Characterisation of α4β7 and DC-SIGN reactivity of HIV-1 subtype C transmitted founder variants compared with chronic infection controls

By Netanya Bernitz

Thesis presented for the

DEGREE OF MASTER OF SCIENCE

In the Department of Molecular and Cell Biology

University of Cape Town

October 2013

Supervisor: Dr Zenda Woodman

Table of contents

Plagiarism declaration

I, Netanya Bernitz, know the meaning of Plagiarism and declare that all of the work in this document, save for that which is properly acknowledged, is my own. Neither the whole work, nor part thereof has been, is being, or is to be submitted for any degree or examination at any other university.

Signature of candidate:_____

Signed on the _____ day of _____, 2013.

Acknowledgments

I am grateful to Prof D. Erle for the pCDM8_α4 and pCEP4_β7 plasmid DNA, Prof V. Learner for the Cy3 antibody and L. Shuping, P. Moore and G. Bandawe for the *env* clones. I thank Prof C. Williamson for the use of her P2+ tissue culture facility and Dr D. Chopera for his assistance with the electroporator. I thank D. Bowers and E. Smit for their assistance with flow cytometry and S. Cooper for her assistance on the fluorescent microscope.

I thank my parents Herman and Zephné for giving me the incredible opportunity to study at UCT and to live in one of the most beautiful cities in the world. Thank you for not asking too many questions and for silently believing that I am capable of great things. I thank Tamara and Peter for their long distance support and love (and help with Photoshop). I thank Christopher for his endless encouragement and support and making me smile and laugh so often. I thank Ana my best friend, statistics guru and greatest fan. I cannot thank you enough for everything you have done for me and continue to do for me every day, you are going to become an epic scientist I am sure of that.

I thank the Woodman laboratory for their help and shared love of HIV research, specifically Lilly who was always so willing to help me at the drop of a hat. You have become one of my closest friends.

I thank my supervisor Dr Z. Woodman for her motivation throughout my undergraduate and postgraduate degrees. I am grateful for her guidance and the time and effort she spent editing my thesis while on maternity leave.

I thank the NRF and UCT by whom this work was funded.

List of abbreviations

α	alpha
β	beta
Δ	delta
µ	micron
µg	micrograms
µL	microlitres
µM	micromolar
AIDS	acquired immune deficiency syndrome
APC	antigen presenting cell
APS	ammonium persulfate
ARP	AIDS Research and Reference Reagent Program
bp	basepairs
BSA	bovine serum albumin
C	constant region
CCR5	C-C chemokine receptor type 5
cDNA	complimentary deoxyribonucleic acid
CHO	Chinese hamster ovary
CLR	C-type lectin receptors
CRF	circulating recombinant forms
CXCR4	C-X-C chemokine receptor type 4
dATP	deoxyadenosine triphosphate
DC-SIGN	dendritic cell-specific ICAM-3 grabbing non-integrin
dH_20	distilled water
DMEM	Dulbecco's modified Eagle medium
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphates
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid

ELISA	enzyme-linked immunosorbent assay
Endo H	endoglycosidase H
Env	envelope
F	forward
FACS	fluorescence-activated cell sorting
FCS	foetal calf serum
G	glutamine
g	times gravitational
GALT	gut-associated lymphoid tissue
gp	glycoprotein
His-tag	polyhistidine-tag
HEV	high endothelial venules
HIV	human immunodeficiency virus
HRP	horseradish peroxidase
ICAM	intercellular adhesion molecule
IIDMM	Institute of Infectious Disease and Molecular Medicine
IgG	immunoglobulin G
IL	interleukin
IMC	infectious molecular clones
kDa	kiloDaltons
LFA-1	lymphocyte function-associated antigen 1
LTR	long terminal repeat
M	molar
MAdCAM-1	mucosal addressin cell adhesion molecule-1
MCS	multiple cloning site
mg	milligrams
MHC	major histocompatibility complex
mL	millilitres
ms	milliseconds
MVB	multivesicular body
N	asparagine

ng	nanograms
NICD	National Institute for Communicable Diseases
NIH	National Institutes of Health
nm	nanometres
nM	nanomolar
NRF	National Research Fund
nt	nucleotide
PAGE	polyacrylamide gel electrophoresis
PEI	polyethylenimine
PBMC	peripheral blood mononuclear cells
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PMSF	phenylmethylsulfonyl fluoride
PNG	potential N-glycan
PVDF	polyvinylidene difluoride
R	reverse
RIPA	radioimmunoprecipitation assay
RNA	ribonucleic acid
RLU	relative light units
RPMI	Roswell Park Memorial Institute medium-1640
RT	room temperature
SAB	sample application buffer
SDM	site-directed mutagenesis
SDS	sodium dodecyl sulfate
SGA	single genome amplification
SIV	simian immunodeficiency virus
STET	sodium chloride, tris, ethylenediaminetetraacetic acid, Triton X-100
STI	sexually transmitted infections
T/F	transmitted founder
T75	75 cm^2 tissue culture flask
Tat	transcriptional transactivator protein

TBS	tris-buffered saline
Tris	Tris (hydroxymethyl) aminomethane
tRNA	transfer RNA
U	units
UCT	University of Cape Town
UNAIDS	The Joint United Nations Programme on HIV/AIDS
USA	United States of America
V	variable region
WT	wild-type
°C	degrees Celsius

Abstract

Given the high genetic diversity of HIV, identifying genotypic signatures and phenotypic characteristics common to all transmitted founder (T/F) variants, irrespective of subtype, could aid the design of vaccine immunogens specific for these motifs. Recently, two putative T/F phenotypes were identified: 1) high affinity interactions between T/F glycoprotein (gp) 120 and the α4β7 integrin that might facilitate the binding and infection of CD4+ T cells in the gut and 2) enrichment of oligomannose type N-glycans on the surface of T/F gp120 that may be involved in binding to the dendritic cell-specific intercellular adhesion molecule-3 (ICAM-3)-grabbing non-integrin (DC-SIGN), a C-type lectin expressed on dendritic cells. DC-SIGN could facilitate the transfer of HIV from dendritic cells in the genital mucosa to CD4+ T cells in the lymph nodes and thus play an important role in transmission. As T/F variants seem to have distinct N-glycosylation patterns compared with variants found later in infection, this study suggests that T/F variants may carry an optimum arrangement of N-glycans that enables interactions with α4β7 and DC-SIGN that are essential for successful transmission in the genital tract. We aimed to investigate the role of α4β7 and DC-SIGN in HIV subtype C sexual transmission using matched single genome amplification (SGA)-derived T/F and chronic infection envelope (*env*) clones from five CAPRISA 002 study participants. In order to compare α4β7 reactivity we transfected HEK293T cells with the cDNA of the integrins. However, due to sub-cloning strategy we were unable to transfect HEK293T cells and therefore we utilised HeLa and TZM-bl cells that endogenously express $\alpha 4\beta 7$ instead. However, the Act-1 monoclonal antibody did not inhibit gp120-α4β7 interactions on HeLa or TZM-bl cells suggesting that binding of pseudovirus was non-specific. Pseudovirus produced in HEK293T cells may have had low α4β7 reactivity or α4β7 may not occur in its activated conformation on HeLa and TZM-bl cells. In order to test whether T/F variants were better transferred to CD4+ cells than the chronic infection controls a DC-SIGN mediated *trans*-infection assay was utilised. In contrast to recent reports we found no difference between T/F and chronic infection variants if we controlled for the effect of Env entry efficiency. Using site-directed mutagenesis (SDM) we deleted the potential N-glycans (PNGs) at positions N386 and N392, believed to carry oligomannose type N-glycans and thought to be

involved in DC-SIGN binding. The deletion of both PNGs not only resulted in a decrease in DC-SIGN reactivity compared with wild-type (WT) for both T/F and chronic infection Env pseudotyped viruses but also lowered Env entry efficiency, suggesting that these PNGs are most likely important for Env structure and function. However, when only a single PNG was deleted the transfer of HIV to CD4+ cells was not affected, suggesting that the N-glycosylation at one site is sufficient to support *trans*-infection. The deletion of the PNG at position N392 rendered the chronic infection Env clone non-functional suggesting that the effect of N-glycans could be clone specific. Therefore, instead of specific N-glycans interacting with DC-SIGN it is possible that the transfer of HIV to CD4+ cells requires an optimum arrangement of oligomannose type N-glycans. As the entry efficiency of the Env clones could play a role in determining the level of *trans*-infection *in vivo*, the presence of the PNGs at positions N386 and N392 could contribute to the overall success of transmission, explaining the conservation of these sites across HIV subtype C variants. Further study to confirm the role of the N-glycans at positions N386 and N392 in HIV transmission could provide invaluable information for the overall understanding of HIV transmission and the design of an HIV vaccine.

Chapter 1: Literature Review: HIV transmission

1.1 Introduction

Human immunodeficiency virus (HIV) is a ribonucleic acid (RNA) virus that targets and destroys CD4+ cells of the human immune system: CD4+ T cells, macrophages and dendritic cells (Dalgleish et al. 1984; Maddon et al. 1986; Duncan & Sattentau 2011; Loré et al. 2005). The UNAIDS 2012 Global Report estimated that ~34 million people are infected with HIV worldwide, ~23.5 million of those infected live in sub-Saharan Africa and ~5.6 million of those live in Southern Africa (Joint United Nations Programme on HIV/AIDS 2012). HIV type 1, hereinafter referred to as "HIV", occurs globally, subdivided into nine distinct subtypes: A, B, C, D, F, G, H, J and K (Hemelaar et al. 2006) with subtype C being the dominant subtype in Southern Africa (Van Harmelen et al. 1999). HIV type 2 is localised to West Africa (Clavel et al. 1987) and is less pathogenic (Levy 1995). Recombination between viral subtypes has resulted in the existence of mosaic viruses known as circulating recombinant forms (CRF) (Chow et al. 2013; Thomson et al. 2005).

Env, the most variable HIV protein can differ by 25-35% within an infected individual (Shaw & Hunter 2012) and up to 50% between viral subtypes (Johnston & Fauci 2007; Gao et al. 2005). This extensive genetic diversity enables the disguise of Env epitopes through N-glycosylation, trimerisation and conformational flexibility (Quiñones-Kochs et al. 2002; Eggink et al. 2010; Liu et al. 2008). HIV's high genetic diversity, the existence of multiple subtypes, the structural elasticity of viral proteins, especially Env and the ability of the virus to escape the immune response are major challenges to effective vaccine development. Antiretroviral (ARV) drugs were the first successful step in HIV management but drug resistance is a major shortcoming (Johnson et al. 2011; Marconi et al. 2008).

As 80% of infections are as a result of a single HIV variant the low diversity at transmission could provide a brief window of opportunity to prevent HIV sexual transmission (Abrahams et al. 2009; Derdeyn et al. 2004; Keele et al. 2008; Haaland et al. 2009; Salazar-Gonzalez et

al. 2008). Identifying the transmission motif(s) that allow(s) these T/F variants to be selectively transmitted is pivotal to the design of vaccine immunogens that target these variants before they diversify, thereby preventing HIV infection.

As T/F variants have fewer PNGs than those from chronic infection (Derdeyn et al. 2004; Chohan et al. 2005) and recently it was shown that the Env PNGs of T/F variants were enriched with oligomannose type N-glycans (Go et al. 2011), the N-glycosylation of Env could comprise the transmission motif. The binding between Env and α4β7 on CD4+ T cells and DC-SIGN on dendritic cells has been suggested to play a role in HIV transmission and these interactions are affected by gp120 N-glycosylation (Geijtenbeek, Kwon, et al. 2000; Feinberg et al. 2007; Hong et al. 2007; Eggink et al. 2010; Nawaz et al. 2011). The potential role of Env N-glycosylation in HIV transmission in relation to the receptors α4β7 and DC-SIGN will be the focus of this review.

1.2 The HIV transmission genetic bottleneck

The UNAIDS in 2009 reported that sexual transmission is the leading cause of new HIV infections with ~80% of new infections worldwide being caused by homo- and hetero-sexual transmission (Joint United Nations Programme on HIV/AIDS 2009). The rate of sexual transmission across a healthy mucosa has been calculated to be a mere 0.02% per coital act (Wawer et al. 2005), suggesting that the immune system, in general is efficient at preventing HIV infection. Studies on sero-discordant couples (Gray et al. 2001; Wawer et al. 2005) and sex workers (Sagar et al. 2004) demonstrated that multiple exposures were required before infection occurred confirming that there are stringent physical and physiological barriers in place that HIV needs to breach before coming into contact with its target CD4+ cells. These barriers include the multi-layered stratified epithelium of the vagina (Haaland et al. 2009), mucous secretions containing antimicrobial peptides and defensins (Tjabringa et al. 2005), the dendrites of Langerhans cells that project into the lumen of the genital tract (de Witte et al. 2007), the presence of dendritic cells in the sub-mucosa expressing C-type lectin receptors (CLR) (Geijtenbeek, Torensma, et al. 2000) and the paucity of CD4+ T cells in the genital mucosa (Pudney et al. 2005).

When HIV transmission is successful, a genetic bottleneck is observed whereby a single variant from the diverse donor viral population initiates productive infection in the recipient in eight out of ten cases (Figure 1.1) (Abrahams et al. 2009; Derdeyn et al. 2004; Keele et al. 2008; Salazar-Gonzalez et al. 2008; Haaland et al. 2009).

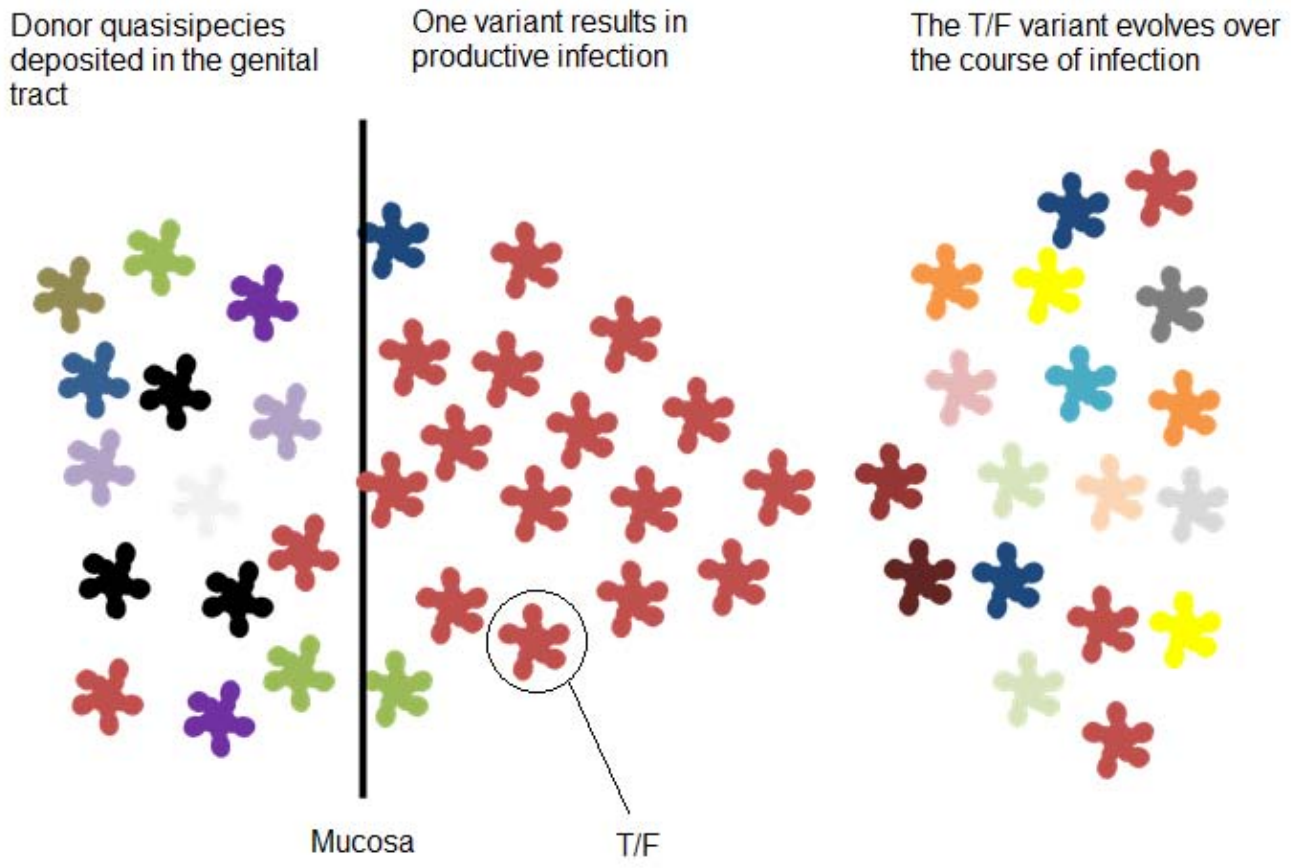


Figure 1.1 Schematic representation of the HIV genetic bottleneck during sexual transmission. HIV quasispecies is deposited in the genital tract. Although more than one viral variant may cross the mucosal barrier only a single variant, termed the transmitted founder (T/F) variant, is successful in causing productive infection. The T/F population evolves over the course of infection, which includes the alteration of the N-glycan repertoire on the surface of glycoprotein (gp) 120 to escape the humoral response.

The integrity of the mucosa appears to be a major contributing factor and aetiology of the genetic bottleneck as sexually transmitted infections (STI) (Horbul et al. 2011; Haaland et al. 2009), the phase of the menstrual cycle (Patton et al. 2000; Wira & Fahey 2008), certain contraceptives (Sagar et al. 2004) and microbicides (Dayal et al. 2003; Cone et al. 2006; Phillips et al. 2000) compromise the integrity of the mucosa and result in the transmission of multiple HIV variants (Abu-Raddad & Longini 2008; Abrahams et al. 2009; Sagar et al. 2004; Haaland et al. 2009).

The HIV genetic bottleneck results in the transmission of a single variant, the T/F virus, which evolves over time due to the high recombination frequency (Rhodes et al. 2003) and mutation rate of the virus (Abram et al. 2010). Although many studies have characterised the early stages of HIV infection, it was only with the advent of SGA that T/F variants were accurately identified and thus characterised (Keele et al. 2008; Salazar-Gonzalez et al. 2008; Lee et al. 2009).

SGA involves limited dilutions of viral complimentary deoxyribonucleic acid (cDNA) followed by two rounds of polymerase chain reaction (PCR) in order to ensure the presence of a single template molecule and to prevent PCR induced recombination (Simmonds et al. 1990; Palmer et al. 2005). The ability to unambiguously identify the T/F variant has enabled studies to elucidate the identity of the transmission motif(s), a distinct characteristic(s) that facilitates viral movement across the genital mucosa and/or its evasion of the immune response.

1.3 Genotypic signatures and phenotypic characteristics of T/F variants

It is well established that T/F variants have a strong preference for the CCR5 co-receptor (Keele et al. 2008; Salazar-Gonzalez et al. 2009) used for viral attachment and entry into the host cell. The CCR5-Δ32 polymorphism is a well-defined CCR5 allelic variation where the deletion of 32 basepairs (bp) completely inhibits viral entry (Samson et al. 1996) so that homozygous CCR5-Δ32 individuals are resistant to infection by R5 tropic variants (Marmor et al. 2001). It was recently observed that subtype C T/F variants were more sensitive to the CCR5 antagonist Maraviroc than chronic infection variants when using cells with high levels of the CCR5 co-receptor (Parker et al. 2013), in contrast to earlier studies that found no difference (Wilen et al. 2011; Parrish et al. 2012). This suggests that restricted use of certain CCR5 conformations could be involved in HIV transmission.

Potential transmission motifs in the CCR5 and CD4 binding sites, the signal peptide and the cytoplasmic domain of Env were recently identified (Gnanakaran et al. 2011). These include a histidine residue at position 12 within the signal peptide, previously shown to be

associated with transportation of the nascent peptide in the endoplasmic reticulum, expression, incorporation and infectivity of the viral particle (Asmal et al. 2011) and the absence of a PNG at position N415, previously shown to be associated with escape from broadly neutralizing antibodies (Gnanakaran et al. 2010). It was suggested that the presence of a histidine residue and/or the absence of a PNG at position 12 and N415, respectively may provide an advantage during transmission.

T/F variants have been shown to preferentially infect CD4+ T cells and not macrophages (Salazar-Gonzalez et al. 2009; Hladik et al. 2007; Brenchley et al. 2004) and to have enhanced neutralization sensitivity (Derdeyn et al. 2004) although studies have contradicted the latter finding (Frost et al. 2005; Keele et al. 2008; Rusert et al. 2005). The discrepancies between findings thus far could be due to the challenges facing these types of studies which include sample size, subtype variation, behavioral risk profiles of participants and differences in demographics, cloning strategy (SGA approach), the generation of pseudovirus and/or infectious molecular clones (IMC) and the use of T/F variants with unmatched chronic infection controls.

1.4 Overall structure and function of the HIV Envelope

HIV is an RNA virus comprised of nine genes: two structural genes *env* and *gag,* an enzymatic gene *pol* and six accessory genes *tat, rev, nef, vif, vpr* and *vpu* which make up the ~10 kb viral genome (Figure 1.2) (Varmus 1988; Frankel & Young 1998).

The structural *env* is the focus of many studies as it encodes the viral Env that interacts with host cell receptors to allow for viral attachment and entry into target CD4+ cells (Kwong et al. 1998). Env comprises of gp160, the precursor gene product that is cleaved into two parts, gp120 and gp41 (Stein & Engleman 1990; Hallenberger et al. 1992). The gp160 precursor protein, although non-functional and requires cleavage by the host protease furin to gp120 and gp41 (Moulard et al. 1999), can still be incorporated into infectious viral particles (Moore et al. 2006).

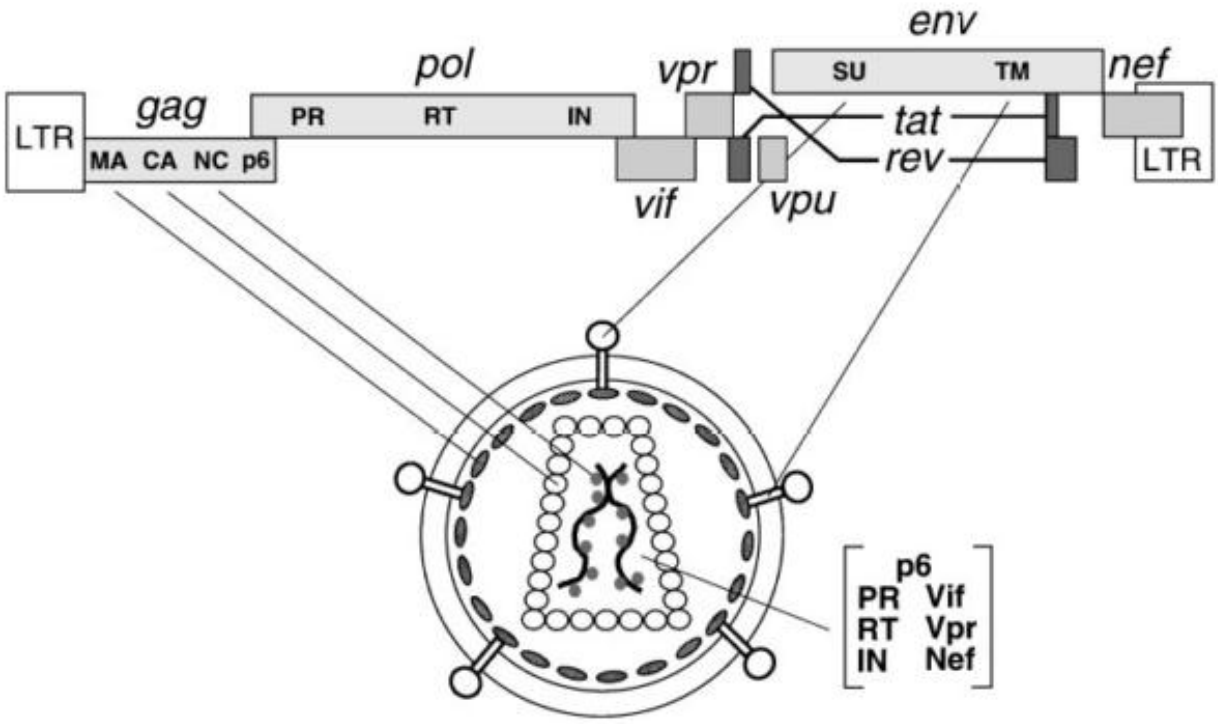


Figure 1.2 HIV genome and virion structure. The enzymatic (*pol*) and structural (*gag* and *env*) genes encode nine proteins: the protease (PR) the reverse transcriptase (RT), the integrase (IN) (Pol proteins); the matrix (MA), the capsid (CA), the nucleocapsid (NC) and p6 (Gag proteins); the surface/gp120 (SU) and the transmembrane gp41 (TM) (Env proteins). The six accessory genes *vif, vpr, vpu, tat, rev* and *nef* encode their own proteins. A 5' and 3' long terminal repeat (LTR) flanks the ~10 kb genome. The Gag and Env proteins make up the structural core and outer envelope of HIV, respectively and the three Pol proteins and three accessory proteins Vif, Vpr and Nef are incorporated within the viral particle. Tat and Rev are gene regulatory proteins expressed after integration into the host genome (Frankel & Young 1998).

Gp120 and gp41 form heterodimers through non-covalent bonds (Staropoli et al. 2000) and trimerise to form the HIV Env spike (Liu et al. 2008; Zanetti et al. 2006; Zhu et al. 2006; Caffrey et al. 1998; Chen et al. 2005). Gp120 is the extracellular receptor binding domain of Env that initially interacts with the CD4 receptor and gp41 is the transmembrane domain that anchors gp120 to the viral membrane and mediates fusion between the viral and host cell membranes (Weissenhorn et al. 1997).

The binding of gp120 to CD4 results in a conformational change that reveals a previously concealed co-receptor (CXCR4 or CCR5) binding site. The co-receptor binding event results in further conformational changes that allow gp41 to penetrate the host cell membrane and fuse the viral and host cell membranes (Checkley et al. 2011; Wilen et al. 2012; Liu et al. 2008; Kwong et al. 1998; Myszka et al. 2000). After membrane fusion, the viral RNA is reverse transcribed and integrated into the host cell's genome using the host's genetic

machinery. The integrated provirus generates new viral progeny thereby allowing further propagation and proliferation of the virus (Figure 1.3).

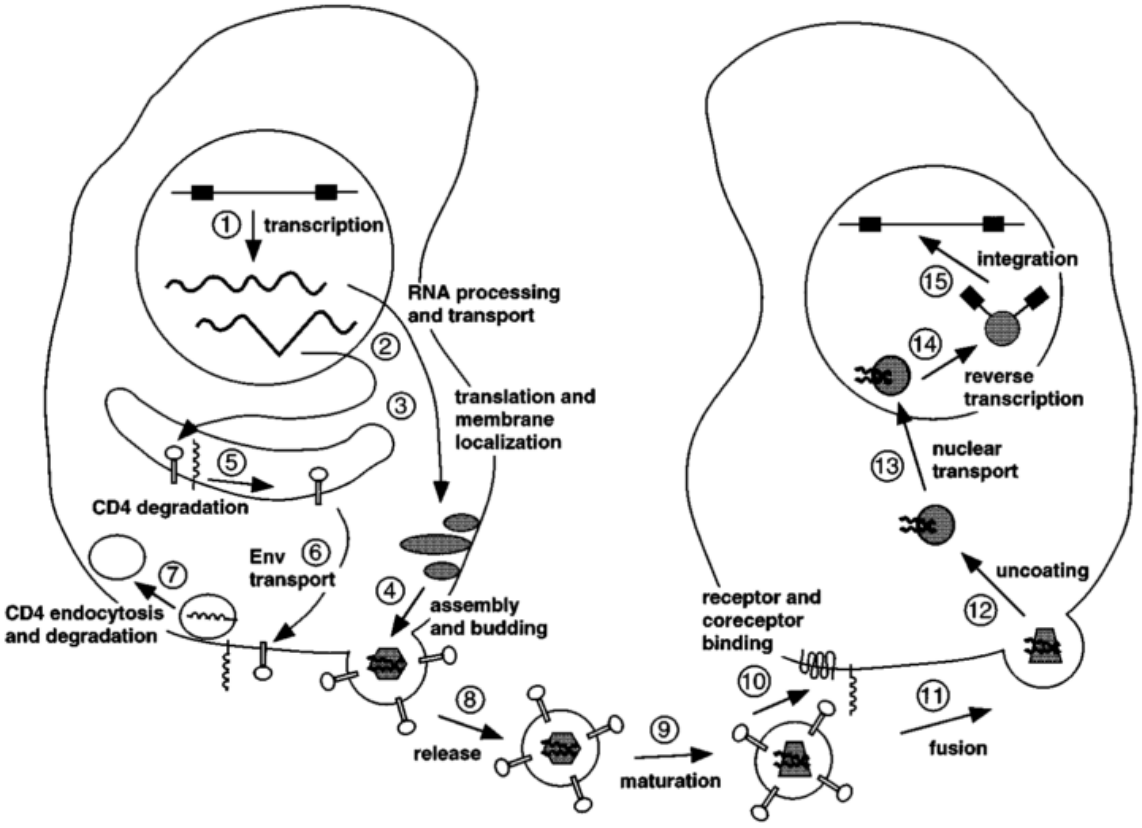


Figure 1.3 The HIV life cycle 1) Transcription of the integrated viral genome through transactivation of the 5' long terminal repeat (LTR) promoter by Tat 2) Spliced viral RNA is transported to the cytoplasm 3) Gag and Pol proteins are localised to the cell membrane 4) Assembly and budding of the viral particle is initiated 5) Env and CD4 are translated in the endoplasmic reticulum 6) Env is transported to the cell surface 7) CD4 is endocytosed and degraded 8) Immature viral particles containing Env are released from the host cell surface 9) Immature viral particles undergo maturation 10) Mature viral particles bind to target cells via CD4 and a co-receptor (CCR5 or CXCR4) 11) Fusion of the host cell and viral membranes 12) The viral contents is released into the host cell 13) Viral RNA is transported to the nucleus of the host cell 14) The viral RNA is reverse transcribed into complementary DNA (cDNA) 15) The viral cDNA is integrated into the host chromosome and the cycle begins again (Frankel & Young 1998).

1.5 Gp120, the receptor binding domain

1.5.1 The structure of gp120

The structure of gp120 has been extensively studied using enzymatic cleavage (Leonard et al. 1990), computer (Modrow et al. 1987) and oligomeric modelling (Kwong et al. 2000), X-ray crystallography (Weissenhorn et al. 1997) and cryo-electron microscopy tomography

(Zhu et al. 2006; Liu et al. 2008; Zanetti et al. 2006). The structure of the second and third variable loops were only recently solved (McLellan et al. 2011) as gp120 is highly N-glycosylated making it difficult to determine the complete structure.

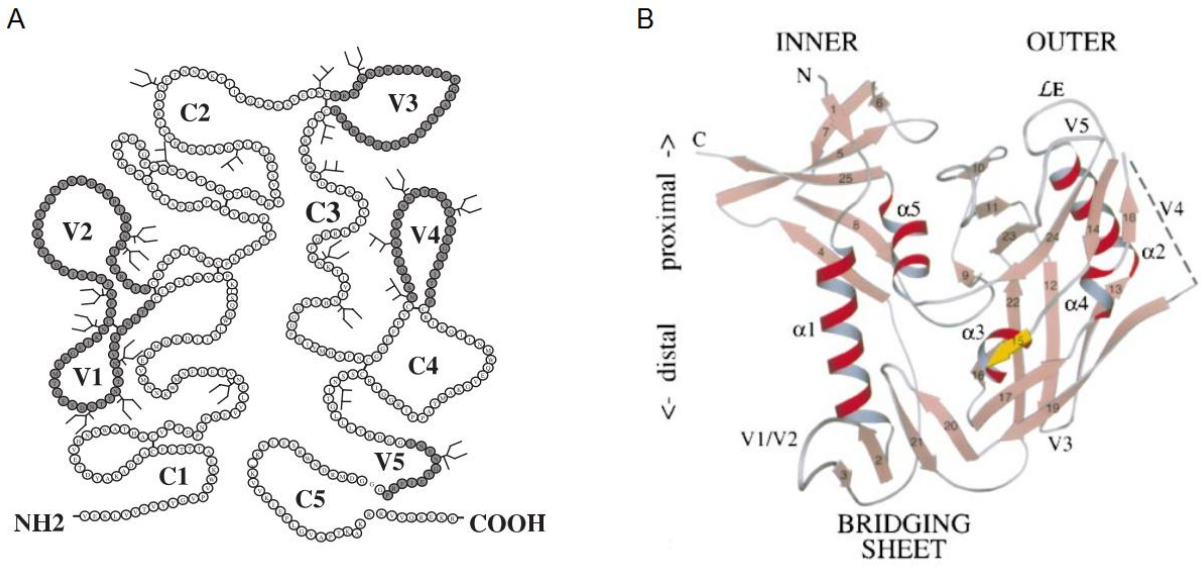


Figure 1.4 The structure of gp120 A) A schematic representation of the HIV glycoprotein (gp) 120 variable (V) and constant (C) regions. Gp120 carries an array of N-glycans (oligomannose or hybrid or complex and unknown) believed to shield antigenic epitopes from neutralizing antibodies. Adapted from McCaffrey et al. (2003) and Leonard et al. (1990) B) A ribbon diagram representing the conformation of gp120 upon binding to CD4. The V1/V2, V3 and V4 loops, the V5 region and the four β-stranded bridging sheet are indicated (Kwong et al. 1998).

Gp120 is comprised of five variable and five conserved regions (Starcich et al. 1986; Srinivasan et al. 1987; Willey et al. 1989). While the variable regions are highly elastic between isolates, the conserved regions show little or no variation (Schønning et al. 1996; Milich et al. 1993; Fenouillet et al. 1990). The variable regions, with the exception of variable region (V) five, are joined by nine disulphide bonds forming distinctive loops (Leonard et al. 1990; Chen et al. 2005) which are critical for Env assembly (Benham et al. 2000). The variable regions are positioned externally on the surface of gp120 and the conserved regions (C) are concealed within the core (Chen et al. 2005; Modrow et al. 1987; Kwong et al. 1998) (Figure 1.4).

The CD4 binding site lies at the interface of the inner and the outer domain (Kwong et al. 1998; Wyatt et al. 1998) and upon CD4 binding a conformational change causes the rearrangement of two pairs of double stranded β sheets, one from the inner and one from

the outer domain, to form a bridging sheet (Kwong et al. 1998). The formation of a bridging sheet reveals a previously concealed co-receptor binding site and the gp41 transmembrane stalk (Wyatt et al. 1995). These events lead up to the fusion of the viral and host cell membranes (Weissenhorn et al. 1997).

1.5.1.1 N-linked glycosylation

N-linked glycosylation is the attachment of an oligosaccharide to the asparagine of an Asn-X-Ser/Thr (NXT/S) sequon, where X is any amino acid but proline (Marshall 1974; Gavel & von Heijne 1990; Apweiler et al. 1999). N-glycans are flexibly attached to the surface of a glycoprotein and extend into the extracellular space as they are relatively large molecules. N-glycans are able to adopt a number of positions around the linkage to the asparagine residue similar to the way a human hand is able to rotate at its wrist, with the carbohydrate branches resembling the fingers (Schwarz & Aebi 2011). The N-glycan adopts a final equilibrium position based on steric and hydrophobic or hydrophilic interactions between other N-glycans and neighbouring amino acid side chains (Wormald & Dwek 1999).

N-linked glycans are highly non-immunogenic and disguise antigenic epitopes on the surface of immunogenic species (Wyatt et al. 1998). An N-linked glycan can either be one of three types: 1) oligomannose, 2) complex, or 3) hybrid, varying in size and complexity (Schwarz & Aebi 2011).

N-glycosylation occurs posttranslationally in the endoplasmic reticulum where initially oligomannose type N-glycans are added (Helenius et al. 2002; Kelleher & Gilmore 2006). The protein is then translocated to the Golgi apparatus where terminal N-glycosylation occurs which trims and modifies the oligomannose residues resulting in complex and/or hybrid types (Hallenberger et al. 1992; Moulard et al. 1999; Helenius et al. 2002). Glycoproteins with oligomannose type N-glycans do not undergo terminal N-glycosylation.

Any NXT/S sequon is considered a PNG as it could potentially, but not necessarily carry an N-linked glycan (Go et al. 2011). The likelihood of a PNG carrying an N-glycan can be

determined by mass spectrometry and enzymatic digestion using specific glycosidases (Go et al. 2013; Go et al. 2011). Petrsecu et al. (2004), Apweiler et al. (1999) and Kelleher and Gilmore (2006) observed PNG sites proximal to disulphide bonds were less likely to be occupied by N-glycans.

N-glycosylation is essential for accurate protein folding, structural maintenance, adhesion to receptors and protein metabolism. Improper N-glycosylation of a protein may result in an incorrectly folded and non-functional protein (Fenouillet & Jones 1995; Fenouillet et al. 1989) as observed when gp120 N-glycosylation was inhibited by producing virus in the absence of a signal peptide and in the presence of tunicamycin, an antibiotic that inhibits the synthesis of N-linked glycoproteins (Fenouillet et al. 1989). Likewise, pseudovirus produced in the presence of tunicamycin had incorrectly folded aggregates (Land et al. 2003) and reduced infectivity (Montefiori et al. 1988).

The HEK293T and CHO cell lines have been successfully used to produce functional gp120 (Bonomelli et al. 2011; Scanlan et al. 2002; Leonard et al. 1990). This suggests that HEK293T and CHO cells have the correct functional machinery to ensure accurate gp120 processing and N-glycosylation. However, differences in N-glycosylation were observed by Go et al. (2013) between cell lines when gp120 produced in CHO cells carried oligomannose type N-glycans at positions N386 and N392 while gp120 produced in HEK293T cells carried a combination of oligomannose and complex N-glycans at these positions. Raska et al. (2010) also demonstrated that the N-gycosylation of gp120 is highly dependent on cell type when they compared gp120 N-glycosyation in HepG2 (hepatocytes), Jurkat (T cells), Dakiki (B cells), HT 1080 (fibrosarcoma cells) HEK293T and CHO cells. It was shown that these differences were dependent on the metabolic state of the cells (Raska et al. 2010).

1.5.1.2 N-linked glycosylation of gp120

Gp120 carries a heterogeneous array of N-linked glycans to protect HIV against proteolytic degradation and to shield antigenic epitopes from antibody recognition (Sagar et al. 2006; Binley et al. 2010; Reitter et al. 1998). On average, gp120 carries 25 PNGs (Zhang et al.

2004). Although N-glycans serve to mask gp120 and thus prevent immune detection, they paradoxically also facilitate immune recognition as it was shown that HIV has an oligomannose patch that may render HIV susceptible to immune cells that express lectins (Bonomelli et al. 2011). Furthermore, the N-glycan composition of gp120 was observed to balance DC-SIGN mediated degradation of HIV with the facilitation of HIV infection (Van`Montfort et al. 2011). Env enriched for oligomannose N-glycans demonstrated enhanced DC-SIGN binding but enhanced endocytic degradation (Van`Montfort et al. 2011).

1.5.2 Gp120 co-receptor binding and viral tropism

For viral attachment and entry into a host cell, CD4 and a second chemokine co-receptor (CXCR4 or CCR5) are required (Deng et al. 1996; Zhang et al. 1999; Sattentau et al. 1986). The particular co-receptor used by HIV determines whether the virus is R5 (utilises CCR5 only), X4 (utilises CXCR4 only) or dual tropic (utilises both CCR5 and CXCCR4). It has been well documented that T/F variants are R5 tropic (Keele et al. 2008; Salazar-Gonzalez et al. 2008).

Although the gp120 V3 loop is highly variable, numerous studies have verified that it carries the co-receptor binding site (Pastore et al. 2006; Polzer et al. 2001; Berger et al. 1999; Cocchi et al. 1996; Kwong et al. 2000; Speck et al. 1997; Delobel et al. 2013; Rizzuto 1998; Valenzuela et al. 1997). A study using two neutralizing antibodies (110-4 and N11-20) against the V3 loop of gp120 demonstrated viral entry inhibition (Valenzuela et al. 1997) and when the V3 loop of an R5 tropic variant was replaced with one from an X4 tropic variant, the former was able to infect cells carrying CXCR4 (Polzer et al. 2001; Speck et al. 1997). Furthermore, the charge on the amino acid side chains (Delobel et al. 2013; Jensen & van ’t Wout 2010) and the N-glycans (Zhang et al. 2004; Li et al. 2005) present on the V3 loop have been shown to determine viral tropism of HIV variants and the conformation of the V3 loop was shown to change upon co-receptor binding (Kwong et al. 1998; Chen et al. 2005; Zanetti et al. 2006). Together these findings corroborate the role of the V3 loop in co-receptor binding.

1.6 Receptors that may play a role in HIV transmission

1.6.1 Dendritic cells and the DC-SIGN lectin

1.6.1.1 Dendritic cells

Dendritic cells comprise of a number of cell types with myeloid and plasmacytoid the major subset. Dendritic cells are antigen presenting cells (APC) that capture antigens in peripheral tissues, process them into peptides and migrate to the lymph nodes where they activate naïve CD4+ T cells upon presentation of the antigen peptide within the context of a major histocompatibility complex (MHC) class 2 molecule (Geijtenbeek, Kwon, et al. 2000; Kwon et al. 2002; Pope et al. 1995; Granelli-Piperno et al. 1999; Holl et al. 2010). HIV and other glycosylated pathogens have evolved a mechanism to subvert the immune function of dendritic cells and exploit the migratory nature of these sentinel cells to encounter naïve CD4+ T cells in the lymph nodes (Chehimi et al. 2003; Geijtenbeek et al. 2009). Glycosylated pathogens are captured by dendritic cells via one of several CLRs [mannose receptor, DC-SIGN, dendritic cell immunoreceptor, dendritic cell-associated lectin-1 and macrophage galactose-type lectin (Geijtenbeek et al. 2004)]. The dendritic cell matures and migrates to the lymph nodes where it transfers the pathogen over to naïve CD4+ T cells for infection, a term referred to as *trans*-infection (Halary et al. 2002; Geijtenbeek, Kwon, et al. 2000).

Two competing models of HIV *trans*-infection exist (Yu et al. 2008). The first suggests that HIV is captured by a CLR on the surface of a dendritic cell triggering the endocytosis of the virus which is packaged into an endosomal multivesicular body (MVB). The MVB is non-endocytic and is released as an exosome which is able to bind to and infect a target CD4+ cell (Izquierdo-Useros et al. 2009). The second model suggests that dendritic cell *trans*-infection occurs via surface-accessible HIV. As soluble CD4 inhibited the transfer of HIV from dendritic cells to CD4+ cells (Cavrois et al. 2007), it was suggested that HIV is not internalised by dendritic cells but is held on the surface until the dendritic cell encounters a CD4+ cell (Cavrois et al. 2007).

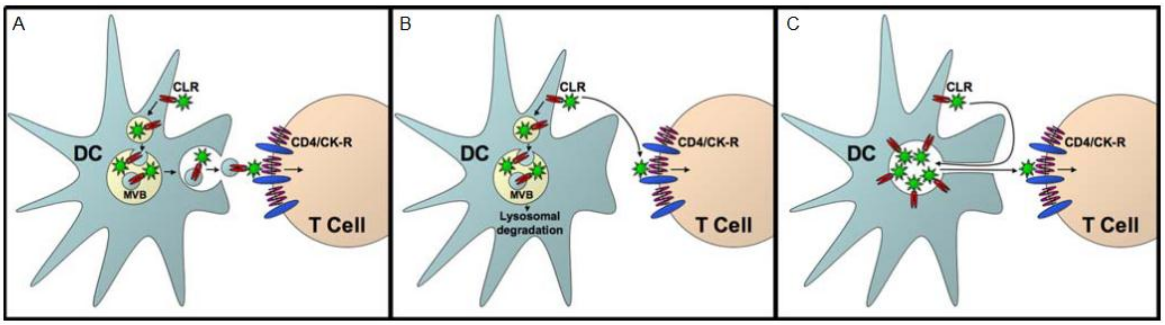


Figure 1.5 Existing models of HIV transfer from dendritic cells to CD4+ cells. Three models exist which may explain how HIV is transferred from dendritic cells to CD4+ cells via a carbohydrate binding lectin receptor (CLR) A) The endosomal multivesicular body (MVB)/exosome model B) The surface-accessible HIV model and C) A combination of the MVB/exosome and surface-accessible HIV model (Yu et al. 2008).

