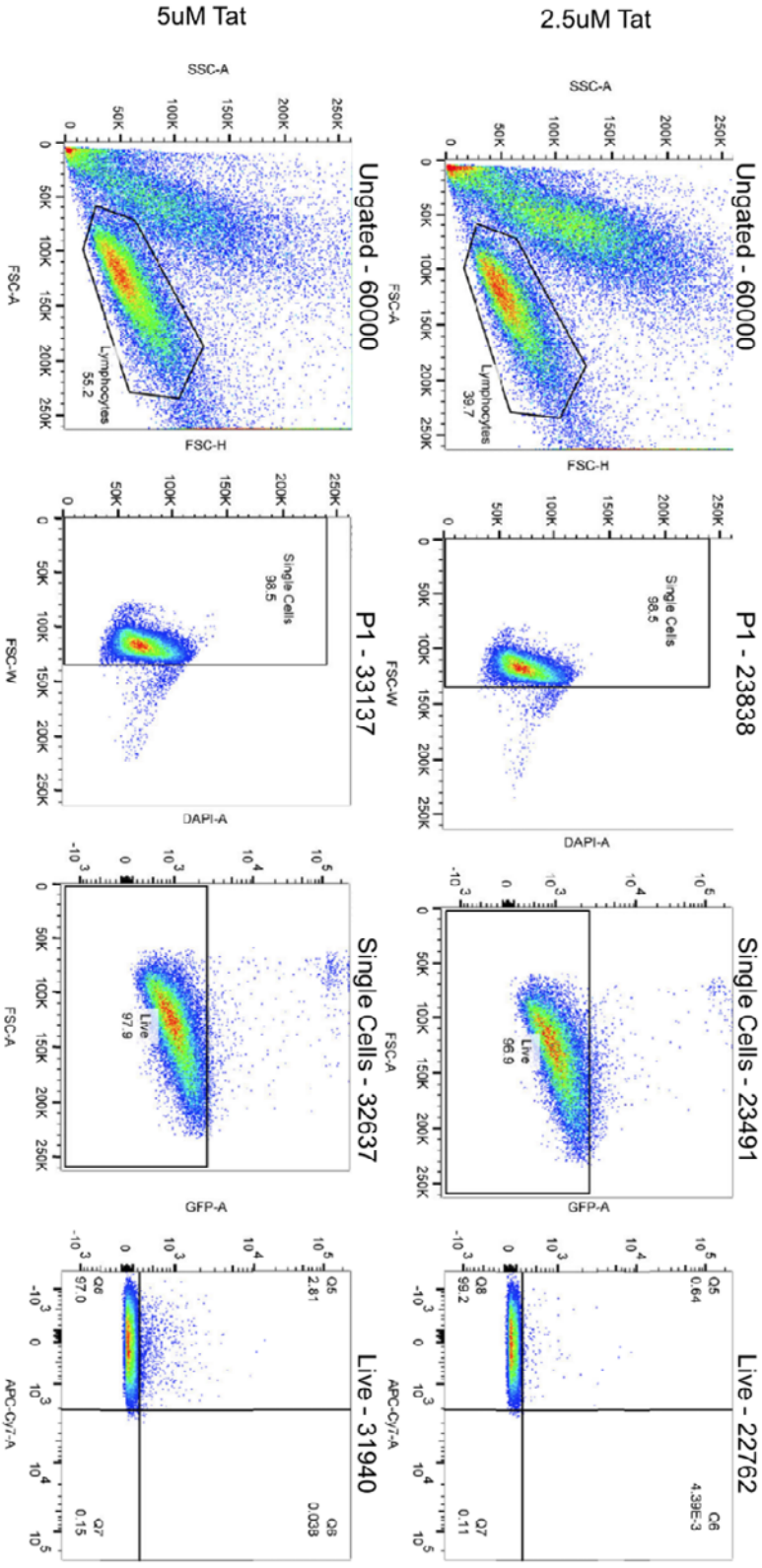


Figure Appendix C.15 FACS Assay to Test J-Lat 8.4 Cells – Treated with Tat (20uM & 10uM)