The role of HIV-1 subtype B Envelope transmission motifs in subtype C variant infectivity

Bahiah Meyer

Dissertation submitted in fulfilment of the requirements for the degree of

MSc (Med) in Medical Biochemistry

in the Department of **Integrative Biomedical Sciences**

Faculty of Health Sciences

UNIVERSITY OF CAPE TOWN

February 2018

Supervisor: Dr Zenda Woodman

DECLARATION

I, Bahiah Meyer, hereby declare that the work on which this dissertation/thesis is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university.

I have used the referencing style from the journal *Nature*. Each significant contribution to, and quotation in, this dissertation from the work(s) of other people has been attributed, cited and referenced.

I empower the university to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signature: Signed by candidate

Date: 16 February 2018

CONTENTS

DECLARATION **2**

ACKNOWLEDGEMENTS **6**

ABSTRACT **7**

SUMMARY **8**

LIST OF ABBREVIATIONS **11**

CHAPTER 1: BACKGROUND AND LITERATURE REVIEW **14**

1.1 Introduction **14**

1.2 Diversity, genome organisation, life cycle and tropism of HIV **15**

1.3 Structure and function of gp120 ..... **17**

1.4 Processing of Envelope **19**
1.4.1 Role of the signal peptide in directing processing of Env 22
1.4.2 N-glycosylation in the ER and compartments of the Golgi 27

1.5 Env transmission motifs **31**
1.5.1 Genetic bottleneck associated with HIV transmission 31
1.5.2 Characteristics of the TF Env 32

1.6 Concluding statements **35**

1.7 Study rationale and research objectives **36**
1.7.1 Aim 37
1.7.2 Objectives 37

CHAPTER 2: MATERIALS AND METHODS **38**

2.1 Envelope constructs **38**

2.2 Cloning and site-directed mutagenesis **38**
2.2.1 Site-directed mutagenesis 38
2.2.2 Bacterial cell culture, plasmid transformation and screening 39
2.2.3 Restriction enzyme digestion 40
2.2.4 Agarose gel electrophoresis 40
2.2.5 Sequence analysis 41

2.3 Mammalian cell culture **41**

2.4 Transient expression, secretion and N-glycosylation analysis of Env clones 42
2.4.1 Plasmid preparation and transfection 42
2.4.2 Cell lysis and measuring transfection efficiency 43
2.4.3 SDS-PAGE and Western blotting 43
2.4.4 Secretion, cleavage efficiency and N-glycosylation analysis of Env 45

2.5 Pseudovirus production and p24 assay 46
2.5.1 Transfections 46
2.5.2 p24 enzyme-linked immunosorbent assay (ELISA) 46

2.6 Entry efficiency assay 47

2.7 Production of chimeric Infectious Molecular Clones (IMCs) 48
2.7.1 PCR amplification of the *env* gene 48
2.7.2 Yeast gap-repair assay 49
2.7.3 Colony PCR 50
2.7.4 Transformation of electro-competent *E. coli* cells 51
2.7.5 Testing envelope functionality of chimeric IMCs 51

2.8 Replication assay 52
2.8.1 Isolation of peripheral blood monocytic cells (PBMCs) 52
2.8.2 Transfection to produce infectious virus 53
2.8.3 Tissue Culture Infectious Dose (TCID50) for titration of infectious virus 53
2.8.4 Testing PBMC donor selection 54
2.8.5 Viral expansion in activated donor PBMCs 54
2.8.6 Concentrating IMCs by ultra-centrifugation 55
2.8.7 Infection of activated PBMCs for replication kinetics 55

2.9 Statistical analyses 55

CHAPTER 3: IMPACT OF SUBTYPE B TRANSMISSION MOTIFS ON SUBTYPE C ENVELOPE FUNCTION 57

3.1 Introduction 57

3.2 Results 59
3.2.1 Site-directed mutagenesis of the TF *env* signal peptide 59
3.2.3 Envelope expression 62
3.2.4 Envelope secretion 66
3.2.5 Envelope cleavage 68
3.2.6 Envelope N-glycosylation 70
3.2.7 Pseudovirion entry efficiency 74
3.2.8 Entry efficiency of chimeric Infectious Molecular Clones 77

3.3 Discussion 78

3.4 Conclusion 86

CHAPTER 4: IMPACT OF SUBTYPE B TRANSMISSION MOTIFS ON SUBTYPE C REPLICATION 87

4.1 Introduction 87

4.3 Results 88
4.3.1 Amplification of the *env* gene from CAP210 E8 constructs 88
4.3.2 Screening yeast for homologous recombination by colony PCR 93
4.3.3 Testing functionality of chimeric viruses using entry efficiencies into TZM-bl cells 94
4.3.4 Testing donor PBMCs that can support replication of HIV-$1_{pNL4.3}$ 94
4.3.5 Replication of expanded IMCs using donor PBMCs 95

4.4 Discussion 97

CHAPTER 5: SUMMARY AND CONCLUSIONS 100

REFERENCES 102

APPENDIX I - SUPPLEMENTARY FIGURES 125

APPENDIX II - SOLUTIONS 127

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to:

My supervisor, Dr Zenda Woodman for her guidance throughout my undergraduate years and this project, her enthusiastic approach to doing science and her continuous support and understanding.

Dr Melissa-Rose Abrahams and Samuel Kariuki from Prof. Carolyn Williamson's group in IIDMM, UCT for help with the replication assay and use of their ultracentrifuge.

Lab members Evelyn Ngwa and Shatha Omar for their guidance in the lab and for their willingness to teach.

Members of Prof. Virna Leaner's group for all their support both in the form of reagents or much needed coffee breaks. Mr Robert "Robbie" Samuels for the laughs, prayers and early morning chats for encouragement. Riley Traviss for the friendship and helping to create a pleasant, engaging working environment in the lab.

My close friends and family for their consistent love and support in every form. My gratitude is endless.

The Winifred McKinnon Trust, the National Research Foundation (NRF), the NRF/DAAD Scholarship, the Poliomyelitis Research Fund (PRF), and the UCT/Carnegie Scholarship for the financial support.

"If you are grateful, I will surely increase you [in favour]" – *Surah Ibrahim*, 14:7

ABSTRACT

Transmitted founders (TF) might carry motifs that provide a phenotypic advantage that enables human immunodeficiency virus type-1 (HIV-1) to overcome immune barriers within the female genital tract. One study compared over 5000 subtype B TF and mismatched chronic infection *envelope* (*env*) sequences and identified two putative transmission motifs: Histidine at position 12 of the signal peptide (His12) and a potential N-glycan site (PNG) at position 413-415. Although, His12 was shown to be important for subtype B Env expression and viral infectivity, in our own sequence analysis subtype C variants did not carry the transmission motifs and the aim of this study was to determine whether His12 and PNG413 was important for subtype C Env expression, processing, function and viral replication. Mutagenesis of a subtype C Env clone indicated that His12 decreased pseudovirion (PSV) entry efficiency without influencing Env expression, secretion and cleavage with no changes in the N-glycosylation profile. This suggested that His12 had a fitness cost and was thus selected against. However, His12 significantly enhanced the entry efficiency of infectious molecular clones (IMCs), suggesting that it might be beneficial for *in vivo* replication. The variation between the PSV and IMC entry of TZM-bl cells could be due to differences in assay conditions. On the other hand deletion of PNG413 enhanced Env expression, secretion, cleavage and PSV and IMC entry efficiency of TZM-bl cells. This would suggest that subtype C TFs carrying a PNG at 413-413 would have lower viral replicative capacity due to poor expression and processing of Env. The benefit of this phenotype on HIV-1 subtype C transmission needs to be further investigated. Unfortunately, PSV and IMC entry of TZM-bl cells could not be confirmed by IMC replication in peripheral blood monocytes because the clones could not replicate to measurable levels in these cells over the culture period. Overall, this study has shown that amino acid residues at positions 12 and 415 do play a role in modulating Env processing and function however the actual mechanism by which these polymorphisms impact viral fitness most likely differ to that of subtype B, explaining why His12 is absent and PNG413 is present in subtype C TFs.

SUMMARY

It has been reported that in approximately 80% of HIV-1 transmission cases, clinical infection is established by a single variant, the transmitted founder (TF). These authors suggested that TFs might carry phenotypes advantageous for transmission. A study which analysed over 5000 *env* sequences of TFs and chronic from subtype B variants circulating in the USA and identified two robust potential transmission motifs: one in the signal peptide and another site that coincided with a potential N-glycosylation site (PNG) at position 413-415 of Envelope (Env), the glycoprotein responsible for binding to host receptors and mediating entry into target cells. Subtype B TFs carried a histidine at position 12 (His12) of the signal peptide of Env (Env-SP) and lacked a PNG at position 413-415 (PNG413). Two follow-up studies suggested that His12 increased expression levels of Env and thus infectivity. An increase in Env expression has been previously associated with an alteration in the N-glycosylation profile of Env, such that predominantly high mannose glycoforms are added. These changes in N-glycosylation on Env could alter the binding between Env and host receptors which could then facilitate transmission. However, in subtype C Env, a Glutamine was the preferred amino acid at position 12 (Q12) and PNG413 was present. This suggested that subtype C TF might not be under the same selective pressure as subtype B variants during transmission. In this study, a subtype C TF variant was mutated by site-directed mutagenesis to resemble a subtype B transmitted variant: The Glutamine in the Env-SP was changed to either a Histidine (Q12H) or Alanine at position 12 (Q12A) and the PNG at 413 was deleted by changing the Threonine at 415 to Isoleucine (T415I). Env expression, secretion, cleavage and N-glycosylation of the mutants were compared to wild-type (WT) by measuring cell-associated gp160 and gp140 or gp140 and gp120 secreted into the culture medium by Western blotting. Pseudovirus (PSV) and infectious molecular clones (IMC) generated by yeast homologous recombination and that carried WT or mutated Env were used to compare Env entry efficiency of TZM-bl cells by luciferase reporter assay. Western blots showed that after transient transfection of HEK 293T cells, there was no significant increase in expression or secretion of Q12H. When Glutamine was replaced with another neutral amino acid, there was also no change in expression or secretion, suggesting that changes in amino acids at this site was not influencing retention of Env within the endoplasmic reticulum (ER) as previously suggested. Env is highly glycosylated

and it has been suggested that changes in ER retention influenced the level of gp160 mannosylation. When gp140 N-glycosylation was compared using endoglycosidase digestion, there was no difference in mannosylation of Q12H and Q12A relative to WT, indicating that Env-SP mutations were not influencing processing within the ER. This was expected as expression and secretion were also not affected. However, Q12A showed a 2-fold decrease in gp140 cleavage relative to WT although this did not reach significance. This could indicate that a small non-polar amino acid at position 12 within the signal peptide influenced gp160 cleavage in the *trans*-Golgi without altering ER-associated processing and trafficking to the plasma membrane. Alternatively, gp140 could be cleaved and processed differently compared to gp160 and thus cleavage of gp140 might not represent that of full-length Env. As we were unable to detect cell-associated gp120 we could not confirm cleavage differences between Q12H and Q12A with WT.

On the other hand, the T415I mutant showed a significant rise in Env expression, secretion and gp140 cleavage compared to WT. This suggested that either an N-glycan at position 413 or the presence of a Threonine at positon 415 was important for Env processing. Changes in Env post-translational processing could be very important for the production and trafficking of native trimers to the cell surface, altering incorporation of functional Env and thus viral infectivity. When PSVs, normalised to 250ng/ml p24 infected TZM-bl cells, Q12H and Q12A had significantly lower entry efficiency than WT whereas T415I entered the reporter cell line with significantly better competency. Therefore, T415I higher expression, secretion and cleavage could have enhanced entry efficiency by increasing the number of functional Env incorporated into PSVs. Similarly, the decreased cleavage of Q12H and Q12A might have lowered PSV entry efficiency by the incorporation of more non-functional gp160. As PSV entry efficiency might not mimic that of viruses able to undergo multiple rounds of infection, TZM-bl cells were also infected with chimeric IMCs. All IMC mutants had greater entry efficiency than WT, although only that of Q12H and Q12A reached significance. Data of T415I described a potential pathway of how changes at this site increased cell-associated levels which led to enhanced trafficking so that more cleaved, functional trimers were incorporated into both PSVs and IMCs. However, the entry efficiency of Q12H and Q12A differed when PSVs and IMCs were compared. A potential reason for this discrepancy could be experimental. To determine whether this difference could be due to the limitation of the TZM-bl cell line to accurately

reflect *in vivo* infection, IMCs were replicated in PBMCs from 2 donors. However, only the positive control, pNL4 and the signal peptide mutant, Q12A showed detectable virus titres after p24 ELISA on days 7, 10 and 14. Therefore a comparison between entry efficiency in the TZM-bl cell line versus replication in donor PBMCs could not be made.