Yu et al. (2008) suggested a combination of the two *trans*-infection models whereby HIV is captured by a CLR on the surface of a dendritic cell and internalised into a non-endocytic pocket-like invagination of the plasma membrane. The virus is not completely internalised and is therefore still susceptible to inhibition by soluble CD4 but remains more protected than cell free virus (Figure 1.5).

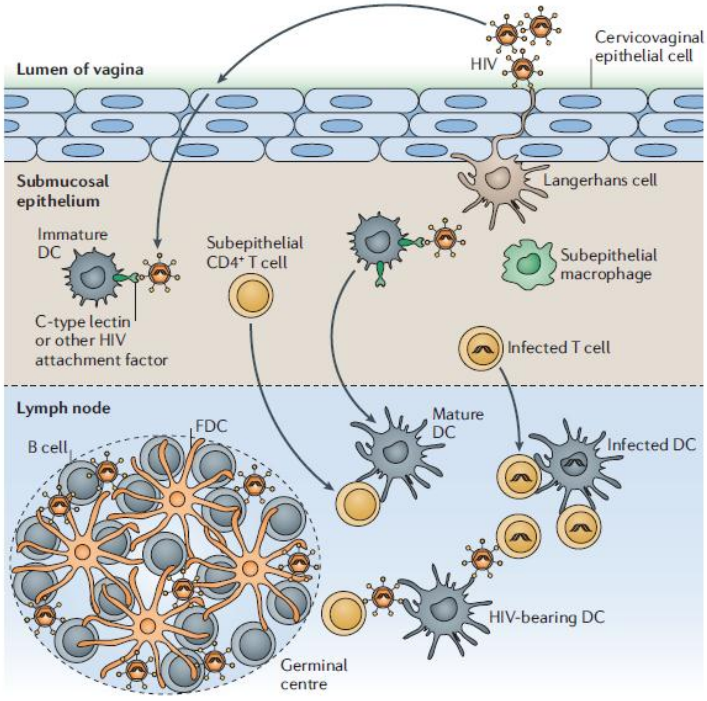


Figure 1.6 Schematic representations of DC-SIGN mediated *cis*- and *trans*-infection of CD4+ cells. HIV is deposited on the cervicovaginal epithelium after coitus. HIV enters the sub-mucosal epithelium where immature dendritic cells bind HIV via a C-type lectin or an HIV attachment factor. Dendritic cells expressing DC-SIGN, CD4 and a co-receptor are infected with HIV (*cis*-infection) while dendritic cells only expressing DC-SIGN mediate *trans*-infection of CD4+ cells. Dendritic cells capture HIV and migrate to the lymph nodes where they mature and encounter HIVs target cells; CD4+ T cells. Infected dendritic cells produce viral progeny that infect target cells or dendritic cells transfer bound HIV over to the CD4+ T cells thereby facilitating HIV dissemination (Wu & KewalRamani 2006).

Dendritic cells can be directly infected by HIV if they express CD4 and a co-receptor. Dendritic cells are infected by HIV in the genital mucosa, migrate to the lymph nodes and produce viral progeny that is released and is able to infect the reservoir of naïve CD4+ T cells (Turville et al. 2003) in a process referred to as *cis*-infection (Lee et al. 2001) (Figure 1.6).

1.6.1.2 DC-SIGN

DC-SIGN is a CLR endogenously expressed on myeloid and plasmacytoid dendritic cells, activated B cells and subsets of macrophages (Geijtenbeek, Torensma, et al. 2000; Rappocciolo et al. 2006). DC-SIGN is a tetrameric type II transmembrane protein (Figure 1.7) and its natural ligand is ICAM-3 expressed on CD4+ T cells (Geijtenbeek, Kwon, et al. 2000; Geijtenbeek et al. 2002; Curtis et al. 1992; Hong et al. 2007; Requena et al. 2008).

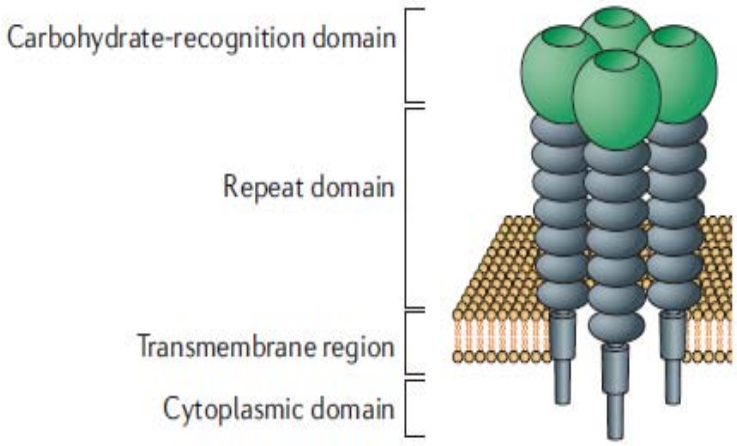


Figure 1.7 Schematic representation of DC-SIGN. DC-SIGN is a C-type lectin receptor expressed on dendritic cells, made up of an N-terminal cytoplasmic domain, a transmembrane domain and an ectodomain. The ectodomain has two components: a repeat region making up the neck and the C-terminal carbohydrate binding domain (CRD) (Wu & KewalRamani 2006).

DC-SIGN has a strong affinity for ICAM-3 (Geijtenbeek, Torensma, et al. 2000) and the highly N-glycosylated HIV gp120 (Geijtenbeek, Kwon, et al. 2000). Anti-DC-SIGN antibodies (AZN-D1 and AZN-D2) inhibited DC-SIGN-ICAM-3 as well as DC-SIGN-gp120 interactions suggesting closely associated (Geijtenbeek, Kwon, et al. 2000) but not overlapping ICAM-3 and gp120 binding sites (Geijtenbeek et al. 2002). DC-SIGN binds HIV gp120 oligomannose residues via the carbohydrate recognition domain (CRD) (Chung et al. 2010; Hong et al. 2002).

1.6.1.3 The role of DC-SIGN in HIV transmission

As there is a paucity of CD4+ T cells in the vaginal mucosa (Pudney et al. 2005) and an abundance of CD4+ T cells in the lymph nodes (Garside et al. 1998) it would be advantageous for HIV to be selectively transferred from the vaginal mucosa to the lymph nodes. It has been suggested that upon sexual transmission, HIV is able to circumvent dendritic cell antigen presentation and can be transferred to the lymph nodes by dendritic cells via DC-SIGN to infect target CD4+ cells (Trumpfheller et al. 2003; Pope et al. 1995; Geijtenbeek, Kwon, et al. 2000; Kwon et al. 2002; Arrighi et al. 2004; Granelli-Piperno et al. 1999).

In vivo experiments have shown that HIV can be safely retained inside dendritic cells for as long as six days however, after 24 hours the variants transferred were shown to be viral progeny and not the original virus (Turville et al. 2004). HIV cell to cell spread is suggested to be the predominant way HIV is disseminated (Chen et al. 2007) and has been shown to be more efficient than cell free replication (Carr et al. 1999). Therefore, it has been suggested that T/F variants carry markers that make them more susceptible to DC-SIGN mediated transfer to target cells.

As DC-SIGN has been shown to have a high affinity for oligomannose type N-glycans (Mitchell et al. 2001), it preferentially binds HIV variants enriched with this type of carbohydrate (Van`Montfort et al. 2011; Feinberg et al. 2001; Kwon et al. 2002; Curtis et al. 1992; Feinberg et al. 2007). HIV has been shown to have an oligomannose N-glycan patch (Bonomelli et al. 2011) with numerous Env N-glycans resistant to oligomannose trimming (Leonard et al. 1990) resulting in an enriched oligomannose repertoire of N-glycans.

Pseudovirus produced in HEK293T cells lacking the GnTI enzyme (HEK293T-/-GnTI) and therefore only carrying oligomannose type N-glycans demonstrated increased DC-SIGN binding to Raji-DC-SIGN cells as well as immature dendritic cells (Eggink et al. 2010; Van`Montfort et al. 2011). Pseudovirus produced in the presence of Kifunensine, an enzyme

that inhibits the trimming of oligomannose type N-glycans, similarly showed an increase in DC-SIGN binding (Eggink et al. 2010; Van`Montfort et al. 2011). Endoglycosidase digestion of oligomannose type N-glycans revealed that DC-SIGN-bound gp120 migrated further on an sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) than gp120 bound to CD4 (Lin et al. 2003), indicating that variants that bound DC-SIGN carried more oligomannose type N-glycans than those that bound to CD4.

Furthermore, the lectins griffithsin, cyanovirin-N and scytovirin that bind oligomannose residues inhibited HIV binding to DC-SIGN and the transfer of HIV to CD4+ cells (Alexandre et al. 2012), supporting previous studies that gp120 oligomannose residues is important for DC-SIGN-mediated transfer of HIV to CD4+ cells. DC-SIGN has been suggested to be the principal CLR involved in capturing, internalising and transferring HIV from dendritic cells to naïve CD4+ T cells as it was demonstrated that approximately 30-50% of viral input was captured by DC-SIGN when immature monocyte-derived dendritic cells were used (Van`Montfort et al. 2007). Together these findings suggest that oligomannose type N-glycans are essential for interactions with DC-SIGN and could comprise the transmission motif. In fact, Go et al. (2011) demonstrated that T/F variants carried more oligomannose type N-glycans than chronic infection controls. As T/F variants were observed to have shorter variable loops with fewer PNGs (Derdeyn et al. 2004; Chohan et al. 2005) and complex N-glycans were shown to occur in the variable loops while oligomannose type carbohydrate structures were located in the constant regions (Go et al. 2011; Leonard et al. 1990), T/F variants may carry an optimal arrangement of oligomannose type N-glycans on the constant regions of gp120 to increase their affinity to DC-SIGN.

T/F variants may be more efficient at engaging DC-SIGN, more efficient at being transferred to CD4+ cells and more efficient at productive infection compared with chronic infection variants due to an optimal arrangement of oligomannose N-glycans. T/F variants were demonstrated to bind to immature monocyte-derived dendritic cells 1.6-fold more efficiently than chronic infection variants (Parrish et al. 2013). This supports the suggestion that the enhanced binding of T/F viruses to DC-SIGN via oligomannose N-glycans could lead to enhanced DC-SIGN mediated *trans*-infection of CD4+ T cells.

1.6.2 The α4β7 integrin and the gut associated lymphoid tissue

1.6.2.1 The α4β7 integrin

The α4β7 integrin, a heterodimer consisting of the non-covalently bound α4 and β7 integrins (Buck & Horwitz 1987), is highly expressed on CD4+ T cells of the genital mucosa (McKinnon et al. 2011; Kelly & Rank 1997) and gut (Cicala et al. 2009). Mucosal addressin cell adhesion molecule-1 (MAdCAM-1) is the natural ligand of α4β7 and is selectively expressed on the high endothelial venules (HEV) of Peyer's patches, mesenteric lymph nodes and the venules of the lamina propria (Berlin et al. 1995). α4β7 mediates the migration of CD4+ T cells from Peyer's patches and mesenteric lymph nodes to the lamina propria in response to the expression of MAdCAM-1 (Bargatze et al. 1995).

1.6.2.2 The role of α4β7 in HIV sexual transmission

The α4β7 binding site that mimics the integrin's natural ligand binding site (Zeller et al. 2001) is found within the V1/V2 loop on the apex of the gp120 trimer spike (Liu et al. 2008). The α4β7 binding site, a tripeptide (Leu-Asp-Val), is conserved across HIV subtypes A, B, C and D (Arthos et al. 2008). It has previously been shown that T/F variants carry gp120 with more compact and less N-glycosylated V1/V2 loops (Derdeyn et al. 2004; Chohan et al. 2005) suggesting that the V1/V2 loop could be important in HIV transmission possibly due to the presence of the α4β7 binding site.

When N-glycans were removed from the V1/V2 loop, α4β7 reactivity increased suggesting that the loss of N-glycans, such as on T/F variants could improve access to this site (Nawaz et al. 2011). This finding as well as the structural mimicry of the binding site to that of its natural ligand, the conservation of the binding site within gp120 and across HIV strains, the observation that viral variants had decreased replication when the α4β7 binding site domain was mutated (Arthos et al. 2008) and the fact that CD4+ T cells with high α4β7 levels are prone to highly productive infection (Cicala et al. 2009) all suggest that α4β7 may be important in HIV infection and transmission.

HIV replicates more efficiently in activated CD4+ T cells than naïve CD4+ T cells (Stevenson et al. 1990) and α4β7 expression was observed on activated cervical CCR5+ CD4+ T cells (McKinnon et al. 2011; Cicala et al. 2009). α4β7, CD4 and CCR5 are found in a complex on the cell membrane seen by confocal microscopy of stained cells from gut biopsies (Cicala et al. 2009). In this complex, it has been suggested that α4β7 is the most prominent receptor of the three receptors. CD4 at 7 nm is three times smaller than α4β7 (22 nm) (Figure 1.7) and the CCR5 binding site on gp120 is effectively hidden prior to the gp120-CD4 interaction (Weissenhorn et al. 1997).

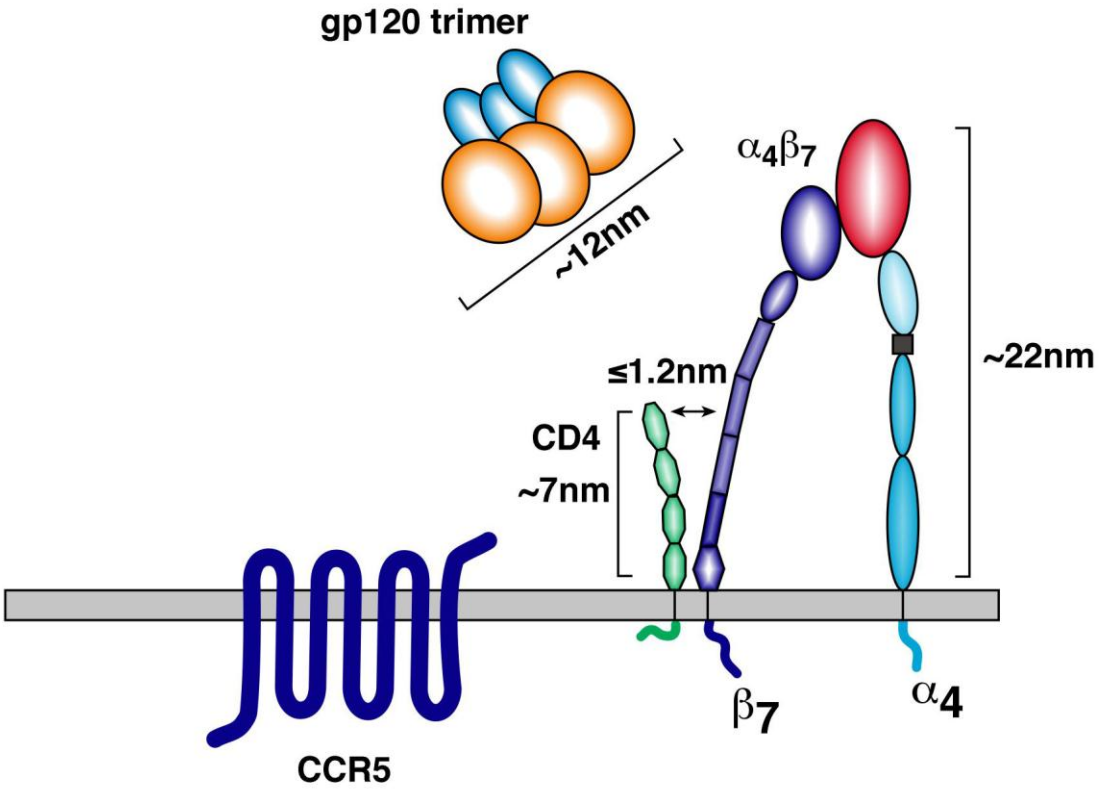


Figure 1.8 A schematic representation of α4β7 expressed alongside the CD4 receptor and CCR5 co-receptor. The α4 and β7 integrins are non-covalently bound to form the α4β7 heterodimer that acts as a homing molecule for CD4+ T cells. α4β7 is closely associated with CD4 (≤1.2 nm) while α4β7 (22 nm) is three times larger than CD4 (7 nm) and CCR5 is effectively hidden prior to CD4-gp120 interaction. α4β7 is suggested to be the prominent receptor seen by the gp120 trimer (12 nm) (Cicala et al. 2009).

As it has been suggested that T/F viruses bind α4β7 better than CD4 (Nawaz et al. 2011), and α4β7 is not only found near the apex of the gp120 trimer spike (Liu et al. 2008) but is also a mere ~1.2 nm from CD4 (Cicala et al. 2009), α4β7 may concentrate virus on the surface of CD4+ T cells thereby facilitating CD4 binding and subsequent infection (Arthos et

al. 2008). Therefore, the R5 tropic nature of T/F variants may be as a result of the co-expression of α4β7 and CCR5 on this subset of susceptible HIV target cells (Cicala et al. 2009).

The gp120-α4β7 interaction was shown to mediate the rapid activation of the lymphocyte function-associated antigen-1 (LFA-1) receptor that facilitates the formation of a virological synapse (Figure 1.9) suggesting that this interaction could facilitate cell to cell spread (Carr et al. 1999; Cicala et al. 2011).

Not all gp120 molecules were observed to bind α4β7 with the same efficiency despite having conserved V2 loops (Nawaz et al. 2011), suggesting that another factor may play a role in mediating this interaction. The type of gp120 N-glycosylation seems to affect α4β7 reactivity as pseudovirus produced in cells that cannot process oligomannose type N-glycans into larger complex N-glycans showed a 100-fold increase in α4β7 reactivity (Nawaz et al. 2011). When the N-glycosylation on the V1/V2 loop was deleted by SDM, α4β7 reactivity increased (Nawaz et al. 2011). Similarly, the removal of N-glycans from chronic infection variants, known to be more heavily glycosylated than T/F variants (Derdeyn et al. 2004) resulted in an increase in α4β7 reactivity (Nawaz et al. 2011). This suggests that early viral isolates with less N-glycosylation, specifically in the V1/V2 loop, may have enhanced α4β7 reactivity as the α4β7 binding site may be more accessible.

Studies using non-human primates corroborated the role of α4β7 in HIV transmission as an anti-α4β7 monoclonal antibody potently inhibited simian immunodeficiency virus (SIV) binding and LFA-1 activation as well as decreased viral dissemination (Ansari et al. 2011). α4β7 CD4+ T cells were preferentially infected in rhesus macaques exposed to SIV (Wang et al. 2009) however recently it was demonstrated that blocking the α4β7 receptor with Act-1, that blocks α4β7-gp120 interactions (Lazarovits et al. 1984), had no effect on viral infectivity negating the previous suggestion that α4β7 was important in HIV transmission and infection (Parrish et al. 2012). This latest finding is contradictory to other studies (Arthos et al. 2008; Cicala et al. 2009; Nawaz et al. 2011) possibly do to the bigger sample size used and the use of IMCs rather than pseudovirus. The use of pseudovirus in previous studies limits α4β7 reactivity to only gp120-α4β7 interactions and not the entire viral structure.

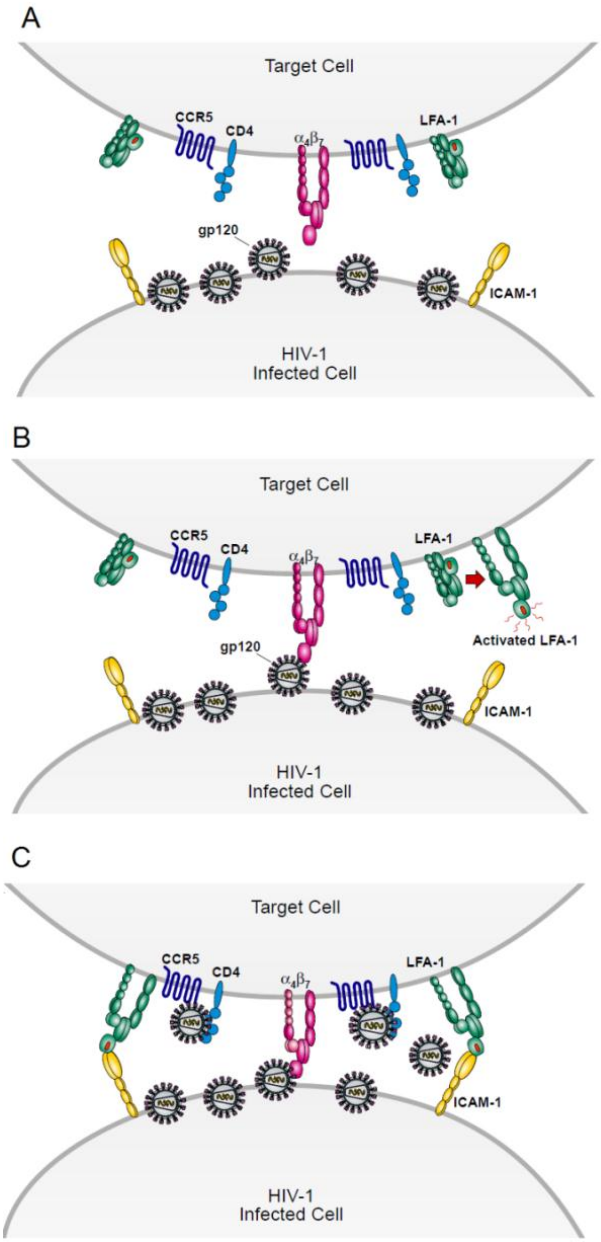


Figure 1.9 A schematic representation of the formation of a virological synapse A) An uninfected target cell expresses CD4, CCR5, α4β7 and an inactivated form of the lymphocyte function-associated antigen-1 (LFA-1) receptor. A nearby cell is infected by HIV and expresses ICAM-1, the natural ligand for LFA-1 B) The gp120 of a virus on the surface of the infected cell interacts with α4β7 on the surface of an uninfected cell which results in the activation of the LFA-1 receptor C) The activated form of the LFA-1 receptor interacts with ICAM-1 on the surface of the infected cell to form a virological synapse which allows gp120 to interact with CD4 and CCR5 to initiate the viral attachment and entry into the uninfected target cell (Cicala et al. 2011).

1.6.2.3 The role of α4β7 in CD4+ T cell depletion in the gut during early infection

It has been established that the gut is preferentially and profoundly affected early on during HIV infection with the near-complete depletion of CD4+ T cells in the gut-associated

lymphoid tissue (GALT) shown to be important in establishing acute infection (Guadalupe et al. 2003; Mehandru et al. 2007; Ansari et al. 2011). The gastro-intestinal tract forms the host-pathogen interface and so has an extensive immune system made up of the GALT which is extremely important in HIV dissemination (Figure 1.10).

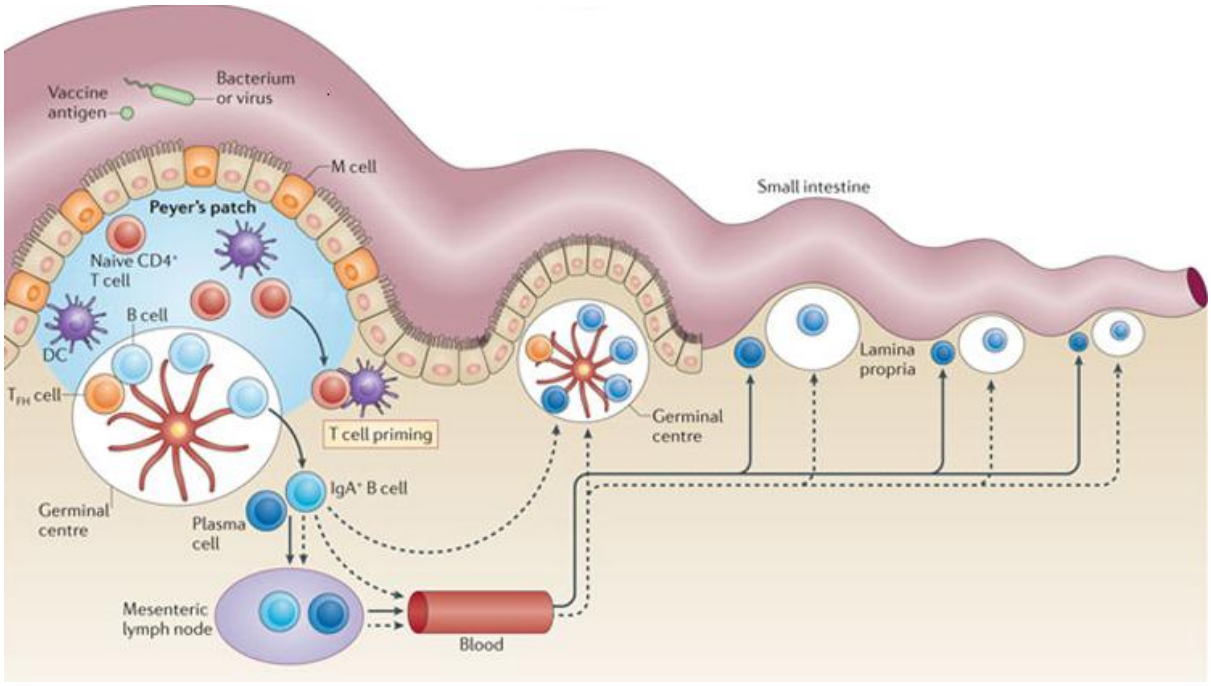


Figure 1.10 A schematic representation of the gastrointestinal associated lymphoid tissue (GALT). The GALT is the largest mass of lymphoid tissue and is found in the gastrointestinal tract. Peyer's patches, mesenteric lymph nodes and the lamina propria are tissues that are highly relevant to HIV pathogenesis (Lycke 2012).

Infected CD4+ T cells migrate from the genital mucosa, the site of infection, to Peyer's patches and the mesenteric lymph nodes within days of sexual transmission (Mehandru et al. 2007). High levels of HIV replication are observed in the lamina propria which proves to be detrimental to the immune system and signals the onset of acute HIV infection and peak viremia (Mehandru et al. 2007). The α4β7 integrin has been implicated in the translocation of HIV from the site of infection, the genital mucosa to the gut to initiate the depletion of CD4+ T cells (Ansari et al. 2011).

α4β7 has been shown to be the homing integrin that recruits CD4+ T cells from the Peyer's patches and mesenteric lymph nodes to the lamina propria upon MAdCAM-1 expression (Berlin et al. 1995; Bargatze et al. 1995). Antibodies to the α4 sub-unit of α4β7 inhibit binding to Peyer's patches (Holzmann & Weissman 1989) and blocking α4β7 with a non

CD4+ T cell depleting antibody, during acute infection of rhesus macaques with SIV, resulted in reduced plasma and GALT viral replication (Ansari et al. 2011) demonstrating the importance of viral replication in the gut to establish acute infection. The α4β7 integrin may therefore be the link between sexual transmission across the genital mucosa and the translocation of infected CD4+ T cells to the gut to initiate near complete depletion of GALT CD4+ T cells.

1.7 Conclusion

Understanding the mechanism behind the genetic bottleneck at HIV transmission will be invaluable to the design of novel anti-HIV drugs and vaccines. Transmission of only a single variant suggests that some viruses may carry a phenotypic marker or transmission motif, allowing the variant to breach the mucosal barrier, escape immune detection and result in clinical infection. A therapy designed to target the transmission motif(s) could effectively prevent the spread of HIV. Env has been extensively studied as the viral protein most likely to carry the transmission motif(s) as it engages in the CD4 receptor and co-receptor essential for viral attachment and entry into the host cell.

Recently, it has been suggested that T/F variants carry gp120 with high affinity for the α4β7 integrin (Nawaz et al. 2011) and are enriched with oligomannose type N-glycans (Go et al. 2011). T/F variants enriched with oligomannose type N-glycans supports the argument that DC-SIGN, a C-type lectin expressed on dendritic cells, may be important in HIV transmission as DC-SIGN has been shown to preferentially bind oligomannose type N-glycans. Therefore it is possible that the N-glycosylation of T/F variants may allow for optimal binding to DC-SIGN, enhanced migration to the lymph nodes and thus selective transmission to CD4+ T cells. Env N-glycosylation has also been shown to affect the interaction of HIV with α4β7. Therefore, T/F variants with high affinity for the α4β7 integrin may also be selectively transmitted, propagated and disseminated as the homing integrin facilitates HIV infection of CD4+ T cells as well as aid the migration of HIV from the genital mucosa to the gut where viral replication can occur unchecked.

Within the context that T/F variants carry more compact gp120 with fewer PNG sites, the dependence of gp120-α4β7 and gp120-DC-SIGN interactions on the presence of PNGs could suggest that Env N-glycosylation may play an important role in HIV transmission. Therefore elucidating the role of Env N-glycosylation in HIV transmission might be instrumental in designing a broadly acting vaccine that could protect against HIV infection.

Study rationale and research objectives

The UNAIDS 2012 Global Report estimated that two thirds of the global population infected by HIV live in sub-Saharan Africa (Joint United Nations Programme on HIV/AIDS 2012) where HIV subtype C is the dominant subtype (Van Harmelen et al. 1999). Previously, studies have focused on subtypes A and B (Wilen et al. 2011; Parrish et al. 2013) and the data from these studies most likely cannot be extrapolated to subtype C due to inter-subtype differences. The use of unmatched T/F variants and chronic infection controls (Wilen et al. 2011) rather than longitudinal sampling, the comparison of non SGA-derived sequences (Derdeyn et al. 2004) rather than SGA-derived sequences (Keele et al. 2008; Salazar-Gonzalez et al. 2009) and the use of pseudovirus (Wilen et al. 2011; Borggren et al. 2008) and not IMC (Parrish et al. 2012) may have contributed to the conflicting results thus far, making it difficult to identify transmission signatures.

This study focuses on five CAPRISA 002 study participants with matched T/F *env* clones and chronic infection controls generated by SGA. This research seeks to first compare the α4β7 and DC-SIGN reactivity of Env clones and then determine the importance of gp120 N-glycosylation for these interactions. We hypothesise that, T/F variants carry gp120 that bind with higher efficiency to α4β7 and DC-SIGN compared to those found later in infection and that this affinity is dependent on the presence of N-glycans.

Study objectives:

1) Develop an α4β7 reactivity assay using transient transfections of mammalian cells to determine whether T/F *env* clones have a higher affinity for the α4β7 integrin than chronic infection controls and whether this is dependent on gp120 N-glycans.
2) Determine whether T/F variants are preferentially *trans*-infected to CD4+ cells via the DC-SIGN CLR compared with those from chronic infection.
3) Determine whether the removal of PNGs, believed to be oligomannose type N-glycans, affect DC-SIGN mediated *trans*-infection of CD4+ cells.

Chapter 2: Methods

2.1 Envelope clones

HIV *env* clones were generated from five CAPRISA 002 study participants (CAP177, CAP239, CAP45, CAP206 and CAP210) at acute (2 to 5 weeks post-infection) and chronic (1.5 to 3 years post-infection) stages of infection.

Table 2.1 Characteristics of the *env* clones

Participant ID (CAP)	Clone ID	Stage of infection	Sampling (Weeks post infection)	Fiebig stage	The vector in which the *env* clone was received	Received from
177	A3	acute	2	I/II	pcDNA™3.1/V5-His TOPO®	NICD
	47	chronic	172	N/A	pcDNA™3.1/V5-His TOPO®	NICD
239	EII	acute	5	I/II	pTARGET™	L. Shuping
	T35	chronic	173	N/A	pcDNA™3.1/V5-His TOPO®	NICD
45	B5	acute	2	I/II	pTARGET™	L. Shuping
	H5	chronic	108	N/A	pTARGET™	L. Shuping
206	H1	acute	4	V	pTARGET™	L. Shuping
	E12-3	chronic	173	N/A	pcDNA™3.1/V5-His TOPO®	NICD
210	E8	acute	5	I/II	pTARGET™	G. Bandawe
	E1	chronic	80	N/A	pTARGET™	L. Shuping
	C7	chronic	80	N/A	pTARGET™	L. Shuping

*Adapted from L. Shuping (unpublished data)

Each study participant had one T/F and one chronic infection *env* clone besides CAP210 which had two chronic infection *env* clones. The *env* clones were either received from P. Moore [National Institute for Communicable Diseases (NICD), South Africa] or cloned by G. Bandawe [Institute of Infectious Disease and Molecular Medicine (IIDMM), University of Cape Town (UCT), South Africa] or L. Shuping (Department of Molecular and Cell Biology, UCT, South Africa).

The entire *env* gene including 5' and 3' flanking sequences were cloned into either the pcDNA™3.1/V5-His TOPO® [Invitrogen, United States of America (USA)] or pTARGET™ (Promega, USA) vector (Table 2.1). The *env* clones were derived from SGA and extensive sequence analysis confirmed that these participants were infected with a single variant

(Abrahams et al. 2009). As the *env* clone selected during acute infection was identical to the consensus sequence, it was assumed to represent the T/F variant. The chronic infection *env* control clones were chosen as they carried the same PNGs found on the consensus sequence.

2.2 Sub-cloning and/or preparation of α4 and β7 cDNA clones

The pCDM8_α4 (Kamata et al. 1995) and pCEP4_β7 (Erle et al. 1991) plasmid DNA was a generous gift from Prof D. Erle, University of California, San Francisco, California.

2.2.1 Sub-cloning of α4 cDNA

The pCDM8 (Invitrogen, USA) mammalian expression vector that the α4 cDNA was originally cloned into utilises the tyrosine transfer RNA (tRNA) suppressor and p3 plasmid system. In order to select pCDM8-containing cells *E. coli* K12 MC1061 [p3] competent cells which contain the p3 plasmid are required. The p3 plasmid contains antibiotic resistance genes interrupted by several non-sense mutations. A tRNA suppressor produced upon expression of the pCDM8 vector suppresses the non-sense mutations thereby initiating expression of the antibiotic resistance genes and allowing antibiotic selection of pCDM8 containing cells. As *E. coli* K12 MC1061 [p3] competent cells were not available and new customs laws restricted the import of certain competent cells into the country at that time, we were unable to make use of this system.

2.2.1.1 PCR

Consequently, we sub-cloned the human α4 cDNA into the pTARGET^{TM} expression system (Promega, USA) using TA cloning. α4 was amplified by PCR with Phusion Hot Start II High Fidelity DNA Polymerase (Thermo Scientific, USA) (Parameters: 95 °C 2 minutes, 20 cycles of 94 °C 20 seconds, 55 °C 20 seconds, 72 °C 8 minutes and 72 °C 5 minutes). The T7 forward (F) primer and the pCDM8 reverse (R) primer (Table 2.2) were used to amplify the 3 096 bp

human α4 cDNA (Accession number NP_034706) as these primer sequences were present in the pCDM8 vector. Each PCR reaction contained 200 µM of each of the dNTP's, 1X Phusion Hot Start II High Fidelity DNA Polymerase buffer, 0.5 µM each oligonucleotide, 0.05 µg template plasmid DNA and 1 U Phusion Hot Start II High Fidelity DNA Polymerase in a total volume of 50 µL. PCR products were visualised by 1% agarose gel electrophoresis (Sambrook & Russell 2001) using a BioDoc-It™ Imaging System (Ultra-Violet Products Ltd., Cambridge, United Kingdom) and the relevant bands were excised and gel purified using the Wizard® SV Gel and PCR clean up system (Promega, USA) as per manufacturer's instructions. PCR products were A-tailed using Supertherm Taq DNA Polymerase (JMR Holdings, USA) (Parameters: 72 °C 10 minutes) and purified using the Wizard® SV Gel and PCR clean up system (Promega, USA) as per manufacturer's instructions. Each A-tailed PCR reaction contained 160 µM dATPs, 1X Supertherm Taq DNA Polymerase buffer, 2 mM $MgCl_2$, 40 µL gel-extracted PCR product and 0.375 U Supertherm Taq DNA Polymerase.

2.2.1.2 Ligation and screening

DNA ligations of the A-tailed PCR products and the pTARGET™ vector were performed using T4 DNA ligase (Fermentas, USA) in a 10 µL reaction at a molar ratio of 4:1 insert: vector. The ligation was done at 4 °C for 24 hours and competent *E. coli* JM109 cells were transformed as per manufacturer's instructions (Promega, USA) and cultured at 30 °C overnight. Transformed cells were isolated on Luria agar plates (1% tryptone, 0.5% yeast extract, 1% NaCl, 1.5% agar) with 100 µg/ml ampicillin.

Colony PCR was used to screen colonies for plasmids containing the insert in the correct orientation using Supertherm Taq DNA Polymerase (JMR Holdings, USA) (Parameters: 95 °C 2 minutes, 20 cycles of 94 °C 20 seconds, 55 °C 20 seconds, 72 °C 5 minutes and 72 °C 5 minutes). Each PCR reaction contained 200 µM each dNTP, 1X Supertherm Taq DNA Polymerase buffer, 2 mM $MgCl_2$ and 0.375 U Supertherm Taq DNA Polymerase. The T7 F primer, found within the pTARGET™ vector sequence and the 12-1691 α4 R2 primer, designed from nucleotide (nt) 2 383 to nt 2 406 and synthesised by Pei-Yin Leibrich (Department of Molecular and Cell Biology, UCT, South Africa) (Table 2.2), were used to

amplify a ~2 400 bp DNA fragment. Small scale plasmid DNA isolation was performed on positive colonies using the Bioflux Miniprep system (Bioer Technology Co., Ltd) according to manufacturer's instructions.

2.2.1.3 DNA sequencing

The α4 F1 primer designed from nt 646 to nt 669, the α4 R1 primer designed from nt 2 383 to nt 2 406 and the α4 R2 primer designed from nt 3 082 to nt 3 097 were used together with the T7 F and pCDM8 R primers to sequence the entire α4 cDNA (Table 2.2). The Big Dye® Terminator v3.1 Cycle Sequencing kit (Applied Biosystems, USA) and Half-dye (Bioline, United Kingdom) was used in the sequencing reactions performed at the Central Analytical facility of the University of Stellenbosch, South Africa.

Table 2.2 The primers that were used to sub-clone the α4 cDNA into the pTARGET™ vector, perform colony PCR on transformed colonies and to sequence positive pTARGET™_α4 clones

*Primer	*Direction	Sequence	Size
T7 F	F	5'-TAA TAC GAC TCA CTA TAG GG-3'	20
12-1962 α4 F1	F	5'-CCA GGA TCA TCT TAC TGG ACT GGC-3'	24
12-0696 pCDM8 R	R	5'-CCT CTA GAG TCG CGG CCG CGA CCT GCA G-3'	28
12-1691 α4 R1	R	5'-CAC CAT GCA CGT TTC AGG CTC ATT-3'	24
12-0786 α4 R2	R	5'-TTA ATC ATC ATT GCT TTT-3'	16

*F=forward, R=reverse

2.2.2 Preparation of pCEP4_β7

E. coli JM109 bacterial cells were transformed with pCEP4_β7 plasmid DNA before being cultured at 30 °C overnight. Transformed cells were isolated on Luria agar plates with 100 µg/ml ampicillin and small scale plasmid DNA isolation was performed on colonies using the Bioflux Miniprep system (Bioer Technology Co., Ltd) according to manufacturer's instructions. The pCEP F primer and a set of three designed primers were used to sequence the 2 767 bp β7 cDNA (Accession number NP_000880). The β7 F1 primer designed from nt 557 to nt 578, the β7 F2 primer designed from nt 2 236 to nt 2 258, the β7 F3 primer

designed from nt 2 598 to nt 2 620, the β7 R1 primer designed from nt 2 379 to nt 2 397 and the β7 R2 primer designed from nt 24 to nt 45 were used (Table 2.3).

Table 2.3 The primers that were used to sequence pCEP4_β7 clones

*Primer	*Direction	Sequence	Size
12-0677 pCEP F	F	5'-AGA GCT CGT TTA GTG AAC CG-3'	20
12-2941 β7 F1	F	5'-CCC ATT CTG TGC GCA TTG GTT T-3'	22
12-2942 β7 F2	F	5'-TAC CGG CTC TCG GTG GAA ATC TA-3'	23
13-1567 β7 F3	F	5'-CCA CCC AAG TAT ACA ATA AAG TC-3'	23
12-0893 β7 R1	R	5'-TCA GAG AGT GGG ACT GTC T-3'	19
13-1566 β7 R2	R	5'-GCT CAG GAC CAG CAG CAA AAC A-3'	22

*F=forward, R=reverse

High quality plasmid DNA was prepared using the PureYield™ Plasmid Midiprep system (Promega, USA) for transfection and the concentration of plasmid DNA was determined spectrophotometrically using a Nanodrop 200 (Thermo Scientific, USA).

2.2.3 Restriction enzyme digest to confirm pTARGET^{TM}_α4 and pCEP4_β7

pTARGET^{TM}_α4 plasmid DNA was screened using *Xho* I and *Sal* I (Fermentas, USA) restriction enzymes and pCEP4_β7 plasmid DNA was screened using *Kpn* I and *BamH* I (Fermentas, USA) restriction enzymes. Restriction enzyme digests were performed using 1X restriction enzyme buffer, 500 ng plasmid DNA and 1 U restriction enzyme made up with nuclease free water to a total volume of 20 µL. The restriction enzyme reactions were incubated in a 37 °C water bath for 1 hour. The relevant insert and vector bands were visualised by 1% agarose gel electrophoresis.

2.3 Protein expression

2.3.1 Cell lines

Several cell lines were used in this study, as well as cells isolated from human blood. HEK293T, CHO, CHO NL4-3 gp160, HeLa, TZM-bl, Raji-DC-SIGN and Raji-WT cell lines were

obtained from the National Institutes of Health (NIH) acquired immune deficiency syndrome (AIDS) Research and Reference Reagent Programme (ARP), USA.

HEK293T is a human embryonic kidney derived cell line that is easily transfectable and commonly used to generate pseudovirus as HEK293T cells have the correct N-glycosylation machinery to accurately N-glycosylate HIV Env (Go et al. 2011; Go et al. 2013). The Chinese hamster ovary (CHO) cell line is frequently used in the expression of Env (Go et al. 2013; Crispin et al. 2006; Kamata et al. 1995; Scanlan et al. 2002). CHO NL4-3 gp160 is a positive control CHO cell line that stably expresses HIV gp160.

HeLa is a cervical cancer derived cell line that endogenously expresses the α4 and β7 integrins. Although no reference was available, HeLa whole cell lysates (Santa Cruz Biotechnology Inc., USA) are recommended as a positive control for Western blotting using antibodies specific for these integrins (*http://datasheets.scbt.com/sc-365209.pdf*, *http://datasheets.scbt.com/sc-166031.pdf*). The TZM-bl cell line is a HeLa derived cell line that is stably transfected with the CD4, CCR5 and CXCR4 receptors that are required for HIV attachment and entry. TZM-bl is an HIV reporter cell line as it carries a luciferase gene downstream of the viral LTR which is inducible by the HIV transactivator, Tat.

HEK293T, CHO, CHO NL4-3 gp160, HeLa, and TZM-bl are adherent cell lines that were maintained in Dulbecco's modified Eagle medium (DMEM) supplemented with 10% foetal calf serum (FCS) (Sigma Aldrich, USA), 100 U/mL penicillin and 100 U/mL streptomycin (Sigma Aldrich, USA). Cells were routinely passaged every 3 to 4 days by washing them with 1X phosphate-buffered saline (PBS), lifting them with 0.25% Trypsin/0.1% EDTA (Lonza, South Africa) and transferring them to fresh 75 cm^2 tissue culture flasks (T75) (Nest Biotechnology Co., Ltd).

Raji is a Burkitt's lymphoma derived B cell line. Raji-DC-SIGN cells are Raji-WT cells stably transfected to express the human DC-SIGN receptor (Wu et al. 2004). Raji-DC-SIGN and Raji-WT are non-adherent cell lines that were maintained in Roswell Park Memorial Institute medium-1640 (RPMI) (Thermo Scientific, USA) supplemented with 10% FCS (Sigma Aldrich,

USA). Cells were routinely passaged every 5 to 6 days by replacing 10 mL cell culture with 10 mL fresh RPMI with 10% FCS.

All cell lines were routinely tested for Mycoplasma contamination as described previously (Lu et al. 1997). Slides were viewed on a Zeiss Axiovert 200 inverted fluorescent microscope (Carl Zeiss, Germany) with the assistance of S. Cooper (Department of Human Biology, UCT, South Africa).

CD4+ T cells were isolated from human blood received from the Western Province Blood Transfusion Service (Pinelands, South Africa). Peripheral blood mononuclear cells (PBMC) were isolated using Ficoll and monocytes were removed using the adherence methodology. The remaining cells were induced with interleukin (IL) -2 (NIH ARP, USA) every 2 days for 6 days in RPMI with 10% FCS and thereafter used for experiments. All cells were maintained at 37 °C in 5% CO_2.

2.3.2 Comparison of PEI, PolyFect® and Electroporation for transfection of plasmid DNA into mammalian tissue culture cells

Polyethylenimine (PEI) (Sigma Aldrich, USA), PolyFect® (QIAGEN, USA) and Electroporation (BIO-RAD, USA) were compared using two HIV *env* clones (CAP177_A3 and CAP206_E12-3) that were previously shown to have good expression using Western blot (Z. Woodman, unpublished data). A total of 7 µg plasmid DNA was introduced into $4X10^5$ HEK293T cells in a 6 well tissue culture dish with 22.5 µL PEI (1 mg/mL) and 21 µL PolyFect® (2 mg/mL) reagents, respectively in a final volume of 2 mL. The culture medium was changed 6 hours after transfection to remove the PEI-DNA complexes.

The Electroporation of $4X10^5$ HEK293T cells with 7 µg plasmid DNA (250 V for 25 ms, capacitance 2 000, resistance 1 500) in a total volume of 100 µL was performed at the IIDMM, UCT, South Africa. The Gene Pulser MXcell™ Electroporation System (BIO-RAD, USA) and a 96 well Electroporation plate were used. After Electroporation, the cells were diluted in 2 mL culture medium and plated into a 6 well tissue culture dish.

Cells transfected using PEI, PolyFect® and Electroporation were grown at 37 °C in 5% CO_2 for 48 hours. Transfected cells were centrifuged at 3 000 times gravitational (g) for 5 minutes at 4 °C before lysed on ice with radioimmunoprecipitation assay (RIPA) buffer (50 mM Tris, 150 mM NaCl, 0.1% SDS, 0.5% sodium deoxycholate, 1% Triton X 100) containing 0.1 g/mL phenylmethylsulfonyl fluoride (PMSF) protease inhibitor (Thermo Scientific, USA). Cell lysates were stored at -20 °C for up to a week until required.

2.3.3 Transfection of HEK293T cells with pTARGET™_α4 and pCEP4_β7 using PEI

pTARGET™_α4 (8 766 bp) and pCEP4_β7 (12 767 bp) plasmid DNA was transfected separately (Table 2.4 and Table 2.5) and co-transfected (Table 2.6) into $4X10^5$ HEK293T cells in a 6 well tissue culture dish.

A ratio of 1 µg plasmid DNA to 3.5 µL 1 mg/mL PEI was used when increasing concentrations of plasmid DNA were transfected. In order to identify the optimal concentration of plasmid DNA for integrin expression, the amount of pTARGET™_α4 plasmid DNA was increased from 0.4 to 2 pmol while pCEP4_β7 plasmid DNA was increased from 0.2 to 2 pmol. The molar ratio of pTARGET™_α4 to pCEP4_β7 plasmid DNA in co-transfection experiments was 1:1. The total plasmid DNA concentration was not kept constant with the addition of pcDNA™3.1/V5-His TOPO® empty vector.

The culture medium was changed 6 hours after transfection to remove the PEI-DNA complexes. Cells were grown at 37 °C in 5% CO_2 for 48 hours in culture medium before the cells were either lysed with RIPA buffer containing PMSF for Western blot analysis or lifted with 1% EDTA in 1X PBS (Ca/Mg++ free) for flow cytometry analysis. A Bradford protein assay (BIO-RAD, USA) was used to determine the protein concentration of each sample and 60 µg protein was loaded per well for SDS-PAGE and Western blot analysis.

Table 2.4 The relative molar amounts of pTARGET™_α4 plasmid DNA that were used in transfections of HEK293T cells

pTARGET™_α4 (pmol)	pTARGET™_α4 (μg)	pCEP4_β7 (μg)
0.4	2.1	0
0.8	4.2	0
1.2	6.4	0
1.6	8.6	0
2	10.8	0

Table 2.5 The relative molar amounts of pCEP4_β7 plasmid DNA that were used in transfections of HEK293T cells

pCEP4_β7 (pmol)	pTARGET™_α4 (μg)	pCEP4_β7 (μg)
0.2	0	1.4
0.4	0	2.96
0.8	0	5.92
1.2	0	8.9
1.6	0	11.9
2	0	14.9

Table 2.6 The relative molar amounts of pTARGET™_α4 and pCEP4_β7 plasmid DNA that were used in co-transfections of HEK293T cells

pTARGET™_α4 and pCEP4_β7 (pmol)	pTARGET™_α4 (μg)	pCEP4_β7 (μg)
0.1	0.53	0.74
0.4	2.1	2.96
0.8	4.2	5.92
1	5.3	7.4

2.3.4 Transfection of CHO cells with pTARGET™_α4 and pCEP4_β7 using PEI and Electroporation

CHO cells were co-transfected with pTARGET™_α4 and pCEP4_β7 plasmid DNA as well as transfected individually with each integrin's cDNA to test whether the cell type used for transfection may impact protein expression. The same PEI and Electroporation methodologies of transfection were used as before and total plasmid DNA was kept constant to control for the effect of high concentrations of plasmid DNA on the transfection efficiency. The α4β7 recombinant protein (5397-A3, R and D systems, England) was used as a positive control and pcDNA™3.1/V5-His TOPO® empty vector was used as a negative control.

Two HIV Env clones, CAP177_A3 (T/F) and CAP206_E12-3 (chronic) were electroporated into HEK293T and CHO cells to confirm the Electroporation methodology. The gp120 recombinant protein (NIH ARP, USA) and the positive control cell line CHO NL4-3 gp160 (NIH ARP, USA) cell lysates (for the CHO transfections only) were used as positive controls. The pcDNA^{TM}3.1/V5-His TOPO$^{®}$ empty vector was used as a negative control.

2.3.5 Using the pGL4 Luc reporter gene to determine transfection efficiency

The luciferase reporter gene construct, pGL4 Luc (0.25 µg) and pTARGET^{TM}_α4 or pCEP4_β7 plasmid DNA (8 µg) were introduced into 4X10^{5} HEK293T cells in a 6 well tissue culture dish using PEI as described. The culture medium was changed 6 hours after transfection to remove the PEI-DNA complexes and cells were grown at 37 °C in 5% CO_2. pGL4 Luc (0.25 µg) was also co-transfected with increasing amounts (0.5 to 1.3 pmol) of pTARGET^{TM}_α4 and pCEP4_β7 plasmid DNA added in a 1:1 molar ratio. Two HIV *env* clones, CAP177_A3 (T/F) and CAP206_E12-3 (chronic) (5 µg) that were previously shown to have high expression levels by Western blot were used as positive controls. Untransfected HEK293T cells and cells transfected with pcDNA^{TM}3.1/V5-His TOPO$^{®}$ empty vector were used as negative controls. Luciferase was read after 48 hours using 50 µL Brite-Glo lysis buffer with the Modulus™ Microplate Luminometer (Promega, USA).

2.4 Protein detection

2.4.1 PAGE and Western blot

2.4.1.1 SDS-PAGE

Cell lysates with equivalent total protein concentrations were mixed with 5X sample buffer (5% SDS, 20% glycerol, 0.1% bromophenol blue, 5% β-mercaptoethanol, pH 6.8) and boiled for 5 minutes. Samples were separated by SDS-PAGE using the mini PROTEAN^{TM} III system (BIO-RAD, USA) with 10% resolving (0.1% SDS, 375 mM Tris, pH 8.8) and 5% stacking (0.1%

SDS, 125 mM Tris, pH 6.8) gels in a Tris-glycine-SDS running buffer (25 mM Tris, 250 mM glycine, 0.1% SDS, pH 8.3) at 25 mA/gel. Gels were used for Western blotting as described below.

2.4.1.2 Native PAGE

Samples were mixed with sample buffer lacking SDS and β-mercaptoethanol (20% glycerol, 125 mM Tris, 0.1% bromophenol blue, pH 6.8). Non-denaturing PAGE was performed in the same way as an SDS-PAGE, except SDS was excluded from the resolving gel, stacking gel and Tris-glycine-SDS running buffer. Gels were run at 25 mA/gel.

2.4.1.3 Western blot

After SDS-PAGE, protein was transferred to a polyvinylidene difluoride (PVDF) membrane (BIO-RAD, USA) in transfer buffer (20 mM Tris, 150 mM glycine, 20% methanol) at 10 V for 1 hour using a semi-dry protein transfer system (Cleaver Scientific, USA). Gels were stained with Coomassie Blue [50% methanol, 10% acetic acid, 0.05% Brilliant Blue R-250, 40% distilled water (dH_20)] and destained (70% dH_20, 10% acetic acid, 20% methanol) to ensure complete transfer of the protein. After transfer, the membranes were incubated in blocking buffer overnight at 4 °C or shaking for 2 hours at room temperature (RT).

Table 2.7 Primary and secondary antibodies used for the detection of the α4 and β7 integrins

Company	*P/S	Antibody
Sigma Aldrich, USA	Primary	Mouse anti-human integrin α4 WH0003676M1
Sigma Aldrich, USA	Primary	Mouse anti-human integrin β7 WH0003695M1
Santa Cruz Biotechnology Inc., USA	Primary	Mouse anti-human integrin α4 sc-365209
Santa Cruz Biotechnology Inc., USA	Primary	Mouse anti-human integrin β7 sc-166031
NIH ARP	Primary	Mouse anti human integrin α4β7 Act-1 Cat#11718
Santa Cruz Biotechnology Inc., USA	Primary	Mouse anti-human β-actin sc-47778
Santa Cruz Biotechnology Inc., USA	Primary	Goat anti-mouse IgG-HRP conjugated sc-2005

For protein detection, membranes were incubated for 1 hour with shaking in blocking buffer [Tris-buffered saline (TBS), 0.1% tween-20, 4% skim-milk] containing 1:100 dilution of the respective primary antibody (Table 2.7). The monoclonal antibody Act-1, raised against the

native human α4β7 integrin, was used to detect protein under non-reducing conditions. The membrane was washed four times for 15 minutes using blocking buffer before incubation with 1:2 000 dilution of the goat anti-mouse immunoglobulin G (IgG)-horseradish peroxidase (HRP) conjugated secondary antibody for 1 hour at RT with shaking and then washed as before with 1X TBS-T (TBS, 0.1% tween-20). The chemiluminescence system together with autoradiography was used for signal detection as per manufacturer's instructions (KPL, USA). A recombinant human integrin α4β7 (R and D systems, England) (Figure 2.7) was used as a positive control at concentrations of up to 100 ng. A 1:2 000 dilution of the mouse anti-human β-actin monoclonal antibody was used to ensure equivalent amounts of total protein were loaded.

2.4.2 Flow cytometry

HEK293T cells co-transfected with 1 pmol pTARGET^{TM}_α4 and pCEP4_β7 plasmid DNA using the PEI transfection methodology were lifted with 1% EDTA in 1X PBS. Cells were centrifuged at 2 000 g for 5 minutes at RT, resuspended in 200 µL cold fluorescence-activated cell sorting (FACS) buffer (1X PBS containing 2% FCS) and stained with 0.4 µg/mL Act-1 monoclonal antibody on ice for 1 hour. The cells were washed by centrifugation three times at 2 000 g for 5 minutes at 4 °C with cold FACS buffer. Cells were then incubated with 7.7 µg/mL Cy3$^{®}$ goat anti-mouse IgG in cold FACS buffer on ice in the dark for 1 hour. The cells were washed three times, resuspended in 200 µL FACS buffer and data was acquired by E. Smit from the South African Tuberculosis Vaccine Initiative laboratory (IIDMM, UCT, South Africa) on a FACS Fortessa^{TM} flow cytometer (BD Biosciences, USA). Control samples of unstained cells and secondary antibody-only stained cells were included. HeLa and TZM-bl cells that endogenously express the α4 and β7 integrins were used as positive controls. Data was analysed by D. Bowers (IIDMM, UCT, South Africa) using FloJo software (v 8.8.2, Tree Star, Stanford, CA).

2.5 Binding assay

2.5.1 Pseudovirus production

Two HIV *env* clones, CAP45_B5 (T/F) and CAP177_47 (chronic), were used to produce pseudovirus for the subsequent binding assays. HIV backbone plasmid DNA, HIV pSG3Δ*env* (5 µg) (Accession number L02317), together with either *env* clone (2.5 µg) were introduced into $4X10^5$ HEK293T cells in a 6 well tissue culture dish using the PEI transfection methodology as described. The culture medium was changed 6 hours after transfection to remove the PEI-DNA complexes and cells were grown at 37 °C in 5% CO_2. Pseudovirus was harvested from the supernatants and clarified through a 0.2 µm filter prior to storage in 20% FCS at -80 °C. The amount of pseudovirus was quantified using a p24 enzyme-linked immunosorbent assay (ELISA).

2.5.2 p24 ELISA

Lyophilized sheep anti-HIV p24 gag, affinity purified, coating antibody (5.5 ng/mL) (Aalto Bio reagents #D7320) in 1X $NaHCO_3$ (pH 8.5) buffer was used to coat high binding Porvair 96 well plates (WhiteSci, USA) overnight at RT. The plates were washed three times with 1X TBS and blocked with 100 µL/well 5% bovine serum albumin (BSA) in 1X TBS for 1 hour at RT. The plates were washed nine times with 1X TBS and dried. Eight p24 standards (16, 8, 4, 2, 1, 0.5, 0.25 and 0 ng/mL) were prepared using lyophilized recombinant HIV-p24 (Aalto Bio reagents #AG6054) diluted in 1% Empigen in 1 XTBS. Pseudovirus was inactivated with 1.25% Empigen in 1X PBS for 1 hour before the p24 ELISA was performed.

The inactivated pseudovirus samples were first diluted 1:50, 1:100, 1:200, 1:400 and 1:800 in 1% Empigen in 1X TBS and then 100 µL/well was added to the plates. The p24 standards were added (100 µL/well) in triplicate before the plates were incubated for 3 hours at RT. Empty rows were filled with 1X TBS to minimize evaporation. The plates were washed nine times with 1X TBS and dried before incubation with 100 µL secondary conjugate EH12AP (Aalto Bio reagents #BC-1071-AP) diluted 1:64 000 in 1X TBS, 20% sheep serum, 0.05%

Tween. The plates were incubated for 1 hour at RT and then washed nine times with 0.1% Tween in 1X TBS and dried. The plates were washed an additional two times with 1X TROPIX buffer (Applied Biosystems, USA) and dried. TROPIX CPD Star/Sapphire II (Applied Biosystems, USA) was diluted 1:4 in 1X TROPIX buffer and 50 µL/well was added to the plates. The plates were incubated for 5 minutes before luminescence was measured on a Modulus™ Microplate Luminometer (Promega, USA).

2.6 Sequencing the flanking regions of α4 in pTARGET^{TM}_α4

The sequence upstream and downstream of the α4 insert in the pTARGET^{TM} vector was sequenced to determine whether the α4 cDNA insert was inserted into the vector in the correct orientation during sub-cloning. The α4 F2 primer designed from nt 2 877 to nt 2 900 and the α4 R3 primer designed from nt 778 to nt 800 were synthesised and used to determine the sequence upstream of the α4 cDNA start codon and the sequence downstream of the α4 cDNA stop codon (Table 2.8).

Table 2.8 The primers used to sequence the flanking regions of α4 in pTARGET^{TM}_α4

*Primer	*Direction	Sequence	Size
12-1692 α4 F2	F	5'-TGC GCA TGT TCT ACT GGA AGG ACT-3'	24
12-1963 α4 R3	R	5'-TCG GTA GTA TGC TGG CTC CGA AA-3'	23

*F=forward, R=reverse

2.6.1 Binding of high and low concentrations of pseudovirus to α4β7 expressing HeLa cells in the presence and absence of increasing concentrations of Act-1

One day before binding 8X10^5 HeLa cells, endogenously expressing α4 and β7, were plated in a 6 well tissue culture dish. HeLa cells were incubated with 0 nM, 33 nM, 66 nM or 99 nM Act-1 for 1 hour. Unbound Act-1 was removed when cells were washed twice with 1X PBS. Approximately 200 ng and 5 ng p24 CAP45_B5 (T/F) and CAP177_47 (chronic) pseudovirus was added to the HeLa cells in triplicate. Binding was allowed for 2.5 hours at 37 °C in 5% CO_2. Unbound pseudovirus was removed and cells were washed by centrifugation three

times with 1X PBS at 2 000 g for 5 minutes at RT. Cells were lysed on ice at 4 °C with 300 µL RIPA buffer containing PMSF. The concentration of p24 in the cell lysates were determined by a P24 ELISA to measure the amount of pseudovirus bound to the cells.

2.6.2 Binding of pseudovirus to TZM-bl cells in the presence and absence of Act-1

TZM-bl cells were tested in the same binding assay as described with 33 nM Act-1. A day before binding, $8X10^5$ TZM-bl cells were plated in a 6 well tissue culture dish and were incubated with 0 nM and 33 nM Act-1 for 1 hour. Unbound Act-1 was removed when cells were washed twice with 1X PBS. Approximately 200 ng p24 pseudovirus of all five T/F and chronic infection Env clones was added to the TZM-bl cells in triplicate. Binding was allowed for 2.5 hours at 37 °C in 5% CO_2. Unbound pseudovirus was removed and cells were washed by centrifugation three times with 1X PBS at 2 000 g for 5 minutes at RT. Infection was read after 48 hours using 50 µL Brite-Glo lysis buffer with the Modulus™ Microplate Luminometer (Promega, USA).

2.6.3 Binding of pseudovirus to CD4+ T cells isolated from blood in the presence and absence of Act-1

Buffy packs of human blood from healthy donors were received from the Western Province Blood Transfusion Service, Cape Town, South Africa. Ficoll Histopaque®-1077 (Sigma Aldrich, USA) (15 mL) was added to Leucosep® (50 mL) tubes and centrifuged at 2 000 g for 1 minute at RT so that the Ficoll moved below the porous separation disc. The blood was diluted 2:3 1X PBS:blood, added (30 mL) to the Leucosep® tubes and centrifuged at 3 000 g for 10 minutes at RT. The porous disc insured that the interface between the Ficoll and the blood was not disturbed upon adding the blood.

The plasma and PBMC layer had a lower density than the Ficoll and therefore remained above the Ficoll while the high density erythrocytes moved below the Ficoll layer. The plasma layer (~5 mL) on the surface was discarded and the PBMC layer seen as a thin white 'buffy' layer was removed and washed twice with 1% FCS in 1X PBS. Cells were resuspended

in RPMI with 10% FCS and monocytes were allowed to adhere for 4 hours at 37 °C in 5% CO_2. Thereafter the culture medium was removed and placed into a T75. The cells were induced with 10 U/mL IL-2 (NIH ARP, USA) every 2 days for 6 days in RPMI. On the seventh day cells were counted and $2.5X10^6$ cells/mL were incubated with 0 nM and 33 nM Act-1 for 1 hour at 37 °C in 5% CO_2 (Parrish et al. 2012). Cells were washed twice with centrifugation at 3 000 g for 3 minutes at RT. Cells were incubated with 100 ng p24 CAP45_B5 (T/F) and CAP177_47 (chronic) pseudovirus and allowed to bind for 5 hours before unbound pseudovirus was removed and the cells were washed four times with RPMI at 3 000 g for 3 minutes at RT. Cells were lysed with 1.25% Empigen in 1X PBS and a p24 ELISA was performed to measure the amount of bound pseudovirus.

2.7 DC-SIGN mediated *trans*-infection

2.7.1 Flow cytometry to determine DC-SIGN expression on Raji-DC-SIGN cells

One million Raji-DC-SIGN cells were centrifuged at 2 000 g for 5 minutes at RT, resuspended in 200 µL cold FACS buffer (1X PBS containing 2% FCS) and stained with either 5 µg/mL PerCP-Cy™5.5 anti-human CD209 (DC-SIGN) antibody (BD Pharmingen™, San Jose, USA) or 5 µg/mL PerCP-Cy™5.5 Mouse IgG2b, κ Isotype control (BD Pharmingen™, San Jose, USA) for 45 minutes at 4 °C in the dark. Cells were then washed four times with 1 mL cold FACS buffer and centrifuged at 2 000 g for 5 minutes at 4 °C. Cells were resuspended in 200 µL cold FACS buffer and D. Bowers (IIDMM, UCT, South Africa) acquired the data on a FACS Caliber™ flow cytometer (BD Biosciences, USA) and analysed the data using FloJo software (v 8.8.2, Tree Star, Stanford, CA). As a control the same procedure was performed on Raji-WT cells not expressing DC-SIGN.

2.7.2 DC-SIGN mediated *trans*-infection assay

The DC-SIGN mediated HIV *trans*-infection assay was conducted as described previously (Alexandre et al. 2012). Pseudovirus was normalised on p24 values determined by a p24

ELISA and 7.5X10^4 Raji-DC-SIGN cells were incubated with 10 ng p24 pseudovirus in a rolling incubator at 37 °C for 2.5 hours. After incubation the Raji-DC-SIGN cells were washed three times with culture medium using centrifugation at 2 000 g for 5 minutes at RT. Raji-DC-SIGN cells were resuspended in culture medium and co-cultured in triplicate with 6X10^4 TZM-bl cells that were plated the day before. Infection was read after 48 hours using 50 µL Brite-Glo lysis buffer (Promega, USA) with the Modulus™ Microplate Luminometer (Promega, USA). As a control for the presence of DC-SIGN, the same procedure was performed on Raji-WT cells not expressing DC-SIGN.

2.7.3 Site-directed mutagenesis

CAP239_E11 (T/F) and CAP239_T35 (chronic) *env* clones were chosen to undergo SDM at two PNG sites. The PNGs (NXS/T) at position N386 and N392 (according to HXB2 numbering) were mutated by changing the codon corresponding to the asparagine to that of a glutamine (N386Q and N392Q), thereby abolishing N-glycosylation at these sites. Similarly, double mutants lacking both PNGs were generated (N386Q,N392Q).

In total four single mutants (239_E11_N386Q, 239_T35_N386Q, 239_E11_N392Q, 239_T35_N392Q) and two double mutants (239_E11_N386Q,N392Q and 239_T35_N386Q,N392Q) were generated. Mutagenesis was performed using a PCR based methodology adapted from Quickchange® site-directed mutagenesis protocol (Stratagene La Jolla, CA, USA). Complementary oligonucleotides containing the desired mutation were designed using WATCUT (*http://watcut.uwaterloo.ca/watcut* designed by Michael Palmer, University of Waterloo, Canada) with restriction analysis and silent mutation analysis software.

Restriction enzyme sites were introduced through silent mutations to facilitate screening. A *BsaW* I (New England Bio labs, USA) restriction enzyme site was introduced into both N386Q mutants and an *Acl* I (Fermentas, USA) restriction enzyme site was introduced into both N392Q mutants. The double mutants contained both restriction enzyme sites. The primers used in SDM are indicated in Table 2.9 with the asparagine to glutamine transition at the N-

glycosylation sites at position N386 and N392 highlighted in bold and the introduced silent mutation incorporating a restriction enzyme recognition site underlined.

Mutagenesis was performed with Phusion Hot Start II High-Fidelity DNA Polymerase (Thermo Scientific, USA) (Parameters: 94 °C 5 minutes, 20 cycles of 94 °C 30 seconds, 55 °C 30 seconds and 72 °C 12 minutes followed by 72 °C 12 minutes). Aliquots of the PCR products were visualised by 1% agarose gel electrophoresis, while the remainder was *Dpn* I (Thermo Scientific, USA) digested overnight to fragment methylated template DNA. *E. coli* JM109 competent cells were transformed with the *Dpn* I digested PCR products and colonies were selected on Luria agar plates with 100 µg/ml ampicillin (Sigma Aldrich, USA).

Table 2.9 The site-directed mutagenesis primers used in generating CAP239_E11 (T/F) and CAP239_T35 (chronic) N-glycan mutants

*Primer	*Direction	*Sequence	Size
239_N386Q F	F	5'-GAG GAG AAT TTT TCT ATT GC**C** A**G**A CAT CCG GTC TGT TTA ATG G-3'	43
239_N386Q R	R	5'-CCA TTA AAC AGA CCG GAT GT**C** T**G**G CAA TAG AAA AAT TCT CCT C-3'	43
239_N392Q F	F	5'-CTA TTG CAA TAC AT CAG GCC TGT TT**C** A**G**G GAA CGT TTA ATG GTA C-3'	45
239_N392Q R	R	5'-GTA CCA TTA AAC GTT CC**C** T**G**A AAC AGG CCT GAT GTA TTG CAA TAG-3'	45
239_N386Q,N392Q F	F	5'-CTA TTG C**C**A **G**AC ATC CGG TCT GTT T**C**A **G**GG AAC GTT TAA TGG TAC-3'	45
239_N386Q,N392Q R	R	5'-GTA CCA TTA AAC GTT CC**C** T**G**A AAC AGA CCG GAT GT**C** T**G**G CAA TAG-3'	45

*F=forward, R=reverse
*Restriction enzyme sites introduced by silent mutation are underlined and nucleotide changes are highlighted in bold

Plasmid DNA was isolated from overnight cultures using the crude sodium chloride, tris, ethylenediaminetetraacetic acid, Triton X-100 (STET) boiling minilysate preparation methodology. Overnight cultures (2 mL) were centrifuged at 13 000 g for 2 minutes at RT. The supernatant was discarded and the cell pellet was resuspended in 250 µL STET buffer (8% Sucrose, 5% Triton X-100, 50 mM EDTA, 50 mM Tris pH 8.0) containing 1 mg/mL lysozyme. The samples were vortexed, boiled for 1 minute and centrifuged at 13 000 g for 8 minutes at RT. The pellet was discarded and an equal volume of isopropanol was added. After an additional 13 000 g centrifugation for 8 minutes at RT the supernatant was discarded and the air dried pellet was resuspended in 30 µL nuclease free water.

The plasmid DNA was precipitated by using 1/10th volume 3 M sodium acetate (pH 5.2) and 2.5X total volume cold absolute ethanol. After mixing very well, the plasmid DNA was incubated at -20 °C for 30 minutes before centrifugation at 13 000 g for 30 minutes at RT. The plasmid DNA pellet was washed with 1 mL 70% ethanol by centrifugation for 15 minutes at RT. The supernatant was discarded and the pellet was resuspended in nuclease free water. The STET prepared plasmid DNA was screened by restriction enzyme digestion for positive clones and visualised by 1% agarose gel electrophoresis where loading dye (0.25% bromophenol blue, 0.25% xylene cyanol, 30% glycerol) contained RNase. The same approach was used to generate the single and double N-glycan mutants however the single mutants were used as the template for the double mutant PCR reactions. A Bioflux Miniprep kit (Bioer Technology Co., Ltd) was used to isolate plasmid DNA for sequencing of the entire *env*, using primers listed in Table 2.10, thus confirming the presence of the mutations.

Table 2.10 The primers used for sequencing *env* N-glycan mutants

Primer	*Direction	Sequence	Size
For14	F	5'-TAT GGG ACC AAA GCC TAA AGC CAT GTG-3'	27
For16	F	5'-TTT AAT TGT GGA GGA GGA TTT TTC TA-3'	26
EF00	F	5'-GGG AAA GAG CAG AAG ACA GTG GCA ATG A-3'	28
EF200	F	5'-GGG ATA ACA TGA CCT GGA TGC AGT GGG-3'	27
EF260	F	5'-TTC AGC TAC CAC CGA TTG AGA GAC T-3'	25
EF15	R	5'-CTT GCT CTC CAC CTT CTT CTT C-3'	22
EF55	R	5'-GCC CCA GAC CGT GAG TTG CAA CAT ATG-3'	27
Rev15	R	5'-CTG CCA TTT AAC AGC AGT TGA GTT GA-3'	26
Rev19	R	5'-ACT TTT TGA CCA CTT GCC ACC CAT-5'	24

*F=forward, R=reverse

2.7.4 DC-SIGN mediated *trans*-infection of N-glycan mutants

The *trans*-infection assay was repeated as outlined above with WT *env* clones and CAP239 T/F and chronic infection N-glycan SDM mutants.

2.7.5 DC-SIGN binding

The DC-SIGN binding assay was performed as previously described (Van`Montfort et al. 2011). Approximately 10 ng p24 pseudovirus was added to $1X10^6$ Raji-DC-SIGN cells and incubated in a rolling incubator at 37 °C for 2.5 hours. After incubation the Raji-DC-SIGN cells were washed three times with culture medium using centrifugation at 2 000 g for 5 minutes at RT. Cells were lysed in 130 µL 1% Empigen in 1X TBS and stored at 4 °C overnight. The amount of cell-bound pseudovirus was quantified using a p24 ELISA. As a control for the presence of DC-SIGN, the same assay was performed with Raji-WT cells not expressing DC-SIGN.

2.8 Statistical analysis

All statistical analysis was performed on GraphPad Prism version 5.00 for Windows, GraphPad Software, La Jolla California USA, *www.graphpad.com*, analysed with in-built statistical software and represented graphically. A non-parametric Mann-Whitney two-tailed t-tests was performed and the results are indicated as statistically significant with a * when $p<0.05$ and ** when $p<0.01$.

Chapter 3: Development of an α4β7 reactivity assay using transient integrin expression in mammalian cell lines

3.1 Introduction

The α4β7 integrin is a heterodimer consisting of the non-covalently attached α4 and β7 integrins (Buck & Horwitz 1987). The α4β7 integrin is a homing receptor expressed on CD4+ T cells that mediates the migration of CD4+ T cells from gut inductive sites to the gut (Cicala et al. 2009). The α4β7 integrin binds to gp120 via a tripeptide in the V2 loop of HIV that is highly conserved across subtypes A, B, C and D (Arthos et al. 2008) suggesting that the gp120-α4β7 interaction provides a certain advantage (Arthos et al. 2008; Cicala et al. 2009). Although, the presence of CD4 and CCR5 is sufficient for HIV infection, α4β7, due to its size (22 nm), might act as an anchor, concentrating virus on the cell's surface and enhancing viral attachment to CD4 (7 nm) and a co-receptor (Cicala et al. 2009).

CD4, CCR5 and α4β7 form close complexes on the cell membrane (Cicala et al. 2009) and this proximity suggests that variants exhibiting optimal α4β7 reactivity may be able to engage the CD4 and CCR5 receptors better than those with low α4β7 reactivity. In fact, HIV variants with higher α4β7 reactivity were shown to bind CD4 better than those with lower reactivity (Cicala et al. 2009). Variants with high α4β7 reactivity should preferentially infect CCR5+ CD4+ T cells carrying α4β7, which may explain why T/F viruses seem to utilise CCR5 exclusively (Keele et al. 2008). Subtype A and C T/F variants have been shown to bind α4β7 better than chronic infection variants (Nawaz et al. 2011) although this finding was not supported by a recent study (Parrish et al. 2012).

HIV T/F variants are distinct in that they are R5 tropic, have shorter variable loops (Chohan et al. 2005) and carry less N-linked glycans in the V2/V4 domains compared with chronic infection variants (Derdeyn et al. 2004). The introduction of N-glycans over the course of infection is a viral strategy used to evade humoral immune responses as the carbohydrate structures shield antibody epitopes from recognition (Binley et al. 2010; Reitter et al. 1998;

Sagar et al. 2006). The difference in T/F N-glycosylation compared to viruses found later in infection was predominantly observed in the V2 loop (Derdeyn et al. 2004), the α4β7 binding site domain, where the presence of fewer N-glycans might facilitate T/F Env binding to α4β7. Nawaz et al. (2011) demonstrated that PNGs in the V1/V2 loop as well as the more distal C3/C4 domain influenced α4β7 reactivity but the authors concluded that this interference was due to changes in gp120 conformation and not due to steric hindrance. Overall, T/F variants might carry an optimal arrangement of N-glycans that facilitate enhanced α4β7 reactivity, suggesting that α4β7 may play a central role in HIV transmission. We hypothesise that T/F variants carry α4β7 binding sites with higher affinity for α4β7 possibly due to the presence or absence of specific PNGs that allow them to interact with the α4β7 integrin more efficiently compared with their chronic infection counterparts.

Decreased binding to α4β7 was associated with lower viral loads (Ansari et al. 2011) and variants with mutated α4β7 binding sites demonstrated reduced viral replication (Arthos et al. 2008). These findings support the hypothesis that α4β7 reactivity may be a distinguishing feature of T/F viruses. However, recently it was shown that the infectivity of subtype C T/F and chronic infection variants was not inhibited by blocking α4β7 with a monoclonal antibody (Parrish et al. 2012), suggesting that α4β7 might not play a role in HIV transmission. These studies did not utilise longitudinal samples to compare T/F with chronic infection variants and comparison of pseudovirus with IMC might have led to conflicting results with regard to gp120-α4β7 interactions. Therefore, this study aimed to investigate α4β7 reactivity by comparing HIV subtype C T/F with chronic infection variants from the same participants.

3.2 Results

3.2.1 Sub-cloning of α4 cDNA and preparation of pCEP4_β7

Although other studies have utilised α4β7+ CD4+ T cells to measure α4β7 reactivity (Nawaz et al. 2011; Cicala et al. 2009), this project aimed to express α4β7 in a mammalian cell line in order to circumvent potential PBMC donor variation. Furthermore, in order to measure the

reactivity between α4β7 and trimeric Env, pseudovirus was generated carrying our subtype C T/F and chronic infection *env* clones. We planned to transiently express α4β7 in HEK293T cells and then compare the binding ability of Env pseudotyped variants. The integrin's cDNA, in expression cassettes pCDM8_α4 and pCEP4_β7 were a gift from Prof D. Erle, University of California, USA (Kamata et al. 1995; Erle et al. 1991).

The pCDM8 vector carrying the α4 cDNA utilised the tyrosine tRNA suppressor and p3 plasmid system for plasmid selection. As this methodology was not set up in our lab and due to constraints on importing competent cells at the time we sub-cloned the α4 cDNA into the pTARGET™ expression vector using TA cloning. Primers were designed and synthesised to amplify the α4 cDNA (3 096 bp) using the pCDM8_α4 plasmid DNA as a template. The PCR product was A-tailed and ligated to pTARGET™, a linear T-tailed vector specifically designed for TA cloning.

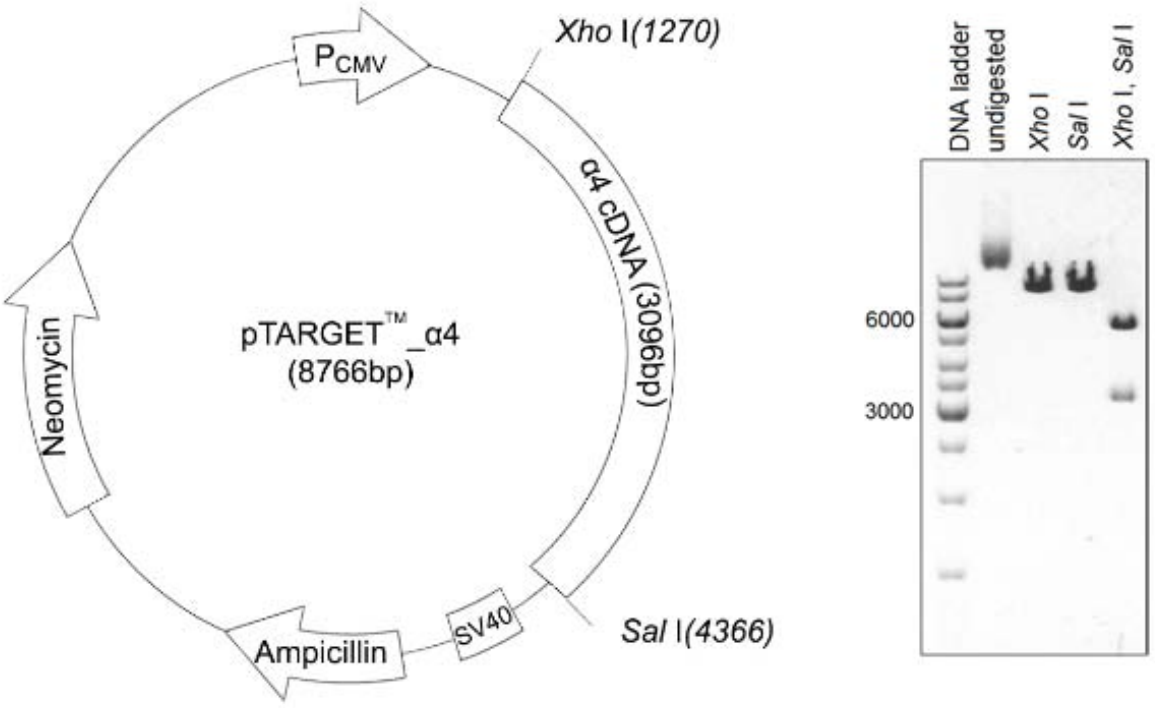


Figure 3.1 Restriction enzyme digestion confirmed pTARGET™_α4 plasmid DNA. A single and double restriction enzyme digest using the *Xho* I and *Sal* I restriction enzymes confirmed the presence of the α4 cDNA insert (3 096 bp) in the pTARGET™ vector (5 670 bp), visualised by 1% agarose gel electrophoresis.

In order to confirm that the PCR product of α4 was correctly cloned into the pTARGET™ vector, plasmid DNA extracted from colonies screened to carry the insert in the correct

orientation was digested with *Xho* I and *Sal* I. A single *Xho* I and *Sal* I restriction enzyme digest of pTARGET^{TM}_α4 resulted in a 8 766 bp band that corresponded to the length of the linearized pTARGET^{TM}_α4 plasmid DNA. A double *Xho* I and *Sal* I restriction enzyme digest of pTARGET^{TM}_α4 resulted in two bands at 5 670 bp and at 3 096 bp that corresponded to the pTARGET^{TM} vector and the α4 cDNA insert, respectively (Figure 3.1).

The pCEP4_β7 plasmid DNA was transformed into *E. coli* JM109 competent cells and amplified using antibiotic selection. Single and double restriction enzyme digestions with *Kpn* I and *BamH* I were carried out to confirm that the correct plasmid DNA had been isolated.

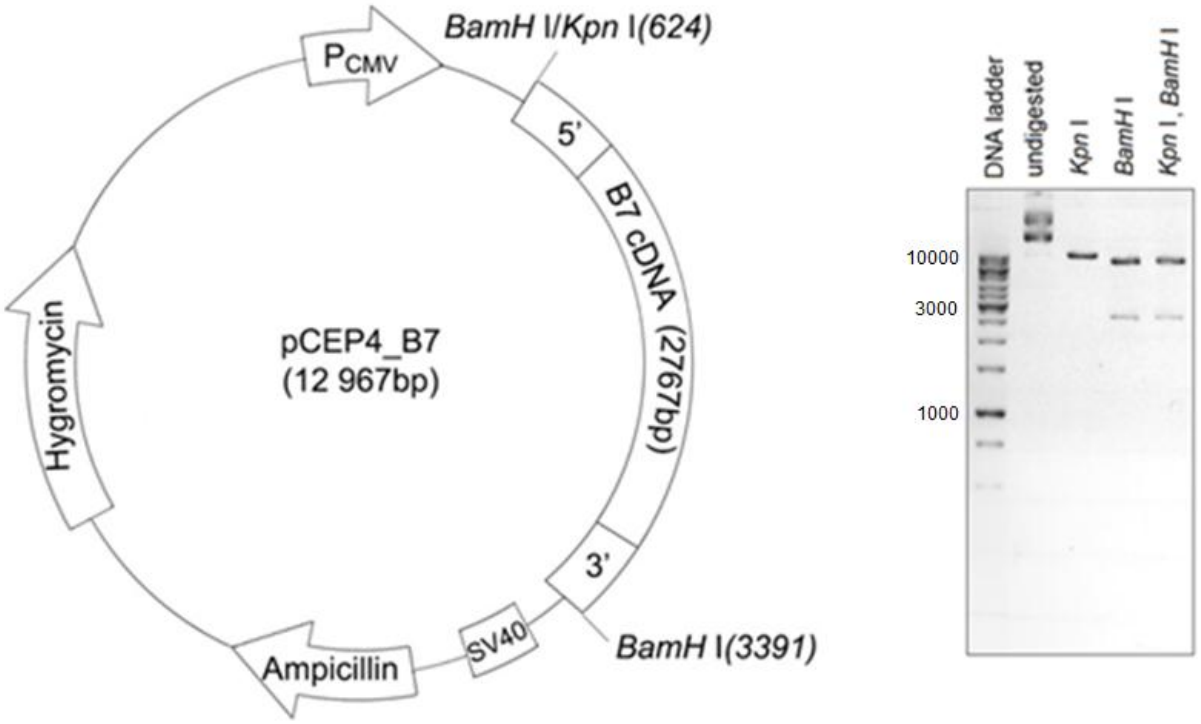


Figure 3.2 Restriction enzyme digests confirmed pCEP4_ β7 plasmid DNA. A single and double restriction enzyme digest using the *Kpn* I and *BamH* I restriction enzymes confirmed the presence of the β7 cDNA insert (2767 bp) in the pCEP4 vector (10 200 bp), visualised by 1% agarose gel electrophoresis.

A single *Kpn* I restriction enzyme digest of pCEP4_β7 resulted in a ≥ 10 000 bp band that corresponded to the length of the linearized pCEP4_β7 plasmid DNA (12 967 bp). A single *BamH* I and a double *Kpn* I and *BamH* I restriction enzyme digest resulted in two bands at ≥ 10 000 bp and at 2 767 bp that corresponded to the pCEP4 vector (10 200 bp) and the β7 cDNA insert, respectively (Figure 3.2).

Single and double restriction enzyme digests visualised by agarose gel electrophoresis confirmed the presence of the α4 and β7 cDNA inserts in the pTARGET^{TM} and pCEP4 vectors, respectively.

3.2.2 Sequence analysis of the α4 and β7 integrin cDNAs

Although, restriction enzyme digestion had confirmed pTARGET^{TM}_α4 and pCEP4_β7 plasmid DNA, we confirmed the sequence identity of both constructs by designing and synthesising primers to sequence the α4 and β7 inserts (Figure 3.3).

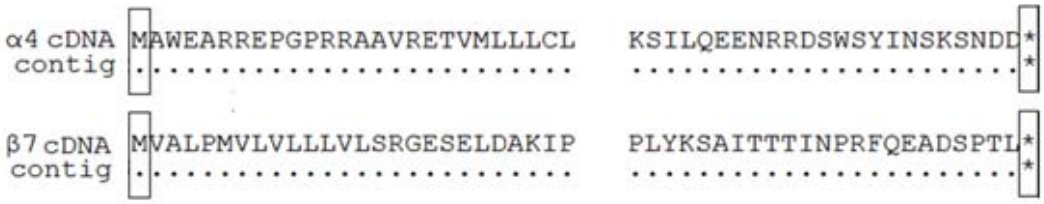


Figure 3.3 DNA sequencing confirmed the identity of the α4 and β7 cDNAs cloned into the pTARGET^{TM} and pCEP4 vectors, respectively. The contigs were generated in ChromasPro using overlapping sequences from primers complimentary to the insert sequences. In order to indicate that the entire cDNA of α4 and β7 was cloned, only the 5' and 3' ends of both α4 and β7 are indicated aligned to the GenBank sequences accession numbers NP_000876 and NP_000880, respectively. Matched nucleotides are indicated with dots. The start of the open reading frames (methionine, M) of the integrins is indicated by a rectangular box in each alignment and the stop codon represented by a * is indicated by a rectangular box.