Irrespective of differences between PSVs and IMCs for the signal peptide mutants, our findings did not support those of the previous studies. These authors reported that His12 was required for increased subtype B TF Env expression, incorporation and PSV/IMC infectivity confirming its role as a transmission motif. Our findings suggest that unlike the Histidine at position 12 of subtype B TF Env signal peptides, this amino acid is not involved in subtype C expression, secretion and N-glycosylation. However, the absence of the PNG413 which led to increased viral entry of the subtype C TF variant, suggests that this motif may play a role in transmission. However, it is unclear why subtype C TFs would carry a PNG at position 413 unless low Env expression is an advantage during transmission. Taken together, Env transmission motifs of subtype C HIV-1 might differ from that of subtype B, suggesting that vaccines that protect against one subtype might not protect against another. Furthermore, immune pressure might drive the loss or gain of transmission motifs which might alter viral fitness and impact disease progression of some individuals.

LIST OF ABBREVIATIONS

~	Approximately
°C	Degrees Celsius
µg	Micrograms
µL	Microliters
µm	Micrometre
µM	Micromolar
bp	Base pair
hr(s)	Hour(s)
kb	Kilobases
kDa	Kilo-Daltons
mg	Milligrams
min(s)	Minute(s)
mL	Milliliters
mM	Millimolar
ng	Nanograms
pg	Picograms
sec	Second(s)
U	Unit(s)
V	Volts
w/v	Weight per volume
α	Alpha
β	Beta
Δ	Delta
BSA	Bovine serum albumin
$CaCl_2$	Calcium chloride
CAPRISA	Centre for the AIDS Programme of Research in South Africa
CCR5	C-C chemokine receptor type 5 CD4/8 Cluster of differentiation 4/8
CD4	Cluster of differentiation 4
CHO	Chinese hamster ovary cell

CO_2	Carbon dioxide
CTL	Cytotoxic T Lymphocytes
CXCR4	C-X-C chemokine receptor type 4
DCs	Dendritic cells
DC-SIGN	DC-specific ICAM-3-grabbing nonintegrin
dH_2O	Distilled water
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleoside triphosphate
E. coli	*Escherischia coli*
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
Env	Envelope
ER	Endoplasmic reticulum
FBS	Fetal bovine serum
HEK	Human embryonic kidney
HIV-1	Human Immunodeficiency virus type 1
IBMS	Integrative Biomedical Sciences
IgG	Immunoglobulin G
IIDMM	Institute of Infectious Diseases and Molecular Medicine
IMC(s)	Infectious molecular clone(s)
KCl	Potassium chloride
LB	Luria broth
MCB	Department of Molecular and Cell Biology
$MgCl_2$	Magnesium chloride
$MgSO_4$	Magnesium sulphate
MMP	Methyl α-D-Mannose-Pyranoside
nAb	Neutralizing antibodies
NaCl	Sodium chloride
$NaHCO_3$	Sodium bicarbonate
NICD	National Institute of Communicable Disease
PAGE	Polyacrylamide gel electrophoresis

PBMCS	Peripheral blood mononuclear cells
PBS	Phosphate buffered saline PCR Polymerase chain reaction
PEI	Polyethylenimine
PNGS	Potential N-linked glycosylation site(s)
RLU	Relative light units
rpm	Revolutions per minute
SDS	Sodium dodecyl sulfate
SGA	Single genome amplification
SIV	Simian immunodeficiency virus
S. cerevisae	*Saccharomyces cerevisae*
T/F	Transmitted/ founder
UCT	University of Cape Town
USA	United States of America

CHAPTER 1: BACKGROUND AND LITERATURE REVIEW

1.1 Introduction

Identified over 30 years ago[1], human immunodeficiency virus type-1 (HIV-1) has been extensively characterised with emphasis on uncovering virological traits which contribute to its success in overcoming the protection strategies of the host cell. The major challenge in the design of a potential vaccine is the high degree of genetic diversity that exists between HIV-1 variants, leading to the classification of HIV-1 into at least 9 different subtypes. Subtype B is the most prevalent in Northern America and Europe while subtype C accounts for more than 60% of infections globally[2,3]. The high genetic diversity of HIV-1 circulating globally, suggests that a vaccine against one subtype is unlikely to prevent infection by another.

Recently, however it has been shown that despite the high genetic diversity within the donor[4] in the majority of transmission cases only a single virus, termed the transmitted founder (TF), leads to productive, clinical infection[5]. This genetic bottleneck at transmission has suggested that TFs carries a transmission motif that provides a selective advantage over other variants during transmission[6,7]. The most likely candidate protein to carry this signature[8] would be *env* encoding the viral Envelope (Env) as this glycoprotein mediates the initial contact between the virus and target cells[9–11]. However, this does not exclude the possibility of other viral proteins contributing to the genetic bottleneck during HIV transmission.

Studies aimed at revealing the potential transmission signatures within Env and the underlying molecular events[8,12,13], particularly at the mucosal surface, have been underway for over a decade now, with the aim of bringing researchers closer to uncovering the genetic composition of TFs[5,14] that will ultimately inform vaccine design.

So far, TFs of subtypes A and C have been shown to carry Envs with shorter variable loop (V-loop) length, have reduced potential N-glycosylation sites (PNGS) and increased susceptibility to neutralizing antibodies[15]. Extensive sequence analysis of subtype B *env* identified two potential transmission motifs: a PNG at position 413-415 and a single amino acid at position 12 of the signal peptide. PNG415 was identified to be involved in neutralization sensitivity and

under-represented in subtype B TFs[8]. The same authors also identified that the signal peptide of TF Envs was enriched for positively charged, basic amino-acids commonly Histidine (H) or alternatively Arginine (R) at position 12. The fact that these two polymorphisms differ between TFs and variants at chronic stages of infection suggests that it may play a role in HIV transmission and potentially Env function and/or infectivity. This review will focus on the role of Env and particularly its signal peptide and N-glycosylation, in HIV transmission and pathogenesis.

1.2 Diversity, genome organisation, life cycle and tropism of HIV

HIV-1 can be classified into 4 phylogenetic groups namely M, N, O and P. Of these, M resulted in the pandemic spread of the virus resulting in the emergence of the virus as nine distinct subtypes (A, B, C, D, E, F, G, H, and K). Particular subtypes were later found to be localised within certain regions of the world, with subtype C dominating 50% of all infections in sub-Saharan Africa and subtypes F, H, J and K only accounting for 1% of infections worldwide[16,17].

It is estimated that an average of $10.3x10^9$ virions can be produced per day and that 2.6 days are required for a newly released virion to infect another cell and cause the release of a new population of viral particles[18]. The life-cycle of HIV has a half-life of 1.6 days in productively infected cells[18] and involves a series of steps that include the delivery of two copies of the viral single-stranded RNA (ssRNA) into host cells[19]. Upon release of the ssRNA into the cell, the 9.7 kilobase (kb) viral genome is converted into complimentary DNA (cDNA) by the viral enzyme, reverse transcriptase (RT) before it is integrated into the host chromosomal DNA by the viral integrase. The viral genes (of which there are nine) are then transcribed into messenger RNA (mRNA) followed by synthesis of all the structural and accessory viral proteins (Figure 1.1)[20]. HIV has a total of six accessory proteins, namely Vif, Vpr, Tat, Rev, Vpu and Nef, that function as trans-activators and regulators of gene expression[21].

The three structural proteins, Gag, Pol and Env[19] are crucial for the assembly of infectious particles. The *pol* gene encodes the protease, RT and integrase. The protease is required for the proteolytic processing of polyprotein precursors into mature functional proteins[22,23]. The *gag* gene encodes the polyprotein Gag-p55[24,25] which is cleaved into: 1) the matrix protein

(p17) that associates with the host membrane and initiates the formation of budding viral particles[26]; 2) the capsid protein (p24) that encloses the viral genome and is important in assembly of infectious viral particles[27]; 3) a protein, p6, that mediates incorporation of the accessory protein, Vpr into budding viral particles[28] and 4) the nucleocapsid (p7) that is important for viral RNA transport and packaging as well as ensuring efficient reverse transcription and viral infectivity[29,30].

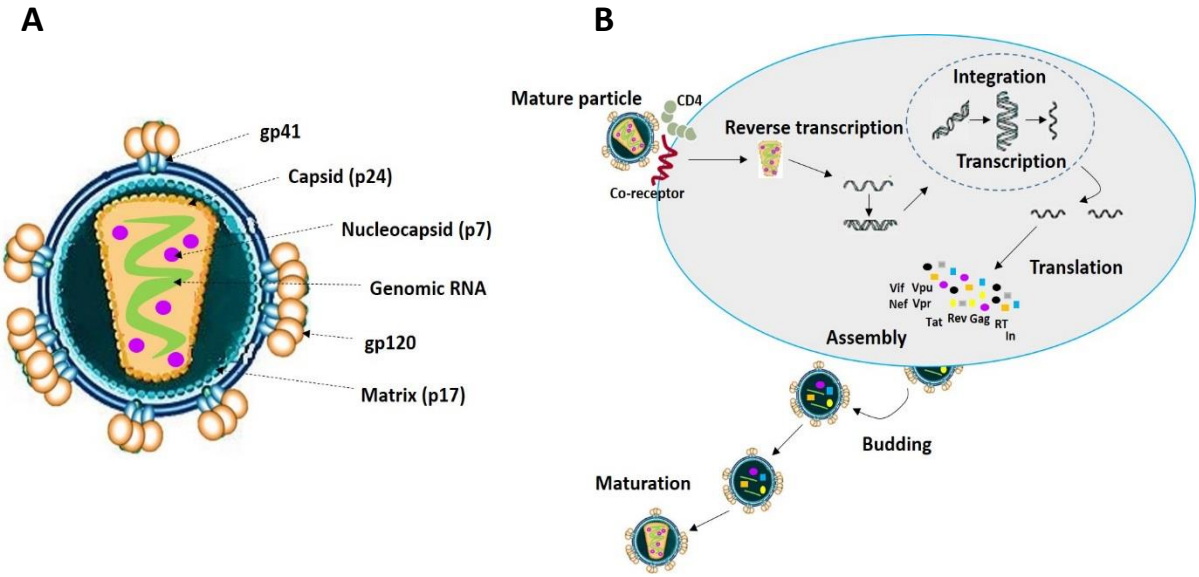

Figure 1.1: HIV-1 structure and life cycle. (A) The Env transmembrane unit (gp41) and surface unit (gp120), the capsid, matrix, nucleocapsid and the viral genome are shown[31]. Picture based on Robinson, 2002. **(B)** Steps involved in the life cycle of HIV: after dual-binding of the incoming mature virus particle to CD4 and a chemokine co-receptor on target cells, the viral RNA genome is released into the host cell and is reverse transcribed into cDNA. The viral cDNA then integrates into the host genome. mRNA transcripts of the viral genes are then transported to the cytoplasm where they are translated. Finally, assembly and budding of the viral particle is initiated and these immature particles undergo maturation[32].

The *env* gene encodes a 160 kDa polyprotein, a precursor to the surface subunit, gp120 and the transmembrane subunit, gp41. These subunits interact non-covalently to form a trimer of heterodimers that make up the mature Env on the surface of the virus (Figure 1.1A)[33,34]. Env is crucial for mediating the initial steps of viral replication in T lymphocytes, i.e. binding to the CD4 receptor and host cell co-receptor and fusion of viral and host cell membranes [35,36]. Env determines viral tropism by recognising host cell-surface receptor, CD4, together with either CCR5 and/or CXCR4 chemokine receptors[21]. Virus strains that only infect *via* CCR5 co-receptors are termed R5-variants whilst those which only infect *via* CXCR4 are termed X4-

variants. Dual tropism has also been observed where Env is able to mediate infection by binding to both CCR5 and CXCR4[37,38]. Gp120 has been shown to interact with other co-receptors such as CCR3 and CX3CR1[36], however CCR5 and CXCR4 have been identified as the major co-receptors involved in transmission, as inhibitors which block interactions by these co-receptors with gp120 also prevent HIV infection[36]. Furthermore, individuals homozygous for the *ccr5*Δ32 polymorphism are resistant to infection by R5 tropic HIV[39].

The dual-binding of gp120 to CD4 and CCR5 or CXCR4, causes a conformational change in gp120 allowing gp41, which has a rod-like structure, to penetrate the cell membrane of the target cell enabling membrane fusion, viral entry ad productive infection[10,11]. On the other hand, dendritic cells (DCs) present in mucosal tissues and form the first line of defence against pathogens, transport the virus to HIV-permissive cells[40,41]. Immature DCs have lower levels of CD4 and CCR5 expression when compared to T-lymphocytes[42] but express high levels of dendritic cell–specific ICAM grabbing non integrin (DC-SIGN) receptor on their surface that binds gp120 without leading to viral entry. Instead, DC-SIGN-gp120 interactions are stable enough for DC migration to transport HIV-1 from the genital tract mucosa to the lymph nodes leading to trans-infection of T-lymphocytes and systemic distribution of the virus[43,44].