Sequence analysis confirmed that the entire cDNA of both α4 and β7 were inserted into the pTARGET^{TM} and pCEP4 vectors, respectively and matched 100% with the GenBank sequence [α4 (NP_000876) and β7 (NP_000880)] (Figure 3.3).

3.2.3 Expression of α4β7 in mammalian cells

3.2.3.1 Determining the optimum transfection methodology

This study aimed to utilise HEK293T cells transiently co-transfected with α4 and β7 expression vectors to measure the α4β7 reactivity of HIV T/F and chronic infection Env clones. In order to determine the transfection methodology that resulted in the most

optimal expression of α4β7, two HIV *env* clones, CAP177_A3 (T/F) and CAP206_E12-3 (chronic) that were previously shown to have high expression levels by Western blot were used to compare the transfection of HEK293T cells using PEI, PolyFect® and Electroporation.

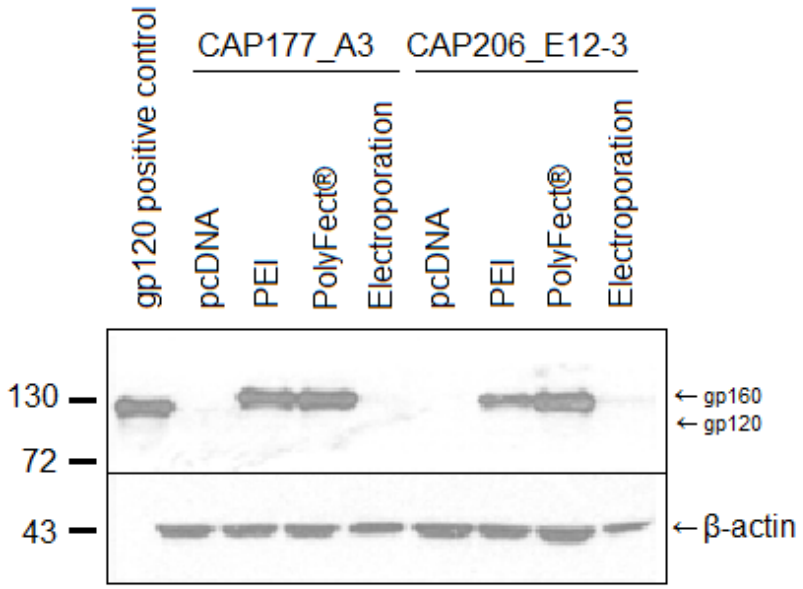


Figure 3.4 Transfections with PEI and PolyFect® resulted in high HIV *env* expression compared with Electroporation. A gp120 Western blot compared *env* expression using PEI, PolyFect® and Electroporation transfection methodologies. CAP177_A3 (T/F) and CAP206_E12-3 (chronic) *env* clones were transfected into HEK293T cells. Equivalent total protein (60 µg) of cell lysates was loaded per well. A 1:3 000 dilution of the anti-gp120 antibody (Mouse antiserum to HIV gp120, NIH ARP) was used. The pcDNA™3.1/V5-His TOPO® empty vector (pcDNA) was included to control for non-specific binding of antibodies. A gp120 soluble recombinant protein (50 ng) (120 kDa) was included to control for successful Western blotting and β-actin (42 kDa) was included as a loading control. The exposure time for this Western blot was 5 minutes. The molecular weight marker is indicated.

Western blot analysis revealed that both PEI and PolyFect® transfections resulted in high Env expression in HEK293T cells compared with Electroporation (Figure 3.4). Even though Electroporation was used to stably express α4 and β7 from pCDM8_α4 and pCEP4_β7 plasmid DNA in K562 human erythroleukemia cells (Abitorabi et al. 1997), Electroporation did not result in high protein expression, but rather poor Env expression. In all the experimental samples a single band was visible corresponding to gp160 as it ran higher than the gp120 band in the positive control, indicating unprocessed HIV Env. It has been previously shown that cell-associated gp160 is present at much higher concentrations than gp120 in cell lysates (Pfeiffer et al. 2006; Asmal et al. 2011).

The PolyFect® CAP206_E12-3 sample showed higher expression than the corresponding PEI sample although the corresponding β-actin band suggested that more total protein was loaded in this lane. As PEI was more cost effective and there were no major variations between PEI and PolyFect®, PEI was used for the expression of α4 and β7.

3.2.3.2 Transfection of HEK293T cells and Western blot analysis with different primary antibodies for the detection of α4β7

HEK293T cells were both singly transfected as well as co-transfected with pTARGET™_α4 and pCEP4_β7 plasmid DNA using PEI (Figure 3.5).

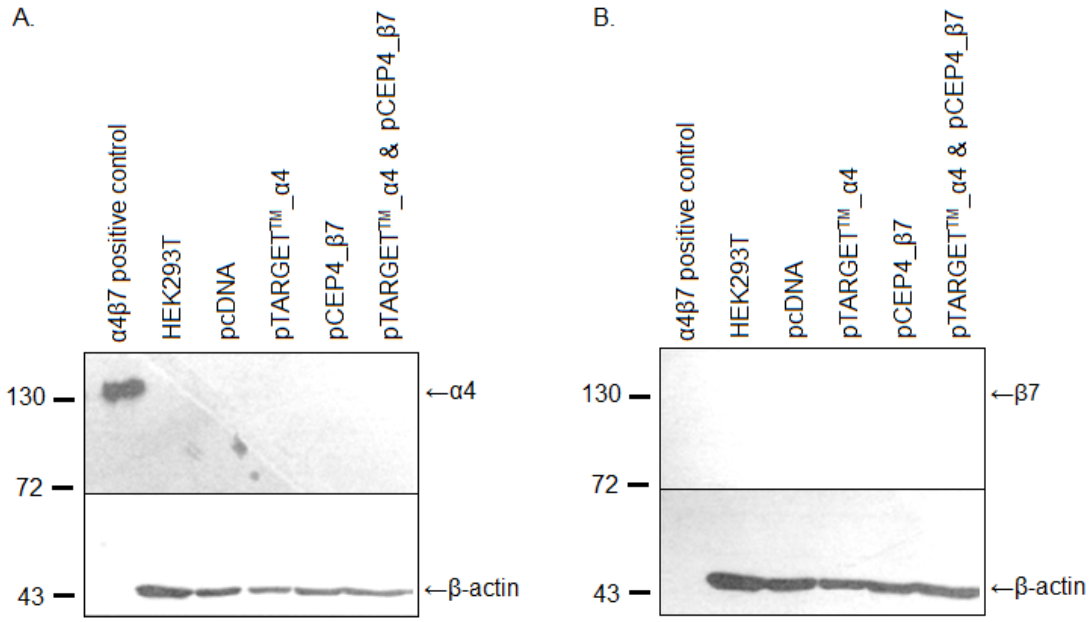


Figure 3.5 Western blot did not detect α4 or β7 integrins in transfected HEK293T cell lysates using antibodies from Sigma Aldrich, USA. pTARGET™_α4 (1 pmol) and pCEP4_β7 plasmid DNA (1 pmol) was transfected separately transfected and co-transfected into HEK293T cells using PEI. Equivalent total protein (60 µg) of cell lysates was loaded per well A) A 1:100 dilution of the anti-α4 antibody (Sigma Aldrich, USA) was used to detect the expression of the α4 integrin and B) A 1:100 dilution of anti-β7 antibody (Sigma Aldrich, USA) was used to detect the expression of the β7 integrin in the transiently transfected HEK293T cells. Cells only (HEK293T) and pcDNA™3.1/V5-His TOPO® empty vector (pcDNA) were included to control for non-specific binding of antibodies. An α4β7 recombinant protein (100 ng) (150 kDa) was included to control for successful Western blotting and β-actin (42 kDa) was included as a loading control. The exposure time for both Western blots was 8 minutes. The molecular weight marker is indicated.

A Western blot using the anti-α4 antibody (Sigma Aldrich, USA) detected the positive recombinant protein at 150 kDa and cell lysates generated from transfection with pcDNA™3.1/V5-His TOPO® empty vector did not yield any bands as expected. However, cells

transfected with 1 pmol pTARGET^{TM}_α4 and co-transfected with 1 pmol pTARGET^{TM}_α4 and pCEP4_β7 showed no α4 expression even though the band corresponding to β-actin indicated comparable loading of samples (Figure 3.5A).

A Western blot using the anti-β7 antibody (Sigma Aldrich, USA) did not detect the positive recombinant protein at 130 kDa, even when 100 ng recombinant protein was loaded. Cell lysates generated from transfection with pcDNA^{TM}3.1/V5-His TOPO$^{®}$ empty vector did not yield any bands as expected while cells transfected with 1 pmol pCEP4_ β7 and co-transfected with 1 pmol pTARGET^{TM}_α4 and pCEP4_ β7 showed no β7 expression even though the band corresponding to β-actin indicated comparable loading of samples (Figure 3.5B).

Transfections and Western blots were repeated and conditions were altered to optimise integrin detection but there was no improvement. As the anti-β7 antibody (Sigma Aldrich, USA) was unable to detect the commercially available α4β7 recombinant protein (R and D systems, England) an alternative antibody (Santa Cruz Biotechnology Inc., USA) was purchased. As it was possible that the anti-α4 antibody was not specific enough to detect low levels of integrin expression, a new commercially available antibody was purchased.

When we acquired the second set of antibodies (Santa Cruz Biotechnology Inc., USA) the data sheet indicated that HeLa cell lysates could be used as a positive control in Western blotting as these cells endogenously express α4 and β7, although no supporting literature was found or could be supplied by Santa Cruz Biotechnology Inc. We included HeLa and TZM-bl (HeLa derived) cell lysates in subsequent Western blots as positive controls.

HEK293T cells were separately transfected and co-transfected with increasing concentrations of pTARGET^{TM}_α4 and pCEP4_β7 plasmid DNA and protein expression was detected using the newly purchased antibodies (Santa Cruz Biotechnology Inc., USA) (Figure 3.6).

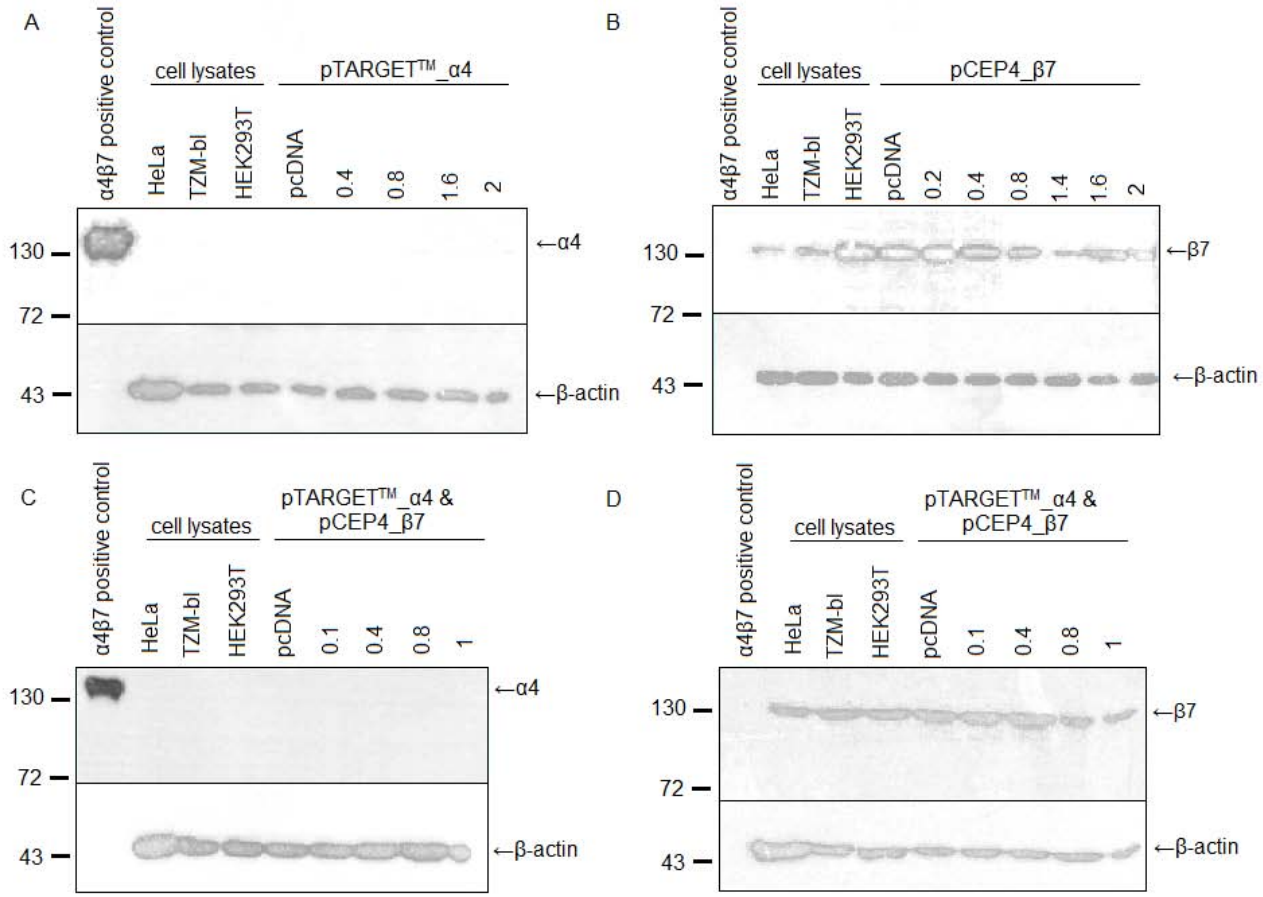


Figure 3.6 Western blot did not detect the α4 and β7 integrins in transfected HEK293T cell lysates using antibodies from Santa Cruz Biotechnology Inc., USA. Increasing amounts of pTARGET™_α4 (0.4, 0.8, 1.6 and 2 pmol) and pCEP4_β7 plasmid DNA (0.2, 0.4, 0.8, 1.4, 1.6 and 2 pmol) were transfected separately into HEK293T cells and increasing amounts of pTARGET™_α4 (0.1, 0.4, 0.8 and 1pmol) and pCEP4_β7 plasmid DNA (0.1, 0.4, 0.8 and 1pmol) were co-transfected into HEK293T cells. A) A 1:100 dilution of anti-α4 antibody (Santa Cruz Biotechnology Inc., USA) was used to detect the expression of the α4 integrin and B) a 1:100 dilution of anti-β7 antibody (Santa Cruz Biotechnology Inc., USA) was used to detect the expression of the β7 integrin in the transiently transfected HEK293T cells and C) A 1:100 dilution of anti-α4 antibody (Santa Cruz Biotechnology Inc., USA) was used to detect the expression of the α4 integrin and D) A 1:100 dilution of anti-β7 antibody (Santa Cruz Biotechnology Inc., USA) was used to detect the expression of the β7 integrin in the transiently co-transfected HEK293T cells. Equivalent total protein (60 µg) of cell lysates was loaded per well. Cells only (HEK293T) and pcDNA™3.1/V5-His TOPO® empty vector (pcDNA) were included to control for non-specific binding of antibodies. An α4β7 recombinant protein (100 ng) (150 kDa) and cells carrying endogenous α4β7 (HeLa and TZM-bl) were included to control for successful Western blotting and β-actin (42 kDa) was included as a loading control. The exposure time for all the Western blots was 8 minutes. The molecular weight marker is indicated.

A Western blot using the newly purchased anti-α4 antibody (Santa Cruz Biotechnology Inc., USA) detected the positive recombinant protein (100 ng) at 150 kDa and cell lysates generated from a transfection with pcDNA™3.1/V5-His TOPO® empty vector did not yield any bands as expected (Figure 3.6A and Figure 3.6C). Cells transfected with pTARGET™_α4 (0.4 to 2 pmol) and co-transfected with pTARGET™_α4 and pCEP4_β7 (0.1 to 1 pmol) showed no α4 expression even though the bands corresponding to β-actin indicated

comparable loading of samples. Unexpectedly the HeLa and TZM-bl cell lysate samples were negative for α4 expression.

A Western blot using the new anti-β7 antibody (Santa Cruz Biotechnology Inc., USA) did not detect the positive recombinant protein at 130 kDa, even when 100 ng recombinant protein was loaded (Figure 3.6B and Figure 3.6D). Cell lysates generated from transfection with pCEP4_ β7 (0.2 to 2 pmol) and co-transfection with pTARGET^{TM}_α4 and pCEP4_β7 (0.1 to 1 pmol) indicated a band that corresponded to the molecular weight of the β7 integrin (130 kDa). This band was present in HEK293T cell lysates not transfected with the integrin plasmid DNA as well as cell lysates transfected with pcDNA^{TM}3.1/V5-His TOPO$^{®}$ empty vector. HEK293T cells do not endogenously express β7 (Qi et al. 2012) suggesting that the DNA band did not represent the β7 integrin but that the anti β7-antibody was binding to a non-specific protein that had a similar molecular weight to β7. The intensity of the 130 kDa band varied with the amount of transfected pCEP4_β7 plasmid DNA (Figure 3.6B) although it was not as expected as transfection with 1.4 pmol pCEP4_β7 resulted in lower expression than transfection with 0.2 pmol of plasmid DNA. The observed difference in expression was not due to loading inconsistent amounts of total protein as the intensity of β-actin was comparable between samples.

This experiment was designed to determine the optimum concentration of plasmid DNA for integrin expression, and therefore the total plasmid DNA concentration was not kept constant with the addition of pcDNA^{TM}3.1/V5-His TOPO$^{®}$ empty vector. It is therefore possible that higher concentrations of plasmid DNA affected the transfection efficiency and/or cell viability.

Dot blots of the α4β7 recombinant protein (R and D systems, England) were performed to determine the optimal concentration of antibody for integrin detection but the recombinant protein (100 ng) was not detected by the anti-β7 antibodies (Sigma Aldrich, USA and Santa Cruz Biotechnology Inc., USA) at the highest recommended dilution (1:100). Western blot parameters including transfer system (wet and dry), time taken to transfer (1 to 2 hours), time taken to block (1 to 48 hours), number of washes (2 to 8 15-minute washes), incubation time of primary and secondary antibodies (1 to 24 hours) and exposure time to

X-ray film (30 seconds to 30 minutes) were tested but neither α4 nor β7 was detected in HEK293T cell lysates after transfection (data not shown).

To ensure that degradation of the plasmid DNA was not responsible for the lack of detectable cell-associated α4β7, plasmid DNA was retransformed, fresh plasmid DNA was prepared and the integrity and quality of the plasmid DNA was confirmed by Nanodrop 260/230 and 260/280 ratios as well as agarose gel electrophoresis (data not shown). A second aliquot of the pCEP4_β7 plasmid DNA was obtained from Prof D. Erle but when this was used in HEK293T transfection no expression could be detected.

In order to confirm the transfection efficiency of the PEI methodology, pTARGET^{TM}_α4 and pCEP4_β7 plasmid DNA were co-transfected together with the pGL4 luciferase reporter gene (Figure 3.7).

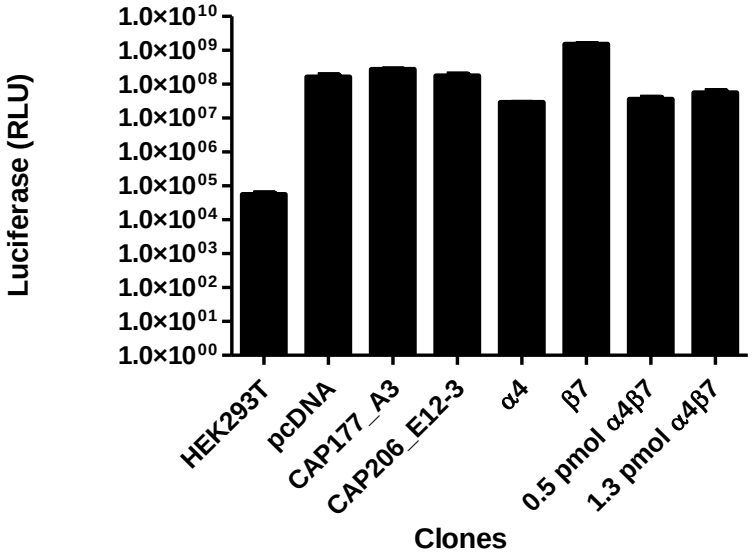


Figure 3.7 The efficiency of the PEI transfection methodology was confirmed by a luciferase reporter gene. HEK293T cells were transfected with pGL4 Luc (1.2 pmol) and either 1.2 pmol pTARGET^{TM}_α4 (α4) or pCEP4_β7 (β7) or co-transfected with 0.5 and 1.3 pmol pTARGET^{TM}_α4 and pCEP4_β7 (α4β7) plasmid DNA. The pGL4 Luc construct carrying a luciferase reporter gene was an indicator of how well the cells took up the plasmid DNA. A negative cells only (HEK293T) control was included. In order to control for the effect of the vector only, pGL4 was co-transfected with pcDNA^{TM}3.1/V5-His TOPO$^{®}$ empty vector (pcDNA). Two HIV *env* clones previously shown to have good expression using Western blot (CAP177_A3 and CAP206_E12-3) were also transfected with pGL4 Luc. The degree of transfection efficiency is indicated by relative light units (RLU) on a $\log_{10}$ scale.

Co-transfection with the pGL4 luciferase reporter gene indicated a 1000-fold increase in relative light units (RLU) compared to when HEK293T cells were mock transfected, indicating that irrespective of the construct being co-transfected, the transfection efficiency was

similar (Figure 3.7). The pcDNA^{TM}3.1/V5-His TOPO$^{®}$ empty vector and two HIV *env* clones, CAP177_A3 (T/F) and CAP206_E12-3 (chronic) that were previously shown to have high expression levels by Western blot (Figure 3.4) were used to compare the transfection efficiency of separately transfected and co-transfected pTARGET^{TM}_α4 and pCEP4_β7 plasmid DNA using PEI.

Similar levels of luciferase activity were evident however, the cell lysates were not analysed by Western blot and it is therefore not possible to confirm that this level of transfection efficiency was sufficient for the detection of CAP177_A3, CAP206_E12-3 and the α4β7 integrin. Previously, Western blots were carried out four times to determine the expression of the *env* clones and the integrin. The expression levels of the Env clones were constantly similar suggesting consistent transfection efficiencies across all experiments.

3.2.3.3 Transfection of α4β7 in CHO cells

In order to determine whether α4β7 expression was restricted in HEK293T cells, an alternative cell line was tested. CHO cells were both separately transfected and co-transfected with pTARGET^{TM}_α4 and pCEP4_β7 plasmid DNA using PEI and Electroporation. Western blots of CHO cell lysates after α4β7 PEI and Electroporation transfections did not indicate integrin expression (data not shown). Similar to what was seen with transfected HEK293T cells, the positive α4β7 recombinant protein was detected when the anti-α4 antibody (Santa Cruz Biotechnology Inc., USA) was used but not the anti-β7 antibody (Santa Cruz Biotechnology Inc., USA). The pcDNA^{TM}3.1/V5-His TOPO$^{®}$ empty vector control was negative and β-actin indicated comparable loading of samples.

In order to control for the ability to electroporate CHO cells, the CAP177_A3 and CAP206_E12-3 *env* clones were electroporated into HEK293T and CHO cells, as these clones were previously shown to have been expressed in electroporated HEK293T cells, although poorly (Figure 3.4).

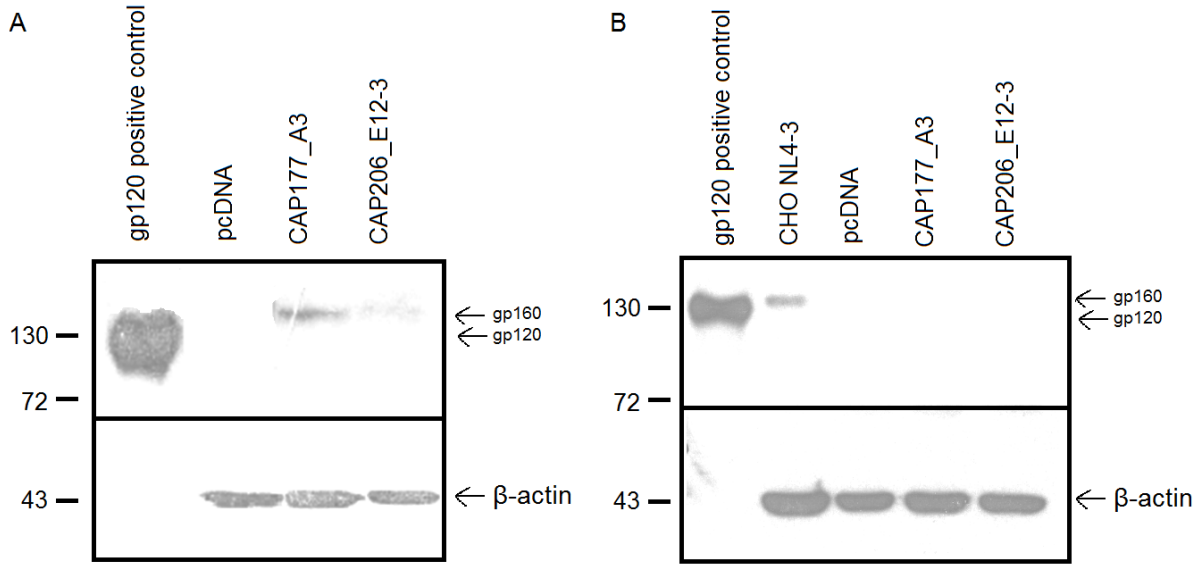


Figure 3.8 Electroporation of HEK293T but not CHO cells resulted in HIV *env* expression A) HEK293T and B) CHO cells were transfected with CAP177_A3 (T/F) and CAP206_E12-3 (chronic) *env* clones using Electroporation. A 1:3 000 dilution of the anti-gp120 antibody was used (Mouse antiserum to HIV gp120, NIH ARP). Equivalent total protein (60 µg) cell lysates was loaded per well. The pcDNA™3.1/V5-His TOPO® empty vector (pcDNA) was included to control for non-specific binding of antibodies. A gp120 soluble recombinant protein (50 ng) (120 kDa) and CHO NL4-3 gp160 (CHO NL4-3) cell lysates (only in B) were included to control for successful Western blotting and β-actin (42 kDa) was included as a loading control. The exposure time for the first Western blot was 6 minutes and the second 5 minutes. The molecular weight marker is indicated.

Western blot analysis revealed that the Electroporation of HEK293T (Figure 3.8A) but not CHO (Figure 3.8B) cells resulted in discernible cell-associated gp160 expression. CHO NL4-3 gp160 (a CHO cell line stably expressing gp160) positive control cell line indicated a single band that corresponded to gp160 as it ran higher than the recombinant gp120 positive control (Figure 3.8B). The pcDNA™3.1/V5-His TOPO® empty vector control was negative and β-actin indicated comparable loading of samples in both Western blots. This result suggests that the electroporation of CHO cells was not successful as expression of the *env* clones was observed when HEK293T cells were transfected, validating our Electroporation protocol. As CHO cells seemed to be resistant to Electroporation we did not continue with the CHO cell line.

3.2.3.4 Detection of α4β7 using flow cytometry

Flow cytometry is a more sensitive detection system than Western blotting and therefore we utilised a fluorescently labelled secondary antibody to detect Act-1 stained

pTARGET™_α4 and pCEP4_β7 PEI co-transfected HEK293T cells. Unstained cells and cells stained with secondary antibody only were used to determine background fluorescence.

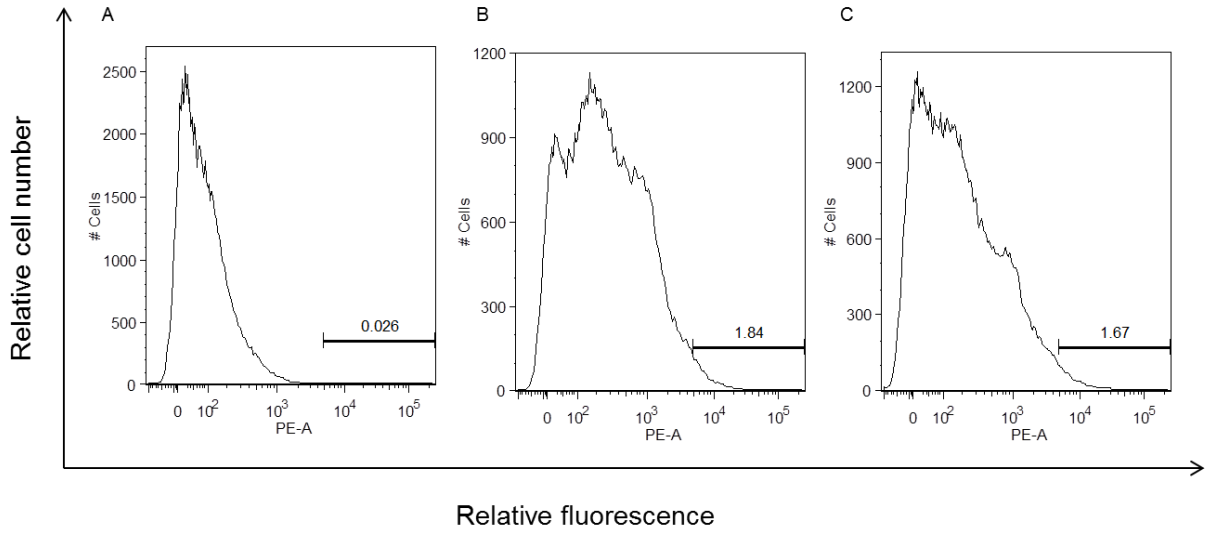


Figure 3.9 Flow cytometry did not detect α4β7 integrin on PEI co-transfected HEK293T cells. HEK293T cells were co-transfected with 1 pmol pTARGET™_α4 and pCEP4_β7 plasmid DNA and stained for flow cytometry analysis A) Unstained transfected HEK293T cells to determine HEK293T-inherent cell fluorescence B) Cells stained with Cy3® goat anti-mouse secondary antibody only to determine background fluorescence due to non-specific antibody binding C) Cells stained with Act-1 and the Cy3® goat anti-mouse antibody.

HEK293T cells co-transfected with pTARGET™_α4 and pCEP4_β7 plasmid DNA did not show α4β7 integrin expression as the fluorescent signal was no different compared to when cells were stained with secondary antibody only (Figure 3.9).

As HeLa and TZM-bl cells endogenously express α4 and β7 they were included as positive controls. Flow cytometry indicated that 57.6% and 77.4% of the TZM-bl (Figure 3.10A to Figure 3.10C) and HeLa (Figure 3.10D to Figure 3.10F) cell populations, respectively were positive for the α4β7 integrin. No Isotype control antibody was available and so cells stained with secondary antibody only were used to take into account background fluorescence due to non-specific binding of the antibody.

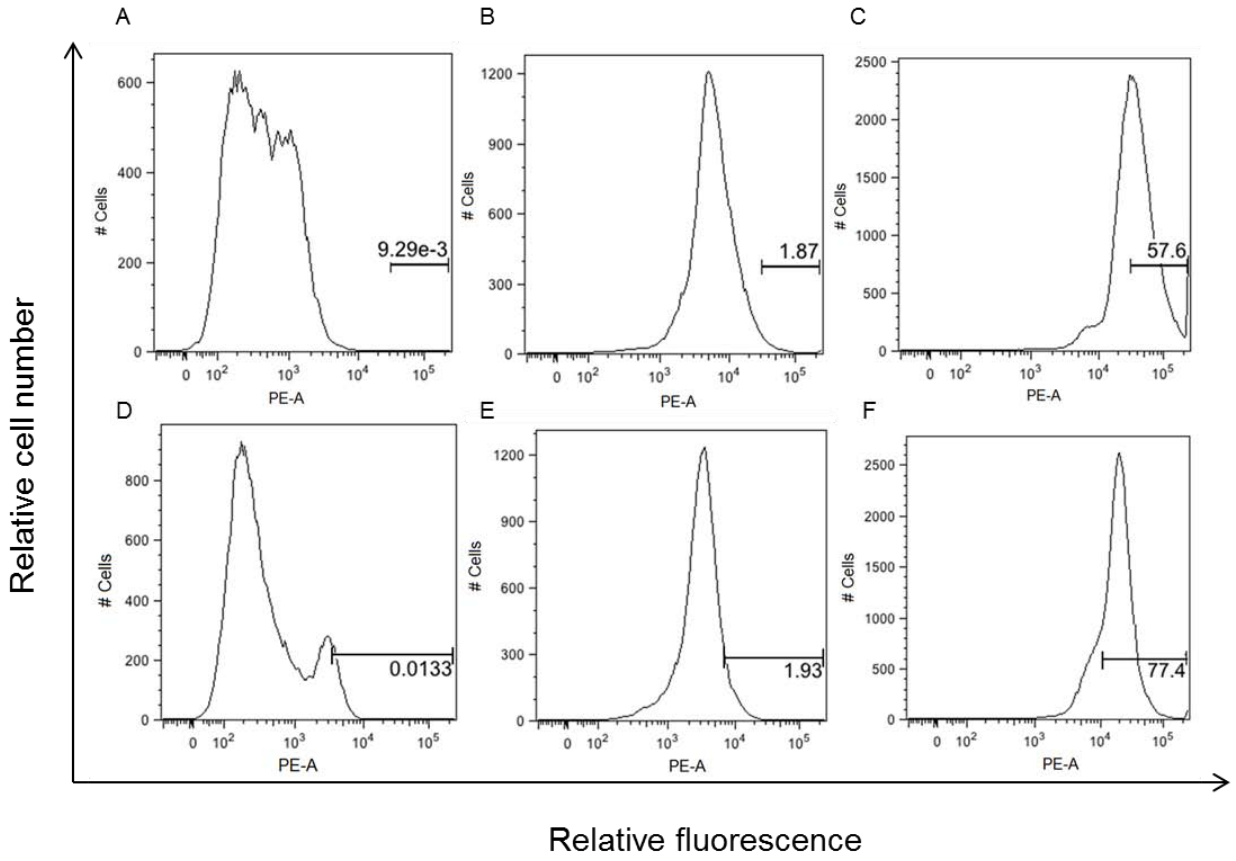


Figure 3.10 Flow cytometry detected the α4β7 integrin on TZM-bl and HeLa cells. One million TZM-bl and HeLa cells were stained for flow cytometry analysis A) Unstained TZM-bl cells to determine TZM-bl-inherent cell fluorescence B) TZM-bl cells stained with Cy3® goat ant-mouse secondary antibody only to determine background fluorescence due to non-specific antibody binding C) TZM-bl cells stained with Act-1 and Cy3® goat anti-mouse antibodies D) Unstained HeLa cells to determine HeLa-inherent cell fluorescence E) HeLa cells stained with Cy3® goat ant-mouse secondary antibody only to determine background fluorescence due to non-specific antibody binding F) HeLa cells stained with Act-1 and Cy3® goat anti-mouse antibodies.

α4β7 was detected on HeLa and TZM-bl (HeLa derived) cells by flow cytometry but not by Western blot. We concluded that either the antibodies or Western blot technique as a whole was not sensitive enough to detect low levels of integrin expression especially noting that the α4 integrin was only detected when high amounts of recombinant protein were used with the lowest recommended dilution of antibody.

As both flow cytometry and Western blot analysis could not detect α4β7 on transfected HEK293T cells, we were concerned that the inclusion of pCDM8 vector sequence in our α4 cloning strategy might have affected expression of α4 and we thus sequenced the plasmid DNA flanking the α4 cDNA.

3.2.4 Determining the 5' and 3' flanking sequences of the α4 cDNA in the pTARGET™ vector

Primers were designed to the 5' and 3' ends of the α4 cDNA insert in order to sequence the plasmid DNA flanking the α4 integrin in the pTARGET™ vector. Sequence analysis revealed that the α4 insert was inserted into the pTARGET™ vector in the incorrect orientation as the *Xho* I and *Sma* I restriction enzyme sites upstream and downstream of the pTARGET™ MCS, respectively were instead downstream and upstream of the α4 cDNA insert, respectively (Appendix A.1.1).

The T7 F primer that occurs in the pCDM8 vector sequence upstream of the pCDM8 MCS was used to PCR out the α4 cDNA from pCDM8_α4 while the T7 F primer also occurs in the pTARGET™ vector sequence upstream of the pTARGET™ MCS and was used during colony PCR to determine whether the α4 cDNA was inserted into the pTARGET™ vector in the correct orientation. Therefore the choice of the T7 F primer for both cloning and colony PCR affected the ability to differentiate whether the α4 cDNA insert was cloned into pTARGET™ in the correct orientation. As our cloning strategy resulted in the selection of a clone with the α4 cDNA in the incorrect orientation, explaining the lack of expression in transiently transfected HEK293T cells, we altered the design of our α4β7 reactivity assay to utilise HeLa cells that endogenously express the integrin as shown by flow cytometry (Figure 3.10.D to Figure 10.F).

3.2.5 α4β7 binding assay

3.2.5.1 Binding of pseudovirus to HeLa cells that endogenously express α4β7

In order to develop an assay to measure α4β7 reactivity, pseudovirus was bound to HeLa cells in the presence and absence of the Act-1 monoclonal antibody that was previously used to block gp120 binding to α4β7 (Arthos et al. 2008; Parrish et al. 2012). In order to

optimise the assay, we only tested the reactivity of two HIV Env clones CAP45_B5 (T/F) CAP177_47 (chronic). A range of Act-1 concentrations were tested to confirm that the binding of pseudovirus to HeLa cells was specifically via α4β7.

Previous DC-SIGN binding assays utilised low equivalent p24 concentrations of input pseudovirus (Alexandre et al. 2012) and so to ensure that this was not impacting the binding of pseudovirus to α4β7, binding was performed with both high (200 ng) and low (5 ng) p24 concentrations of input pseudovirus (Figure 3.11).

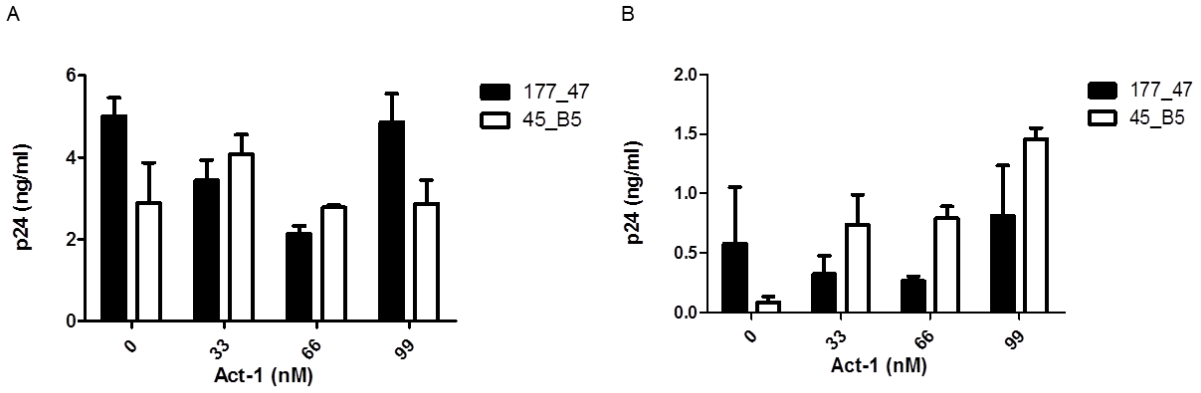


Figure 3.11 Act-1 does not consistently inhibit binding of pseudovirus to HeLa cells endogenously expressing α4β7. Pseudovirus equivalent to A) 200 ng and B) 5 ng p24 carrying Env clones CAP177_47 and CAP45_B5 was added to cells pre-incubated with 0 nM, 33 nM, 66 nM and 99 nM Act-1 monoclonal antibody. Bound pseudovirus was measured by a p24 ELISA. This data represents two independent experiments performed in triplicate.

In the absence of Act-1, 2.5% and 1.4% of virus pseudotyped with the Env clones CAP177_47 and CAP45_B5, respectively bound to HeLa cells when 200 ng p24 input pseudovirus was added (Figure 3.11A). This was lower than when 5 ng p24 input pseudovirus was added as 11% and 1.6% of pseudovirus carrying Env clones CAP177_47 and CAP45_B5, respectively bound to HeLa cells (Figure 3.11B). The overall low binding of pseudovirus to HeLa cells suggests that either the levels of α4β7 were too low and too easily saturated to allow for high numbers of pseudovirus to bind to HeLa cells or the Env clones were not binding to the integrin efficiently.

Pseudovirus carrying Env clone, CAP177_47 bound more efficiently to HeLa cells than those carrying CAP45_B5 Env clone at both concentrations of input virus, suggesting that pseudovirus binding to HeLa cells was clone specific. In fact, the binding of pseudovirus

carrying CAP45_B5 Env clone was not altered by changing the input concentration whereas that of CAP177_47 increased 4-fold. With an input of 200 ng p24 pseudovirus, CAP177_47 had lower binding in the presence of 33 nM and 66 nM Act-1 (1.7% and 1.1%, respectively) although binding in the presence of 99 nM Act-1 (2.4%) was the same as binding in the absence of Act-1. This trend was similar when the viral input was decreased to 5 ng although the experimental error was higher due to low p24 values.

CAP45_B5 binding did not follow the same trend and there was no obvious difference in binding in the absence or presence of increasing concentrations of Act-1. Therefore, there was no consistent inhibition of pseudovirus binding to HeLa cells endogenously expressing α4β7 in the presence of increasing concentrations of Act-1 and binding was never reduced by 100% even by the highest concentration of Act-1, suggesting that binding of pseudovirus to HeLa cells might not have been via α4β7.

3.2.6 Infection of TZM-bl cells in the presence and absence of the Act-1

The CD4 receptor and α4β7 are within 1.2 nm of each other as determined by co-precipitation using a cross linking reagent with a spacer arm (Cicala et al. 2009). α4β7 (22 nm) is three times larger than CD4 (7 nm) and both are located in close proximity to the CCR5 co-receptor (Cicala et al. 2009). TZM-bl is an HIV reporter cell line that is used to measure HIV infectivity and TZM-bl cells were shown to endogenously express the α4β7 integrin using flow cytometry (Figure 3.10A to Figure 3.10.C). Due to the proximity of CD4 and α4β7, incubation with Act-1 should result in the antibody binding to the integrin which should inhibit the binding of gp120 not only to α4β7 but also to CD4. This in turn should prevent viral entry measured by cell-associated luminescence, abrogating the need to use the p24 ELISA.

The entry efficiency assay, using pseudovirus to directly infect TZM-bl cells was performed in the presence and absence of saturating concentration of Act-1 (33 nM) (Parrish et al. 2012). All 11 T/F and chronic infection *env* clones were tested for entry efficiency in the presence and absence of the Act-1 monoclonal antibody (Figure 3.12).

There was a wide range of entry efficiency ranging from 1.0% to 199.3% higher than background. The entry efficiency of the CAP206 T/F clone was below the cut off value for the assay (2X higher than background) (Fouda et al. 2013), suggesting that this clone was non-functional. However, the 100 ng of pseudovirus added in this assay was likely too low to detect the entry of CAP206 T/F clone as the addition of undiluted pseudovirus resulted in RLU values only 4-fold higher than background (Z. Woodman, unpublished results). Therefore, although this Env clone (CAP206 T/F) is able to mediate infection of TZM-bl cells, higher viral inoculum was required for this assay and therefore it was not included in the analysis.