1.3 Structure and function of gp120

Gp120 consists of variable loops (V1-V5) and constant domains (C1-C5)[45,46]. Following cleavage, gp120 and gp41, interact non-covalently and are assembled in a trimeric complex[47]. The gp120 and gp41 interactive region is made up of a 7-stranded β-sandwich and the termini of gp120 anchor the surface subunit in the viral spike[48,49]. There are a number of highly conserved cysteine residues positioned throughout gp120 and gp41 that form intramolecular bonds important for Env trimer structure[50]. The cysteine residues within gp41 are sites of palmitoylation that ensure trafficking of Env trimers to lipid rafts for incorporation within viral particles. However, as not all Envs carry Cysteines within gp41 cytoplasmic tails, they are not essential for HIV infectivity[51].

Understanding the three-dimensional structure of Env can help identify mechanisms of function and regions ideal for vaccine design. A number of studies have investigated the

crystal structure of gp120 and gp41 as separate subunits, monomers and trimers either alone or complexed with receptors, ligands or inhibitors[35,49,52–57]. The crystal structure of gp120 (Figure 1.2A) shows the organization of gp120 into inner (in orange) and outer (in blue) domains that form the CD4 receptor binding site (in red) at their interface. C1, C3 and C4 of gp120 bind CD4 whilst the variable domains have no known direct role in interactions with the CD4 receptor[58–60]. A 4-stranded β-sheet subdomain termed the bridging sheet (in green) is composed of two β-strands from the outer domain and the V1/V2 stem from the inner domain. The bridging sheet is formed as a consequence of CD4 binding, by the rearrangement of the two pairs of double stranded β sheets, one from the inner and one from the outer domain[35,46]. The formation of this bridging sheet reveals a previously concealed co-receptor binding site (for dual-binding) and the gp41 transmembrane stalk[61]. The gp120 and gp41 themselves interact *via* the 7 β-strands of the inner domain and the extended N/C-termini[62].

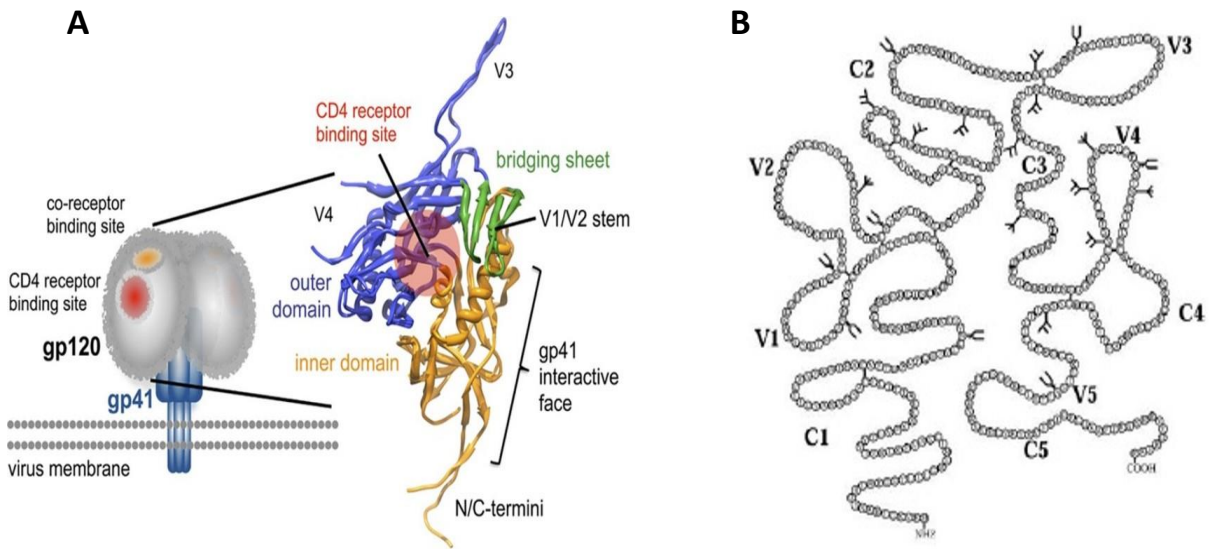

Figure 1.2: The core structure of HIV-1 Envelope glycoproteins, gp120 and gp41. (A) Gp120 on the surface of the virus membrane in the CD4-bound conformation. The ribbon diagram illustrates the hypervariable or V-loops contained within gp120 that can be modified during the course of infection to optimise utilization of CCR5 receptors and enhance resistance against neutralizing antibodies produced in the infected host. The antigenic regions and CD4 receptor binding site (in red) are also indicated here. Taken from Guttman *et al.* 2012. (B) A schematic representation of HIV gp120 variable and constant domains. Gp120 carries an array of N-glycans (Y high oligomannose or Y hybrid or complex and Y unknown) believed to shield antigenic epitopes from neutralizing antibodies. Adapted from McCaffrey *et al*. 2003 and Leonard *et al.* 1990.

The crystal structure of gp41 indicated that the conformational changes induced by gp120 binding to CCR5 triggered the folding of gp41 into a six helical bundle structure that juxtaposes host cell and viral membranes. The fusion peptide then interacts with the plasma membrane, creating a fusion pore which is enlarged by the membrane proximal region (MPER), and allowing membrane fusion[63,64]

The sequence and structural changes in the hypervariable loops of the HIV-1 gp120 proteins have been the subject of a number of studies[15,62,65,66]. It has been suggested that the V2 loop contains a signal tripeptide Leucine (L)-Aspartate (D)-Valine (V) that is important in the infection of memory CD4+ T-lymphocytes and spread of HIV-1 in the gastrointestinal tract[12,67] whilst the V3 loop is important for membrane fusion[68], a prominent target for neutralising antibodies[69] and plays a role in co-receptor interactions[70]. Mutations in the V3 loop have been suggested to influence the switch from R5 to X4 tropism by altering the overall positive charge of this region and thus allowing for the interaction with the negatively charged CXCR4 co-receptor[71–73]. The biological properties that these changes may influence such as efficiency of transmission and infectivity of the virus has been extensively studied, particularly in response to the pressure exerted by host immune responses[74–76].

1.4 Processing of Envelope

Processing of the Env precursor occurs both co-translationally and at the post-translational level. The precursor is synthesised as gp160 monomers in the endoplasmic reticulum (ER) and modified by Asparagine (N)-linked glycosylation. These N-glycans are then further modified in the compartments of the Golgi after trimerisation and before cleavage and incorporation of functional trimers into virions[77]. Env is targeted to the rough endoplasmic reticulum (RER) by a signal peptide comprising 30 amino acids long with an N-terminal region containing five positive amino acids[78]. Approximately 20-35 potential N-glycosylation sites (PNGS) have been identified in gp160[79] and N-linked glycosylation contributes to approximately 50% of the molecular weight of gp120[58,80]. N-glycosylation occurs simultaneously with translation[81] along with O-linked glycosylation[82] by membrane-bound oligosaccharyl transferases found in the ER[83] (Figure 1.3) generating gp160 trimers carrying high mannose N-glycans. The high mannose residues are then modified in the Golgi network and gp160 trimers are cleaved into

gp120 and gp41 in the *cis* or *medial* Golgi before insertion into the plasma membrane for incorporation, virion assembly, maturation and budding[84,85]. Cleavage of gp160 trimers is essential for viral infectivity seeing that it releases the fusion peptide, essential for membrane fusion[86]. The mechanism of Env trimer incorporation within infectious virions is not well characterised however the cytoplasmic tail of gp41 and the Gag matrix is thought to play a role[87]. It has been estimated that approximately 10 Env trimers are incorporated per virion due to reduced Env levels within the plasma membrane because of rapid endocytosis of Env trimers and shedding of gp120 into the culture medium[88]. The potential advantage to this low number is possibly to limit detection by immune responses especially noting that increasing the number of Env spikes above a particular threshold did not lead to increased HIV infectivity[89].

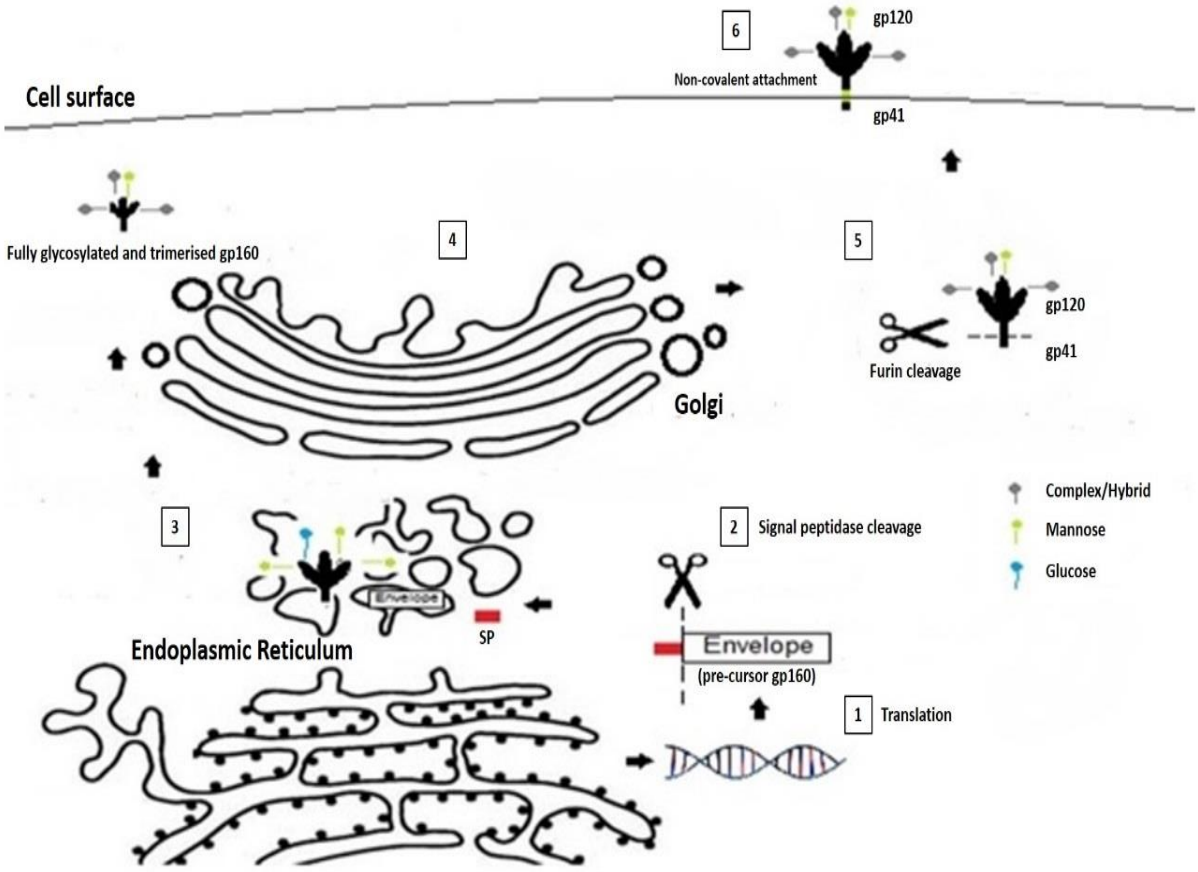

Figure 1.3: Schematic diagram of the overall processing and trafficking of HIV-1 Envelope. 1) During translation, the precursor Env (gp160) is directed to the rough endoplasmic reticulum (RER) by its signal peptide (SP) found on the N-terminal domain of the nascent polypeptide. 2) The SP is cleaved by a signal peptidase co-translationally and Env is then released into the lumen of the RER. 3) During folding, precursor N-glycan structures are added, followed by mannose addition and and trimerization. 4) The Env then translocates through the various compartments of the Golgi where its glycosylation is further modified in a sequence of reactions. 5) The fully glycosylated and trimerized Env is then cleaved by host cell proteases within the *trans*-Golgi network into gp120 and gp41 subunits or inside *trans*-Golgi-derived secretory vesicles. The surface unit, gp120, and transmembrane unit, gp41 associate *via* non-covalent interactions. 6) The cleaved trimers accumulate at the cell membrane of the host cell where they await the other viral proteins for assembly and release of the virus particle[85,90–92].

1.4.1 Role of the signal peptide in directing processing of Env

1.4.1.1 Overview of the mammalian secretion system

The eukaryotic secretion system involves the translocation of proteins into a specialised intracellular organelle, the ER and then to the cell surface *via* the Golgi apparatus[93]. In the ER, proteins can be modified either during (co-translationally) or after synthesis (post-translationally). Modifications include: proline hydroxylation, N-linked glycosylation and/or disulphide bond formation for stability upon release into the extracellular environment[94]. Misdirected or misfolded proteins on the other hand are secreted to the cytosol (away from the ER) where they are degraded[94,95].

1.4.1.2 Structure and function of the signal peptide in mammalian secretion systems

Signal peptides (SPs) are usually situated at the N-terminal end of nascent secretory and membrane proteins[96]. SPs function to direct the transport of proteins destined for the plasma membrane to the ER[97] as well as other post-targeting functions[98]. It is commonly comprised of 16-30 amino acids and is necessary in the recognition of the nascent protein by signal recognition protein (SRP). It has a tripartite structure (Figure 1.4) consisting of a hydrophilic and usually positively charged N-terminal region, a central hydrophobic region important for targeting to and insertion into the ER membrane[99], and a C-terminal region with the cleavage site for the signal peptidase (SPase)[100]. For cleavage to occur the -1 position must contain an amino acid residue with a short side chain whilst position -3 should not contain any charged amino acids[99]. The SP is cleaved in the ER by the SPase, made up of five polypeptides forming a hetero-oligomeric complex, following transport to the ER. Even with these common characteristics, signal peptides do not share any sequence similarity and may differ in length as well[101] with the N-terminal region displaying the greatest diversity in amino acid composition[102].