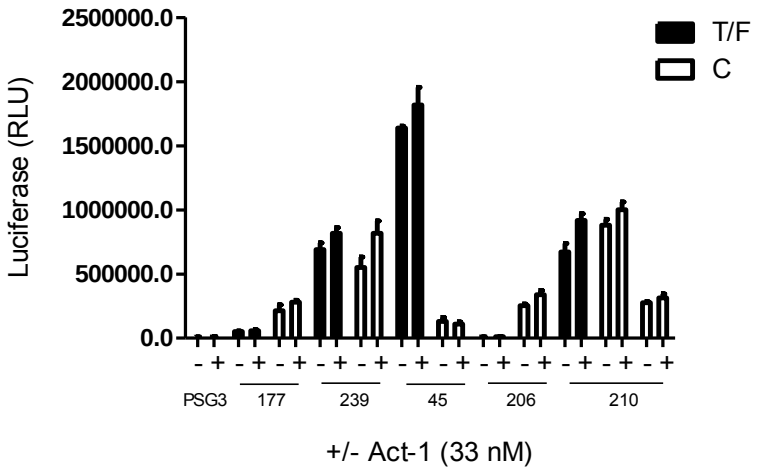


Figure 3.12 Act-1 does not inhibit infection of TZM-bl cells. The five study participants T/F and chronic infection *env* clones, CAP177 (177), CAP239 (239), CAP45 (45), CAP206 (206) and CAP210 (210) were used to generate pseudovirus to infect TZM-bl cells in the presence and absence of 33 nM Act-1 monoclonal antibody. The T/F and chronic infection Env clones of each participant are represented by black and white bars, respectively. CAP210 has one T/F *env* clone and two chronic infection *env* clones. A pSG3.1Δ*env* (pSG3) was included as a negative control as pseudovirus was generated in the absence of an *env* clone. A non-parametric Mann-Whitney statistical test was performed. This data represents an average of three independent experiments performed in triplicate.

There was no statistical difference between entry efficiency in the absence and presence of Act-1 for all nine Env clones from each of the four study participants, suggesting that pseudovirion entry into TZM-bl cells was independent of gp120-α4β7 interactions.

3.2.7 Infection of CD4+ T cells in the presence and absence of the Act-1 monoclonal antibody

Previous studies utilised CD4+ T cells isolated from PBMC to demonstrate gp120-α4β7 interactions (Nawaz et al. 2011; Cicala et al. 2009; Arthos et al. 2008; DeNucci et al. 2010). In order to determine whether the use of CD4+ T cells would allow us to detect the α4β7 reactivity of our T/F and chronic infection Env clones PBMC isolated from human blood were activated with IL-2.

CD4+ T cells incubated with retinoic acid, as used by Nawaz et al. (2011), Cicala et al. (2009), Arthos et al. (2008) and DeNucci et al. (2010), resulted in the activation of the α4β7 integrin (Iwata et al. 2004). However due to project time constraints cells were differentiated in the absence of retinoic acid and flow cytometry was not carried out to determine the level of α4β7 expression. However, although we anticipated that only a subset of cells would express the integrin (Cicala et al. 2009), we still wanted to determine whether Act-1 inhibition could distinguish Env α4β7 reactivity.

CD4+ T cells pre-incubated in the presence or absence of 33 nM Act-1 monoclonal antibody were incubated with CAP177_A3 (T/F) and CAP206_E12-3 (chronic) pseudovirus (100 ng). After washing, bound pseudovirus was measured by p24 ELISA. No binding of pseudovirus was detected suggesting that neither CD4 nor α4β7 was expressed on the surface of the differentiated PBMC and due to the lack of availability of Act-1 this experiment could not be repeated.

3.3 Discussion

We planned to test the α4β7 reactivity of five subtype C T/F and chronic infection Env clones by transiently expressing α4β7 in HEK293T cells. We decided to transiently express the integrin in a cell line as it is more cost effective and the use of CD4+ T cells from human blood introduces donor variation as seen by Cicala et al. (2009). By transiently expressing

α4β7 in a cell line, cell surface integrin expression between experiments could also be kept fairly constant.

Abitorabi et al. (1997) stably expressed α4 and β7 from pCDM8_α4 and pCEP4_β7 plasmid DNA in K562 human erythroleukemia cells using Electroporation and Qi et al. (2012) transiently expressed α4 and β7 cloned into pcDNA3.1/Hygro (-) vectors in HEK293T cells for purification purposes. This suggests that transiently transfected cell lines are a good model to yield high levels of integrin expression.

The integrins cDNA's were received in two expression cassettes, pCDM8_α4 and pCEP4_β7, and we subsequently sub-cloned the α4 cDNA from the pCDM8 vector into the pTARGET^{TM} vector. However, when we transfected HEK293T cells with the two constructs, we initially could not detect α4 nor β7 in our experimental samples nor the positive control, recombinant α4β7 protein (R and D systems, England). It was only after probing high amounts of recombinant protein with high concentrations of antibody that we could detect α4, suggesting that poor antibody binding in combination with low expression of our integrin was the problem. When we changed the detection system from Western blot to flow cytometry, we were able to detect α4β7 on HeLa and TZM-bl (HeLa derived) cells but not transfected HEK293T cells, suggesting that either our expression levels after transient transfection were below the detection limit of flow cytometry or there was no integrin expression.

Information on the plasmid DNA sequences or the cloning methodology was not available when we received the cDNA's of the integrins and we thus sequenced the inserts using vector based primers as well as internal primers designed to the integrin sequence. In order to save time, we also sub-cloned α4 cDNA into pTARGET^{TM} by amplifying the entire insert using pCDM8 primers as this did not require α4 sequence information and would guarantee the sub-cloning of all sequences required for expression as the original constructs were previously used to express α4β7 (Abitorabi et al. 1997). However, when transient expression of α4β7 in HEK293T cells was not detected using neither Western blot nor flow cytometry, we considered the possibility that our cloning methodology had affected the expression of α4.

In order to determine the sequence of the plasmid DNA flanking the α4 insert, we designed reverse and forward primers to the 5' and the 3' ends of α4, respectively to sequence the upstream and downstream regions, respectively. Sequencing indicated that the α4 cDNA was cloned in the incorrect orientation downstream of the CMV promoter, explaining the lack of α4β7 expression.

The lack of detection of β7 expression by Western blot was likely due to the poor binding of the β7 antibodies as the positive α4β7 recombinant control was never detected using either antibody. The addition of a C-terminal polyhistidine-tag (his-tag) could have aided the detection of the β7 integrin although, a his-tag has been shown to decrease the solubility of protein (Busso et al. 2003) which negatively impacts protein expression.

Due to time constraints we did not repeat the sub-cloning of α4 nor add a his-tag to β7 but instead decided to utilise HeLa cells that endogenously express α4β7 in the reactivity assay. The premise for the assay was that Env pseudotyped viruses would bind to HeLa cells with differential efficiency due to differences in α4β7 reactivity and that addition of Act-1, would distinguish α4β7-Env binding from non-specific HeLa cell-pseudovirus interactions.

Pseudovirus carrying CAP45_B5 (T/F) and CAP177_47(chronic) Env clones bound to HeLa cells but with low efficiency and saturating levels of Act-1 did not inhibit binding, suggesting that binding was non-specific. It is unlikely that poor binding was due to low levels of α4β7 as flow cytometry confirmed that 77.4% of HeLa cells carried α4β7 on the cell surface, comparable to 50% of α4β7+ CD4+ T cells demonstrated by Nawaz et al. (2011). However, the ability of the Env clones to bind to α4β7 could have been affected by the activation state of the integrin as Arthos et al. (2008) indicated that gp120 bound α4β7 in its activated form. Furthermore, pseudovirus has been shown to bind to HeLa cells non-specifically (Mondor et al. 1998) and Nawaz et al. (2011) showed that gp120's produced in HEK239T and HEK293F cells have low α4β7 reactivity. The authors suggest that cell specific post-translational differences might account for the altered reactivity of the Env clones. However, when we attempted to transiently express Env in CHO cells, we could not detect protein expression using Western blot. As we were unable to transiently transfect CHO cells using different transfection methodologies, we could not test whether pseudovirus generated in CHO cells

would have bound to HeLa cells endogenously expressing α4β7 and whether this binding was inhibited by Act-1. It is therefore possible that the HEK293T cell derived pseudovirus did not bind specifically to α4β7 but non-specifically to the HeLa cells. As the binding assay relied on a p24 ELISA with a limiting detection (0.5 ng/ml) we repeated the assay using TZM-bl cells.

Arthos et al. (2008) showed that the presence of Act-1 inhibited the binding of gp120 to α4β7 and consequently viral replication. If α4β7 was present on the surface of TZM-bl cells, the addition of Act-1 prior to adding pseudovirus should inhibit gp120-CD4 interactions and therefore viral replication. TZM-bl cells, a HeLa derived reporter gene cell line that was also confirmed to endogenously express α4β7 using flow cytometry, were infected with pseudovirus in the presence and absence of Act-1. Cell associated luciferase activity which is more sensitive than the p24 ELISA was used to quantify entry of pseudovirus into TZM-bl cells. If α4β7 (22 nm) was expressed on TZM-bl cells, Act-1 would inhibit pseudovirus binding to CD4 (7 nm) and consequently infection, which could be determined by cell associated luciferase. However, no difference in TZM-bl infection was observed in the presence and absence of Act-1.

We cannot confirm the activation state of α4β7 on the surface of HeLa and TZM-bl cells and it is therefore possible that Env pseudotyped viruses were unable to bind due to incorrect integrin conformation. As retinoic acid results in the activation of α4β7 on CD4+ T cells (Cicala et al. 2009) it is possible that HeLa cells could have been treated similarly and tested with a monoclonal antibody (J19) that specifically recognizes the activated form of α4β7 (Qi et al. 2012). However, due to time constraints, this strategy was not tested. It is also possible that the close proximity of CD4 and α4β7 might be cell-specific so that HeLa cells do not support CD4 and α4β7 complex formation. Co-immunoprecipitation studies as outlined by Cicala et al. (2009) might indicate whether α4β7 is within 1.2 nm of CD4 and CCR5 on HeLa cells. Therefore, a α4β7 binding assay utilising transfected cell lines, although more cost effective might not have been the best choice unless we could confirm the activation state of the integrin. Furthermore, production of pseudovirus in HEK293T cells most likely altered the processing and N-glycosylation of Env so that they did not bind to α4β7, supporting a previous study (Nawaz et al. 2011). Production of pseudovirus in CHO cells and

binding to α4β7+ CD4+ T cells and retinoic-treated HeLa cells will confirm whether the experimental approach of this study is still feasible.

We attempted to differentiate PBMC into CD4+ T cells that are known to express α4β7 (DeNucci et al. 2010) but pseudovirus binding was not different in the presence and absence of Act-1 as seen with HeLa and TZM-bl cells. We however did not induce α4β7 expression using retinoic acid and thus the levels of integrin might also have been too low to accurately measure pseudovirus binding. Unfortunately we were unable to confirm the phenotype of CD4+ T cells with flow cytometry due to the lack of availability of the Act-1 antibody and therefore these experiments could not be repeated.

3.4 Conclusion

As integrin expression was detected on HeLa and TZM-bl cells these cell lines were used to compare the binding and entry of pseudoviruses carrying subtype C T/F and chronic infection Env clones in the presence and absence of the Act-1 antibody. However, pseudovirus binding and entry was not inhibited by Act-1, suggesting that viral-cell interactions in this system were not mediated by α4β7. This could have been due to the activation state of the integrin and/or the N-glycosylation profile of Env produced in HEK293T cells. The production of pseudovirus in an alternative cell line other than HEK293T and testing the activation state of the integrin using monoclonal antibodies able to distinguish the two forms will allow comparisons between subtype C HIV T/F and matched chronic infection *env* clones providing insight into the role of α4β7 reactivity in HIV transmission.

Chapter 4: Comparing the role of Env N-glycans of T/F and chronic infection variants in DC-SIGN mediated *trans*-infection of CD4+ cells

4.1 Introduction

The N-glycans on gp120 assist HIV in evading antibody immune detection through masking antigenic epitopes. Furthermore, it has been suggested that the virus has evolved a mechanism whereby the N-glycans facilitate binding to dendritic cells in order to enhance HIV infection (Van`Montfort et al. 2011). Gp120 N-glycans are targets for the C-type lectin DC-SIGN expressed on dendritic cells of the genital mucosa and this lectin has been shown to facilitate HIV infection (Curtis et al. 1992; Geijtenbeek, Torensma, et al. 2000; Lee et al. 2001). Upon antigen recognition, dendritic cells migrate to the lymph nodes where they present processed antigens to naïve CD4+ T cells. Therefore, HIV, once bound to DC-SIGN is transferred to the lymph nodes where it encounters abundant CD4+ T cells, allowing rapid viral replication.

DC-SIGN preferentially binds oligomannose type N-glycans (Feinberg et al. 2007; Geijtenbeek, Torensma, et al. 2000; Mitchell et al. 2001; Van`Montfort et al. 2011), and it has recently been shown that subtype B and C T/F variants are enriched with oligomannose type N-glycans compared with viruses circulating during chronic infection (Go et al. 2011). Furthermore, earlier studies have indicated that a genetic bottleneck occurs at transmission and that the Envs of T/F variants are more compact with fewer N-glycans (Chohan et al. 2005; Derdeyn et al. 2004), suggesting that variants could be preferentially selected due to their array of N-glycans.

We hypothesise that T/F variants carry an optimum arrangement of oligomannose type N-glycans that facilitates DC-SIGN mediated *trans*-infection which then allows for successful transmission in the genital tract. Furthermore, we suggest that the accumulation of Env

PNGs over the course of infection hinders the *trans*-infection capacity of variants circulating during chronic infection.

In order to investigate whether DC-SIGN is able to mediate the *trans*-infection of subtype C HIV T/F variants better than viruses circulating during chronic infection, a DC-SIGN mediated *trans*-infection assay was optimized for pseudovirus carrying the Env clones of all five study participants. The *env* clones were generated from SGA-derived amplicons by L. Shuping. The T/F clones were carefully selected based on identity to consensus sequence and the chronic infection clones were selected based on consensus N-glycosylation patterns (Table 4.1).

Table 4.1 Characteristics of the Env clones

Participant ID (CAP)	Clone ID	Stage of infection	Sampling (Weeks post infection)	*Functionality (-fold entry efficiency)	N386	N392
177	A3	acute	2	*35	+	-
	47	chronic	172	19	+	+
239	Ell	acute	5	50	+	+
	T35	chronic	173	5	+	+
45	B5	acute	2	10	+	-
	H5	chronic	108	*7	+	-
206	H1	acute	4	*10	+	-
	E12-3	chronic	173	35	-	+
210	E8	acute	5	120	+	+
	E1	chronic	80	128	+	+
	C7	chronic	80	37	+	+

*+: PNG present, -: no PNG present at positions N386 and N392
*Functionality of the *env* clones was determined by infecting TZM-bl cells with 100 ng pseudovirus normalised on p24 concentration except for CAP177_A3, CAP206_H1 and CAP45_H5 clones where the virus was not diluted. The -fold increase in entry efficiency above background was 2.6-, 1.2- and 1.1-fold, respectively for these clones when 100 ng of pseudovirus was used to infect TZM-bl cells (Adapted from L. Shuping, unpublished data).

The entry efficiency of the Env clones varied from 1.1- to 128-fold higher than background when the equivalent of 100 ng p24 pseudovirus was used to infect TZM-bl cells. The entry efficiency of 100 ng p24 CAP177_A3, CAP45_H5 and CAP206_HI clones was below cut off for functionality for this assay (2X higher than background) (Fouda et al 2010) but when undiluted virus was tested, these Env clones were 35-, 7- and 10-fold higher than background, respectively indicating that they are functional but have poor entry efficiency.

The sequences flanking the Env clones with low entry efficiency were analysed for correct *rev*, *vpu* and *env* open reading frames, confirming that *rev-vpu* fusion was not the reason for reduced Env function (Kraus et al. 2010). As the CAP177_A3 *env* clone sequence was identical to the consensus of 20 SGA-derived sequences and CAP177_A3 was infected with a single variant, this clone should represent the T/F and therefore its low entry efficiency should be sufficient for productive HIV infection. However, as CAP206_H1 and CAP45_H5 clones had one and two amino acid differences, respectively compared to consensus, it is possible that these mismatches altered Env structure and function.

Go et al. (2011) and Go et al. (2013) used mass spectrometry to identify two PNGs, N386 and N392 (numbered according to HXB2), that are believed to carry oligomannose N-glycans. These sites not only differed between T/F and chronic infection variants with the latter carrying processed carbohydrates but have also been shown to be involved in DC-SIGN binding (Hong et al. 2007). We therefore hypothesised that these sites might be essential for DC-SIGN mediated *trans*-infection and might comprise a motif involved in the selection of specific variants for transmission. In order to investigate the role of the oligomannose type N-glycans at PNGs N386 and N392 in subtype C DC-SIGN mediated *trans*-infection of HIV, we generated N-glycan mutants of both T/F and chronic infection variants and compared their *trans*-infection capacity.

4.2 Results

4.2.1 DC-SIGN mediated transfer of T/F and chronic infection Env clones to CD4+ cells

A Burkitt's lymphoma cell line engineered to stably express DC-SIGN, Raji-DC-SIGN, was used for the *trans*-infection assay (Nabatov et al. 2006; Van`Montfort et al. 2007; Balzarini et al. 2007; Liao et al. 2011; Eggink et al. 2010; Alexandre et al. 2012).

Flow cytometry using a DC-SIGN specific antibody indicated that 24.8% of Raji-DC-SIGN cells expressed DC-SIGN (Figure 4.1D to Figure 4.1F) confirming the presence of the DC-SIGN receptor whereas Raji-WT cells did not express the lectin (Figure 4.1A to Figure 4.1C).

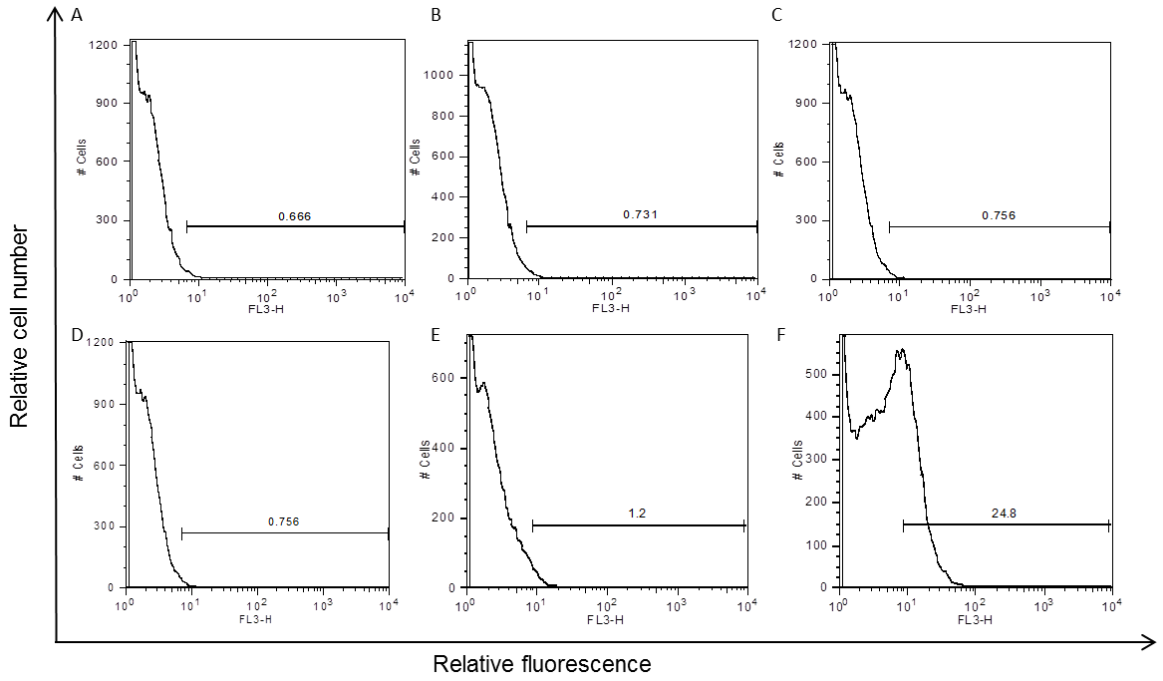


Figure 4.1 Flow cytometry confirmed DC-SIGN expression on Raji-DC-SIGN cells. One million Raji-DC-SIGN and Raji-WT cells were stained for flow cytometry analysis A) Unstained Raji-WT cells to determine Raji-WT-inherent cell fluorescence B) A stained Isotype control (PerCP-Cy™5.5 Mouse IgG2b, κ) of Raji-WT cells to determine background fluorescence due to non-specific antibody binding C) Raji-WT cells stained cells with PerCP-Cy™5.5 anti-human CD209 (DC-SIGN) antibody D) Unstained Raji-DC-SIGN cells to determine Raji- DC-SIGN-inherent cell fluorescence E) A stained Isotype control (PerCP-Cy™5.5 Mouse IgG2b, κ) of Raji-DC-SIGN cells to determine background fluorescence due to non-specific antibody binding F) Raji-DC-SIGN cells stained cells with PerCP-Cy™5.5 anti-human CD209 (DC-SIGN) antibody.

Pseudovirus, normalised on p24 was added to Raji-DC-SIGN cells before being added to TZM-bl cells. The level of TZM-bl cell-associated luciferase activity was an indicator of how well the Raji-DC-SIGN cells were able to mediate *trans*-infection of the CCR5+ CXCR4+ CD4+ cell line. We hypothesised that pseudoviruses carrying the Env of T/F variants would be transferred more efficiently to CD4+ cells compared with those carrying the Env of chronic infection viruses if T/F variants are dependent on DC-SIGN mediated *trans*-infection for successful transmission in the genital tract.

Raji-WT cells were unable to transfer pseudovirus to TZM-bl cells, indicating that DC-SIGN is required for *trans*-infection of CD4+ cells (data not shown). When *trans*-infection data was not normalised on entry efficiency, two participant's (CAP45 and CAP210) T/F Env clones showed enhanced *trans*-infection compared with the chronic infection Env clones, two participant's (CAP177 and CAP206) chronic infection Env clones showed enhanced *trans*-

infection compared with the T/F Env clones and one participant (CAP239) showed no difference in *trans*-infection between the T/F and chronic infection Env clones (Figure 4.2A).

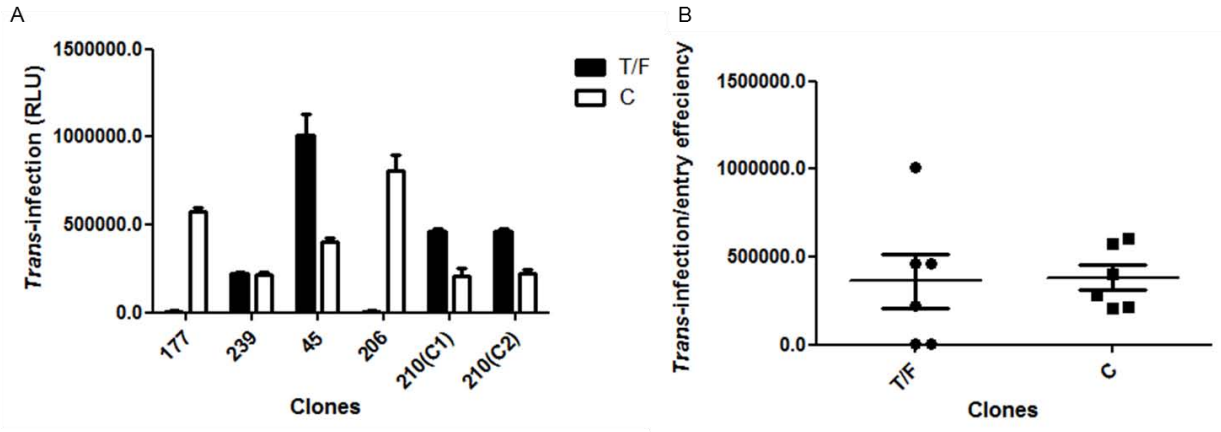


Figure 4.2 There was no difference between the *trans*-infection of pseudovirus carrying the Env of T/F and chronic infection variants. T/F and chronic infection (C) Env clones from five subtype C study participants were tested for DC-SIGN mediated transfer to CD4+ cells. Participants CAP177, CAP239, CAP45, CAP206 and CAP210 are indicated by 177, 239, 45, 206 and 210, respectively. Two chronic infection *env* clones were used for CAP210, indicated by 210(C1) and 210(C2) A) DC-SIGN mediated t*rans*-infection data was not normalised to the entry efficiency of each clone and *trans*-infection data is represented in relative light units (RLU). This data represents one of two independent experiments performed in triplicate B) DC-SIGN mediated *trans*-infection data was normalised to the entry efficiency of each clone. Two independent experiments performed in triplicate were pooled and compared using a Mann Whitney statistical T test (p=0.8099). Horizontal lines represent the medians. Data is represented on a $\log_{10}$ scale.

The *trans*-infection assay involves the transfer of pseudovirus from Raji-DC-SIGN cells to TZM-bl cells which are then infected. Therefore, the entry efficiency of the Env clones might skew the *trans*-infection assay if pseudovirus infected TZM-bl cells with varied efficiencies. In order to take into account the potential influence that Env entry efficiency might have on the *trans*-infection assay, the *trans*-infection data was normalised to the entry efficiency of the Env clones. When normalised *trans*-infection data of the Env clones of the T/F variants were compared to variants found later in infection, there was no significant difference between the two medians (p=0.8099) (Figure 4.2B).

4.2.2 Site-directed mutagenesis of the T/F and chronic infection CAP239 *env* clones

SDM was performed on the T/F (CAP239_E11) and chronic infection (CAP239_T35) *env* clones of participant CAP239 by mutating the asparagine (N) to a glutamine (Q) at positions N386 and N392 thereby preventing N-glycosylation at these sites. T/F and chronic infection

CAP239 *env* clones were mutated at both PNGs to generate single and double *env* N-glycan mutants (Figure 4.3).

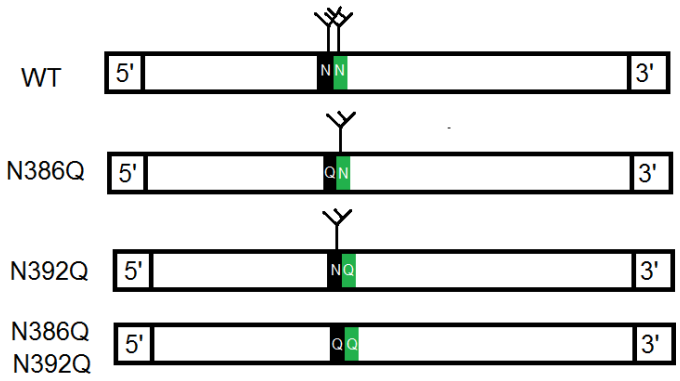


Figure 4.3 A graphical representation of the N-glycan mutants compared with WT. The WT CAP239 T/F and chronic infection *env* clones underwent mutagenesis at two PNGs, N386 and N392 (according to HXB2 numbering) to generate single (N386Q and N392Q) and double (N386Q N392Q) N-glycan mutants. The black rectangle represents position 386 and the green rectangle represents position 392. An asparagine (N) to glutamine (Q) mutation resulted in the disruption of the PNG NXT/S sequon and thus prevented the attachment of an N-glycan (Y).

Our laboratory compared the N-glycosylation of Env clones using Endoglycosidase H (Endo H) to remove oligomannose type N-glycans and SDS-PAGE. Preliminary data obtained by L. Shuping showed that partially purified CAP239 WT T/F gp120 protein treated with Endo H migrated further on an SDS-PAGE compared with the CAP239 WT chronic infection clone treated similarly (Lin et al. 2003). This suggested that the T/F clone carried more oligomannose type N-glycans than the chronic infection clone as Endo H treatment resulted in a greater shift in molecular weight. CAP239 *env* clones were selected to generate the Env N-glycan mutants and these were tested for DC-SIGN mediated *trans*-infection and compared with the WT Env clones. We hypothesised that if the oligomannose residues at positions N386 and N392 were essential for DC-SIGN binding then deletion of the sequons would reduce DC-SIGN mediated *trans*-infection of CD4+ cells.

SDM primers were designed, synthesised and used to generate the respective single and double *env* N-glycan mutants. The sequences of the T/F and chronic infection CAP239 *env* clones were identical on either side of both PNGs so a single primer pair was used for each site to generate both T/F and chronic infection N-glycan mutants. Primers were designed to insert the recognition site for either the *BSaW* I or *Acl* I restriction enzyme by silent mutation alongside N386Q and N392Q, respectively.

WT

N386Q

N392Q

N386Q,N392Q

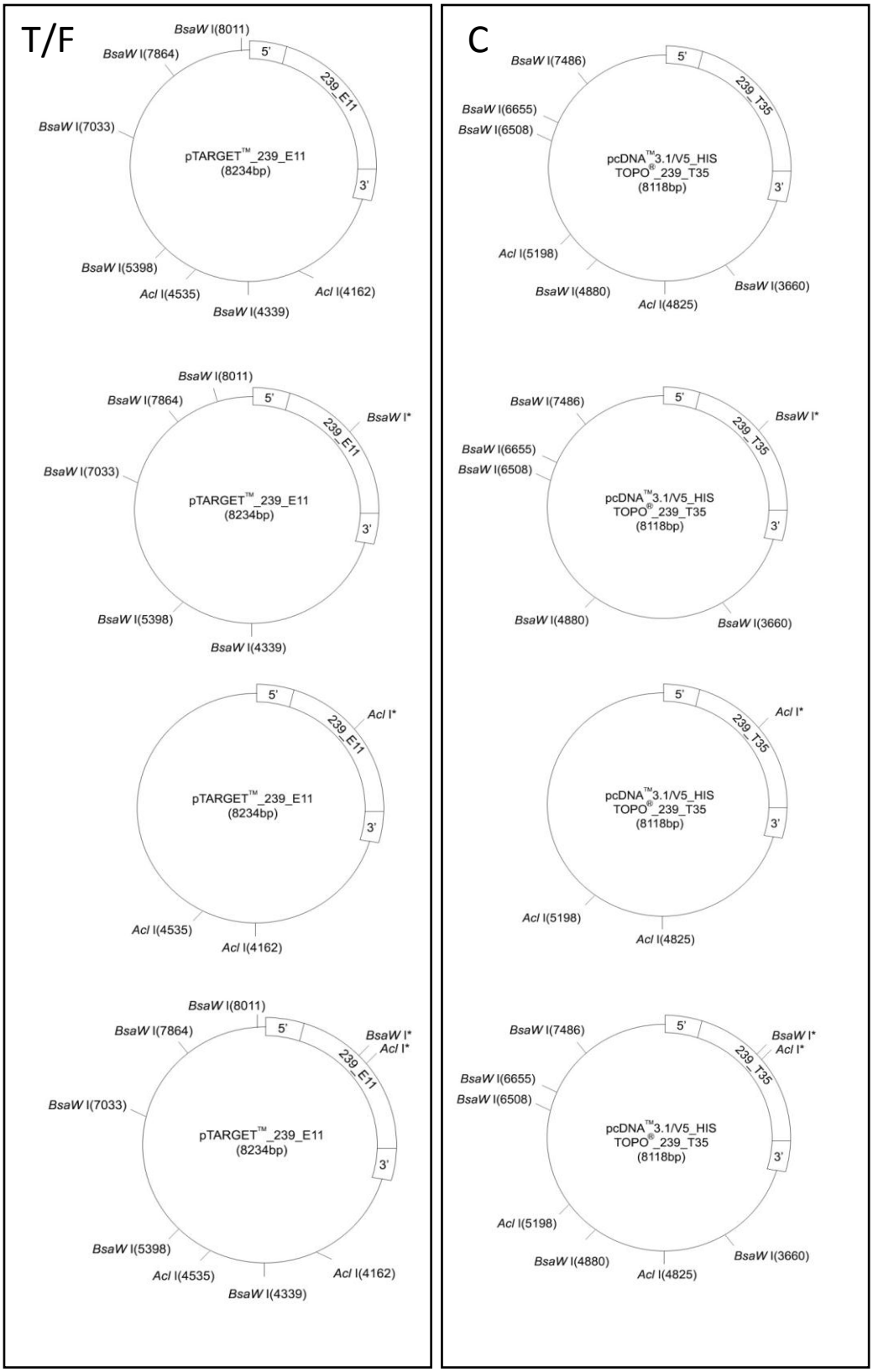


Figure 4.4 DNA plasmid vector maps of wild-type, single and double N-glycan mutants. The restriction enzyme sites followed with a * indicates that the site was inserted by mutagenesis at position 1 423 for *BsaW* I and 1 433 for *Acl* I. No * indicates that the site was present in the wild-type (WT) sequence. Both *envs* are flanked by 5' and 3' sequences. The pTARGET™_239_E11 (T/F) and pcDNA™3.1/V5_HIS TOPO®_239_T35 (chronic) *env* clones were inserted into mammalian expression vectors pTARGET™ and pcDNA™3.1/V5_HIS TOPO®, respectively.

Restriction enzyme digestions were used to screen for the presence of the mutations and were visualised by agarose gel electrophoresis. The single mutants (N386Q and N392Q) were screened with their respective enzyme and the double mutants (N386Q,N392Q) with both restriction enzymes.

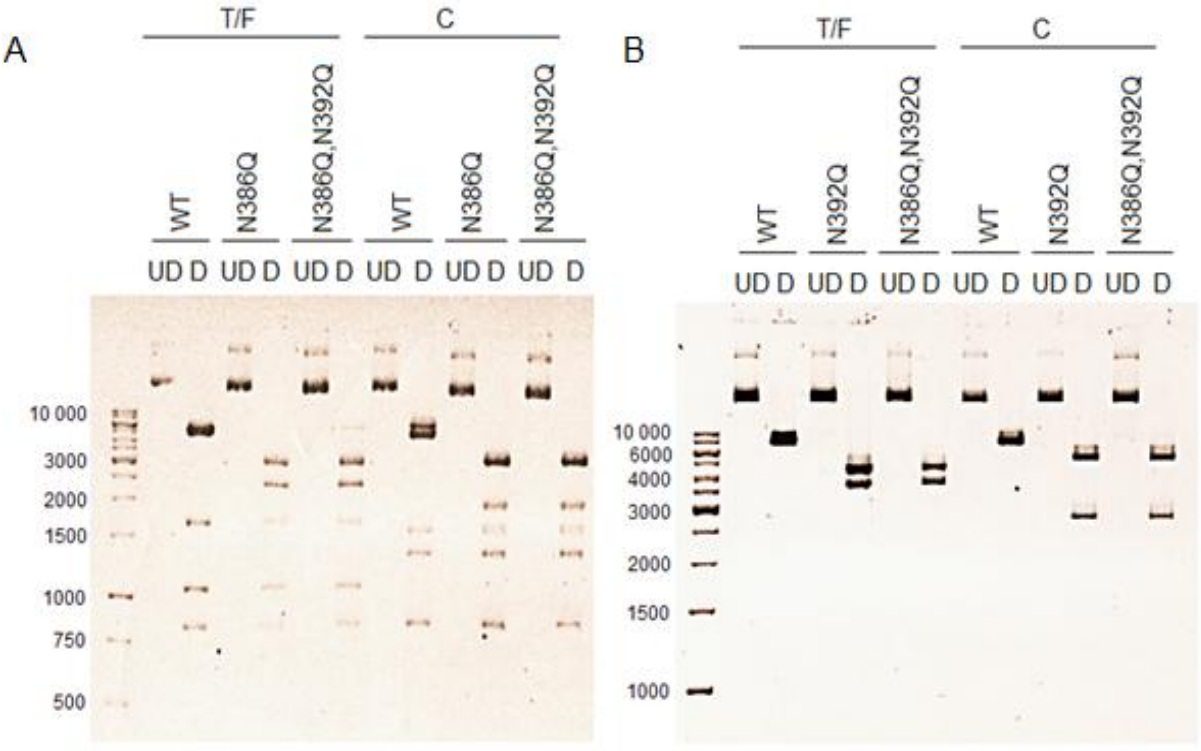


Figure 4.5 Restriction enzyme digestion of wild-type *env* to and the N-glycan mutants to confirm site-directed mutagenesis A) T/F and chronic infection (C) CAP239 wild-type (WT) *env* clones, single N386Q and double N386Q,N392Q N-glycan mutants either undigested (UD) or digested (D) with the *BsaW* I restriction enzyme and B) T/F and chronic infection (C) CAP239 WT *env* clones, single N392Q and double N386Q,N392Q N-glycan mutants either undigested (UD) or digested (D) with the *AcI* I restriction enzyme.

A restriction enzyme digestion of the CAP239 WT T/F *env* clone with *BsaW* I generated DNA fragments with the molecular weight of 4 562 bp, 1 635 bp, 1 059 bp, 831bp and 147 bp. The deletion of the PNG at position N386 resulted in the loss of the 4 562 bp band with the concomitant appearance of a 3 078 bp and a 2 462 bp band instead of the expected 2 462 bp and 2 100 bp bands (Figure 4.4). The appearance of the 3 078 bp band could be due to incomplete digestion at sites *BsaW* I (8 011) and *BsaW* I (7 864) resulting in a larger fragment comprising of the 831 bp, 147 bp and 2 100 bp fragments (Figure 4.5).

A restriction enzyme digestion of the CAP239 WT T/F *env* clone with *Acl* I generated DNA fragments with the molecular weight of 7 861 bp and 373 bp. The deletion of the PNG at position N392 resulted in the loss of the 7 861 bp band with the concomitant appearance of a 4 600 bp and a 3 800 bp band (Figure 4.5) instead of the expected 5 132 bp and 2 729 bp bands (Figure 4.4). However, incomplete digestion could not explain the difference between the expected and observed DNA banding patterns.

A restriction enzyme digestion of the CAP239 WT chronic infection *env* clone with *BsaW* I generated DNA fragments with the molecular weight of 4 292 bp, 1 628 bp, 1 220 bp, 831 bp and 147 bp. The deletion of the PNG at position N386 resulted in the loss of the 4 292 bp band with the concomitant appearance of a 2 702 bp and 1 600 bp band (Figure 4.5) instead of the expected 2 237 bp and 2 055 bp bands (Figure 4.4). Once again incomplete digestion could not explain the difference between expected and observed DNA banding patterns.

A restriction enzyme digestion of the CAP239 WT chronic infection *env* clone with *Acl* I generated DNA fragments with the molecular weight of 7 745 bp and 373 bp. The deletion of the PNG at position N392 resulted in the loss of the 7 745 bp band with the concomitant appearance of a 5 145 bp and 2 600 bp band (Figure 4.5), confirming that the observed banding pattern matched that of the expected band sizes (Figure 4.4).

The digestion of the CAP239 WT T/F and chronic infection double mutants with both *BsaW* I and *Acl* I generated the same banding patterns as their respective single mutants (Figure 4.5) but as restriction enzyme digestion did not confirm that SDM was successful for CAP239 WT T/F *env* clone and in order to confirm that spurious mutations were not introduced during PCR, all the putative PNG mutants were sequenced from the start to the stop codon of Env using a set of nine primers (Appendix A.2). The raw sequencing data was analysed using ChromasPro version 1.5 (*http://www.technelysium.com.au/ChromasPro.html*) and aligned to WT sequences using BioEdit Sequence Alignment Editor version 7.0.9 (Hall 1999) and the Clustal*W* multiple sequence alignment algorithm (Larkin et al. 2007).

Sequence analysis confirmed that the glutamines at positions N386 and N392 were successfully introduced together with the silent restriction enzyme sites (Figure 4.6A).

Sequencing confirmed that mutagenesis did not result in the introduction of any spurious PCR-derived mutations as mutants were identical to the WT clonal sequences except where the intentional mutations were introduced. Therefore, the N-glycan Env mutants only differed from WT due to the absence of a single or double PNG (Figure 4.6B).

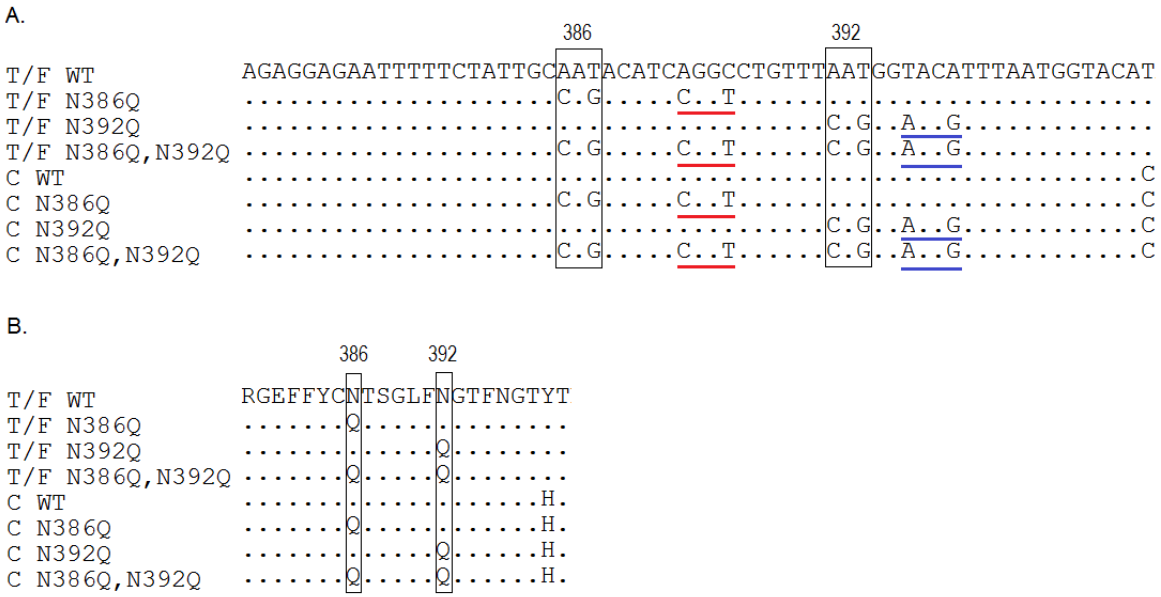


Figure 4.6 Sequence alignment of the N-glycan *env* mutants relative to wild-type confirm deletion of PNGs A) DNA sequences of wild-type (WT) CAP239_E11 T/F and WT CAP239_T35 chronic infection (C) *env* clones were obtained from P. Moore, (NICD, South Africa) and were aligned to N-glycan mutant sequences using Clustal*W*. Identical nucleotides are indicated by dots. The codons for N386 and N392 are outlined by a rectangular box. AAT codes for an asparagine and is replaced by a CAG which codes for a glutamine in the respective mutants. The *BSa*W I restriction enzyme site (CGGT) inserted in the N386Q mutants is underlined in red and the *Acl* I restriction enzyme site (AACG) in the N392Q mutants are indicated in blue. Both double mutants (T/F N386Q,N392Q and C N386Q,N392Q) have both asparagine (AAT) to glutamine (CAG) mutations and carry both restriction enzyme sites B) Protein sequence alignment of the N-glycan mutants compared with WT *env* clones. The asparagine (N) to glutamine (Q) amino acid changes are outlined by a rectangular box at positions N386 and N392.

4.2.3 Entry efficiency of N-glycan mutants

The deletion of one or two N-glycans on gp120 can impact the overall conformation of Env thereby affecting its function (Eggink et al. 2010; Fenouillet & Jones 1995; Fenouillet et al. 1989; Land et al. 2003; Montefiori et al. 1988; Pal et al. 1989) and the entry efficiency of the virus (Liao et al. 2011). In order to ensure that the loss of either PNG had no effect on the entry efficiency of the N-glycan mutants relative to WT, the entry efficiency of the WT clones and the N-glycan mutants were compared. Pseudovirus, normalised on p24 was used

to infect TZM-bl cells and the level of TZM-bl cell-associated luciferase activity indicated entry efficiency of each virus (Figure 4.7).

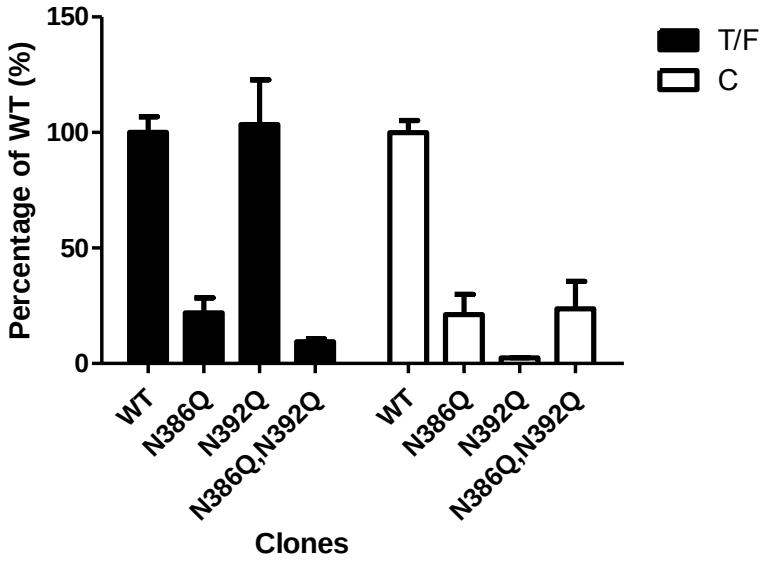


Figure 4.7 The entry efficiency of N-glycan mutants differ when compared with wild-type. Single (N386Q and N392Q) and double (N386Q,N392Q) T/F and chronic infection (C) CAP239 N-glycan mutants were tested for entry efficiency. The level of entry efficiency was determined as a percentage of wild-type (WT) and error bars represent standard deviation. This data represents an average of two independent experiments performed in triplicate.

The T/F N392Q mutant was the only N-glycan mutant that maintained similar entry efficiency to that of its respective WT Env clone. The other N-glycan mutants demonstrated a significant reduction in entry efficiency compared with WT (range: 75%-98% reduction) (Figure 4.7). Of the five N-glycan mutants with reduced entry efficiency compared with WT, one (chronic infection N392Q) was termed non-functional with this amount of input virus as RLU readings were only 1.5-fold higher than background (results not shown). The other four N-glycan mutants had RLU readings 14- to 156-fold higher than background (results not shown) and were thus above the 2-fold cut-off (Fouda et al. 2013).

4.2.4 *Trans*-infection of N-glycan mutants

In order to determine whether the loss of PNGs at positions N386 and N392 impacts the *trans*-infection of Env clones of T/F and chronic infection variants, pseudovirus carrying the N-glycan Env mutants were bound to Raji-DC-SIGN cells and then added to TZM-bl cells and luminescence, as a marker for infection, was measured (Figure 4.8A). As a result of the

reduction in the entry efficiency of the N-glycan Env mutants, the DC-SIGN mediated *trans*-infection assay data was also normalised on the entry efficiency of the Env clones (Figure 4.8B).

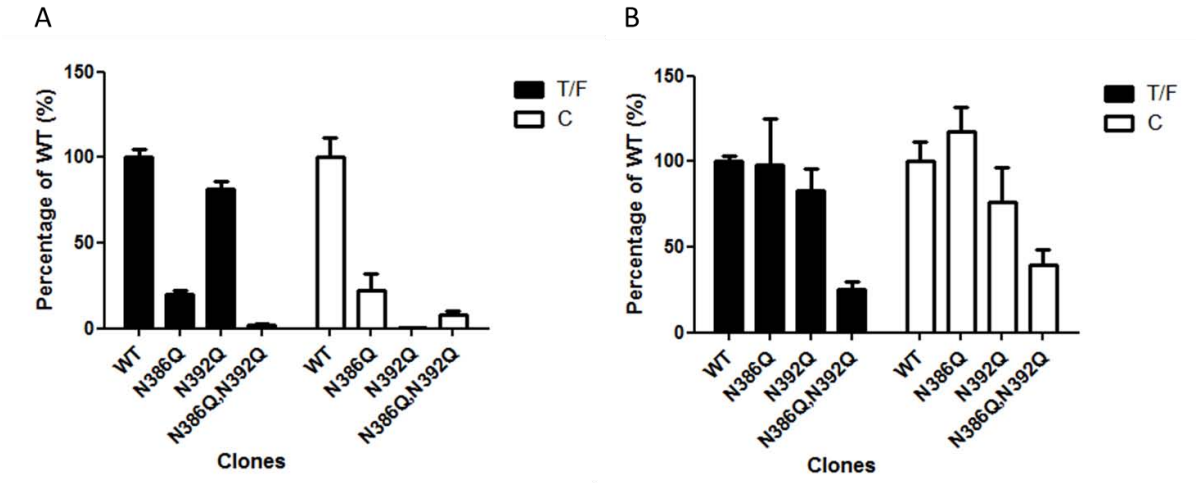


Figure 4.8 Deletion of PNGs at N386 and N392 reduced DC-SIGN mediated *trans*-infection when compared with WT. Single (N386Q and N392Q) and double (N386Q,N392Q) T/F and chronic infection (C) CAP239 N-glycan mutants were tested for DC-SIGN mediated *trans*-infection. The level of *trans*-infection was determined as a percentage of WT and error bars indicate standard deviation A) *Trans*-infection data was not normalised on the entry efficiency of each mutant B) *Trans*-infection data was normalised on the entry efficiency of each mutant. This data represents an average of two independent experiments performed in triplicate.

When the cell-associated luciferase activity values were not normalised on entry efficiency, all the T/F and chronic infection single and double N-glycan mutants demonstrated reductions in HIV transfer from DC-SIGN expressing cells to CD4+ cells (Figure 4.8A). The chronic infection N392 N-glycan mutant had the lowest *trans*-infection values (1-fold above Raji-WT control cells), indicating that this clone was unlikely transferred to CD4+ T cells as it was non-functional.

When cell-associated luciferase activity values were normalised on entry efficiency, the DC-SIGN mediated *trans*-infection of the T/F and chronic infection single N-glycan mutants (N386Q and N392Q) were no longer reduced while the double N-glycan mutants (N386Q,N392Q) remained unchanged, suggesting that the removal of both PNGs abrogated transfer to CD4+ cells. The chronic infection N392 N-glycan mutant had an apparent increase in *trans*-infection when the data was normalised on entry efficiency similar to WT (Figure 4.8B), but this was likely due to the poor entry efficiency of this clone inflating the result.

4.2.5 DC-SIGN binding of N-glycan mutants

Virus that binds DC-SIGN more efficiently is expected to be transferred more effectively to CD4+ cells (Parrish et al. 2013) and so a DC-SIGN binding assay was performed. Pseudovirus carrying the N-glycan mutants and their corresponding WT clones, normalised on p24 concentration was added to Raji cells expressing DC-SIGN and the amount of bound virus was measured by a p24 ELISA.

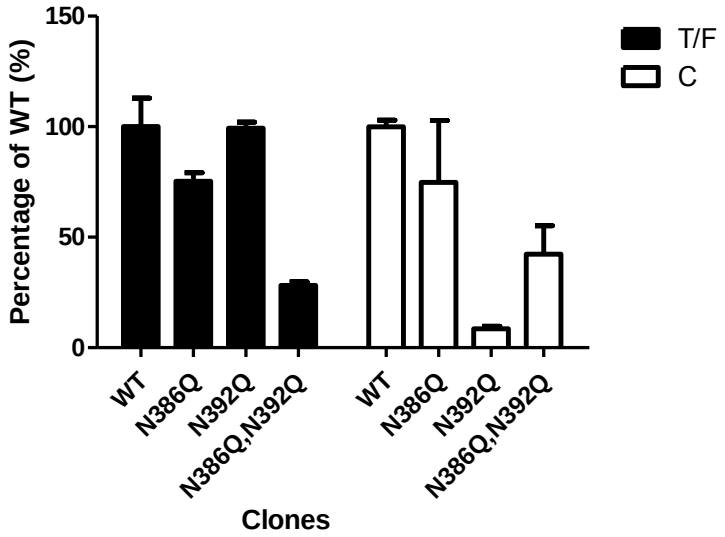


Figure 4.9 Some of the N-glycan mutants have reduced binding to DC-SIGN compared with wild-type A) Single (N386Q and N392Q) and double (N386Q,N392Q) T/F and chronic infection (C) CAP239 mutants were tested for DC-SIGN binding using Raji-wild-type (WT) and Raji-DC-SIGN cells. The level of DC-SIGN binding of the N-glycan mutants was determined as a percentage of WT and error bars represent standard deviation. This data represents an average of two independent experiments performed in triplicate.

Deletion of the PNG at position N392 of the chronic infection Env clone resulted in the greatest reduction in DC-SIGN binding, confirming that the loss of the PNG most likely altered Env conformation. Therefore, taken together this data suggests that removal of the PNG at position N392 rendered the chronic infection CAP239 N-glycan mutant non-functional. The T/F and chronic infection double N-glycan mutants (N386Q,N392Q) demonstrated reductions in DC-SIGN binding compared with WT while the single N-glycan mutants demonstrated no differences (Figure 4.9). These results supported the normalised *trans*-infection data (Figure 4.8 B).

4.3 Discussion

It is well established that binding to DC-SIGN is mediated by oligomannose type N-glycans on the surface of gp120 (Geijtenbeek et al. 2002; Geijtenbeek, Kwon, et al. 2000) and it was recently shown that subtype B and C T/F viruses are enriched with oligomannose type N-glycans (Go et al. 2011). Therefore, dendritic cells expressing DC-SIGN could bind HIV within the genital tract and migrate to the lymph nodes where HIV could be transferred to CD4+ T cells via *trans*-infection (Curtis et al. 1992; Geijtenbeek, Torensma, et al. 2000; Lee et al. 2001). Therefore, DC-SIGN may bind oligomannose type N-glycans on the surface of T/F viruses in the genital mucosa and facilitate the selective transmission, dissemination and expansion of these variants.

In order to determine whether pseudovirus carrying the Env of T/F variants are transferred to CD4+ cells more efficiently than those carrying Env clones from chronic infection variants, five subtype C T/F and chronic infection matched *env* clones were compared in *trans*-infection and DC-SIGN binding assays.

Wilen et al. (2011) demonstrated that there was no difference in DC-SIGN mediated *trans*-infection of 24 subtype B T/F and chronic infection variants while Parrish et al. (2013) demonstrated that subtype B (but not C) T/F variants displayed greater DC-SIGN mediated *trans*-infection compared with chronic infection variants. King et al. (2013) demonstrated that chronic infection variants displayed greater DC-SIGN mediated *trans*-infection than T/F variants whereas Borggren et al. (2008) showed that chronic infection variants displayed greater DC-SIGN mediated *trans*-infection than end-stage (AIDS) HIV variants.

The conflicting results thus far could be due to differences in study design with some groups comparing IMC (Parrish et al. 2013), and others pseudovirus (Wilen et al. 2011; Borggren et al. 2008), or generating pseudovirus from PBMC (Borggren et al. 2008) or engineered cell lines (Wilen et al. 2011), or using DC-SIGN expressing primary cells (King et al. 2013) or stably transfected DC-SIGN cell lines (Borggren et al. 2008) for the *trans*-infection assay. Furthermore, these groups did not compare T/F and chronic infection variants from the

same participant and chronic infection ranged from 1 to 15 years (King et al. 2013; Wilen et al. 2011).

We hypothesised that if T/F variants are dependent on DC-SIGN mediated *trans*-infection for successful transmission in the genital tract, they would be more easily transferred to CD4+ cells than the chronic infection variants. Raji-DC-SIGN cells used in the DC-SIGN mediated *trans*-infection assay were confirmed to express the DC-SIGN lectin (24%). Raji-WT control cells were shown not to express DC-SIGN and did not bind pseudovirus nor transfer pseudovirus to CD4+ cells compared with Raji cells expressing DC-SIGN, confirming the role of the lectin in gp120 binding and virus transfer.

When the *trans*-infection of the Env clones were compared there was no trend between pseudoviruses carrying T/F and chronic infection Env clones, suggesting that T/F variants would not be preferentially transferred to CD4+ T cells compared to those found later in infection (Wilen et al. 2011).

It is proposed that DC-SIGN mediated *trans*-infection is a two-step process in which DC-SIGN first binds to N-glycans on the surface of HIV, followed by the subsequent transfer of HIV to CD4+ T cells (Parrish et al. 2013). Assays used to measure *trans*-infection rely on the detection of infection of a reporter cell line, which can mask DC-SIGN reactivity, the actual role of DC-SIGN in the process. In order to investigate the efficacy of DC-SIGN reactivity of different Env clones many studies normalise *trans*-infection data on viral infectivity (Wilen et al. 2011; Borggren et al. 2008) or add equivalent TCID50 (median tissue culture infective dose) viral titres to dendritic cells (King et al. 2013) in order to take into account the impact that differences in Env entry efficiency could have on *trans*-infection results. Some studies however prefer to take into account the role that Env entry efficiency would have on DC-SIGN mediated *trans*-infection as it more likely represents what happens *in vivo* (Parrish et al. 2013; Alexandre et al. 2012). We did not initially normalise our data on entry efficiency as we wanted to determine whether the poor *trans*-infection of certain clones could be due to poor infection of TZM-bl cells.

Previously in our laboratory it was shown that CAP177_A3, CAP45_H5 and CAP206_H1 had very low entry efficiency. As only pseudovirus equivalent to 100 ng p24 was tested for *trans*-infection and unbound virus was removed before mixing with TZM-bl cells, the low *trans*-infection of CAP177_A3 and CAP206_H1 is likely due to very low entry efficiency. However, CAP45_H5 Env clone had good *trans*-infection efficiency, suggesting that this clone might have enhanced reactivity with DC-SIGN that may rescue the low entry efficiency of this clone. Therefore, poor entry efficiency did not always translate into poor *trans*-infection. When *trans*-infection values were normalised on entry efficiency the T/F clones still did not have enhanced transfer compared to those from chronic infection, confirming that the transmitted founder variants did not have better DC-SIGN reactivity.

One drawback to this study was the small sample size as this study focussed on only five participants. Furthermore, given the high genetic diversity at chronic infection, the selection of a single Env clone as a representative of the quasispecies has its limitations. Wilen et al. (2011) selected chronic infection clones randomly while Parrish et al. (2013) looked for clusters of recent clonal expansions identify an Env clone representative of chronic infection. Parker et al. (2012) selected certain chronic infection clones randomly and others from clusters of recent clonal expansions. Borggren et al. (2008) used seven matched participants similar to our five matched T/F and chronic infection participants. It is thus possible that wider sampling with an increased number of clones from chronic infection might reveal differences between T/F and chronic infection variants. However, although the sample size of this study was limited, selection of the clones was thorough and this approach is important to validate findings of larger, less detailed cohorts.

In this study the chronic infection clones of each T/F-chronic infection pair were chosen based on identical N-glycan profiles compared with the consensus sequence, except for one of the chronic infection clones of CAP210. The two chronic infection CAP210 *env* clones performed similarly despite the one clone, CAP210_E1 having consensus N-glycosylation pattern with two amino acid mismatches and the other, CAP210_C7 differing by nine amino acids and one PNG compared to the consensus. This suggests that a single Env clone with an N-glycosylation profile identical to consensus might be sufficient when comparing DC-SIGN

binding and *trans*-infection although participants infected with viruses with higher diversity, additional clones might need to be tested.

Hong et al. (2007) identified PNGs at positions N295, N386 and N392 of subtype B Envs that were involved in binding to DC-SIGN. As N295 was only found in the T/F clone of one study participant, CAP239, it is unlikely to be important in subtype C Env-DC-SIGN interactions. However, one or both PNGs at N386 and N392 were present in all five study participants suggesting that they might play an important role in subtype C gp120-DC-SIGN binding. Furthermore, recently, Go et al. (2011) determined that these sites carried oligomannose N-glycans, carbohydrates shown to bind DC-SIGN (Geijtenbeek, Kwon, et al. 2000; Geijtenbeek, Torensma, et al. 2000). We therefore hypothesised that either one or both of these sites might play an important role in binding to DC-SIGN on dendritic cells. SDM was performed on the CAP239 T/F and chronic infection *env* clones to replace the NXS/T sequon with a QXS/T motif, thus preventing N-glycosylation at these sites. Sequence analysis indicated that the N-glycan Env mutants were identical to WT except for the asparagine to glutamine amino acid change.

N-glycans are large, flexible molecules that project from the surface of the protein and move around in the extracellular space, fixed only at the point of attachment (an asparagine residue). An N-glycan adopts a final position based on steric, hydrophobic and hydrophilic interactions between neighbouring N-glycans and amino acid side chains (Petrescu et al. 2004). The posttranslational removal of Env N-glycosylation by enzymatic treatment with glycosidases has shown no effect on the overall structure and function of Env (Fenouillet et al. 1990; Depetris et al. 2012) while the removal of Env N-glycosylation prior to protein processing has been shown to affect Env structure and function (Eggink et al. 2010; Fenouillet et al. 1989; Fenouillet & Jones 1995; Montefiori et al. 1988; Land et al. 2003; Pal et al. 1989).

The entry efficiency of each N-glycan mutant was compared with its respective WT clone to determine whether the removal of one or both PNGs at position N386 and N392 affected the structure and function of the Env clone and therefore the infectivity of the corresponding pseudovirus. Hong et al. (2007) and van Montfort et al. (2011) showed no

differences in entry efficiency between WT and N-glycan mutants whereas Liao et al. (2011) showed that two N-glycan mutants had enhanced entry efficiency and one had reduced entry efficiency compared with WT. This study demonstrated a reduction in entry efficiency for all N-glycan mutants compared with WT, except for the T/F N392Q N-glycan mutant. While removal of the PNG at position N392 of the chronic infection clone resulted in loss of Env function, the entry efficiency of the T/F N392Q N-glycan mutant was not affected, suggesting that the effect of alterations in N-glycosylation are clone specific. Our results suggest that for 5/6 N-glycan mutants the structure and function of gp120 was altered when PNGs were deleted at positions N386 and N392. It has been shown that the N-glycan at position N386 is not essential for gp120 protein folding (Sanders et al. 2008) and therefore we anticipated that the removal of this PNG would not affect the entry efficiency of the pseudovirus. However, as the effect of PNG removal seems to be Env clone specific, functional dependence on certain N-glycans is likely due to overall differences in Env structure.

To compare the N-glycan mutants to one another, the level of *trans*-infection was calculated relative to that of the WT Env clones. The loss of PNGs at positions N386 and N392 resulted in a reduction of *trans*-infection of both Env clones with the most marked reduction occurring when both sites were deleted. This suggests that these PNGs are important for DC-SIGN mediated *trans*-infection. In order to distinguish the extent to which DC-SIGN reactivity and Env entry efficiency might contribute to *trans*-infection, we normalised DC-SIGN mediated *trans*-infection of Env clones with Env entry efficiency. Once the entry efficiency of the Env clones was taken into account, only the double mutants of both clones had reduced *trans*-infection, suggesting that the presence of one of these sites were required for interaction with DC-SIGN similar to WT levels. This result was corroborated when pseudovirus was tested for binding to DC-SIGN as only the double mutants had reduced binding compared to WT. The binding of the chronic infection N-glycan N392Q mutant to DC-SIGN although greatly reduced compared to WT, suggested that DC-SIGN might still bind to non-functional envelopes.

Even though the loss of PNGs at either one or both sites lowered or abrogated the entry efficiency of the Env clones, the pseudovirus carrying these clones could still interact with

DC-SIGN and could be transferred to CD4+ cells if functional. In the case of functional single N-glycan mutants this suggested that although the structure of Env could no longer mediate entry as efficiently, this did not affect the ability of DC-SIGN to bind to gp120 and possibly transfer the virus to TZM-bl cells.

Liao et al. (2011) observed that 5/7 single N-glycan mutants had similar DC-SIGN mediated *trans*-infection compared with WT while all three double N-glycan mutants and one triple N-glycan mutant showed a reduction in DC-SIGN mediated *trans*-infection compared with WT. Similarly, when both PNGs at N386 and N392 were removed, there was a reduction in *trans*-infection of the T/F and chronic infection mutants compared with their respective WT clones. This suggested that the presence of one of these sites could compensate for the absence of the other but when both are removed, *trans*-infection is reduced.

The removal of a single N-glycan may be partially or completely masked by the presence of adjacent N-glycans while the removal of two or more N-glycans may produce more noticeable effects. The N-glycans at positions N386 and N392 are extremely close to one another and consequently the removal of one may be masked by the presence of the remaining so that differences in DC-SIGN mediated *trans*-infection are not detected. The reduction in DC-SIGN mediated *trans*-infection observed when both PNGs were removed suggests that our CAP239 WT clones carry N-glycans at positions N386 and N392 and are not only essential for Env function but also important for DC-SIGN binding. This suggestion is corroborated by the fact that all of the four other Env clones in this study have either one or both of these PNGs and Go et al. (2011) found that the PNGs at positions N386 and N392 were conserved in the T/F *env* clones but not the chronic infection clones. Furthermore, when we analysed 103 unrelated subtype C sequences, the PNGs at positions N386 and N392 were conserved suggesting that these sites provide a replication or survival advantage.

As it was determined that the PNGs at positions N386 and N392 carried oligomannose type N-glycans (Go et al. 2011) when pseudovirus was produced in HEK293T cells and we demonstrated that the deletion of both of these sites lowered DC-SIGN mediated *trans*-infection when pseudovirus was generated in HEK293T cells then it is possible that

oligomannose type N-glycans could play an important role in HIV transmission. However Go et al. (2013) demonstrated that pseudovirus generated in CHO cells carry oligomannose N-glycans at positions N386 and N392 while pseudovirus generated in HEK293T cells carry a combination of oligomannose and complex type N-glycans. Furthermore, DC-SIGN was shown to bind to complex N-glycans (Hong et al. 2007; Liao et al. 2011) in addition to oligomannose type N-glycans (Van`Montfort et al. 2011; Geijtenbeek, Kwon, et al. 2000; Geijtenbeek, Torensma, et al. 2000) and the PNG at position N386 contrary to Go et al. (2011) and consistent with Go et al. (2013) was demonstrated to carry a complex N-glycan (Liao et al. 2011). We can therefore not confidently assume, without mass spectrometry analysis, that our CAP239 WT T/F and chronic infection *env* clones carry oligomannose type N-glycans at N386 and N392 and therefore the reduction in DC-SIGN mediated *trans*-infection was due to the removal of these types of N-glycans.

Efficient binding between DC-SIGN and gp120 N-glycans should translate into enhanced DC-SIGN mediated *trans*-infection of CD4+ cells (Borggren et al. 2008; Hong et al. 2007; Parrish et al. 2013). However van Montfort et al. (2011) and Alexandre et al. (2012) demonstrated that enhanced DC-SIGN binding does not correlate with enhanced DC-SIGN mediated *trans*-infection of CD4+ cells. DC-SIGN binding of our N-glycan *env* mutants were tested and increased DC-SIGN mediated *trans*-infection normalised for Env entry efficiency was associated with increased DC-SIGN binding for all T/F and chronic infection N-glycan mutants with the exception of the chronic infection N392Q N-glycan mutant. The low entry efficiency of this mutant due to its lack of function inflated the normalised *trans*-infection data. The lack of correlation between DC-SIGN binding and *trans*-infection may be due to certain groups not normalising their *trans*-infection data (Alexandre et al. 2012) as Parrish et al. (2013) observed a correlation after normalisation although, care must be taken for clones with very low entry efficiency.

Du et al. (2012) recently demonstrated that DC-SIGN binding and DC-SIGN mediated *trans*-infection are directly correlated when using Raji-DC-SIGN expressing cells but not directly correlated when using immature monocyte-derived dendritic cells. Although DC-SIGN is the principle receptor involved in DC-SIGN mediated *trans*-infection of CD4+ cells (Alexandre et al. 2011; Balzarini et al. 2007), a number of complex interactions involving other receptor

ligand pairs and/or signalling molecules and pathways are likely to collectively orchestrate DC-SIGN mediated *trans*-infection of CD4+ T cells *in vivo* (Alexandre et al. 2011; Van`Montfort et al. 2011).

Our assay utilised Raji cells expressing DC-SIGN as surrogate dendritic cells and TZM-bl cells expressing CD4, CCR5 and CXCR4 as surrogate CD4+ T cells. As we are only taking into account these specific receptors our assay may be missing as yet undiscovered key elements that are involved in DC-SIGN mediated *trans*-infection of CD4+ T cells *in vivo*. The direct correlation observed between DC-binding and DC-SIGN *trans*-infection (with exception of the chronic N392Q N-glycan mutant) is consistent with what Du et al. (2012) observed while using genetically engineered DC-SIGN expressing cells.

Furthermore, Raji cells are a Burkitt's lymphoma derived B cell line and a number of human B cell lines have shown to facilitate HIV transmission to CD4+ T cells (Bobardt et al. 2003; Jakubik et al. 2000; Moir et al. 2000). Although the use of Raji-DC-SIGN cells in DC-SIGN binding and DC-SIGN mediated *trans*-infection assays as an alternative for monocyte-derived dendritic cells is well established (Geijtenbeek, Kwon, et al. 2000; Geijtenbeek, Torensma, et al. 2000; Geijtenbeek et al. 2002), the use of Raji cells may need to be re-evaluated with respect to their B cell lineage.

Upon N-glycan analysis of the CAP239 T/F and chronic infection *env* clones an equal number of PNGs were identified on both gp120 sequences however L. Shuping showed that the CAP239 WT T/F *env* clone treated with Endo H migrated further on a SDS-PAGE compared with the CAP239 WT chronic infection clone (unpublished data), suggesting that the T/F clone carries more oligomannose N-glycans than the chronic infection clone despite the equal number of PNGs. The reason there is no observable difference in DC-SIGN mediated *trans*-infection of T/F and chronic infection Env clones after the removal of two N-glycan sites comprising of the putative DC-SIGN binding site could be due to the fact that all chronic infection variants do not automatically lose the ability to be transferred to CD4+ cells by DC-SIGN. Evolution over the course of infection most likely involves the random introduction of PNGs, some of which will abrogate DC-SIGN mediated *trans*-infection while others will have no impact.

Increasing the sample size and sampling at later time points might allow for the detection of variants that have DC-SIGN reactivity due to the accumulation of PNGS. This study has provided a detailed analysis of the *trans*-infection capacity of five paired subtype C Env clones and it has indicated that PNGs at positions N386 and N392 might be important for DC-SIGN interactions and further sampling might detect differences between T/F and chronic infection variants.

4.4 Conclusion

We hypothesised that T/F variants carry an optimum arrangement of oligomannose type N-glycans if they are dependent on DC-SIGN mediated *trans*-infection for successful transmission in the genital tract. Five subtype C T/F and matched chronic infection Env clones demonstrated no difference in DC-SIGN mediated *trans*-infection of CD4+ cells consistent with previous findings. In order to determine whether PNGs at positions N386 and N392 reported to carry oligomannose N-glycans influenced the DC-SIGN binding and DC-SIGN *trans*-infection of CD4+ cells, N-glycan mutants were generated. These mutations affected the ability of Env to mediate infection of TZM-bl cells, suggesting they play an important role in protein structure and function although the extent differed between clones. No difference in *trans*-infection was observed when only one of the PNGs was removed but the removal of both PNGs resulted in a reduction in DC-SIGN mediated *trans*-infection when normalised for Env entry efficiency. The presence of one of the PNGs compensated for the absence of the other but when both were removed DC-SIGN mediated *trans*-infection was reduced. DC-SIGN binding correlated with DC-SIGN mediated *trans*-infection with the exception of the single chronic infection N392Q N-glycan mutant that was non-functional. It can be concluded that the removal of both N-glycans at position N386 and N392 reduced DC-SIGN mediated *trans*-infection of CAP239 T/F and chronic infection *env* clones.

Chapter 5: Conclusion

As T/F variants have distinct N-glycosylation patterns compared with chronic infection variants, this study suggested that T/F variants carry an optimum arrangement of N-glycans that enable interactions with α4β7 and DC-SIGN that are essential for successful transmission in the genital tract. The first objective of this study was to develop a reactivity assay to determine whether T/F *env* clones have a higher affinity for the α4β7 integrin than chronic infection controls. Due to experimental design error we utilised HeLa cells that endogenously express the α4β7 integrin to compare the binding and entry of pseudoviruses in the presence and absence of the Act-1 monoclonal antibody. Pseudovirus-HeLa cell interactions were not inhibited by Act-1, suggesting that α4β7 was not responsible for mediating these interactions, possibly due to the activation state of the integrin and/or the N-glycosylation profile of the Env clones. As we were not able to optimise the α4β7 reactivity assay, we were not able to compare the binding of pseudovirus carrying T/F Env clones to that of the matched chronic infection controls and determine the role of Env N-glycosylation in this process.

The second objective was to determine whether T/F variants are preferentially transferred to CD4+ cells via the DC-SIGN CLR. There was no difference in DC-SIGN mediated *trans*-infection of CD4+ cells of subtype C T/F and matched chronic infection *env* clones, although the small sample size could have influenced the outcome of the study.

The third and final objective of this study was to determine whether the removal of PNGs, previously shown to carry oligomannose type N-glycans that interact with DC-SIGN affect the *trans*-infection of CD4+ cells. When N-glycan mutants were generated by deleting PNGs at positions N386 and N392, the entry efficiency and DC-SIGN binding of the mutants decreased in an Env clone specific manner, suggesting that N-glycans at the same site impact Env structure and function differently. When *trans*-infection data was normalised to entry efficiency there was no difference in *trans*-infection of the N-glycan mutants when only one of the PNGs was removed but the removal of both PNGs resulted in a reduction in

DC-SIGN mediated *trans*-infection. This would suggest that N-glycosylation at one site can compensate for the absence at another. However, as the removal of certain N-glycans affected Env entry efficiency, the presence and absence of PNGs could impact viral transmissibility as certain variants may not be able to replicate as well as others after transfer to CD4+ T cells.

Future work could involve testing whether retinoic acid activates HeLa cell integrins and whether producing pseudovirus in CHO cells would be more beneficial than producing pseudovirus in HEK293T cells when testing α4β7 reactivity. The sample size will be expanded to include more chronic infection control clones in order to ensure sampling of all variants at this timepoint. The PNGs at positions N386 and N392 will also be deleted in additional Env clones and alternative N-glycans will be removed via site directed mutagenesis to determine whether these impact DC-SIGN interactions similar to that of the CAP239 Env.

Appendix

A.1. Sequencing pTARGET™_α4 and pCEP4_β7

A.1.1 pTARGET™_α4

The pTARGET™_α4 plasmid DNA was sequenced to confirm the presence of the entire α4 cDNA inserted into the pTARGET™ vector in the correct orrientaton.

Sma I RE site T7 F primer ATG start codon of α4 cDNA

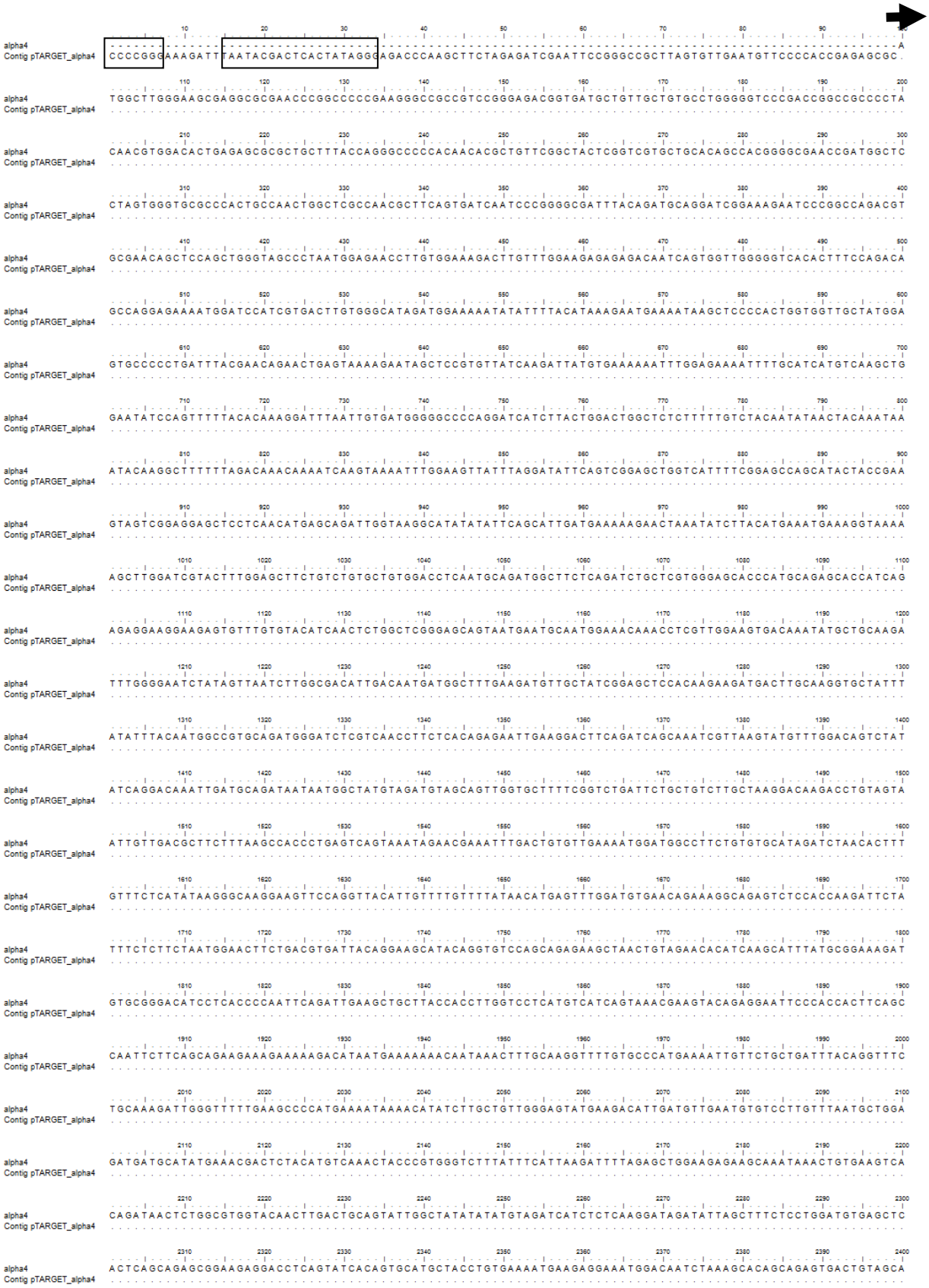


Figure A.1.1. Nucleotide sequence alignment demonstrates that the α4 cDNA was inserted into the pTARGET^{TM} vector in the incorrect orrentation. The α4 cDNA nucleotide sequence (alpha4) (accesion number NP_034706) was aligned to a contig sequence (contig pTARGET_alpha4) generated by sequencing the pTARGET^{TM}_α4 construct. The *Sma* I restriction enzyme site (RE), the T7 forward (T7 F) primer, the ATG start codon, the *Xho* I RE site and the TAA stop codon is indicated. Dots indicate 100% nucleotide matches.

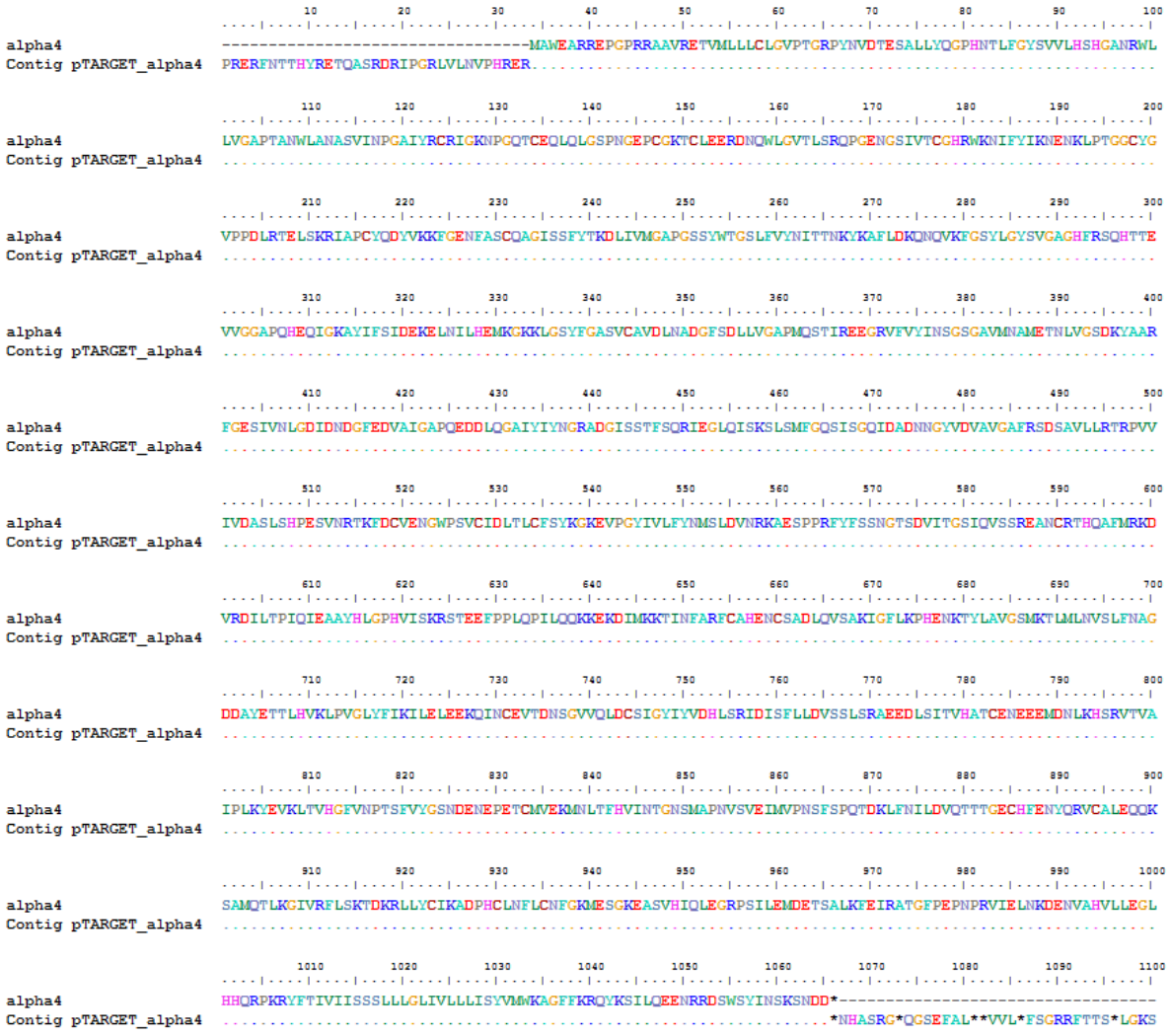


Figure A.1.2. Protein sequence alignment confirms the α4 cDNA was inserted into the pTARGET™ vector in the incorrect orientation. The α4 cDNA amino acid sequence (alpha4) (accesion number NP_034706) was aligned to a contig sequence (contig pTARGET_alpha4) generated by sequencing the pTARGET™_α4 construct. A methionine (M) indicates the start codon and a * indicates the stop codon. Dots indicate 100% amino acid matches.

A.1.2 pCEP4_β7

The pCEP4_β7 plasmid DNA was sequenced to confirm the presence of the entire β7 cDNA in the pCEP4 vector.

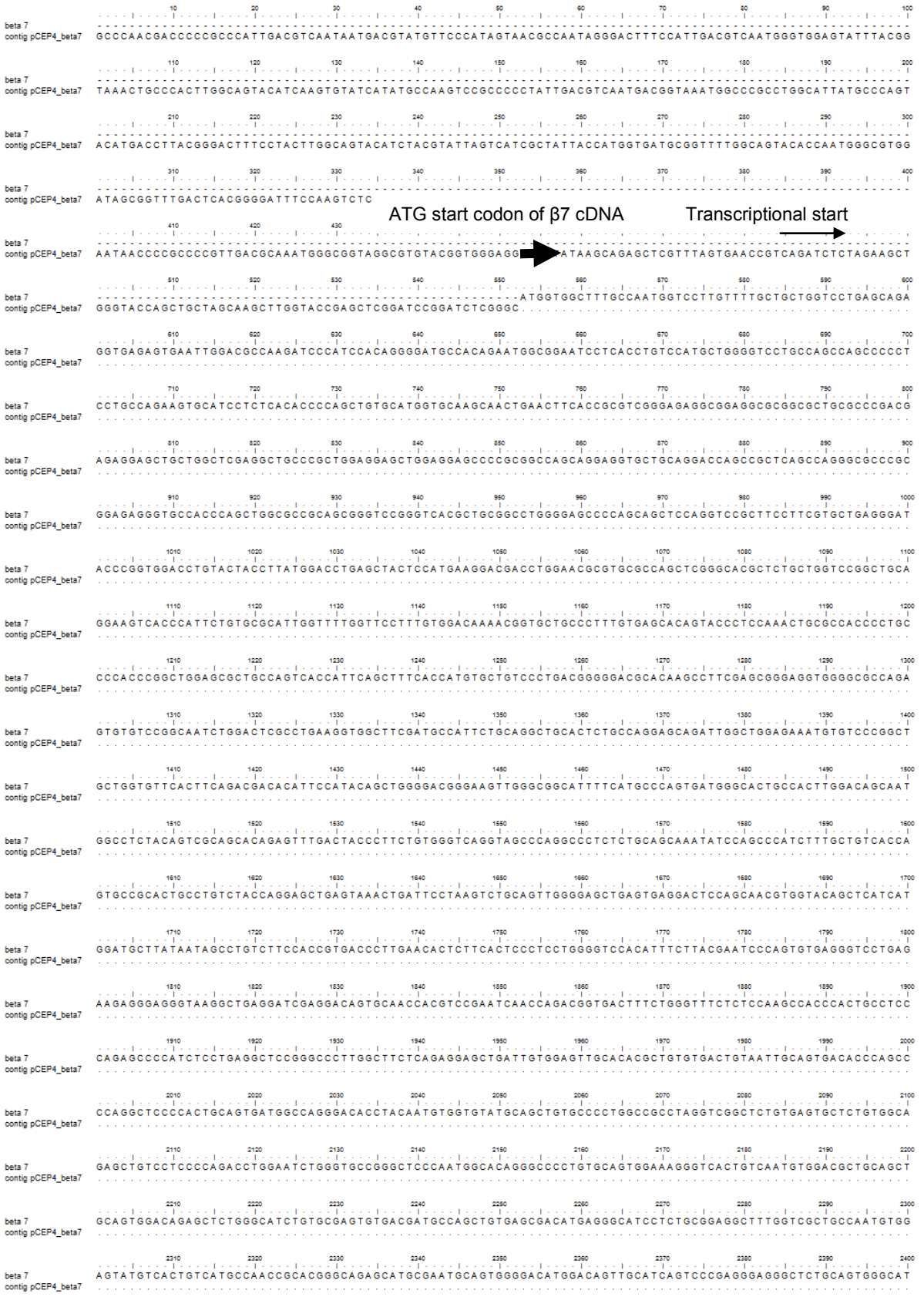

ATG start codon of β7 cDNA
Transcriptional start
beta 7
contig pCEP4_beta7

Figure A.1.3. Nucleotide sequence alignment confirms complete β7 cDNA in pCEP4 vector. The β7 cDNA nucleotide sequence (beta7) (accesion number NP_000880) was aligned to a contig sequence (contig pCEP4_beta7) generated by sequencing the pCEP4_β7 construct. The transcriptional start, the ATG start codon and the TGA stop codon is indicated. Dots indicate 100% nucleotide matches.