Figure 1.4: Tripartite structure of the N-terminal signal sequence found on nascent secretory and membrane proteins. The signal sequence is comprised of a hydrophilic and usually positively charged N-terminal region (n-region), a central hydrophobic region (h-region) and a C-terminal region with the cleavage site for the signal peptidase (c-region)[98].

Upon arrival at the ER, the signal sequence inserts into the membrane in a loop-like fashion where the N-terminal portion remains on the cytosolic side and the C-terminal part emerging into the lumen of the ER[103–105]. The signal sequence is usually rapidly cleaved off co-translationally in the ER lumen. However in some instances where the signal sequence influences the timing of folding and association with chaperones in the ER, cleavage occurs several minutes after the appearance of the protein in the ER[106]. This slow cleavage of the signal sequence has thus been identified as a key determinant for the retention of certain proteins in the ER during the folding pathway[107].

Short signal sequences function only in the targeting of nascent proteins for insertion into the ER membrane and can be rapidly degraded. Good examples of these are the signal sequences of serum albumin and vesicular stomatitis virus G-protein (VSV‑G), considered as minimal signal sequences. SPs have demonstrated post-targeting functions as well by functioning either from within the membrane, as a released peptide or as a fragment generated by intra-membrane proteolysis. These are structurally different from conventional signal sequences as they have extended N-terminal regions or two hydrophobic regions. So far, all examples of stably membrane-integrated signal peptides are of viral origin and are derived from precursors of viral envelope glycoproteins[98].

1.4.1.3 Mechanism of mammalian secretory pathways

Entry into the ER begins with the recognition of the protein's N-terminal signal sequence by SRP[108] which consists of six polypeptides[109] and a small cytoplasmic RNA (7sL RNA) [110]. Interaction between the signal sequence and SRP determines the efficiency with which the

nascent polypeptide enters the secretory pathway[111,112]. The ribosome-nascent polypeptide chain-SRP complex binds to the ER membrane *via* the SRP receptor on the surface of the ER[113,114] (Figure 1.5). The SRP acts to inhibit further elongation of the nascent polypeptide chain until binding to the SRP receptor occurs. Binding to the receptor releases the SRP from both the ribosome and the signal sequence of the growing polypeptide chain[115,116]. The Sec61 translocon is a heterotrimeric protein complex consisting of Sec61α, Sec61β and Sec61γ chains[117]. It acts as a proteinaceous pore in the membrane which accepts the ribosome-nascent polypeptide complex. This allows for the translocation of the chain across the membrane and into the ER[94] and elongation continues[118]. This process is coupled to the hydrolysis of GTP[119]. SRP and the Sec61 translocon are also potential targets for regulation in translocation to the ER as both recognise the signal sequence in the growing polypeptide chain[120]. Folding of the polypeptide takes place in the lumen of the ER and as folding occurs, exposure to modifying enzymes create a temporal relationship between folding and post-translational modifications[94].

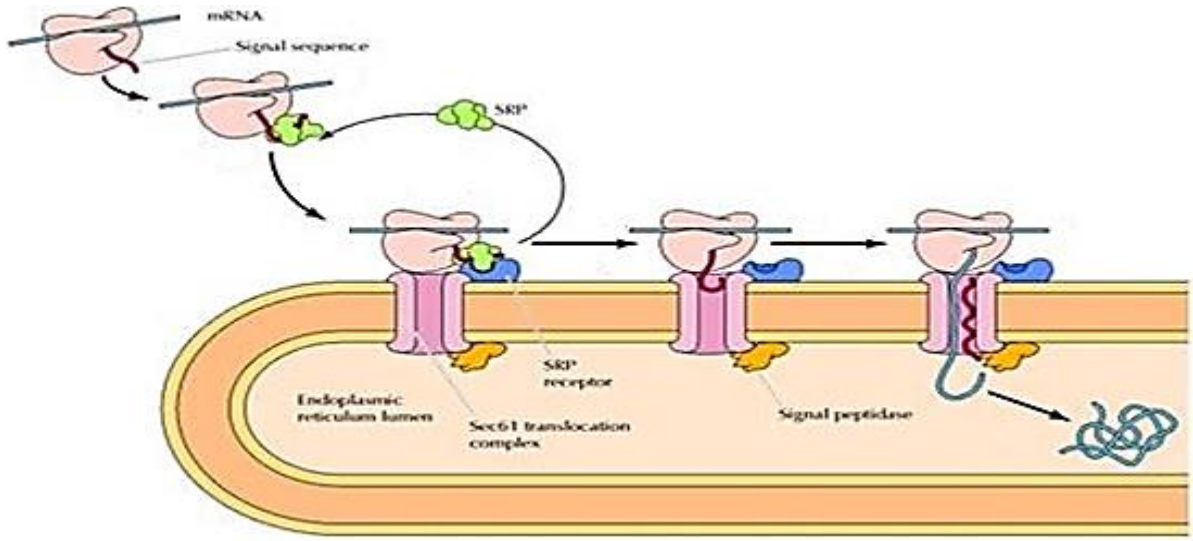

Figure 1.5: Co-translational translocation of proteins into the ER. Translocation into the ER occurs with the recognition of the nascent signal peptide (SP) as it emerges from the ribosome and is recognised by the signal recognition protein (SRP). Translation halts temporarily as the ribosome-SP-SRP complex is directed to the ER membrane where it is recognised by the SRP receptor. The ribosome then binds the Sec61 translocon pore for the movement of the nascent polypeptide into the ER lumen. Upon recognition of the cleavage site, the signal peptidase cleaves the signal peptide and the nascent polypetide is then bound by molecular chaperones (not shown) in order for folding and glycosylation to occur. Illustration accessed from: http://mol-biol4masters.masters.grkraj.org/html/Co_and_Post_Translational_Events3-Eukaryotic_System.htm

1.4.1.4 Secretion of Env

Whilst still inside the ER lumen, the precursor gp160 associates with lectin chaperones that assist in its folding as it translocates through the ER on its way to the Golgi apparatus. These chaperones, calnexin (a type I membrane protein) and calreticulin (soluble in the ER), associate with the N-glycosylated gp160[121] and are involved in the retention of misfolded or unfolded glycoproteins in the ER[122]. Other chaperones include the ER resident protein 57 (Erp57), a soluble thiol-disulfide oxidoreductase, which associates with calnexin and calreticulin in order to introduce disulfide bonds in the folding glycoprotein[123]. The oxidising environment of the ER lumen itself also facilitates the conversion of the cysteine residues in the glycoprotein into disufide bonds[124]. Binding immunoglobulin protein (BiP), is a chaperone also residing in the ER lumen that seals the Sec61 translocon on the luminal side of the ER and is responsible for recognising exposed regions of unfolded proteins as well as targeting them for degradation[125].

1.4.1.5 Timing of signal peptide cleavage is crucial for Env structure and function

The Env signal peptide (Env-SP) is removed post-translationally[107] by ER membrane-bound peptidases. Folding of gp160 is a requirement for cleavage of the Env-SP as lack of disulfide bonds leads to loss of signal peptide cleavage[107,126]. Once properly folded, trimerised and partially glycosylated, precursor gp160 moves to the translational ER (tER) where it is packaged into vesicular tubular complexes (VTCs) which then fuse to the *cis*-Golgi. The precursor gp160 contained within these VTCs are then transported through the stacks of the Golgi and enter into the *trans*-Golgi network (TGN) where complex sugars are added to the high-mannose oligosaccharide side chains[86,127]. Inside the Golgi, gp160 is proteolytically cleaved by furin or furin-like proteases at a highly conserved Lysine (K)/N-X-K/N-N motif[128,129]. Gp160 proteolytic cleavage into gp120 and gp41 subunits can also occur in *trans*-Golgi-derived (TGN) secretory vesicles[130]. Proteolytic processing of gp160 is an absolute requirement for the activation of Env fusogenic activity. After cleavage, gp120 and gp41 remain associated *via* non-covalent interactions[47].

1.4.1.6 Functional determinants of signal peptide function

Expression and secretion levels of protein is associated with the rate of Env-SP cleavage which is influenced by the presence of positively charged amino acids at the N-terminus[13,107,131]. The hydrophobic core has been suggested to be involved in the initiation of secretion[132] although substitution of non-polar with charged amino acids did not abrogate translocation and N-glycosylation[133]. On the other hand, the positively charged amino acids have been suggested to play a role in determining the efficiency of protein secretion[132,134]. Env-SP has an unusually high number of positively charged amino acids in the N-terminal region compared to typical eukaryotic SPs[78,107]. Env expression and secretion levels are associated with the rate of Env-SP cleavage which is influenced by the presence of positively charged amino acids at the N-terminus[13,107,131]. Sequential substitution of the positively charged residues with neutral amino acids was found to enhance protein expression and secretion by reducing the SP cleavage time and thus the time spent in the ER lumen[107]. The slow cleavage of the SP (and low secretion) is an intrinsic property of Env irrespective of the expression system[102] and it has been suggested that the secretion block is as a result of the presence of a highly charged N-terminal portion of the signal sequence. As a result, this limits the transport of gp120 from the ER to the compartments of the Golgi[107].

1.4.1.7 Role of the signal peptide in Env expression and processing

Early research showed that the signal peptides of honey bee mellitin and murine interleukin 3 increased efficient folding, intracellular transport, and secretion of gp120. A similar result was observed when the positively charged amino acids in the native gp120 signal sequence was replaced with neutral amino acids, highlighting the importance of the SP on Env processing[107]. SP exchange experiments suggested that when the Env-SP was replaced with a heterologous one, expression was altered[135,136]. Furthermore, deletion of the endogenous SP resulted in the nuclear accumulation of non-glycosylated Env as entry into the ER was prevented[107]. Changes in Env expression might lead to changes in its N-glycosylation profile as transfection with increased ratios of Env:backbone plasmids lead to an increase in mannosylation[137]. Slower cleavage of the Env-SP has been suggested to lead to an accumulation of Env in the ER[84,107,124] which could lead to the production of “junk” gp160 compromising predominantly of unprocessed oligomannose-type N-glycans[138]. The

accumulation of Env inside the ER may overwhelm the cellular machinery of the organelle[139], potentially influencing downstream processing such as its N-glycosylation.

1.4.2 N-glycosylation in the ER and compartments of the Golgi

1.4.2.1 N-glycosylation in mammalian cells

Over 70% of the eukaryotic proteome is thought to be N-glycosylated and different glycoproteins are modified to different extents so that secreted or membrane-bound proteins can emerge from the Golgi with a variety of different N-linked oligosaccharides[140]. Within the ER, N-glycosylation starts with the formation of the core oligosaccharide comprising of three glucose (Glc), nine mannose (Man) and two N-acetylglucosamine (GlcNac) residues synthesized on a lipid dolichol carrier anchored in the ER membrane[141,142]. As the growing polypeptide chains are translocated into ER, the $Glc_3Man_9GlcNac_2$ core oligosaccharide is transferred from the dolichol carrier to Asparagine residues within the sequon N-X-Threonine (T)/Serine (S) (where "X" is any amino acid except proline)[143–145]. Addition of glycans at each of these motifs does not always occur, as sequons are variably utilized even on the same protein[146]. It has been postulated that sequons proximal to the C termini of glycoproteins are unlikely to be glycosylated[144,147] and that structural constraints, such as proximity to a disulfide bond for example, may influence sequon occupation[142,145,148]. ER glucosidases and *α*-mannosidase I, sequentially remove the three Glc residues in a site-specific manner along with a mannose residue yielding $Man_8GlcNAc_2$[149]. The removal of glucose residues is thought to be essential for proper folding and transport of proteins to the Golgi apparatus[127].

1.4.2.2 Processing of high mannose core oligosaccharides in the Golgi

Once the protein reaches the *cis*-Golgi compartment, mannose moieties are removed from $Man_8GlcNAc_2$ to give $Man_5GlcNAc_2$, the high-mannose N-glycan that is the substrate for the enzymes found in the compartments of the Golgi for further modifications (Figure 1.6).

In the *medial*-Golgi, two more Man residues are removed and three GlcNAc and one fucose are added, giving rise to hybrid type N-glycans. Mature complex N-glycans are formed in the *trans*-Golgi with the addition of more GlcNAc moieties accompanied by other modifications

via the *β*-1,2-N-acetylglucosaminyl-transferase I (GnTI)[127,150,151]. Fully processed Envs that carry a combination of high mannose, hybrid and complex N-glycans are then transported to the cell membrane or shed as gp120 into the extracellular environment.

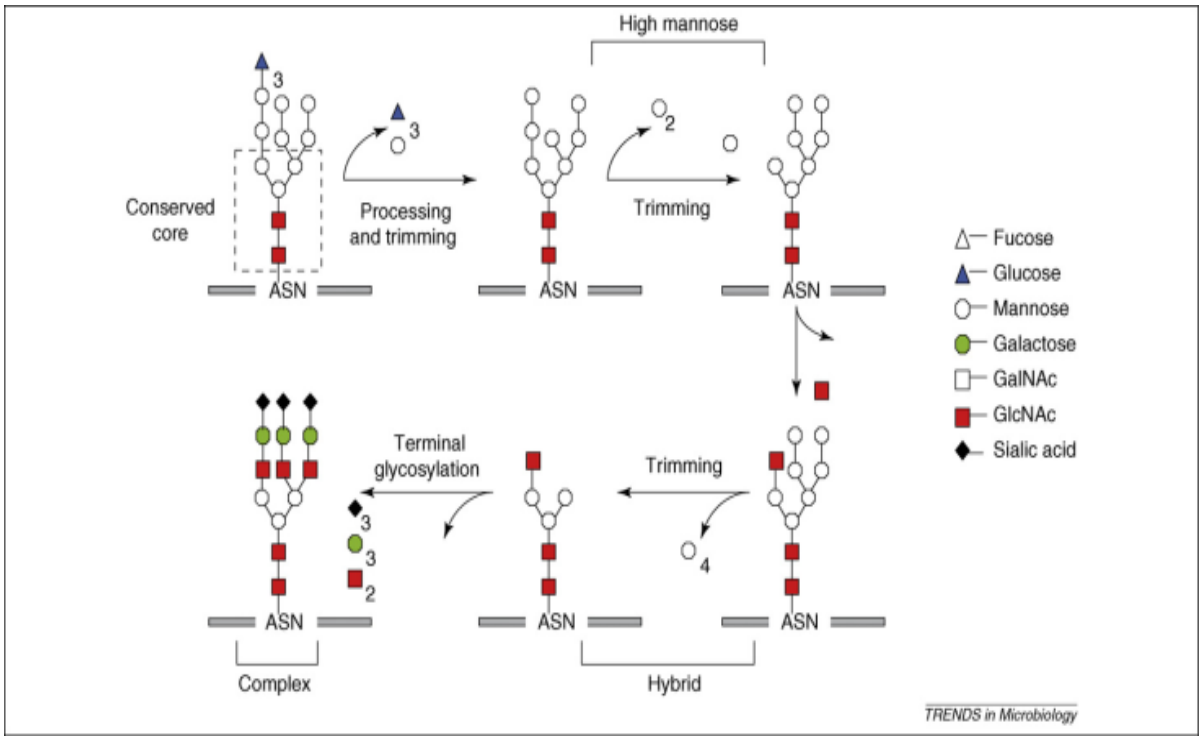

Figure 1.6: N-linked glycosylation of Envelope. As the protein enters the Golgi, a sequence of reactions occur beginning with the trimming of the conserved 14-sugar core yielding the high mannose glycoform. This is followed by the addition of an N-acetylglucosamine, the removal of additional mannose residues to yield hybrid N-glycans. Finally, terminal glycosylation occurs with the addition of three galactose and three sialic acid and/or fucose residues and two N-acetylglucosamines to give rise to the complex type N-glycans [150].

1.4.2.3 N-glycosylation of Envelope

HIV-1 Env is entirely processed by the glycosylation machinery of the host cell[152] and interaction with the spectrum of enzymatic activities present in the secretory pathway determines the types of glycans that will be presented on gp120 at the virion surface or as a recombinant protein. An "intrinsic" high-mannose patch was identified at the surface of recombinant gp120 and virus-derived Envs[153]. Despite sequence variation, location and number of PNGs, the mannose patch is highly conserved over the course of infection. It can thus provide a stable target for vaccine strategies[154] especially as it has been identified as a potential non-self-target for broadly neutralising antibodies (bNAbs)[50,137,153,155,156]. Even

though glycan structures are known to act as a "shield" to protect the polypeptide chain from immune recognition[62], some of the most potent nAb recognise the mannose patch[77,157].

In addition to the mannose patch found on gp120 monomers and viral-associated trimers, virus-associated gp120 carried more high mannose N-glycans (56-79%) than that of recombinant protein (29%). When pseudovirus-associated gp120 was characterised, 98% of total N glycans comprised of oligomannose residues[137]. A number of subsequent studies confirmed that the majority of Env PNGs carried oligomannose N-glycans[77,158]. Last year, Go *et al.* (2017) analysed the N-glycosylation of 11 trimers and identified a "consensus profile" where more than half of the PNGs carried the same type of N-glycosylation i.e. either high mannose residues or complex sugars. However, trimers were still predominantly mannosylated[159].

As mammalian N-glycosylation pathways do not usually produce clusters of oligomannose carbohydrates[148], the apparent high mannosylation of Env might be due to variation in viral processing signals or cell-specific machinery. It has been proposed that the oligomannose-type N-glycans found on gp120[153] could be due to an alternative processing pathway which bypasses the *trans*-Golgi and thus the action of GnTIs[137] potentially because of the accumulation of Env in the ER due to slow signal peptide cleavage[84,107,124]. Alternatively, it has also been suggested that high oligomannose density prevents access of endo-glycosidases to sites of processing, such that complex residues are not added[160]. This hypothesis was supported when virus-associated trimers were shown to carry predominantly high mannose residues and recombinant protein had more complex N-glycosylation[153].

The HEK 293T and CHO cell lines have been successfully used in the production of functional recombinant gp120[81,137] as well as stable Env trimers[161]. This suggests that these cell lines have the functional machinery to ensure the correct processing and glycosylation of gp120 especially noting that the N-glycosylation profiles of Env produced in PBMCs are similar to that expressed in HEK 293T cells. However, differences in N-glycosylation between cell lines have been observed[156] as pseudovirus-associated gp120 are differentially N-glycosylated at specific positions depending on the cell line used[160,162,163]. Furthermore, Raska *et al.* (2010) showed that the influence of cells on Env N-glycosylation also extended to the metabolic state

of the cells[164]. These findings suggest that glycosylation of Env is determined by the structure of Env as well as cell-directed effects[160,162,163]. It is likely that optimum combinations of viral factors and cell-mediated processes are required to reproduce Env N-glycosylation of native trimers. Therefore, selection of Env clones and cell lines for the production of recombinant gp120 and virion-associated Env trimers for vaccine production must be carefully considered as Env N-glycosylation must resemble that of virus infecting CD4+ T cells to elicit broadly neutralising antibodies.

1.4.2.4 Significance of N-glycosylation on Env folding and retention within the ER

It has previously been shown that heavily glycosylated proteins require N-glycans for correct folding and Env is no exception as N-glycosylation has been found to be crucial for the transport of Env from the ER to the Golgi[82,165,166]. N-linked oligosaccharides stabilise the folded domains[167] and make the proteins more soluble[126]. They are needed for recognition by lectin chaperones that assist with Env folding in the ER[121]. When cells were treated with the glycosylation inhibitor tunicamycin[168], misfolding of gp120 occurred[169] and as a result the signal peptide was not removed[124]. The deletion of multiple N-glycosylation sites impacted Env folding, whereas the removal of one or two sites had no effect[166]. However when only three conserved N-glycosylation sites were removed within the gp41 domain, misfolding occurred which resulted in the retention of Env in the ER[170]. This suggests that the influence of N-glycosylation on protein folding might differ according to the PNG under investigation and/or the Env clone.

N-glycans were also important for the functioning of Env as loss of certain PNGS resulted in increased or decreased viral infectivity[171]. Overall changes in Env conformation due to loss of N-glycans might have a domino effect on the spatial rearrangement of domains that facilitate and moderate interactions with host proteins such as binding to co-receptors on target cells[171,172]. Furthermore, it has been shown that the type of N-glycans also influences Env function with high mannose residues favouring dendritic cell interactions while lowering the efficiency at which virus infected TZM-bl cells[173]. These findings suggest that N-glycosylation plays a role in the function of Env that may ultimately impact transmission.

1.5 Env transmission motifs

1.5.1 Genetic bottleneck associated with HIV transmission

The genetic bottleneck during HIV transmission refers to the movement of a single variant to a new host from a donor infected with variants of high diversity[5] (Figure 1.7). Wolinksky and colleagues (1992) were the first to identify that a specific subset of viruses were present following mother-to-infant HIV-1 transmission shortly after transmission[174]. After the analysis of the V3 loop and V4-V5 *env* sequences of mother-to-child transmission pairs, infant *env* sequences were found to be less variable than those found within the respective mother. Later, a study using a quantitative heteroduplex tracking assay provided the first evidence of selection during HIV-1 heterosexual transmission due to the presence of a more homogeneous *env* population in acutely HIV-1 infected subjects compared to those found in the donors[175]. Another study then showed that heterosexual transmission of subtype C HIV-1 in discordant couples was characterised by a genetic bottleneck as only a single variant from a diverse donor quasispecies established productive infection[176]. Subsequent studies using single genome amplification (SGA) *env* amplicons revealed that in approximately 8-% of subtype A, subtype B, subtype C and recombinant HIV-1 infections, productive infection was as a result of a single infectious virus[7,177–180] and initiated the search for the "transmission motif" thought to be found on a subset of variants - transmitted founders (TFs) – that provide a selective advantage during transmission[7].

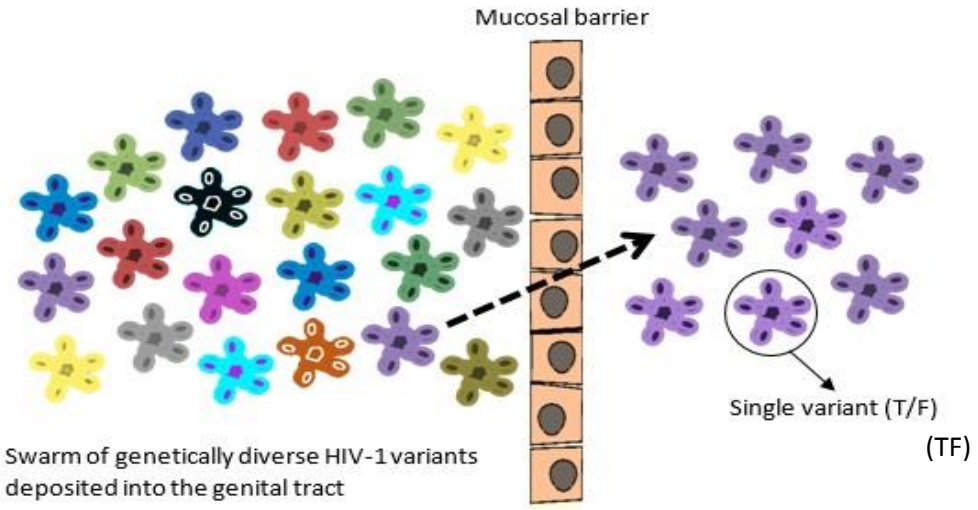

Figure 1.7: HIV-1 genetic bottleneck at transmission. Despite the deposition of a swarm of genetically diverse HIV-1 variants in the genital tract, only a single variant, termed the transmitted founder (TF), crosses the mucosal barrier to initiate productive infection.

Transmission rate per coital act is as low as 1/200 to 1/2000 per coital act for male-to-female (MTF) transmission and female-to-male (FTM) transmissions is 1/200 to 1/10000[181]. These frequencies occur even in regions where there is a high prevalence of HIV-1 infections. This suggests that the genital mucosa is a formidable barrier to HIV-1 transmission. The frequency with which more than one virus is able to lead to productive clinical infection was found to increase dramatically in the presence of chronic ulcerative disease or inflammatory genital infections in the recipient, suggesting that disruption of the barrier might abrogate the selection mechanism[179]. TF *envs* have low diversity[178] and are almost always R5-tropic, supporting the hypothesis that the viral envelope may be responsible for the selection of specific variants in the female genital tract[14,179,182]. Thus, TFs might have distinct characteristics that allow them to overcome the mucosal barrier as well as enhance infection of susceptible cells.

1.5.2 Characteristics of the TF Env

Difficulties in identifying and characterising TF Env were overcome with the advent of single genome amplification (SGA). SGA involves limited dilution of viral cDNA followed by two rounds of polymerase chain reaction (PCR) in order to ensure the presence of a single template molecule and prevent PCR-induced recombination[183,184]. This has provided researchers with the ability to accurately identify the TF variant and reveal distinct characteristics that may facilitate transmission. Genotypic characteristics of TF *envs* appear to be subtype-specific. Analysis of subtype A and C TF *env* sequences showed that these variants had shorter variable loops with fewer PNGS compared to variants isolated during chronic infection[15,185]. This was in contrast to *env* sequences analysed from subtype D TFs which, although had shorter variable loops, did not differ in the number of PNGS compared to chronic infection viruses[186]. There was no consensus on whether subtype B TF *env* sequences differed in variable loop length and PNG number compared to chronic infection Envs as some studies showed no difference[6,187]. whereas others indicated that transmission might select for variants with shorter Env variable loops and fewer PNGS[8]. Although subtype-specific differences in variable loop length and number of PNGS may exist, the transmission of variants carrying gp120s with less PNGS has been consistently identified. This supports the

suggestion that high numbers of N-glycosylated residues accumulate over the course of infection to conceal conserved Env antigenic regions from antibody neutralization[15,62].