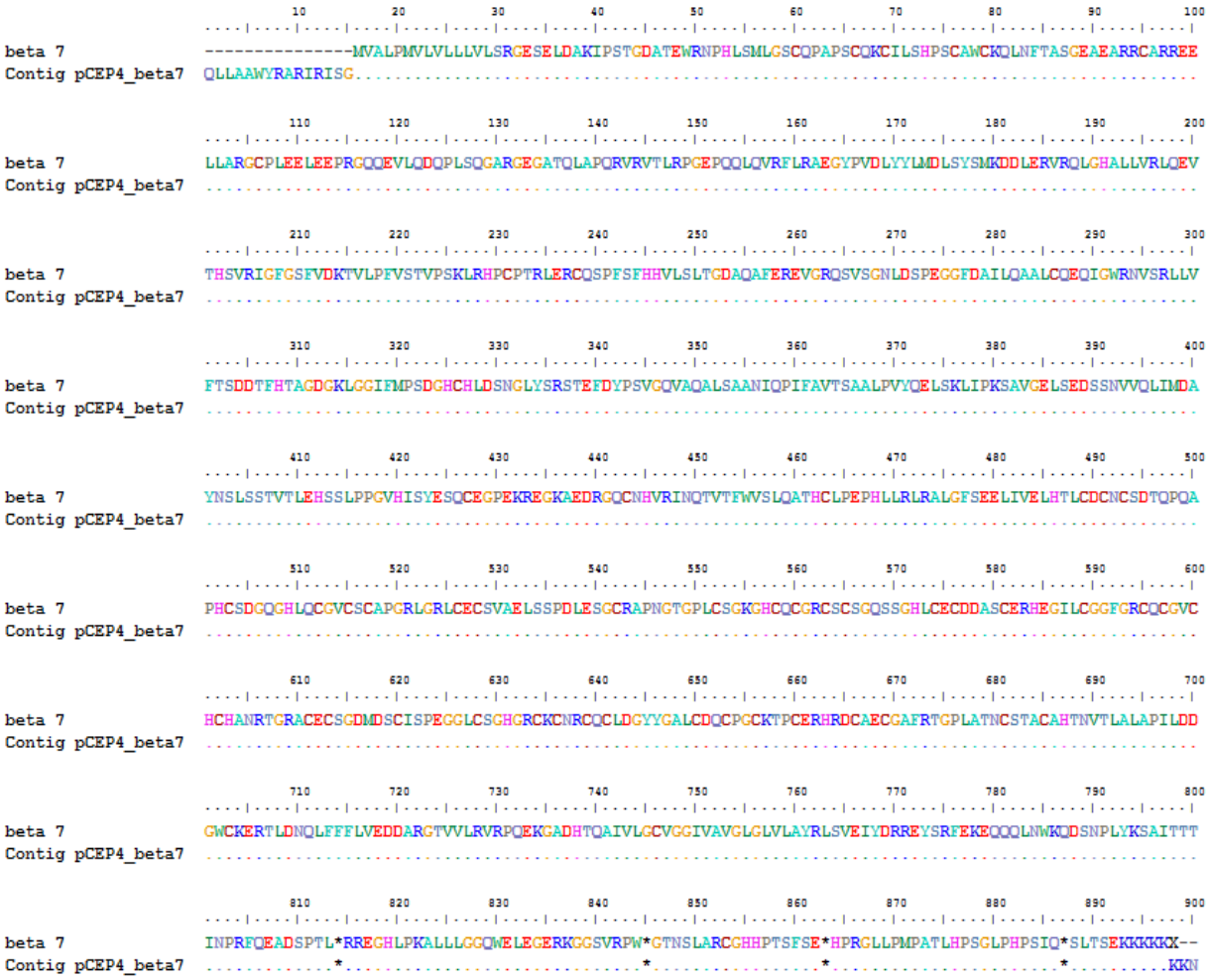


Figure A.1.4. Protein sequence alignment confirms complete β7 cDNA in pCEP4 vector. The β7 cDNA amino acid sequence (beta7) (accesion number NP_000880) was aligned to a contig sequence (contig pCEP4_beta7) generated by sequencing the pCEP4_β7 construct. A methionine (M) indicates the start codon and a * indicates the stop codon. Dots indicate 100% amino acid matches.

A.2. Sequencing the N-glycan Env mutants

The entire *envs* of CAP239 T/F and chronic infection single and double N-glycan mutants were sequenced with a set of nine primers to ensure that mutagenesis was successful. A contig sequence was generated from sequence analysis and was aligned to either the WT CAP239 T/F or chronic infection *env* clone

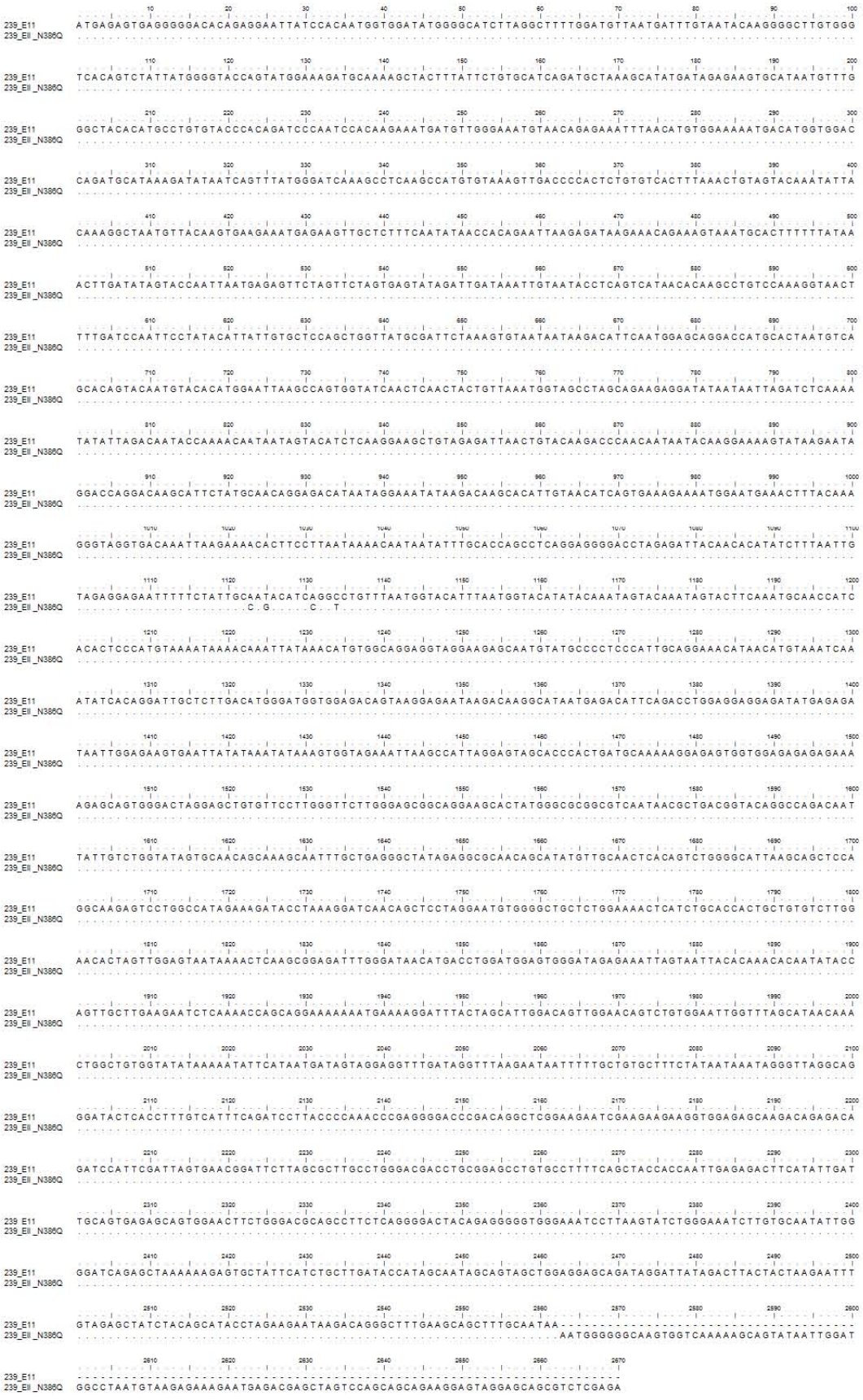

239_E11
239_EII_N386Q

Figure A.1.5. Sequence analysis confirms succesful site-directed mutagensis of the T/F N386Q single N-glycan mutant. The T/F WT sequence (239_E11) was aligned to a contig sequence (239_E11_N386Q) generated by sequencing the entire *env* of the T/F N386Q single N-glycan mutant A) Nucleotide sequence analysis shows an adenine (A) to cytosine (C) and a thymine (T) to guanine (G) mismatch indicating an asparagine (AAT) to glutamine (CAG) mutation. An adenine (A) to cytosine (C) and a cytosine (C) to a thymine (T) mismatch indicates the insertion of a *BsaW* I restriction enzyme site (CCGGT) B) Amino acid sequence analysis shows an asparagine (N) to glutamine (Q) mismatch indicating an NTS (PNG) to a QST (no PNG) sequon mutation. The *BsaW* I restriction enzyme site is silent as no other mismatches are present. A * indicates a stop codon and dots indicate 100% nucleotide/amino acid matches.

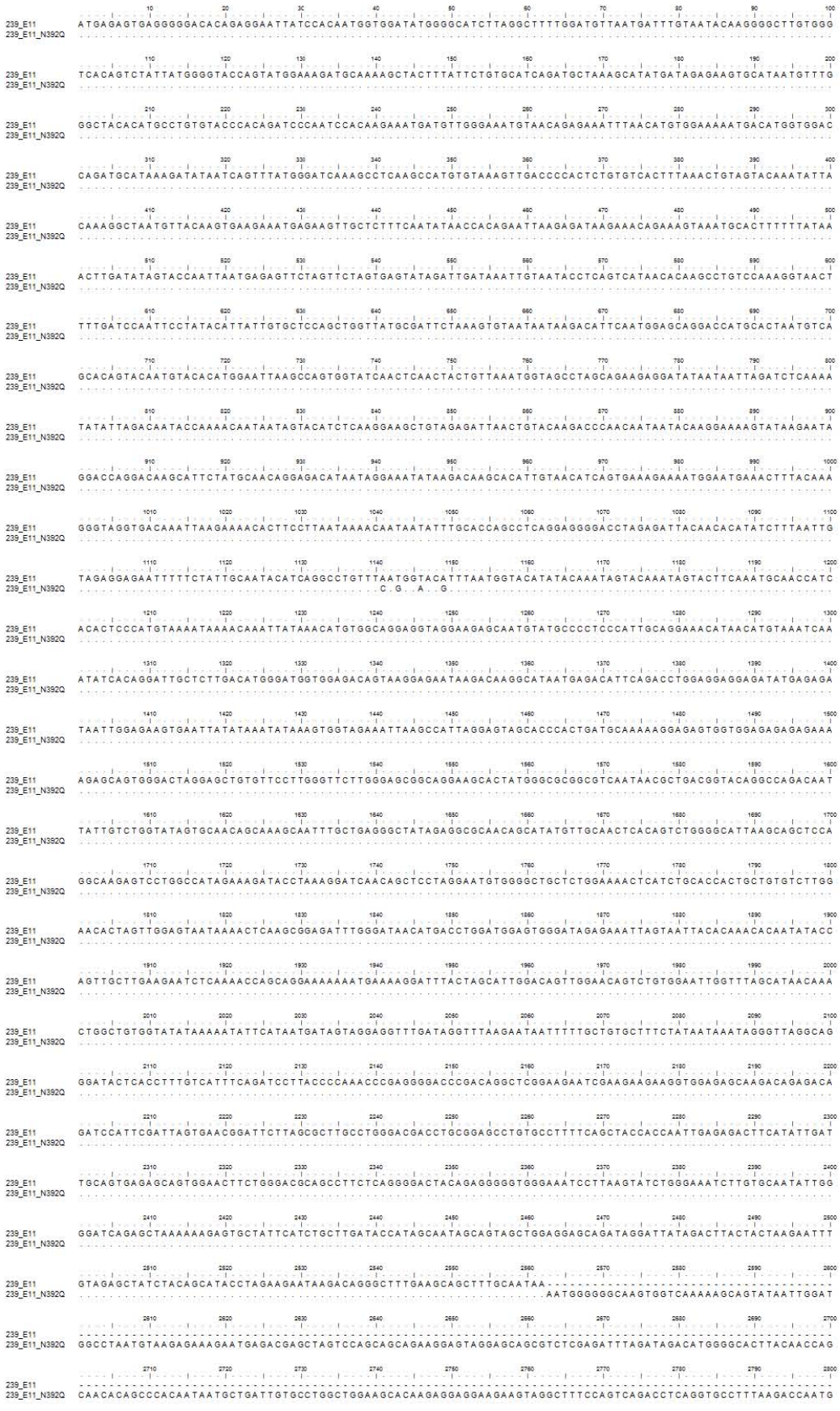

239_E11
239_E11_N392Q
ATGAGAGTGAGGGGGACACAGAGGAATTATCCACAATGGTGGATATGGGGCATCTTAGGCTTTTGGATGTTAATGATTTGTAATACAAGGGGCTTGTGGG
TCACAGTCTATTATGGGGTACCAGTATGGAAAGATGCAAAAGCTACTTTATTCTGTGCATCAGATGCTAAAGCATATGATAGAGAAGTGCATAATGTTTG
GGCTACACATGCCTGTGTACCCACAGATCCCAATCCACAAGAAATGATGTTGGGAAATGTAACAGAGAAATTTAACATGTGGAAAAATGACATGGTGGAC
CAGATGCATAAAGATATAATCAGTTTATGGGATCAAAGCCTCAAGCCATGTGTAAAGTTGACCCCACTCTGTGTCACTTTAAACTGTAGTACAAATATTA
CAAAGGCTAATGTTACAAGTGAAGAAATGAGAAGTTGCTCTTTCAATATAACCACAGAATTAAGAGATAAGAAACAGAAAGTAAATGCACTTTTTTATAA
ACTTGATATAGTACCAATTAATGAGAGTTCTAGTTCTAGTGAGTATAGATTGATAAATTGTAATACCTCAGTCATAACACAAGCCTGTCCAAAGGTAACT
TTTGATCCAATTCCTATACATTATTGTGCTCCAGCTGGTTATGCGATTCTAAAGTGTAATAATAAGACATTCAATGGAGCAGGACCATGCACTAATGTCA
GCACAGTACAATGTACACATGGAATTAAGCCAGTGGTATCAACTCAACTACTGTTAAATGGTAGCCTAGCAGAAGAGGATATAATAATTAGATCTCAAAA
TATATTAGACAATACCAAAACAATAATAGTACATCTCAAGGAAGCTGTAGAGATTAACTGTACAAGACCCAACAATAATACAAGGAAAAGTATAAGAATA
GGACCAGGACAAGCATTCTATGCAACAGGAGACATAATAGGAAATATAAGACAAGCACATTGTAACATCAGTGAAAGAAAATGGAATGAAACTTTACAAA
GGGTAGGTGACAAATTAAGAAAACACTTCCTTAATAAAACAATAATATTTGCACCAGCCTCAGGAGGGGACCTAGAGATTACAACACATATCTTTAATTG
TAGAGGAGAATTTTTCTATTGCAATACATCAGGCCTGTTTAATGGTACATTTAATGGTACATATACAAATAGTACAAATAGTACTTCAAATGCAACCATC
C G A G
ACACTCCCATGTAAAATAAAACAAATTATAAACATGTGGCAGGAGGTAGGAAGAGCAATGTATGCCCCTCCCATTGCAGGAAACATAACATGTAAATCAA
ATATCACAGGATTGCTCTTGACATGGGATGGTGGAGACAGTAAGGAGAATAAGACAAGGCATAATGAGACATTCAGACCTGGAGGAGGAGATATGAGAGA
TAATTGGAGAAGTGAATTATATAAATATAAAGTGGTAGAAATTAAGCCATTAGGAGTAGCACCCACTGATGCAAAAAGGAGAGTGGTGGAGAGAGAGAAA
AGAGCAGTGGGACTAGGAGCTGTGTTCCTTGGGTTCTTGGGAGCGGCAGGAAGCACTATGGGCGCGGCGTCAATAACGCTGACGGTACAGGCCAGACAAT
TATTGTCTGGTATAGTGCAACAGCAAAGCAATTTGCTGAGGGCTATAGAGGCGCAACAGCATATGTTGCAACTCACAGTCTGGGGCATTAAGCAGCTCCA
GGCAAGAGTCCTGGCCATAGAAAGATACCTAAAGGATCAACAGCTCCTAGGAATGTGGGGCTGCTCTGGAAAACTCATCTGCACCACTGCTGTGTCTTGG
AACACTAGTTGGAGTAATAAAACTCAAGCGGAGATTTGGGATAACATGACCTGGATGGAGTGGGATAGAGAAATTAGTAATTACACAAACACAATATACC
AGTTGCTTGAAGAATCTCAAAACCAGCAGGAAAAAAATGAAAAGGATTTACTAGCATTGGACAGTTGGAACAGTCTGTGGAATTGGTTTAGCATAACAAA
CTGGCTGTGGTATATAAAAATATTCATAATGATAGTAGGAGGTTTGATAGGTTTAAGAATAATTTTTGCTGTGCTTTCTATAATAAATAGGGTTAGGCAG
GGATACTCACCTTTGTCATTTCAGATCCTTACCCCAAACCCGAGGGGACCCGACAGGCTCGGAAGAATCGAAGAAGAAGGTGGAGAGCAAGACAGAGACA
GATCCATTCGATTAGTGAACGGATTCTTAGCGCTTGCCTGGGACGACCTGCGGAGCCTGTGCCTTTTCAGCTACCACCAATTGAGAGACTTCATATTGAT
TGCAGTGAGAGCAGTGGAACTTCTGGGACGCAGCCTTCTCAGGGGACTACAGAGGGGGTGGGAAATCCTTAAGTATCTGGGAAATCTTGTGCAATATTGG
GGATCAGAGCTAAAAAAGAGTGCTATTCATCTGCTTGATACCATAGCAATAGCAGTAGCTGGAGGAGCAGATAGGATTATAGACTTACTACTAAGAATTT
GTAGAGCTATCTACAGCATACCTAGAAGAATAAGACAGGGCTTTGAAGCAGCTTTGCAATAA
AATGGGGGGCAAGTGGTCAAAAAGCAGTATAATTGGAT
GGCCTAATGTAAGAGAAAGAATGAGACGAGCTAGTCCAGCAGCAGAAGGAGTAGGAGCAGCGTCTCGAGATTTAGATAGACATGGGGCACTTACAACCAG
CAACACAGCCCACAATAATGCTGATTGTGCCTGGCTGGAAGCACAAGAGGAGGAAGAAGTAGGCTTTCCAGTCAGACCTCAGGTGCCTTTAAGACCAATG

Figure A.1.6. Sequence analysis confirms succesful site-directed mutagensis of the T/F N392Q single N-glycan mutant. The T/F WT sequence (239_E11) was aligned to a contig sequence (239_E11_N392Q) generated by sequencing the entire *env* of the T/F N392Q single N-glycan mutant A) Nucleotide sequence analysis shows an adenine (A) to cytosine (C) and a thymine (T) to guanine (G) mismatch indicating an asparagine (AAT) to glutamine (CAG) mutation. An thymine (T) to an adenine (A) and an adenine (A) to guanine (G) mismatch indicates the insertion of a *AcI* I restriction enzyme site (AACGTT) B) Amino acid sequence analysis shows an asparagine (N) to glutamine (Q) mismatch indicating an NTS (PNG) to a QST (no PNG) sequon mutation. The *AcI* I restriction enzyme site is silent as no other mismatches are present. A * indicates a stop codon and dots indicate 100% nucleotide/amino acid matches.

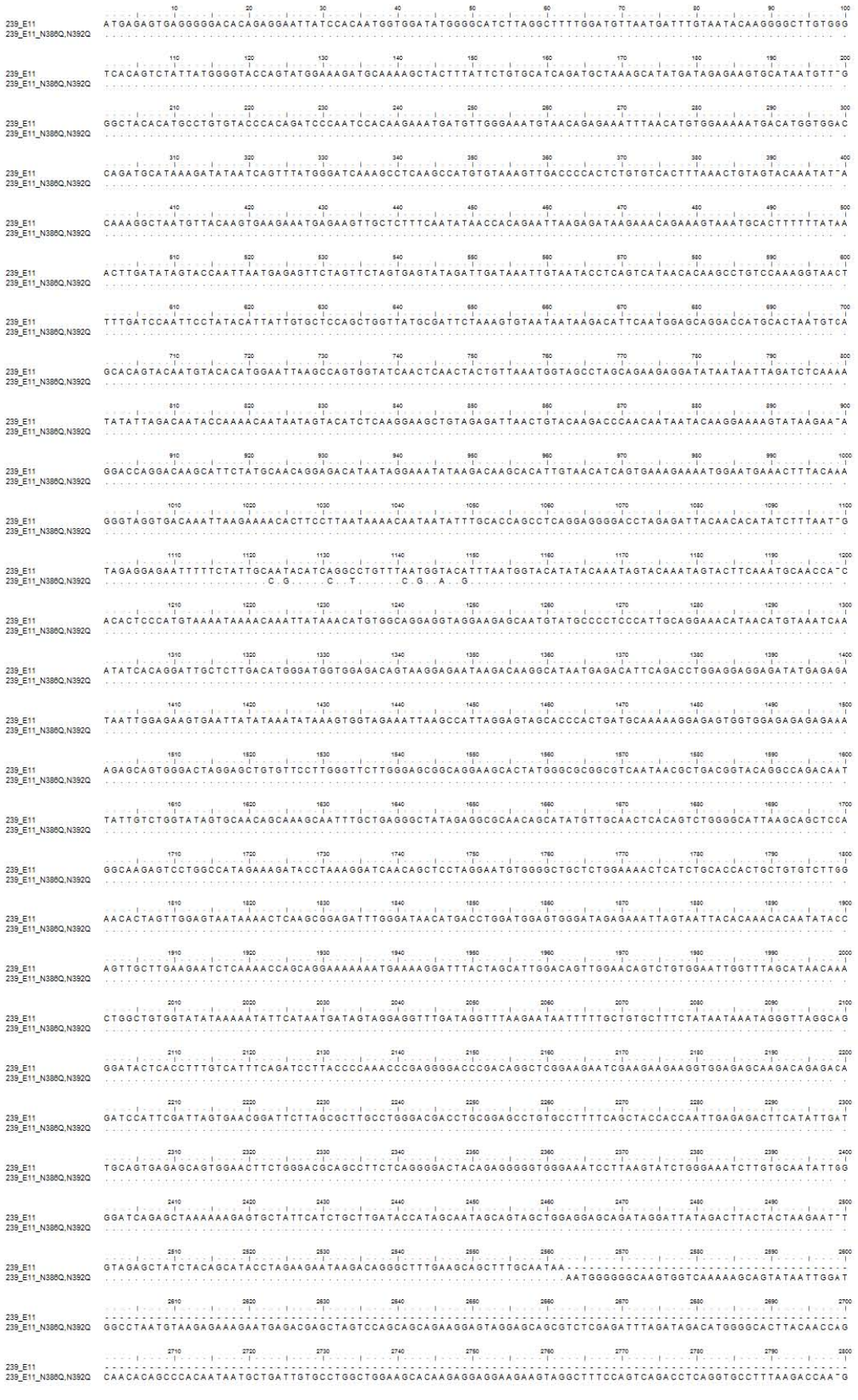
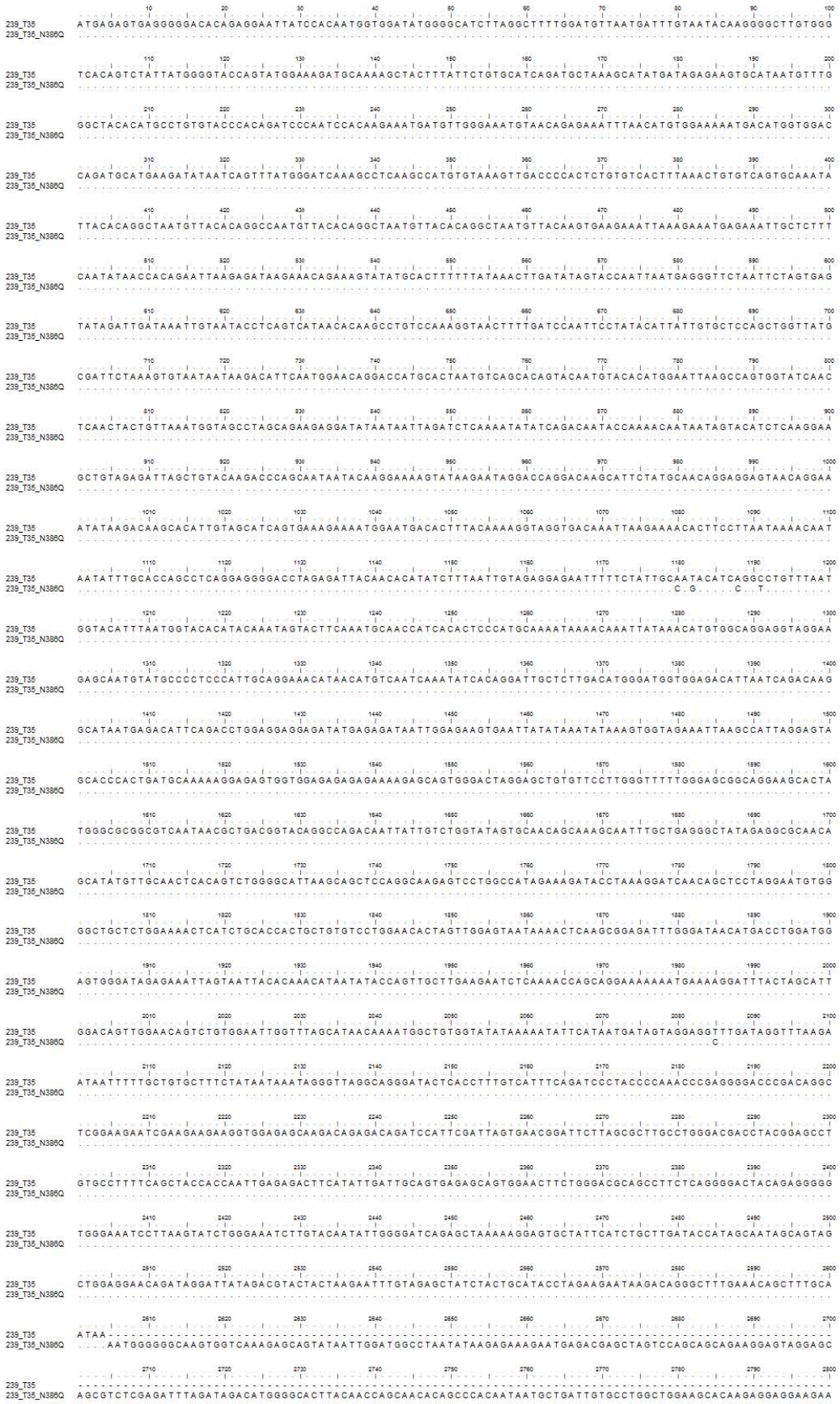

239_E11
239_E11_N386Q,N392Q

Figure A.1.7. Sequence analysis confirms succesful site-directed mutagensis of the T/F N386,N392Q double N-glycan mutant. The T/f WT sequence (239_E11) was aligned to a contig sequence (239_E11_N386Q,N392Q) generated by sequencing the entire *env* of the chronic infection N386Q,N392Q double N-glycan mutant A) Nucleotide sequence analysis shows two adenine (A) to cytosine (C) and thymine (T) to guanine (G) mismatches indicating two asparagine (AAT) to glutamine (CAG) mutations. An adenine (A) to cytosine (C) and a cytosine (C) to thymine (T) mismatch indicates the insertion of a *BsaW* I restriction enzyme site (CCGGT). A thymine (T) to an adenine (A) and an adenine (A) to a guanine (G) mismatch indicates the insertion of an *AcI* I restriction enzyme site (AACGTT) B) Amino acid sequence analysis shows two asparagine (N) to glutamine (Q) mismatches indicating two NTS (PNG) to a QST (no PNG) sequon mutations. The *BsaW* I and *AcI* I restriction enzyme sites are silent as no other mismatches are present. A * indicates a stop codon and dots indicate 100% nucleotide/amino acid matches.


239_T35
239_T35_N386Q

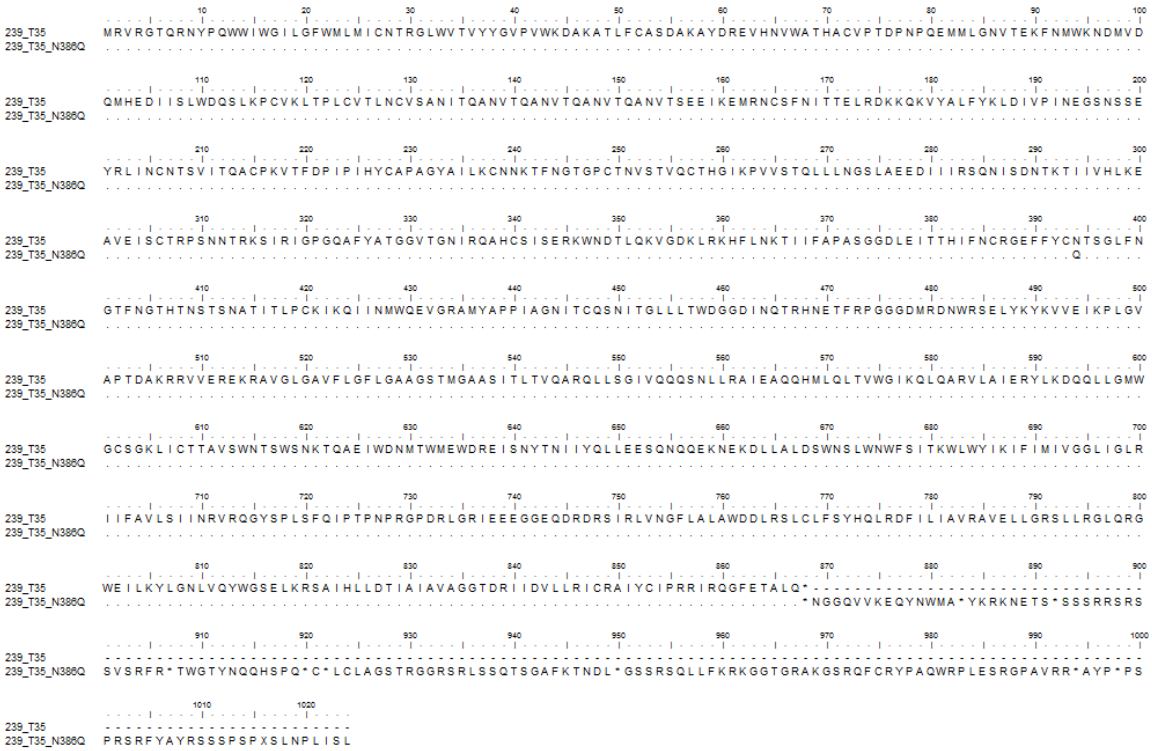


Figure A.1.8. Sequence analysis confirms succesful site-directed mutagensis of the chronic infection N386Q single N-glycan mutant. The chronic infection WT sequence (239_T35) was aligned to a contig sequence (239_T35_N386Q) generated by sequencing the entire *env* of the chronic infection N386Q single N-glycan mutant A) Nucleotide sequence analysis shows an adenine (A) to cytosine (C) and a thymine (T) to guanine (G) mismatch indicating an asparagine (AAT) to glutamine (CAG) mutation. An adenine (A) to cytosine (C) and a cytosine (C) to a thymine (T) mismatch indicates the insertion of a *BsaW* I restriction enzyme site (CCGGT) B) Amino acid sequence analysis shows an asparagine (N) to glutamine (Q) mismatch indicating an NTS (PNG) to a QST (no PNG) sequon mutation. The *BsaW* I restriction enzyme site is silent as no other mismatches are present. A * indicates a stop codon and dots indicate 100% nucleotide/amino acid matches.

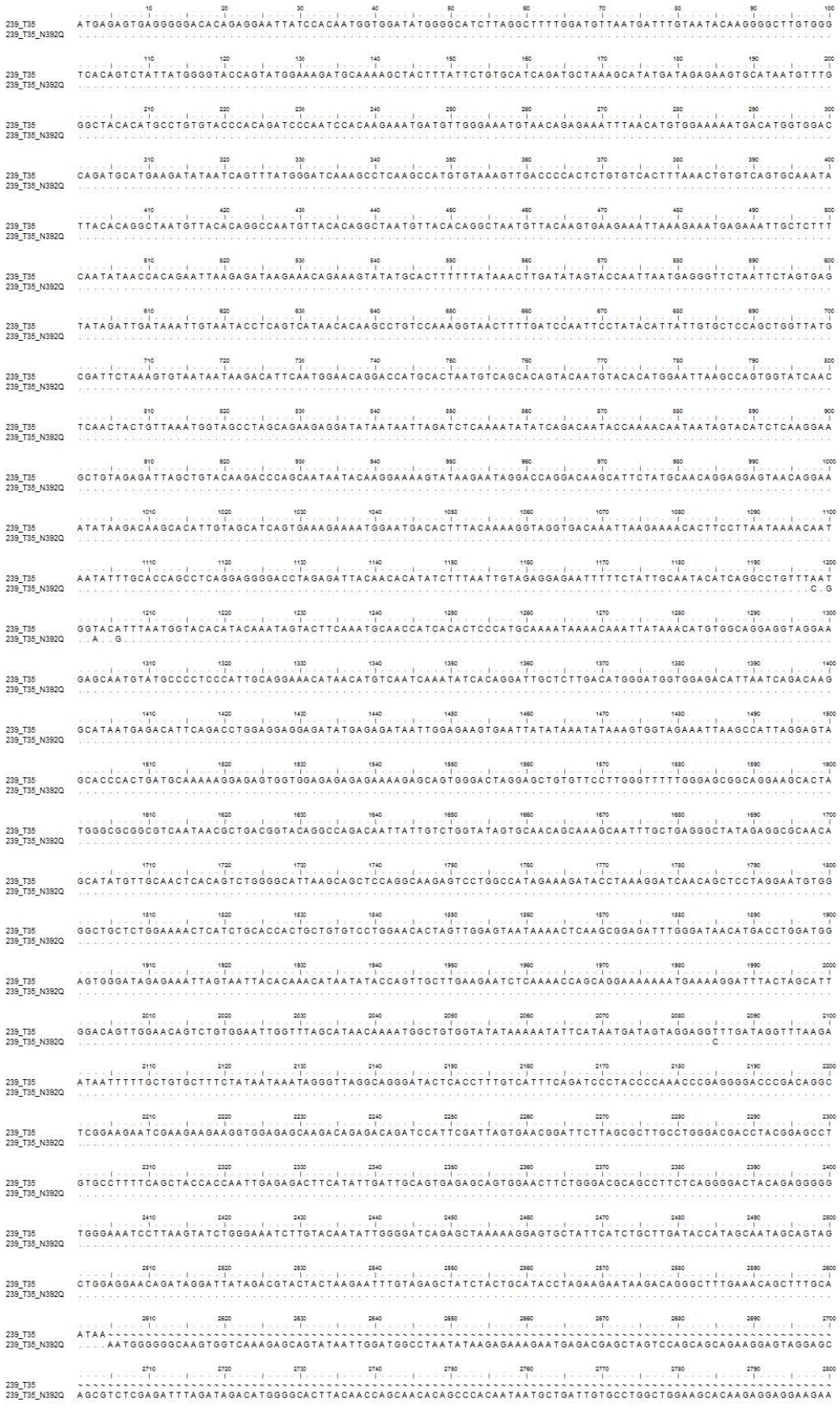

239_T35
239_T35_N392Q
ATGAGAGTGAGGGGGACACAGAGGAATTATCCACAATGGTGGATATGGGGCATCTTAGGCTTTTGGATGTTAATGATTTGTAATACAAGGGGCTTGTGGG
TCACAGTCTATTATGGGGTACCAGTATGGAAAGATGCAAAAGCTACTTTATTCTGTGCATCAGATGCTAAAGCATATGATAGAGAAGTGCATAATGTTTG
GGCTACACATGCCTGTGTACCCACAGATCCCAATCCACAAGAAATGATGTTGGGAAATGTAACAGAGAAATTTAACATGTGGAAAAATGACATGGTGGAC
CAGATGCATGAAGATATAATCAGTTTATGGGATCAAAGCCTCAAGCCATGTGTAAAGTTGACCCCACTCTGTGTCACTTTAAACTGTGTCAGTGCAAATA
TTACACAGGCTAATGTTACACAGGCCAATGTTACACAGGCTAATGTTACACAGGCTAATGTTACAAGTGAAGAAATTAAAGAAATGAGAAATTGCTCTTT
CAATATAACCACAGAATTAAGAGATAAGAAACAGAAAGTATATGCACTTTTTTATAAACTTGATATAGTACCAATTAATGAGGGTTCTAATTCTAGTGAG
TATAGATTGATAAATTGTAATACCTCAGTCATAACACAAGCCTGTCCAAAGGTAACTTTTGATCCAATTCCTATACATTATTGTGCTCCAGCTGGTTATG
CGATTCTAAAGTGTAATAATAAGACATTCAATGGAACAGGACCATGCACTAATGTCAGCACAGTACAATGTACACATGGAATTAAGCCAGTGGTATCAAC
TCAACTACTGTTAAATGGTAGCCTAGCAGAAGAGGATATAATAATTAGATCTCAAAATATATCAGACAATACCAAAACAATAATAGTACATCTCAAGGAA
GCTGTAGAGATTAGCTGTACAAGACCCAGCAATAATACAAGGAAAAGTATAAGAATAGGACCAGGACAAGCATTCTATGCAACAGGAGGAGTAACAGGAA
ATATAAGACAAGCACATTGTAGCATCAGTGAAAGAAAATGGAATGACACTTTACAAAAGGTAGGTGACAAATTAAGAAAACACTTCCTTAATAAAACAAT
AATATTTGCACCAGCCTCAGGAGGGGACCTAGAGATTACAACACATATCTTTAATTGTAGAGGAGAATTTTTCTATTGCAATACATCAGGCCTGTTTAAT
C . G
GGTACATTTAATGGTACACATACAAATAGTACTTCAAATGCAACCATCACACTCCCATGCAAAATAAAACAAATTATAAACATGTGGCAGGAGGTAGGAA
. A . . G .
GAGCAATGTATGCCCCTCCCATTGCAGGAAACATAACATGTCAATCAAATATCACAGGATTGCTCTTGACATGGGATGGTGGAGACATTAATCAGACAAG
GCATAATGAGACATTCAGACCTGGAGGAGGAGATATGAGAGATAATTGGAGAAGTGAATTATATAAATATAAAGTGGTAGAAATTAAGCCATTAGGAGTA
GCACCCACTGATGCAAAAAGGAGAGTGGTGGAGAGAGAGAAAAGAGCAGTGGGACTAGGAGCTGTGTTCCTTGGGTTTTTGGGAGCGGCAGGAAGCACTA
TGGGCGCGGCGTCAATAACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAACAGCAAAGCAATTTGCTGAGGGCTATAGAGGCGCAACA
GCATATGTTGCAACTCACAGTCTGGGGCATTAAGCAGCTCCAGGCAAGAGTCCTGGCCATAGAAAGATACCTAAAGGATCAACAGCTCCTAGGAATGTGG
GGCTGCTCTGGAAAACTCATCTGCACCACTGCTGTGTCCTGGAACACTAGTTGGAGTAATAAAACTCAAGCGGAGATTTGGGATAACATGACCTGGATGG
AGTGGGATAGAGAAATTAGTAATTACACAAACATAATATACCAGTTGCTTGAAGAATCTCAAAACCAGCAGGAAAAAAATGAAAAGGATTTACTAGCATT
GGACAGTTGGAACAGTCTGTGGAATTGGTTTAGCATAACAAAATGGCTGTGGTATATAAAAATATTCATAATGATAGTAGGAGGTTTGATAGGTTTAAGA
C
ATAATTTTTGCTGTGCTTTCTATAATAAATAGGGTTAGGCAGGGATACTCACCTTTGTCATTTCAGATCCCTACCCCAAACCCGAGGGGACCCGACAGGC
TCGGAAGAATCGAAGAAGAAGGTGGAGAGCAAGACAGAGACAGATCCATTCGATTAGTGAACGGATTCTTAGCGCTTGCCTGGGACGACCTACGGAGCCT
GTGCCTTTTCAGCTACCACCAATTGAGAGACTTCATATTGATTGCAGTGAGAGCAGTGGAACTTCTGGGACGCAGCCTTCTCAGGGGACTACAGAGGGGG
TGGGAAATCCTTAAGTATCTGGGAAATCTTGTACAATATTGGGGATCAGAGCTAAAAAGGAGTGCTATTCATCTGCTTGATACCATAGCAATAGCAGTAG
CTGGAGGAACAGATAGGATTATAGACGTACTACTAAGAATTTGTAGAGCTATCTACTGCATACCTAGAAGAATAAGACAGGGCTTTGAAACAGCTTTGCA
ATAA~~~~~~
. . . . . AATGGGGGGCAAGTGGTCAAAGAGCAGTATAATTGGATGGCCTAATATAAGAGAAAGAATGAGACGAGCTAGTCCAGCAGCAGAAGGAGTAGGAGC
AGCGTCTCGAGATTTAGATAGACATGGGGCACTTACAACCAGCAACACAGCCCACAATAATGCTGATTGTGCCTGGCTGGAAGCACAAGAGGAGGAAGAA

Figure A.1.9. Sequence analysis confirms succesful site-directed mutagensis of the chronic infection N392Q single N-glycan mutant. The chronic infection WT sequence (239_T35) was aligned to a contig sequence (239_T35_N392Q) generated by sequencing the entire *env* of the chronic infection N392Q single N-glycan mutant A) Nucleotide sequence analysis shows an adenine (A) to cytosine (C) and a thymine (T) to guanine (G) mismatch indicating an asparagine (AAT) to glutamine (CAG) mutation. A thymine (T) to an adenine (A) and an adenine (A) to a guanine (G) mismatch indicates the insertion of an *AclI* restriction enzyme site (AACGTT) B) Amino acid sequence analysis shows an asparagine (N) to glutamine (Q) mismatch indicating an NTS (PNG) to a QST (no PNG) sequon mutation. The *AclI* restriction enzyme site is silent as no other mismatches are present. A * indicates a stop codon and dots indicate 100% nucleotide/amino acid matches.