A common feature of TF viruses is their dependence on CCR5 co-receptors for infection of T-lymphocytes, indicating that R5-tropism is a characteristic selected during transmission and/or acute infection[7,65]. These variants have also been found to have enhanced neutralization sensitivity[15] although other findings have not supported this suggestion[187,188]. Initial studies found that R5 tropism was linked to macrophage-tropism[175], however recent studies have shown that TFs are unlikely to infect macrophages due to their low levels of CD4 expression[189,190].

Dissemination of the virus into the gut-associated lymphoid tissue (GALT) results in a rapid depletion of the gut CD4+ cells contributing to immune dysfunction[191]. Migration of T-lymphocytes to the GALT is mediated by the homing integrin, alpha-4-beta-7 (α4β7), which is highly expressed on CD4+ T cells of the genital mucosa[192] and gut[193]. α4β7, mediates the migration of CD4+ T cells from Peyer's patches and mesenteric lymph nodes to the lamina propria by binding mucosal addressin cell adhesion molecule-1 (MAdCAM-1) expressed on the endothelial cells of these tissues[194]. α4β7 also binds to a motif identified in the V2 loop of gp120 identical to that of MAdCAM-1 and is conserved across HIV subtypes A, B, C and D[12,193,195]. As HIV-1 infected CD4+ T-lymphocytes that express α4β7 with high efficiency[191], the α4β7 binding site is conserved in the highly variable V2 loop and substitution of this tripeptide resulted in a reduction in the replicative capacity of these viruses[12] it was suggested that α4β7 may be important in HIV infection and transmission. However, when Parrish and colleagues (2012) compared subtype C TF and chronic Envs they found that TF Envs did not utilise α4β7, CD4 or CCR5 more efficiently suggesting that the interaction between gp120 and α4β7 was not essential for transmission across mucosal surfaces[67].

Analysis of 6303 early and chronic SGA Env sequences identified potential transmission motifs in the signal peptide and V4 region of gp120[8]. A Histidine residue at position 12 [(His12) (position 9 according to HXB2 numbering)] within the signal peptide and the absence of a PNG at position 413-415 (PNG413) were identified as robust markers of TF variants[8]. It was

suggested that the presence of His12 and/or the absence of a PNG413, respectively, may provide an advantage during transmission. Recently, a site within Env-SP was also identified as a putative simian immunodeficiency virus (SIV) transmission motif, Arg9, although whether the actual residue aligned with His12 or an amino acid three residues upstream could not be confirmed due to gaps within the sequences[196].

1.5.2.1 PNG 413-415

Contrary to chronic infection variants, TFs tended not to carry a PNG at position 413-415, suggesting that this site is selected against during transmission, or shortly thereafter[8]. The PNG at position 413-415 has also been associated with a conserved sequence motif, RIKQ (HXB2 position 419-422), a few residues upstream. This conserved motif has been found to be involved with CCR5[197] and antibody binding[56]. Since PNG413 is almost always found to be glycosylated when present[198], it has been hypothesised that due to its close proximity to RIKQ, the glycosylation of the PNG physically prevents epitope recognition by host antibodies. This may provide a mechanism for escape from the host immune response, improving the chance of survival[199]. However, escape mutations have been shown to carry a fitness cost and changes in N-glycosylation could influence protein expression, processing and function[150]. Therefore, selection of variants lacking PNG415 during transmission could be beneficial for viral survival post-transmission as these variants could have higher replicative fitness.

1.5.2.2 Histidine at position 12 of the Env signal peptide

Env-SP in retroviruses, particularly that of HIV-1[200] has been shown to influence protein cleavage rate[102], folding[126,201], trafficking and processing[47,92]. Exchange of the endogenous signal peptide with heterologous ones from erythropoietin (EPO) or tissue plasminogen activator (t-PA)[135], increased the expression levels and secretion[202] as well as the incorporation of Env and the infectivity of the virions[136]. His12 in the signal peptide of subtype B Envs has been shown to increase expression of the Env, resulting in better incorporation and thus more infectious viral particles[13]. This suggests that the presence of His at position 12 could influence the replicative fitness of the TF so that its presence is required for successful transmission. In 2018, another study showed that substituting His12 with Arg, Gln or Tyrosine (Tyr), reduced virion- and cell-associated Env levels as well as infectivity of infectious

molecular clones (IMCs) without influencing Env synthesis and cleavage[203]. This finding did not only support Asmal *et al*. (2011) but also showed that the presence of Gln at position 12 was associated with enhanced sensitivity to monoclonal antibodies due to alterations in Env N-glycosylation. Therefore, changes in the signal peptide can influence Env N-glycosylation and antibody sensitivity. Unlike subtype B Envs, subtype C Envs do not carry a His at position 12, suggesting that the replicative fitness of subtype C TF is not dependent on His at this position but that other amino acids could provide a similar advantage during HIV-1 transmission.

Current vaccine strategies require high levels of Env expression[136] for purification that can induce long-lived neutralising antibody responses for a broad range of virus variants[204]. To achieve this, expression systems that can facilitate high-level expression without compromising the structure and function of the protein have to be utilised. If changes within the Env itself could be made to upregulate its expression whilst still maintaining correct folding and subsequent function in mammalian cells, this would facilitate the production of immunogenic Env. As subtype B TF Env naturally carries a Histidine at position 12 within the signal peptide and has been shown to increase Env expression, altering subtype C Env-SP sequence by mutational manipulations might increase expression without compromising Env structure and function. This approach would thus serve as an alternative to the use of heterologous signal peptides that are routinely used to increase Env expression for immunogen production.

1.6 Concluding statements

The enrichment of His12 and absence of PNG413 in Env of TFs would suggest that these sites are enhancing transmission by influencing viral replication. Env phenotypes that might enhance viral replication include increased levels of incorporation of Env trimers into viral particles and/or improved utilisation of CD4, CCR5, DC-SIGN and α4β7, or other unknown host factors that promote HIV infection. It follows that the more virus bound to host receptors the higher the viral replication. Both the signal peptide and N-glycosylation influences Env expression and processing, leading to increased incorporation of high levels of native trimers into infectious viral particles with or without changes in Env conformation that might enhance

interactions with CD4, CCR5 and/or DC-SIGN. We thus hypothesised that the two subtype B transmission motifs might function similarly to enhance transmission fitness by improving Env processing. A review of the literature highlighted the importance of the signal peptide and N-glycans on viral fitness by altering expression, protein folding, function and immune escape. Characterisation of His12 suggested that replacing a His at position 12 with a neutral amino acid may slow down Env-SP cleavage, lower cell-surface expression and assist subtype B variants to evade neutralising antibodies during chronic stages of infection[13]. Following this suggestion, at transmission when there is no immune response, neutral amino acids at position 12 will not be beneficial. If neutral amino acids at position 12 are associated with a fitness cost, it might revert to consensus- a Histidine- in the absence of immune pressure. This suggests that His12 is not itself important for transmission but is required for optimal Env expression in the absence of an immune response. Alternatively, His12 may provide a direct advantage to TFs during transmission by enhancing infectivity. Interestingly, subtype C TFs tend to carry an uncharged Glutamine at position 12 in Env-SP, suggesting that: 1) Gln12 does not have a fitness cost and is thus transmitted, 2) His12 may not be a determinant in the transmission of subtype C variants or 3) TFs have evolved differently so that both Gln and His at position 12 of Env-SP have the same impact on transmission. Overall, studies aimed at elucidating the mechanisms by which these potential transmission motifs facilitate HIV transmission are crucial to our understanding of whether selection of these transmitted variants at the genital mucosa is subtype-specific and what this could mean for vaccine design.

1.7 Study rationale and research objectives

The transmission motif found at position 12 of the signal peptide not only influences subtype B Env expression and processing but also incorporation, contributing to the production of more infectious virus particles[13]. However, subtype C TF is not enriched for His at position 12, suggesting that this position is not required for subtype C transmission. Furthermore, the use of heterologous signal peptides to enhance Env production in vaccine production could alter the structure and immunogenicity of the immunogen. The absence of a PNG at position 413-415 in subtype B Env has been found to provide a mechanism for immune escape during chronic infection by enabling resistance to potent and broadly neutralising host antibodies. Thus it has been suggested that this PNG is selected against in transmitted founder viruses

and may therefore also influence replication[8]. However, subtype C TFs carry a PNG at position 413-415 suggesting that this might not be a transmission motif for subtype C variants. Therefore, this project aims to determine the importance of the Env signal peptide and the absence of PNG413 in subtype C transmission.

1.7.1 Aim

The overall aim of this project was to investigate the role of the transmission motifs identified in the Env of subtype B TF variants, i.e. His at position 12 of the signal peptide and the absence of a PNG at 413-415, in altering the expression, secretion, N-glycosylation, entry efficiency and replication of a subtype C TF variant.

1.7.2 Objectives

1. Generate subtype B TF Env signal peptide motif-carrying constructs using a subtype C variant.
2. Compare expression, secretion and N-glycosylation patterns of Envs (wild-type, signal peptide mutants and T415I mutant).
3. Generate pseudovirus from cloned Envs (wild-type, signal peptide mutants and T417I mutant) and compare entry efficiency using a reporter TZM-bl entry assay.
4. Produce chimeric infectious molecular clones (IMCs) of wild-type, signal peptide mutants and T417I mutant using a yeast gap-repair assay.
5. Replicate the IMCs in peripheral blood monocytic cells (PBMCs) and compare the replicative fitness of the IMCs to each other and compare to the entry efficiency of pseudovirus.

CHAPTER 2: MATERIALS AND METHODS

2.1 Envelope constructs

The subtype C TF *env* clone, CAP210 E8 (donated by Gama Bandawe, IIDMM, UCT) was obtained from the Centre for the AIDS Programme of Research in South Africa (CAPRISA) 002 study cohort, isolated from a participant 5 weeks post infection. In order to characterise the *env* of the TF variant, single genome amplification (SGA) was performed. Viral RNA was extracted from the first plasma sample that tested either viral RNA or antibody positive. The next steps in SGA include cDNA preparation and limiting dilution in order to obtain a single template per PCR reaction followed by direct sequencing of the amplicon[5,7,177]. From the SGA-derived *env* sequences, a consensus *env* sequence was generated which was most representative of the Env of the TF of that patient. The CAP210 E8 TF *env* SGA amplicon was then cloned into a mammalian expression vector, in this case the 5.1 kilobases (kb) long pcDNA3.1 V5-His Topo® (Life technologies™) for further genotypic and phenotypic characterisation. The PNG415 mutant (T417I) of this *env* clone was generated by Riley Traviss (IBMS, UCT) and included in this study. The signal peptide mutants, Q12H and Q12A, were generated by site-directed mutagenesis (Section 2.2.1). The soluble form of CAP210 E8, gp140 was generated by Evelyn Ngwa (MCB, UCT) and was also mutated.

2.2 Cloning and site-directed mutagenesis

2.2.1 Site-directed mutagenesis

A modified Quikchange® site-directed mutagenesis protocol (Stratagene) was used to generate signal peptide mutant CAP210 E8 clones. Complimentary primers carrying the desired mutations were designed and the *Nde*I restriction enzyme site was introduced through silent mutations using the online tool WATCUT (http://watcut.uwaterloo.ca). All primers were 40 nucleotides long with 15 bases on either side of the mutated nucleotides (Table 2.1) and were synthesised at the oligosynthesis laboratory by Pei-Yin Leibrich (MCB, UCT). All PCR reactions were performed in an XP Thermal cycler (Bioer technology). The DNA concentrations were determined using a NanoDrop® spectrophotometer (Thermo Scientific). PCR constituents are listed in Table 2.2.

Table 2.1: Site-directed mutagenesis primers

Name	**Direction**	**Sequence 5'→3'**	**Restriction Enzyme**
CAP210_E8_Q12H_F	Forward	CAG AGG AAT TGG CAG CAC TGG GGC ATA TGG GGC ATC TTA G	*Nde*I
CAP210_E8_Q12H_R	Reverse	CTA AGA TGC CCC ATA TGC CCC AGT GCT GCC AAT TCC TCT G	
CAP210_E8_Q12A_F	Forward	CAG AGG AAT TGG CAG GCA TGG GGC ATA TGG GGC ATC TTA G	
CAP210_E8_Q12A_R	Reverse	CTA AGA TGC CCC ATA TGC CCC ATG CCT GCC AAT TCC TCT G	

*Nucleotide substitutions are bold faced and restriction enzyme recognition sequences are underlined.

Table 2.2: Site-directed mutagenesis PCR mixture

Reaction mix	**Concentration**
Buffer + $MgCl_2$	1X
dNTP mix	0.4 mM
DMSO	5%
Forward primer	0.5 µM
Reverse primer	0.5 µM
Template DNA	50 ng
Phusion Hot Start Polymerase	1 U
Total (made up with nuclease-free dH_20)	50 µL

Site-directed mutagenesis PCR was performed at an annealing temperature of 55 °C. The cycling conditions were as follows: 94 °C for 3 min followed by 25 cycles of 94 °C for 30 sec, 55 °C for 30 sec, 72 °C for 12 min and a final elongation step at 72 °C for 20 min. To verify whether the PCR yielded the desired products, 8 µL of the product was separated on a 1% agarose gel. The template plasmid in the reaction mixture was digested with 20 U of *Dpn*I (Thermo Scientific), a restriction enzyme that recognises methylated DNA so that only newly synthesised unmethylated plasmids were transformed into competent *E. coli* JM109 cells (Promega™).

2.2.2 Bacterial cell culture, plasmid transformation and screening

E. coli JM109 cells (Promega) used for production of plasmid DNA were grown in Luria broth (LB) growth media [1% tryptone, 0.5% yeast extract, 1% NaCl in distilled water and 100 µg/mL carbenicillin disodium salt (Sigma Aldrich®)] overnight at 30 °C with shaking. For plasmid

transformations, competent *E. coli* JM109 cells (Promega) stored at -80 °C were thawed and the desired amount of plasmid DNA was added to 50 µL of cells and incubated on ice for 20-30 min. The cells were heat shocked for 45 seconds (sec) at 42 °C and immediately chilled at 4 °C for 2 min. Cells were allowed to recover in 950 µL LB (without carbenicillin disodium salt) for 1 hour (hr) at 37 °C before centrifugation at 5000 revolutions per minute (rpm) for 5 min and 950 µL LB was removed. The cells were resuspended in the remaining LB, spread on Luria agar (LB, 1.5% w/v agar and 100 µg/mL carbenicillin disodium salt) plates and grown overnight at 30 °C[205]. Transformation efficiency controls were included at 10 pg, 100 pg and 1 ng of pcDNA3.1 (Invitrogen®).

To screen for positive mutants, 5mL LB was inoculated from the LA plates containing transformants and grown O/N at 30 °C. Plasmid DNA from these O/N cultures was then extracted using the STET (8% sucrose, 5% Triton X-100, 50 mM Tris pH 8.0, 50 mM EDTA, 0.5 mg/mL lysozyme) Boiling Preparation method[206], a crude extraction method used initially to screen and identify colonies carrying the desired mutations with restriction enzyme digestion. After screening, glycerol stocks were made of potential positive mutants before colonies were cultured in 5 mL LB O/N for plasmid extraction by the Bioflux™ Plasmid DNA Extraction Kit (Bioer Technologies) according to the manufacturer's instructions. Plasmid DNA was eluted with 50 µL of nuclease free water and sent for sequencing to confirm that the desired mutations were present. 200-300 ng of this DNA was digested (Section 2.2.3) with *Nde*I (Table 2.1).

2.2.3 Restriction enzyme digestion

Restriction enzyme digestions were carried out in 20-50 µL reaction mixtures containing DNA (200-300 ng), enzyme buffer (1X) and 1 unit (U) of the appropriate restriction enzyme. *Nde*I and *Pac*I were purchased from Thermo Scientific™ and digestions with these enzymes were performed at 37 °C for 1 hr.

2.2.4 Agarose gel electrophoresis

Agarose gel (0.8% or 1% w/v) were prepared by adding the appropriate amount of agarose (Whitehead Scientific™) to TBE buffer (2 mM EDTA, 89 mM Tris base and 89 mM boric acid)

boiled until complete dissolution followed by addition of ethidium bromide (0.3 µg/mL). DNA was mixed with loading buffer (5% glycerol and 0.001% bromophenol blue in distilled water) and separated by agarose gel electrophoresis in TBE buffer at 100 volts (V) until the desired bands could be visualised.

2.2.5 Sequence analysis

After screening colonies for SDM mutants using restriction enzyme digestion by *Nde*I, the plasmid DNA was extracted (Section 2.2.3) and sent to the Central Analytical DNA Sequencing Facility at the University of Stellenbosch. All primers used are listed in Table 2.3 below. Upon identification of potential mutants, primers Rev15 and EF175 which span the signal peptide region of the Env (HXB2 numbering) were used to confirm the presence of the mutations. Thereafter the remaining 6 primers which cover the entire *env* were used to check for any additional mutations. All primers used are listed in Table 2.3 below.

Table 2.3: Primers used for sequencing of the full-length HIV-1 *env* gene after site-directed mutagenesis

Primer name	Direction	Primer sequence	Length (bp)	HXB2 position (entire genome)
Rev15	Reverse	CTG CCA TTT AAC AGC AGT TGA GGT GA	26	6990←7012
EF175	Reverse	TTT AGC ATC TGA TGC ACA GAA TAG	24	6378←6398
For14	Forward	TTG CCA ATC AAG GAA GTA GCC TTG TGT	27	6559
EF15	Reverse	CTT GCT CTC CAC CTT CTT CTT C	22	8424←8442
EF00	Forward	GGG AAA GAG CAG AAG ACA GTG GCA ATG A	28	6204→6228
EF55	Reverse	GCC CCA GAC CGT GAG TTG CAA CAT ATG	27	7914←7937
E200	Forward	GGG ATA ACA TGA CCT GGA TGC AGT GGG	27	8095→8118
EF260	Forward	TTC AGC TAC CAC CGA TTG AGA GAC T	25	8523→8544

2.3 Mammalian cell culture

HEK 293T, the human embryonic kidney cell line and TZM-bl cells (obtained through the NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH from Dr. John C. Kappes, Dr. Xiaoyun Wu and Tranzyme Inc)[207], are adherent cell lines that were maintained in growth medium made up of Dulbecco Modified Eagle high glucose medium (DMEM) (Sigma Aldrich®) supplemented with 10% fetal bovine serum (FBS) (Biocom Biotech®), 1 U/mL penicillin and 1 µg/mL streptomycin (Sigma Aldrich®). Cells were incubated and grown at 37 ^{o}C in a water-jacket

incubator (90% humidity and 5% CO_2). Stocks of cells were stored in 10% Dimethyl sulfoxide (DMSO) (Sigma-Aldrich®) in FBS (Biocom Biotech®) at -80 °C. All cells were tested for mycoplasma infection using the Fluorescent Hoechst DNA stain[208] every month.

2.4 Transient expression, secretion and N-glycosylation analysis of Env clones

2.4.1 Plasmid preparation and transfection

Plasmid DNA used for transfections was extracted from transformed *E. coli* cells that were initially inoculated into 5mL LB containing 100 ug/mL carbenicillin disodium salt and placed on a shaker at 30 ^{o}C for 10 hrs. These were either inoculated from picked colonies or glycerol stocks. The 5ml growth culture was transferred to 50 mL LB containing 100 ug/mL carbenicillin disodium salt and incubated at 30 ^{o}C with shaking for 16 hrs. Plasmids were purified using the PureYield™ Plasmid Midiprep System (Promega®) according to the manufacturer's instructions. HEK 293T cell monolayers were lifted with 1% trypsin (Whitehead Scientific™), stained with trypan blue (Lonza®), counted using a haemocytometer. In order to determine the number of viable cells present, cells were resuspended 1 in 10 in trypan blue (Lonza®), a dye which may only be taken up by dead cells that do not possess an intact cell membrane. All cells which appeared blue were excluded from the cell count. Cells were seeded at a density of $4x10^5$ cells/well in a 6 well plate (NEST®) in 2 mL of growth medium and incubated overnight at 37 ^{o}C in a water-jacket incubator (90% humidity and 5% CO_2). When the cell monolayers were 50-60% confluent, they were co-transfected with 3.25 µg of either *env* clones or an empty vector control, pcDNA3.1 V5-His Topo® (Invitrogen®) and 0.5 µg of pGL4-luc+ vector using Polyethylenimine [PEI (Sigma Aldrich®)]. In 400 µL of DMEM, 21 µL of PEI (1 mg/mL) was mixed with 3.25 µg of DNA. The mixture was vortexed vigorously for 15 sec and incubated at room temperature for 10 min to allow for DNA: PEI-complex formation. Growth medium was removed from HEK 293T cells and fresh medium was added to the cells before the DNA: PEI-mix was added drop wise. After incubation for 6 hrs to allow cells to take up the DNA, the growth medium was replaced with fresh growth medium and the cells were incubated for 48-72 hrs for expression of the Env protein.

2.4.2 Cell lysis and measuring transfection efficiency

Growth medium was removed from the HEK 293T cells. Cell associated Env (gp160, gp120 and gp140) was harvested by adding 300 µL of RIPA lysis buffer (10 mM Tris buffer pH 7.5, 2 mM Na_2EDTA pH 8, 150 mM $NaCl_2$, 1% Triton X-100 detergent and 10 mM phenylmethanesulfonyl fluoride (PMSF) [Sigma Aldrich®] per 50 mL buffer solution) to each well of the six-well culture plate. Lysis was carried out for 10 min at 4 °C or on ice and the cell debris was removed by centrifugation of the lysates at 14000 rpm for 2 min at 4 °C. In order to read luciferase luminescence due to transfection control pGL4-luc+ vector, 30 µL of the supernatants were added to an opaque 96 well plate (Porvair®) and 30 µL of BriteGlo® lysis buffer (Promega®) was added and incubated for 2 min at room temperature. Luciferase activity was measured using a luminometer (Glomax® 96 Modulus Microplate) and relative light units (RLU) was used as an indicator of luminescence. The supernatants were stored at -20 °C and used for protein concentration determination and Env expression experiments. For cells transfected with soluble gp140 *env* constructs, the growth medium from the HEK 293T cells was used immediately (Section 2.4.4) as it contained the released soluble protein.

2.4.3 SDS-PAGE and Western blotting

Total protein concentration of the cell lysates was determined using a Bradford assay (Bio-Rad) according to the manufacturer's instructions. A standard curve was constructed using bovine serum albumin (BSA) and used to determine the unknown protein concentration of cell lysates. To compare Env expression levels, duplicates of BSA standards (concentration range of 0-2000 µg/µL) were averaged and the protein concentration of Bradford assay (BioRad®) was calculated from a BSA standard concentration curve, to normalize the total protein concentrations in the relative samples (Microsoft Excel™ 2010) for western blot analysis. The desired amount of protein was mixed with loading buffer (1% SDS, 8% glycerol, 1% β-mercaptoethanol, and 0.01% bromophenol blue) and denatured for 10 min at 90 °C. Proteins were separated by electrophoresis under denaturing conditions using 8% Bis-Tris polyacrylamide gels and transferred to polyvinylidene difluoride (PVDF) membrane (Immuno-Blot®, Bio-Rad) with transfer buffer (25 mM Tris, 192 mM glycine and 20% methanol) at 100 V for 1 hr using a wet transfer system (Bio-Rad). For Env expression experiments, the house-keeping protein β-actin, was used as a loading control. The membrane was cut into two

sections for detection of the higher molecular weight Envs (160 and 120 kDa) and the lower molecular weight β-actin (43 kDa). For Env processing experiments, the gp41 portion of the full-length Env was detected. The membranes were incubated overnight in blocking buffer (5% fat free skim milk and 0.5% Tween-20 in TBS [50 mM Tris, 150 mM NaCl, pH 7.5]) at 4 °C. The membranes were incubated with the appropriate concentration of primary antibodies (Table 2.3) made in blocking buffer for 1 hr at room temperature with shaking. The membranes were then washed four times with rinsing, in 15 min intervals, with 0.5% Tween-20 in TBS and the appropriate secondary antibody (Horse-radish peroxidase conjugated Goat anti-sheep IgG or Goat anti-mouse IgG for gp120 and β-actin, respectively) (Table 2.4) made in blocking buffer was added and incubated for 1 hr at room temperature with shaking. The membrane was then washed four times, with 15 min intervals, with TBS-T and then with TBS only to remove residual 0.5% Tween. Enhanced chemiluminescence using the Lumino Glo® substrate kit (KPL) and autoradiography films (Santa Cruz Biotechnology®) were used for visualization. The density of each of the bands on the autoradiography film was determined using Image J™, which measures the intensity in pixels of each band and generates peaks. The area under each peak is determined automatically and the percentage relative to the total area is given. The percentage for the wild type (reported as a density) was then used to determine the relative densities of each of the samples. These relative density values were then divided by the densities obtained for the actin control, i.e. the adjusted density. These adjusted densities were then used for statistical analyses.