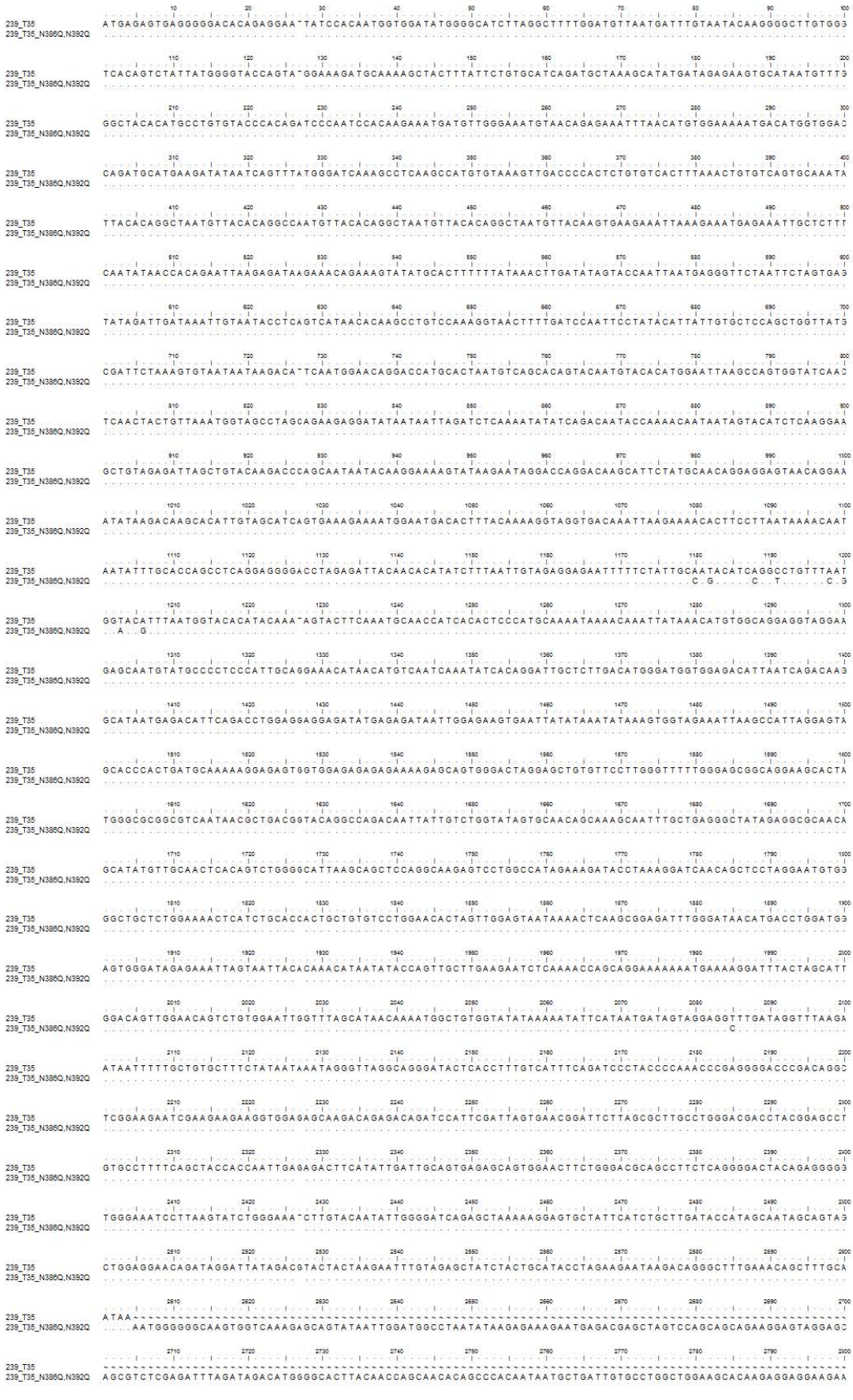

239_T35
239_T35_N386Q,N392Q

Figure A.1.10. Sequence analysis confirms succesful site-directed mutagensis of the chronic infection N386,N392Q double N-glycan mutant. The chronic infection WT sequence (239_T35) was aligned to a contig sequence (239_T35_N386Q,N392Q) generated by sequencing the entire *env* of the chronic infection N386Q,N392Q double N-glycan mutant A) Nucleotide sequence analysis shows two adenine (A) to cytosine (C) and thymine (T) to guanine (G) mismatches indicating two asparagine (AAT) to glutamine (CAG) mutations. An adenine (A) to cytosine (C) and a cytosine (C) to thymine (T) mismatch indicates the insertion of a *BsaW* I restriction enzyme site (CCGGT). A thymine (T) to an adenine (A) and an adenine (A) to a guanine (G) mismatch indicates the insertion of an *AcI* I restriction enzyme site (AACGTT) B) Amino acid sequence analysis shows two asparagine (N) to glutamine (Q) mismatches indicating two NTS (PNG) to a QST (no PNG) sequon mutations. The *BsaW* I and *AcI* I restriction enzyme sites are silent as no other mismatches are present. A * indicates a stop codon and dots indicate 100% nucleotide/amino acid matches.

Bibliography

Abitorabi, M.A. et al., 1997. Presentation of integrins on leukocyte microvilli: a role for the extracellular domain in determining membrane localization. *The Journal of cell biology*, 139(2), pp.563–571.

Abrahams, M.-R. et al., 2009. Quantitating the multiplicity of infection with human immunodeficiency virus type 1 subtype C reveals a non-poisson distribution of transmitted variants. *Journal of virology*, 83(8), pp.3556–3567.

Abram, M.E. et al., 2010. Nature, position, and frequency of mutations made in a single cycle of HIV-1 replication. *Journal of virology*, 84(19), pp.9864–9878.

Abu-Raddad, L.J. & Longini, I.M., 2008. No HIV stage is dominant in driving the HIV epidemic in sub-Saharan Africa. *AIDS*, 22(9), pp.1055–1061.

Alexandre, K.B. et al., 2011. Binding of the mannose-specific lectin, griffithsin, to HIV-1 gp120 exposes the CD4-binding site. *Journal of virology*, 85(17), pp.9039–9050.

Alexandre, K.B. et al., 2012. The lectins griffithsin, cyanovirin-N and scytovirin inhibit HIV-1 binding to the DC-SIGN receptor and transfer to CD4(+) cells. *Virology*, 423(2), pp.175–186.

Ansari, A.A. et al., 2011. Blocking of α4β7 gut-homing integrin during acute infection leads to decreased plasma and gastrointestinal tissue viral loads in simian immunodeficiency virus-infected rhesus macaques. *Journal of immunology*, 186(2), pp.1044–1059.

Apweiler, R., Hermjakob, H. & Sharon, N., 1999. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochimica et biophysica acta*, 1473(1), pp.4–8.

Arrighi, J. et al., 2004. DC-SIGN-mediated infectious synapse formation enhances X4 HIV-1 transmission from dendritic cells to T cells. *The Journal of experimental medicine*, 200(10), pp.1279–1288.

Arthos, J. et al., 2008. HIV-1 envelope protein binds to and signals through integrin alpha4beta7, the gut mucosal homing receptor for peripheral T cells. *Nature immunology*, 9(3), pp.301–309.

Asmal, M. et al., 2011. A signature in HIV-1 envelope leader peptide associated with transition from acute to chronic infection impacts envelope processing and infectivity. *PloS one*, 6(8), pp.1–12.

Balzarini, J. et al., 2007. Carbohydrate-binding agents efficiently prevent dendritic cell-specific intercellular adhesion molecule-3-grabbing nonintegrin (DC-SIGN)-directed HIV-1 transmission to T lymphocytes. *Molecular pharmacology*, 71(1), pp.3–11.

Bargatze, R.F., Jutila, M.A. & Butcher, E.C., 1995. Distinct roles of L-selectin and integrins alpha 4 beta 7 and LFA-1 in lymphocyte homing to Peyer's patch-HEV in situ: the multistep model confirmed and refined. *Immunity*, 3(1), pp.99–108.

Benham, A.M. et al., 2000. The CXXCXXC motif determines the folding, structure and stability of human Ero1-Lalpha. *The EMBO journal*, 19(17), pp.4493–4502.

Berger, E.A., Murphy, P.M. & Farber, J.M., 1999. Chemokine receptors as HIV-1 coreceptors: roles in viral entry, tropism, and disease. *Annual review of immunology*, pp.17657–17700.

Berlin, C. et al., 1995. alpha 4 integrins mediate lymphocyte attachment and rolling under physiologic flow. *Cell*, 80(3), pp.413–422.

Binley, J.M. et al., 2010. Role of complex carbohydrates in human immunodeficiency virus type 1 infection and resistance to antibody neutralization. *Journal of virology*, 84(11), pp.5637–5655.

Bobardt, M.D. et al., 2003. Syndecan captures, protects, and transmits HIV to T lymphocytes. *Immunity*, 18(1), pp.27–39.

Bonomelli, C. et al., 2011. The glycan shield of HIV is predominantly oligomannose independently of production system or viral clade. *PloS one*, 6(8), pp.1–7.

Borggren, M. et al., 2008. Evolution of DC-SIGN use revealed by fitness studies of R5 HIV-1 variants emerging during AIDS progression. *Retrovirology*, 5(28), pp.1–11.

Brenchley, J.M. et al., 2004. CD4+ T cell depletion during all stages of HIV disease occurs predominantly in the gastrointestinal tract. *The Journal of experimental medicine*, 200(6), pp.749–759.

Buck, C.A. & Horwitz, A.F., 1987. Integrin, a transmembrane glycoprotein complex mediating cell-substratum adhesion. *Journal of cell science*, 8, pp.231–250.

Busso, D., Kim, R. & Kim, S.-H., 2003. Expression of soluble recombinant proteins in a cell-free system using a 96-well format. *Journal of biochemical and biophysical methods*, 55(3), pp.233–240.

Caffrey, M. et al., 1998. Three-dimensional solution structure of the 44 kDa ectodomain of SIV gp41. *The EMBO journal*, 17(16), pp.4572–4584.

Carr, J.M. et al., 1999. Rapid and efficient cell-to-cell transmission of human immunodeficiency virus infection from monocyte-derived macrophages to peripheral blood lymphocytes. *Virology*, 265(2), pp.319–329.

Cavrois, M. et al., 2007. In vitro derived dendritic cells trans-infect CD4 T cells primarily with surface-bound HIV-1 virions. *PLoS pathogens*, 3(1), pp.38–45.

Checkley, M.A., Luttge, B.G. & Freed, E.O., 2011. HIV-1 Envelope Glycoprotein Biosynthesis , Trafficking , and Incorporation. *Journal of Molecular Biology*, 410(4), pp.582–608.

Chehimi, J. et al., 2003. HIV-1 transmission and cytokine-induced expression of DC-SIGN in human monocyte-derived macrophages. *Journal of leukocyte biology*, 74(5), pp.757–763.

Chen, B. et al., 2005. Structure of an unliganded simian immunodeficiency virus gp120 core. *Nature*, 433(7028), pp.834–841.

Chen, P. et al., 2007. Predominant mode of human immunodeficiency virus transfer between T cells is mediated by sustained Env-dependent neutralization-resistant virological synapses. *Journal of virology*, 81(22), pp.12582–12595.

Chohan, B. et al., 2005. Selection for human immunodeficiency virus type 1 envelope glycosylation variants with shorter V1-V2 loop sequences occurs during transmission of certain genetic subtypes and may impact viral RNA levels. *Journal of virology*, 79(10), pp.6528–6531.

Chow, W.Z. et al., 2013. Molecular diversity of HIV-1 among people who inject drugs in Kuala Lumpur, Malaysia: massive expansion of circulating recombinant form (CRF) 33_01B and emergence of multiple unique recombinant clusters. *PloS one*, 8(5), pp.1–11.

Chung, N.P.Y. et al., 2010. HIV-1 transmission by dendritic cell-specific ICAM-3-grabbing nonintegrin (DC-SIGN) is regulated by determinants in the carbohydrate recognition domain that are absent in liver/lymph node-SIGN (L-SIGN). *The Journal of biological chemistry*, 285(3), pp.2100–2112.

Cicala, C. et al., 2009. The integrin alpha4beta7 forms a complex with cell-surface CD4 and defines a T-cell subset that is highly susceptible to infection by HIV-1. *Proceedings of the National Academy of Sciences of the United States of America*, 106(49), pp.20877–20882.

Cicala, C., Arthos, J. & Fauci, A.S., 2011. HIV-1 envelope, integrins and co-receptor use in mucosal transmission of HIV. *Journal of translational medicine*, 9(2), pp.1–10.

Clavel, F. et al., 1987. Human immunodeficiency virus type 2 infection associated with AIDS in West Africa. *The New England journal of medicine*, 316(19), pp.1180–1185.

Cocchi, F. et al., 1996. The V3 domain of the HIV-1 gp120 envelope glycoprotein is critical for chemokine-mediated blockade of infection. *Nature medicine*, 2(11), pp.1244–1247.

Cone, R.A. et al., 2006. Vaginal microbicides: detecting toxicities in vivo that paradoxically increase pathogen transmission. *BMC infectious diseases*, 6(90), pp.1–16.

Crispin, M. et al., 2006. Inhibition of hybrid- and complex-type glycosylation reveals the presence of the GlcNAc transferase I-independent fucosylation pathway. *Glycobiology*, 16(8), pp.748–756.

Curtis, B.M., Scharnowske, S. & Watson, A.J., 1992. Sequence and expression of a membrane-associated C-type lectin that exhibits CD4-independent binding of human immunodeficiency virus envelope glycoprotein gp120. *Proceedings of the National Academy of Sciences of the United States of America*, 89(17), pp.8356–8360.

Dalgleish, A.G. et al., 1984. The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. *Nature*, 312(5996), pp.763–767.

Dayal, M.B. et al., 2003. Disruption of the upper female reproductive tract epithelium by nonoxynol-9. *Contraception*, 68(4), pp.273–279.

Delobel, P. et al., 2013. Primary resistance of CCR5-tropic HIV-1 to maraviroc cannot be predicted by the V3 sequence. *The Journal of antimicrobial chemotherapy*, pp.1–9.

Deng, H. et al., 1996. Identification of a major co-receptor for primary isolates of HIV-1. *Nature*, 381(6584), pp.661–666.

DeNucci, C.C. et al., 2010. Control of alpha4beta7 integrin expression and CD4 T cell homing by the beta1 integrin subunit. *Journal of immunology*, 184(5), pp.2458–2467.

Depetris, R.S. et al., 2012. Partial enzymatic deglycosylation preserves the structure of cleaved recombinant HIV-1 envelope glycoprotein trimers. *The Journal of biological chemistry*, 287(29), pp.24239–24254.

Derdeyn, C.A. et al., 2004. Envelope-constrained neutralization-sensitive HIV-1 after heterosexual transmission. *Science*, 303(5666), pp.2019–2022.

Duncan, C.J.A. & Sattentau, Q.J., 2011. Viral determinants of HIV-1 macrophage tropism. *Viruses*, 3(11), pp.2255–2279.

Eggink, D. et al., 2010. Lack of complex N-glycans on HIV-1 envelope glycoproteins preserves protein conformation and entry function. *Virology*, 401(2), pp.236–247.

Erle, D.J. et al., 1991. Complete amino acid sequence of an integrin beta subunit (beta 7) identified in leukocytes. *The Journal of biological chemistry*, 266(17), pp.11009–11016.

Feinberg, H. et al., 2007. Multiple modes of binding enhance the affinity of DC-SIGN for high mannose N-linked glycans found on viral glycoproteins. *The Journal of biological chemistry*, 282(6), pp.4202–4209.

Feinberg, H. et al., 2001. Structural basis for selective recognition of oligosaccharides by DC-SIGN and DC-SIGNR. *Science*, 294(5549), pp.2163–2166.

Fenouillet, E. et al., 1989. Role of N-linked glycans in the interaction between the envelope glycoprotein of human immunodeficiency virus and its CD4 cellular receptor. Structural enzymatic analysis. *The Journal of experimental medicine*, 169(3), pp.807–822.

Fenouillet, E., Gluckman, J.C. & Bahraoui, E., 1990. Role of N-linked glycans of envelope glycoproteins in infectivity of human immunodeficiency virus type 1. *Journal of virology*, 64(6), pp.2841–2848.

Fenouillet, E. & Jones, I.M., 1995. The glycosylation of human immunodeficiency virus type 1 transmembrane glycoprotein (gp41) is important for the efficient intracellular transport of the envelope precursor gp160. *The Journal of general virology*, 76, pp.1509–1514.

Fouda, G.G. et al., 2013. Postnatally-transmitted HIV-1 Envelope variants have similar neutralization-sensitivity and function to that of nontransmitted breast milk variants. *Retrovirology*, 10(1), pp.1–20.

Frankel, A.D. & Young, J.A., 1998. HIV-1: fifteen proteins and an RNA. *Annual review of biochemistry*, 67, pp.1–25.

Frost, S.D.W. et al., 2005. Characterization of human immunodeficiency virus type 1 (HIV-1) envelope variation and neutralizing antibody responses during transmission of HIV-1 subtype B. *Journal of virology*, 79(10), pp.6523–6527.

Gao, F. et al., 2005. Antigenicity and immunogenicity of a synthetic human immunodeficiency virus type 1 group m consensus envelope glycoprotein. *Journal of virology*, 79(2), pp.1154–1163.

Garside, P. et al., 1998. Visualization of specific B and T lymphocyte interactions in the lymph node. *Science*, 281(5373), pp.96–99.

Gavel, Y. & von Heijne, G., 1990. Sequence differences between glycosylated and non-glycosylated Asn-X-Thr/Ser acceptor sites: implications for protein engineering. *Protein engineering*, 3(5), pp.433–442.

Geijtenbeek, T.B., Kwon, D.S., et al., 2000. DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells. *Cell*, 100(5), pp.587–597.

Geijtenbeek, T.B., Torensma, R., et al., 2000. Identification of DC-SIGN, a novel dendritic cell-specific ICAM-3 receptor that supports primary immune responses. *Cell*, 100(5), pp.575–585.

Geijtenbeek, T.B.H. et al., 2002. Identification of different binding sites in the dendritic cell-specific receptor DC-SIGN for intercellular adhesion molecule 3 and HIV-1. *The Journal of biological chemistry*, 277(13), pp.11314–11320.

Geijtenbeek, T.B.H. et al., 2004. Self- and nonself-recognition by C-type lectins on dendritic cells. *Annual review of immunology*, 22, pp.33–54.

Geijtenbeek, T.B.H., den Dunnen, J. & Gringhuis, S.I., 2009. Pathogen recognition by DC-SIGN shapes adaptive immunity. *Future microbiology*, 4(7), pp.879–890.

Gnanakaran, S. et al., 2010. Genetic signatures in the envelope glycoproteins of HIV-1 that associate with broadly neutralizing antibodies. *PLoS computational biology*, 6(10), pp.1–26.

Gnanakaran, S. et al., 2011. Recurrent signature patterns in HIV-1 B clade envelope glycoproteins associated with either early or chronic infections. *PLoS pathogens*, 7(9), pp.1–19.

Go, E.P. et al., 2011. Characterization of glycosylation profiles of HIV-1 transmitted/founder envelopes by mass spectrometry. *Journal of virology*, 85(16), pp.8270–8284.

Go, E.P. et al., 2013. Characterization of host-cell line specific glycosylation profiles of early transmitted/founder HIV-1 gp120 envelope proteins. *Journal of proteome research*, 12(3), pp.1223–1234.

Granelli-Piperno, a et al., 1999. Virus replication begins in dendritic cells during the transmission of HIV-1 from mature dendritic cells to T cells. *Current biology*, 9(1), pp.21–29.

Gray, R.H. et al., 2001. Probability of HIV-1 transmission per coital act in monogamous, heterosexual, HIV-1-discordant couples in Rakai, Uganda. *Lancet*, 357(9263), pp.1149–1153.

Guadalupe, M. et al., 2003. Severe CD4+ T-cell depletion in gut lymphoid tissue during primary human immunodeficiency virus type 1 infection and substantial delay in restoration following highly active antiretroviral therapy. *Journal of virology*, 77(21), pp.11708–11717.

Haaland, R.E. et al., 2009. Inflammatory genital infections mitigate a severe genetic bottleneck in heterosexual transmission of subtype A and C HIV-1. *PLoS pathogens*, 5(1), pp.1–13.

Halary, F. et al., 2002. Human cytomegalovirus binding to DC-SIGN is required for dendritic cell infection and target cell trans-infection. *Immunity*, 17(5), pp.653–664.

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, 41, pp.95–98.

Hallenberger, S. et al., 1992. Inhibition of furin-mediated cleavage activation of HIV-1 glycoprotein gp160. *Nature*, 360(6402), pp.358–361.

Van Harmelen, J.H. et al., 1999. A predominantly HIV type 1 subtype C-restricted epidemic in South African urban populations. *AIDS research and human retroviruses*, 15(4), pp.395–398.

Helenius, J. et al., 2002. Translocation of lipid-linked oligosaccharides across the ER membrane requires Rft1 protein. *Nature*, 415(6870), pp.447–450.

Hemelaar, J. et al., 2006. Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. *AIDS*, 20(16), pp.13–23.

Hladik, F. et al., 2007. Initial events in establishing vaginal entry and infection by human immunodeficiency virus type-1. *Immunity*, 26(2), pp.257–270.

Holl, V. et al., 2010. Stimulation of HIV-1 replication in immature dendritic cells in contact with primary CD4 T or B lymphocytes. *Journal of virology*, 84(9), pp.4172–4182.

Holzmann, B. & Weissman, I.L., 1989. Peyer's patch-specific lymphocyte homing receptors consist of a VLA-4-like alpha chain associated with either of two integrin beta chains, one of which is novel. *The EMBO journal*, 8(6), pp.1735–1741.

Hong, P.W.-P. et al., 2002. Human immunodeficiency virus envelope (gp120) binding to DC-SIGN and primary dendritic cells is carbohydrate dependent but does not involve 2G12 or cyanovirin binding sites: implications for structural analyses of gp120-DC-SIGN binding. *Journal of virology*, 76(24), pp.12855–12865.

Hong, P.W.-P. et al., 2007. Identification of the optimal DC-SIGN binding site on human immunodeficiency virus type 1 gp120. *Journal of virology*, 81(15), pp.8325–8336.

Horbul, J.E. et al., 2011. Herpes simplex virus-induced epithelial damage and susceptibility to human immunodeficiency virus type 1 infection in human cervical organ culture. *PloS one*, 6(7), pp.1–13.

Iwata, M. et al., 2004. Retinoic acid imprints gut-homing specificity on T cells. *Immunity*, 21(4), pp.527–538.

Izquierdo-Useros, N. et al., 2009. Capture and transfer of HIV-1 particles by mature dendritic cells converges with the exosome-dissemination pathway. *Blood*, 113(12), pp.2732–2741.

Jakubik, J.J. et al., 2000. Immune complexes containing human immunodeficiency virus type 1 primary isolates bind to lymphoid tissue B lymphocytes and are infectious for T lymphocytes. *Journal of virology*, 74(1), pp.552–555.

Jensen, M.A. & van 't Wout, A.B., 2010. Predicting HIV-1 coreceptor usage with sequence analysis. *AIDS reviews*, 5(2), pp.104–112.

Johnson, V. a et al., 2011. 2011 update of the drug resistance mutations in HIV-1. *Topics in antiviral medicine*, 19(4), pp.156–164.

Johnston, M.I. & Fauci, A.S., 2007. An HIV vaccine--evolving concepts. *The New England journal of medicine*, 356(20), pp.2073–2081.

Joint United Nations Programme on HIV/AIDS, 2009. Global Report: UNAIDS Report On The Global AIDS Epidemic 2009.

Joint United Nations Programme on HIV/AIDS, 2012. Global Report: UNAIDS Report On The Global AIDS Epidemic 2012.

Kamata, T., Puzon, W. & Takada, Y., 1995. Identification of putative ligand-binding sites of the integrin alpha 4 beta 1 (VLA-4, CD49d/CD29). *The Biochemical journal*, 305, pp.945–951.

Keele, B.F. et al., 2008. Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. *Proceedings of the National Academy of Sciences of the United States of America*, 105(21), pp.7552–7557.

Kelleher, D.J. & Gilmore, R., 2006. An evolving view of the eukaryotic oligosaccharyltransferase. *Glycobiology*, 16(4), pp.47–62.

Kelly, K.A. & Rank, R.G., 1997. Identification of homing receptors that mediate the recruitment of CD4 T cells to the genital tract following intravaginal infection with Chlamydia trachomatis. *Infection and immunity*, 65(12), pp.5198–5208.

King, D.F.L. et al., 2013. Mucosal tissue tropism and dissemination of HIV-1 subtype B acute envelope-expressing chimeric virus. *Journal of virology*, 87(2), pp.890–899.

Kraus, M.H. et al., 2010. A rev1-vpu polymorphism unique to HIV-1 subtype A and C strains impairs envelope glycoprotein expression from rev-vpu-env cassettes and reduces virion infectivity in pseudotyping assays. *Virology*, 397(2), pp.346–57.

Kwon, D.S. et al., 2002. DC-SIGN-mediated internalization of HIV is required for trans-enhancement of T cell infection. *Immunity*, 16(1), pp.135–144.

Kwong, P.D. et al., 2000. Oligomeric modeling and electrostatic analysis of the gp120 envelope glycoprotein of human immunodeficiency virus. *Journal of virology*, 74(4), pp.1961–1972.

Kwong, P.D. et al., 1998. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature*, 393(6686), pp.648–659.

Land, A., Zonneveld, D. & Braakman, I., 2003. Folding of HIV-1 envelope glycoprotein involves extensive isomerization of disulfide bonds and conformation-dependent leader peptide cleavage. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 17(9), pp.1058–1067.

Larkin, M. a et al., 2007. Clustal W and Clustal X version 2.0. *Bioinformatics*, 23(21), pp.2947–2948.

Lazarovits, A.I. et al., 1984. Lymphocyte activation antigens. I. A monoclonal antibody, anti-Act I, defines a new late lymphocyte activation antigen. *Journal of immunology*, 133(4), pp.1857–1862.

Lee, B. et al., 2001. cis Expression of DC-SIGN allows for more efficient entry of human and simian immunodeficiency viruses via CD4 and a coreceptor. *Journal of virology*, 75(24), pp.12028–12038.

Lee, H.Y. et al., 2009. Modeling sequence evolution in acute HIV-1 infection. *Journal of theoretical biology*, 261(2), pp.341–360.

Leonard, C.K. et al., 1990. 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. *The Journal of biological chemistry*, 265(18), pp.10373–10382.

Levy, L.A., 1995. History and epidemiology of acquired immune deficiency syndrome. *Journal of the American Podiatric Medical Association*, 85(7), pp.346–351.

Li, H. et al., 2005. Chemoenzymatic synthesis of HIV-1 V3 glycopeptides carrying two N-glycans and effects of glycosylation on the peptide domain. *The Journal of organic chemistry*, 70(24), pp.9990–9996.

Liao, C. et al., 2011. Identification of the DC-SIGN-interactive domains on the envelope glycoprotein of HIV-1 CRF07_BC. *AIDS research and human retroviruses*, 27(8), pp.831–839.

Lin, G. et al., 2003. Differential N-linked glycosylation of human immunodeficiency virus and Ebola virus envelope glycoproteins modulates interactions with DC-SIGN and DC-SIGNR. *Journal of virology*, 77(2), pp.1337–1346.

Liu, J. et al., 2008. Molecular architecture of native HIV-1 gp120 trimers. *Nature*, 455(7209), pp.109–113.

Loré, K. et al., 2005. Myeloid and plasmacytoid dendritic cells transfer HIV-1 preferentially to antigen-specific CD4+ T cells. *The Journal of experimental medicine*, 201(12), pp.2023–2033.

Lu, Z. et al., 1997. Evolution of HIV-1 coreceptor usage through interactions with distinct CCR5 and CXCR4 domains. *Proceedings of the National Academy of Sciences of the United States of America*, 94(12), pp.6426–6431.

Lycke, N., 2012. Recent progress in mucosal vaccine development: potential and limitations. *Nature reviews. Immunology*, 12(8), pp.592–605.

Maddon, P.J. et al., 1986. The T4 gene encodes the AIDS virus receptor and is expressed in the immune system and the brain. *Cell*, 47(3), pp.333–348.

Marconi, V.C. et al., 2008. Prevalence of HIV-1 drug resistance after failure of a first highly active antiretroviral therapy regimen in KwaZulu Natal, South Africa. *Clinical infectious diseases*, 46(10), pp.158915–158997.

Marmor, M. et al., 2001. Homozygous and heterozygous CCR5-Delta32 genotypes are associated with resistance to HIV infection. *Journal of acquired immune deficiency syndromes*, 27(5), pp.472–481.

Marshall, R.D., 1974. The nature and metabolism of the carbohydrate-peptide linkages of glycoproteins. *Biochemical Society symposium*, 41, pp.17–26.

McKinnon, L.R. et al., 2011. Characterization of a human cervical CD4+ T cell subset coexpressing multiple markers of HIV susceptibility. *Journal of Immunology*, 187(11), pp.6032–6042.

McLellan, J.S. et al., 2011. Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9. *Nature*, 480(7377), pp.336–343.

Mehandru, S. et al., 2007. Mechanisms of gastrointestinal CD4+ T-cell depletion during acute and early human immunodeficiency virus type 1 infection. *Journal of virology*, 81(2), pp.599–612.

Milich, L., Margolin, B. & Swanstrom, R., 1993. V3 loop of the human immunodeficiency virus type 1 Env protein: interpreting sequence variability. *Journal of virology*, 67(9), pp.5623–5634.

Mitchell, D.A., Fadden, A.J. & Drickamer, K., 2001. A novel mechanism of carbohydrate recognition by the C-type lectins DC-SIGN and DC-SIGNR. Subunit organization and binding to multivalent ligands. *The Journal of biological chemistry*, 276(31), pp.28939–28945.

Modrow, S. et al., 1987. Computer-assisted analysis of envelope protein sequences of seven human immunodeficiency virus isolates: prediction of antigenic epitopes in conserved and variable regions. *Journal of virology*, 61(2), pp.570–578.

Moir, S. et al., 2000. B cells of HIV-1-infected patients bind virions through CD21-complement interactions and transmit infectious virus to activated T cells. *The Journal of experimental medicine*, 192(5), pp.637–646.

Mondor, I., Ugolini, S. & Sattentau, Q.J., 1998. Human immunodeficiency virus type 1 attachment to HeLa CD4 cells is CD4 independent and gp120 dependent and requires cell surface heparans. *Journal of virology*, 72(5), pp.3623–3634.

Montefiori, D.C., Robinson, W.E. & Mitchell, W.M., 1988. Role of protein N-glycosylation in pathogenesis of human immunodeficiency virus type 1. *Proceedings of the National Academy of Sciences of the United States of America*, 85(23), pp.9248–9252.

Moore, P.L. et al., 2006. Nature of nonfunctional envelope proteins on the surface of human immunodeficiency virus type 1. *Journal of virology*, 80(5), pp.2515–2528.

Moulard, M. et al., 1999. Processing and routage of HIV glycoproteins by furin to the cell surface. *Virus research*, 60(1), pp.55–65.

Myszka, D.G. et al., 2000. Energetics of the HIV gp120-CD4 binding reaction. *Proceedings of the National Academy of Sciences of the United States of America*, 97(16), pp.9026–9031.

Nabatov, A.A. et al., 2006. Interaction of HIV-1 with dendritic cell-specific intercellular adhesion molecule-3-grabbing nonintegrin-expressing cells is influenced by gp120 envelope modifications associated with disease progression. *The FEBS journal*, 273(21), pp.4944–4958.

Nawaz, F. et al., 2011. The genotype of early-transmitting HIV gp120s promotes α (4) β(7)-reactivity, revealing α (4) β(7) +/CD4+ T cells as key targets in mucosal transmission. *PLoS pathogens*, 7(2), pp.1–14.

Pal, R., Hoke, G.M. & Sarngadharan, M.G., 1989. Role of oligosaccharides in the processing and maturation of envelope glycoproteins of human immunodeficiency virus type 1. *Proceedings of the National Academy of Sciences of the United States of America*, 86(9), pp.3384–3388.

Palmer, S. et al., 2005. Multiple, linked human immunodeficiency virus type 1 drug resistance mutations in treatment-experienced patients are missed by standard genotype analysis. *Journal of clinical microbiology*, 43(1), pp.406–413.

Parker, Z.F. et al., 2013. Transmitted/founder and chronic HIV-1 envelope proteins are distinguished by differential utilization of CCR5. *Journal of virology*, 87(5), pp.2401–2411.

Parrish, N.F. et al., 2013. Phenotypic properties of transmitted founder HIV-1. *Proceedings of the National Academy of Sciences of the United States of America*, 110(17), pp.6626–6633.

Parrish, N.F. et al., 2012. Transmitted/founder and chronic subtype C HIV-1 use CD4 and CCR5 receptors with equal efficiency and are not inhibited by blocking the integrin α4β7. *PLoS pathogens*, 8(5), pp.1–16.

Pastore, C. et al., 2006. Human immunodeficiency virus type 1 coreceptor switching: V1/V2 gain-of-fitness mutations compensate for V3 loss-of-fitness mutations. *Journal of virology*, 80(2), pp.750–758.

Patton, D.L. et al., 2000. Epithelial cell layer thickness and immune cell populations in the normal human vagina at different stages of the menstrual cycle. *American journal of obstetrics and gynecology*, 183(4), pp.967–973.

Petrescu, A. et al., 2004. Statistical analysis of the protein environment of N-glycosylation sites: implications for occupancy, structure, and folding. *Glycobiology*, 14(2), pp.103–114.

Pfeiffer, T. et al., 2006. Effects of signal peptide exchange on HIV-1 glycoprotein expression and viral infectivity in mammalian cells. *FEBS letters*, 580(15), pp.3775–3778.

Phillips, D.M. et al., 2000. Nonoxynol-9 causes rapid exfoliation of sheets of rectal epithelium. *Contraception*, 62(3), pp.149–154.

Polzer, S. et al., 2001. Loss of N-linked glycans in the V3-loop region of gp120 is correlated to an enhanced infectivity of HIV-1. *Glycobiology*, 11(1), pp.11–19.

Pope, M. et al., 1995. Low levels of HIV-1 infection in cutaneous dendritic cells promote extensive viral replication upon binding to memory CD4+ T cells. *The Journal of experimental medicine*, 182(6), pp.2045–2056.

Pudney, J., Quayle, A.J. & Anderson, D.J., 2005. Immunological microenvironments in the human vagina and cervix: mediators of cellular immunity are concentrated in the cervical transformation zone. *Biology of reproduction*, 73(6), pp.1253–1263.

Qi, J. et al., 2012. Identification, characterization, and epitope mapping of human monoclonal antibody J19 that specifically recognizes activated integrin α4β7. *The Journal of biological chemistry*, 287(19), pp.15749–15759.

Quiñones-Kochs, M.I., Buonocore, L. & Rose, J.K., 2002. Role of N-linked glycans in a human immunodeficiency virus envelope glycoprotein: effects on protein function and the neutralizing antibody response. *Journal of virology*, 76(9), pp.4199–4211.

Rappocciolo, G. et al., 2006. DC-SIGN on B lymphocytes is required for transmission of HIV-1 to T lymphocytes. *PLoS pathogens*, 2(7), pp.691–704.

Raska, M. et al., 2010. Glycosylation patterns of HIV-1 gp120 depend on the type of expressing cells and affect antibody recognition. *The Journal of biological chemistry*, 285(27), pp.20860–20869.

Reitter, J.N., Means, R.E. & Desrosiers, R.C., 1998. A role for carbohydrates in immune evasion in AIDS. *Nature medicine*, 4(6), pp.679–684.

Requena, M. et al., 2008. Inhibition of HIV-1 transmission in trans from dendritic cells to CD4+ T lymphocytes by natural antibodies to the CRD domain of DC-SIGN purified from breast milk and intravenous immunoglobulins. *Immunology*, 123(4), pp.508–518.

Rhodes, T., Wargo, H. & Hu, W., 2003. High rates of human immunodeficiency virus type 1 recombination: near-random segregation of markers one kilobase apart in one round of viral replication. *Journal of virology*, 77(20), pp.11193–11200.

Rizzuto, C.D., 1998. A Conserved HIV gp120 Glycoprotein Structure Involved in Chemokine Receptor Binding. *Science*, 280(5371), pp.1949–1953.

Rusert, P. et al., 2005. Virus isolates during acute and chronic human immunodeficiency virus type 1 infection show distinct patterns of sensitivity to entry inhibitors. *Journal of virology*, 79(13), pp.8454–8469.

Sagar, M. et al., 2006. Human immunodeficiency virus type 1 V1-V2 envelope loop sequences expand and add glycosylation sites over the course of infection, and these modifications affect antibody neutralization sensitivity. *Journal of virology*, 80(19), pp.9586–9598.

Sagar, M. et al., 2004. Identification of modifiable factors that affect the genetic diversity of the transmitted HIV-1 population. *AIDS*, 18(4), pp.615–619.

Salazar-Gonzalez, J.F. et al., 2008. Deciphering human immunodeficiency virus type 1 transmission and early envelope diversification by single-genome amplification and sequencing. *Journal of virology*, 82(8), pp.3952–3970.

Salazar-Gonzalez, J.F. et al., 2009. Genetic identity, biological phenotype, and evolutionary pathways of transmitted/founder viruses in acute and early HIV-1 infection. *The Journal of experimental medicine*, 206(6), pp.1273–1289.

Sambrook, J.. & Russell, D.., 2001. *Molecular Cloning: A Laboratory Manual* 3rd ed., Cold Spring Harbor: Cold Spring Harbor Laboratory Press.

Samson, M. et al., 1996. Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. *Nature*, 382(6593), pp.722–725.

Sanders, R.W. et al., 2008. 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), pp.1–15.

Sattentau, Q.J. et al., 1986. Epitopes of the CD4 antigen and HIV infection. *Science*, 234(4780), pp.1120–1123.

Scanlan, C.N. et al., 2002. The broadly neutralizing anti-human immunodeficiency virus type 1 antibody 2G12 recognizes a cluster of alpha1-->2 mannose residues on the outer face of gp120. *Journal of virology*, 76(14), pp.7306–7321.

Schønning, K. et al., 1996. Resistance to V3-directed neutralization caused by an N-linked oligosaccharide depends on the quaternary structure of the HIV-1 envelope oligomer. *Virology*, 218(1), pp.134–140.

Schwarz, F. & Aebi, M., 2011. Mechanisms and principles of N-linked protein glycosylation. *Current opinion in structural biology*, 21(5), pp.576–582.

Shaw, G.M. & Hunter, E., 2012. HIV transmission. *Cold Spring Harbor perspectives in medicine*, 2(11), pp.1–24.

Simmonds, P. et al., 1990. Analysis of sequence diversity in hypervariable regions of the external glycoprotein of human immunodeficiency virus type 1. *Journal of virology*, 64(12), pp.5840–5850.

Speck, R.F. et al., 1997. Selective employment of chemokine receptors as human immunodeficiency virus type 1 coreceptors determined by individual amino acids within the envelope V3 loop. *Journal of virology*, 71(9), pp.7136–7139.

Srinivasan, A. et al., 1987. Molecular characterization of human immunodeficiency virus from Zaire: nucleotide sequence analysis identifies conserved and variable domains in the envelope gene. *Gene*, 52(1), pp.71–82.

Starcich, B.R. et al., 1986. Identification and characterization of conserved and variable regions in the envelope gene of HTLV-III/LAV, the retrovirus of AIDS. *Cell*, 45(5), pp.637–648.

Staropoli, I. et al., 2000. Processing, stability, and receptor binding properties of oligomeric envelope glycoprotein from a primary HIV-1 isolate. *The Journal of biological chemistry*, 275(45), pp.35137–35145.

Stein, B.S. & Engleman, E.G., 1990. Intracellular processing of the gp160 HIV-1 envelope precursor. Endoproteolytic cleavage occurs in a cis or medial compartment of the Golgi complex. *The Journal of biological chemistry*, 265(5), pp.2640–2649.

Stevenson, M. et al., 1990. HIV-1 replication is controlled at the level of T cell activation and proviral integration. *The EMBO journal*, 9(5), pp.1551–1560.

Thomson, M.M. et al., 2005. Identification of a novel HIV-1 complex circulating recombinant form (CRF18_cpx) of Central African origin in Cuba. *AIDS (London, England)*, 19(11), pp.1155–1163.

Tjabringa, G.S. et al., 2005. Host defense effector molecules in mucosal secretions. *FEMS immunology and medical microbiology*, 45(2), pp.151–158.

Trumpfheller, C. et al., 2003. Cell type-dependent retention and transmission of HIV-1 by DC-SIGN. *International immunology*, 15(2), pp.289–298.

Turville, S. et al., 2003. The role of dendritic cell C-type lectin receptors in HIV pathogenesis. *Journal of leukocyte biology*, 74(5), pp.710–718.

Turville, S.G. et al., 2004. Immunodeficiency virus uptake, turnover, and 2-phase transfer in human dendritic cells. *Blood*, 103(6), pp.2170–2179.

Valenzuela, A. et al., 1997. Neutralizing antibodies against the V3 loop of human immunodeficiency virus type 1 gp120 block the CD4-dependent and -independent binding of virus to cells. *Journal of virology*, 71(11), pp.8289–8298.

Van`Montfort, T. et al., 2007. Efficient capture of antibody neutralized HIV-1 by cells expressing DC-SIGN and transfer to CD4+ T lymphocytes. *Journal of immunology*, 178(5), pp.3177–3185.

Van`Montfort, T. et al., 2011. HIV-1 N-glycan composition governs a balance between dendritic cell-mediated viral transmission and antigen presentation. *Journal of immunology*, 187(9), pp.4676–4685.

Varmus, H., 1988. Regulation of HIV and HTLV gene expression. *Genes & development*, 2(9), pp.1055–1062.

Wang, X. et al., 2009. Monitoring alpha4beta7 integrin expression on circulating CD4+ T cells as a surrogate marker for tracking intestinal CD4+ T-cell loss in SIV infection. *Mucosal immunology*, 2(6), pp.518–526.

Wawer, M.J. et al., 2005. Rates of HIV-1 transmission per coital act, by stage of HIV-1 infection, in Rakai, Uganda. *The Journal of infectious diseases*, 191(9), pp.1403–1409.

Weissenhorn, W. et al., 1997. Atomic structure of the ectodomain from HIV-1 gp41. *Nature*, 387(6631), pp.426–430.

Wilen, C.B. et al., 2011. Phenotypic and immunologic comparison of clade B transmitted/founder and chronic HIV-1 envelope glycoproteins. *Journal of virology*, 85(17), pp.8514–8527.

Wilen, C.B., Tilton, J.C. & Doms, R.W., 2012. HIV: cell binding and entry. *Cold Spring Harbor perspectives in medicine*, 2(8), pp.1–13.

Willey, R.L. et al., 1989. Functional interaction of constant and variable domains of human immunodeficiency virus type 1 gp120. *Journal of virology*, 63(9), pp.3595–3600.

Wira, C.R. & Fahey, J. V, 2008. A new strategy to understand how HIV infects women: identification of a window of vulnerability during the menstrual cycle. *AIDS*, 22(15), pp.1909–1917.

De Witte, L. et al., 2007. Langerin is a natural barrier to HIV-1 transmission by Langerhans cells. *Nature medicine*, 13(3), pp.367–371.

Wormald, M.R. & Dwek, R.A., 1999. Glycoproteins: glycan presentation and protein-fold stability. *Structure*, 7(7), pp.155–160.

Wu, L. et al., 2004. Raji B cells, misidentified as THP-1 cells, stimulate DC-SIGN-mediated HIV transmission. *Virology*, 318(1), pp.17–23.

Wu, L. & KewalRamani, V.N., 2006. Dendritic-cell interactions with HIV: infection and viral dissemination. *Nature reviews Immunology*, 6(11), pp.859–868.

Wyatt, R. et al., 1995. Involvement of the V1/V2 variable loop structure in the exposure of human immunodeficiency virus type 1 gp120 epitopes induced by receptor binding. *Journal of virology*, 69(9), pp.5723–5733.

Wyatt, R. et al., 1998. The antigenic structure of the HIV gp120 envelope glycoprotein. *Nature*, 393(6686), pp.705–711.

Yu, H.J., Reuter, M. a & McDonald, D., 2008. HIV traffics through a specialized, surface-accessible intracellular compartment during trans-infection of T cells by mature dendritic cells. *PLoS pathogens*, 4(8), pp.1–14.

Zanetti, G. et al., 2006. Cryo-electron tomographic structure of an immunodeficiency virus envelope complex in situ. *PLoS pathogens*, 2(8), pp.790–797.

Zeller, Y., Mechtersheimer, S. & Altevogt, P., 2001. Critical amino acid residues of the alpha4 subunit for alpha4beta7 integrin function. *Journal of cellular biochemistry*, 83(2), pp.304–319.

Zhang, M. et al., 2004. Tracking global patterns of N-linked glycosylation site variation in highly variable viral glycoproteins: HIV, SIV, and HCV envelopes and influenza hemagglutinin. *Glycobiology*, 14(12), pp.1229–1246.

Zhang, Z. et al., 1999. Sexual transmission and propagation of SIV and HIV in resting and activated CD4+ T cells. *Science*, 286(5443), pp.1353–1357.

Zhu, P. et al., 2006. Distribution and three-dimensional structure of AIDS virus envelope spikes. *Nature*, 441(7095), pp.847–852.