Table 2.4: Antibody specifications

Name	Protein recognised	Dilution	Source
Sheep anti-gp120	Polyclonal serum specific for HIV-1 gp120	1:2500	ARP* Catalogue number 288
Mouse anti-actin	β-Actin	1:4000	Sigma-Aldrich®
Sheep anti-HIV-1-p24 gag (D7320)	HIV-1 p24	1:600	Aalto Bio reagents Ltd.
Alkaline Phosphatase-mouse anti-HIV-1 p24	HIV-1 p24	1:64000	Aalto Bio reagents Ltd
Horse-radish peroxidase-Goat anti-sheep IgG	Secondary sheep antibodies	1:4000	Santa Cruz Biotechnology®
Horse-radish peroxidase-Goat anti-mouse IgG	Secondary mouse antibodies	1:4000	Santa Cruz Biotechnology®
Horse-radish peroxidase-Goat anti-human IgG	Secondary human antibodies	1:4000	Santa Cruz Biotechnology®

*The antibody was obtained through the NIH AIDS Reagent Program (ARP), Division of AIDS, NIAID, NIH: Antiserum to HIV-1 gp120 from Dr. Michael Phelan.

2.4.4 Secretion, cleavage efficiency and N-glycosylation analysis of Env

HEK 293T growth medium (1 mL) (Section 2.4.2) from transient expression of the soluble gp140 Env was centrifuged at 14 000 rpm for 2 min at 4 °C to remove residual cell debris then diluted with 1mL of MES binding buffer (20 mM morpholinoethanesulfonic acid, 130 mM NaCl, 10 mM $CaCl_2$) at 4 °C. For each Env protein extract, 30 µL of *N-Galanthus nivalis* agarose beads (Vectorlabs®) was added and binding between the lectin agarose beads and Env was carried out overnight at 4 °C with gentle mixing on a rolling apparatus in 2 mL Eppendorf® tubes. The agarose beads were centrifuged at 14000 rpm for 3 min at 4 °C and the supernatant was removed. If the pelleted beads could not be observed after centrifugation, the samples were centrifuged once more as before. For endoglycosidase treatments, beads were washed twice with 1 mL phosphate buffered saline (PBS) (Sigma-Aldrich®) at 4 °C, supplemented with 0.9 mM $CaCl_2$ and 0.49 mM $MgCl_2$ and then treated with endo-β-*N*-acetylglucosaminidase H (Endo H) (0.5 U) or peptide *N*-glycosidase F (PNGase-F) (0.5 U) according to the manufacturer's instructions (New England Biolabs®) then analysed by SDS-PAGE and Western blotting. For secretion analyses, beads were washed four times with phosphate buffered saline (PBS) (Sigma Aldrich®) at 4°C and bound protein eluted with 40 µL 1 M Methyl α-*D*-Mannose-Pyranoside (MMP) (Sigma Aldrich®). The total volume of eluted protein was then

halved for SDS-PAGE and Coomassie staining or Western blotting in order to normalise total protein and compare levels of secretion, respectively[209]. Once the endoglycosidase treated Envs were visualized by autoradiography, the molecular weights of the bands were determined using their migration distances. A standard curve of log molecular weight vs. relative migration distance (Rf) was generated, based on the values obtained for the bands of the molecular weight ladder. The Rf was measured from the top of the gel to the centre of the bands. The molecular weight of glycosidase treated and untreated Env bands were then calculated. Differences in N-glycosylation were determined by comparing the molecular weight of the demannosylated Env and deglycosylated Env with the untreated samples in order to determine the degree of mannosylation[210]. To calculate cleavage efficiency of secreted gp140, the densities of cell associated-gp140, gp140 and gp120 in the culture medium after western blotting was used to determine the relative level of gp120 expressed as a percentage of total Env.

2.5 Pseudovirus production and p24 assay

2.5.1 Transfections

To produce pseudoviruses, HEK293T cells were transfected as above (Section 2.4.1) with the following modifications: PEI (21 µg) was mixed with 2.5 µg of each *env* clone and 5 µg of the viral backbones pSG3Δ*env* or pNL4.3-R-E-luc+. After 48 hrs, growth medium containing the pseudoviruses was collected and clarified through a 0.22 µm pore size filter. The pseudovirus stocks were stored at -80 °C in growth medium supplemented with 20% FBS and an aliquot was inactivated with 1% Empigen® (Sigma-Aldrich®) in TBS to determine p24 concentration.

2.5.2 p24 enzyme-linked immunosorbent assay (ELISA)

In order to determine the concentration of pseudoviruses, a chemiluminescent p24 in-house ELISA (Aalto Bio reagents) and the TROPIX® detection system (CDP-Star®, Applied Biosystems) was used. High bind assay plates (Whitehead Scientific™) were coated with 100 µL of sheep anti-HIV-1 p24 gag antibody (Table 2.4) diluted to a final concentration of 2 µg/mL in coating buffer (100 mM $NaHCO_3$ pH 8.5) and incubated overnight at room temperature. Unbound antibody was removed, and the wells were washed three times with 100µL TBS. Plates were blocked with 5% BSA (Biocom Biotech®) in TBS for 1-2 hrs at room temperature and thereafter

stored at -20 °C until needed. Serial dilutions of the inactivated pseudoviruses were made in 1% Empigen®-TBS and 100 µL was added to the antibody coated plates that were washed four times to remove BSA. A purified p24 protein (0-32 ng/mL) (Aalto Bio reagents) was serially diluted in 1% Empigen®-TBS and used to construct a standard curve. After addition of samples and p24 standards in duplicate to the appropriate wells, the plates were incubated for 3 hrs, before unbound p24 was removed and the wells washed four times with 100 µL TBS. An alkaline phosphatase conjugated mouse anti-HIV-1 p24 antibody (Aalto Bio reagents) was diluted with TBS-0.1% Tween-20 (TBS-T) containing 20% sheep serum (Biocom Biotech) and 2% BSA and 100 µL was added to the plates and incubated for 1 hr at room temperature. Unbound antibody was removed and washed with 100 µL TBS-T eight times and then twice with 100 µL 1X TROPIX® buffer (200 mM Tris, 10 mM $MgCl_2$ pH 8.9). The detection reagent was diluted four times with TROPIX® buffer and 50 µL was added to the wells. Luminescence was read using a multi-well plate reader (Turner Biosystems® Modulus Microplate). Relative light units (RLU) of pseudovirion samples were converted to p24 concentration (ng/mL) using a standard curve and non-linear regression analysis using Microsoft Excel™ 2010.

2.6 Entry efficiency assay

To determine the entry efficiency of the pseudovirion stocks, a TZM-bl cell assay was used. The TZM-bl cell line is a genetically engineered HeLa cell clone that expresses CD4 and co-receptors CCR5 and CXCR4. These cells contain a luciferase reporter gene under the control of an HIV-1 long terminal repeat which is induced upon infection with HIV-1 due to the presence of the Tat protein[211,212]. TZM-bl cell monolayers were lifted with 1% trypsin, stained for cell viability as described before (Section 2.4.1) and counted using a haemocytometer and 10^4 cells in 200 µL growth medium were seeded per well in a 96 well plate (NEST®). The cells were incubated overnight at 37 °C in a water-jacket incubator (90% humidity and 5% CO_2). On the day of infection, the cell monolayers were 50-60% confluent and 5-fold serial dilutions of the pseudoviruses was made in growth medium. All pseudovirus stocks were normalised to either 100 ng/mL or 250 ng/mL p24 prior to serial dilution. Half of the growth medium was removed from each well and replaced with 100µL of titrated pseudoviruses in triplicate. Pseudovirus single round infection was carried out for 48 hrs at 37 °C in the presence of pseudovirus. Growth medium (150 µL) with input pseudovirus was removed from the cells

and 50 µL of BriteGlo® lysis buffer (Promega®) was added. Cells were lysed for 2 min at room temperature, mixed gently and 100µL was transferred to an opaque 96 well plate (Porvair®). Luciferase activity was measured using a luminometer (Glomax® 96 Modulus Microplate) as a proxy for single round replication. RLU curves were plotted by subtracting background RLU obtained from cells to which no virus was added from each triplicate value. The standard deviation of the triplicate values was determined (Microsoft Excel™ 2010) and plotted showing the fold change in entry efficiencies normalised to the wildtype were plotted using GraphPad Prism™ 5.

2.7 Production of chimeric Infectious Molecular Clones (IMCs)

2.7.1 PCR amplification of the *env* gene

In order to generate chimeric infectious molecular clones (IMCs), the *env* gene from the CAP210 E8 TF *env* constructs had to be PCR amplified. Primers were designed in order to anneal to regions which flank the *env* gene and are needed for homologous recombination with the subtype B proviral backbone, pCMV-PBS-LTR-NL4-3Δgp160/URA in yeast cells[213,214]. PCR constituents are listed in Table 2.5.

Table 2.5: Conventional *env* primers for homologous recombination in yeast

Name	**Direction**	**Sequence 5'→3'**
Env_IF	Forward	AGA AAG AGC AGA AGA CAG TGG CAA TGA
Env_IR	Reverse	TTT TGA CCA CTT GCC ACC CAT

Table 2.6: *env* amplification PCR mixture

Reaction mix	**Concentration**
Buffer	1X
dNTP mix	0.2 mM
$MgCl_2$ or $MgSO_4$	1.5 mM
Forward primer	0.2 µM
Reverse primer	0.2 µM
Template DNA	1-50 ng
DNA Polymerase	1-2 U
Total (made up with nuclease-free dH_20)	50 µL

Env PCR was performed at an annealing temperature of 55°C using either KAPA HiFi *Taq* DNA Polymerase kit (KAPA Biosystems®), *Phusion* Hotstart Polymerase (ThermoFischer®), Q5 High Fidelity Polymerase (New England BioLabs®) or Platinum *Taq* High Fidelity Polymerase (Invitrogen®). The cycling conditions were as follows: 94-98 °C for 2 min followed by 25 cycles of 94 °C for 30 sec, 55 °C for 30 sec, 68 °C or 72 °C for 90 sec and a final elongation step at 68 °C or 72 °C for 4 min. To verify whether the PCR yielded the desired 3 kb product, 5 µL of the product was separated on a 0.8% agarose gel. Upon confirmation of successful PCR, the remainder of the PCR products were purified using the Wizard® SV Gel and PCR Clean-Up kit using the manufacturer's instructions (Promega®).

2.7.2 Yeast gap-repair assay

The generation of chimeric IMCs carrying *env* clones using the yeast gap-repair system[213] involves the homologous recombination of the PCR amplified *env* and the linearised pCMV-PBS-LTR-NL4-3Δgp160/URA backbone in yeast cells by transformation.

For transformations, *S. cerevisiae* competent S288C cells prepared as previously described[215] stored at -80°C were thawed and 100 µL cells pelleted by centrifugation at 13 000 rpm for 30 sec. The supernatant was discarded and the following reagents added in this order: *env* and linearised pCMV-PBS-LTR-NL4-3Δgp160/URA vector at 1 µg and 0.2 µg, respectively; 10 µL salmon sperm DNA (Invitrogen®); 240 µL 50% w/v polyethylene glycol (PEG); 36 µL 1M lithium acetate and distilled water to a total volume of 370 µL. Cells were incubated at 30 °C for 30 min followed by heat shock at 40 °C for 15 min. Cells were then centrifuged at 13 000 rpm for 30 sec and the supernatant discarded. Cell pellets were resuspended in 100 µL Tris-EDTA (TE) buffer (100 mM Tris, 10 mM EDTA at pH 8.0). Resuspended cells were spread on yeast amino acid dropout agar [2% peptone, 1% yeast extract, 2% w/v agar, 0.67% w/v complete supplement mixture minus leucine (CSM-leu), 2% dextrose, 0.1% w/v 5-fluoro-1,2,3,6-tetrahydro-2,6-dioxo-4-pyrimidine carboxylic acid (FOA; Zymo Research Corporation™) in distilled water] and grown 2-3 days at 30 °C. Transformation efficiency controls were included: 0.2 µg of undigested pCMV-PBS-LTR-NL4-3Δgp160/URA vector. Replica plates of all colonies were made for further screening and following positive identification of homologous recombination were cultured in 2 mL yeast amino acid dropout media [2% peptone, 1% yeast

extract, 0.67% w/v complete supplement mixture minus leucine (CSM-leu), 2% dextrose in distilled water] overnight at 30 °C.

2.7.3 Colony PCR

To screen for positive colonies with *env* inserted in the correct orientation into the pCMV-PBS-LTR-NL4-3Δgp160/URA backbone, colony PCR with the *env* specific forward and reverse primers, FW-Env-Y and RV-Env-Y were used. If the *env* amplicon was not inserted by homologous recombination, then a PCR product of approximately 1 kb would not be amplified using the Q5 High Fidelity Polymerase kit (Inqaba Biosystems®). Single colonies were picked using sterile tips and added to a reaction mixture whose contents are listed in Table 2.7.

Table 2.7: Colony PCR *env* primers for screening yeast

Name	**Direction**	**Sequence 5'→3'**
FW-Env-Y	Forward	AAT GTC AGC ACA GTA CAA TGT ACA CAT GG
RV-Env-Y	Reverse	GGA GCT GTT GAT TTA GGT ATC TTT C

Table 2.8: Colony PCR mixture

Reaction mix	**Concentration**
Buffer	1X
dNTP mix	0.2 mM
FW-Env-Y	0.2 µM
RV-Env-Y	0.2 µM
Q5 High Fidelity Polymerase	1 U
Total (made up with nuclease-free dH_20)	25 µL

The cycling conditions were as follows: 98 °C for 2 min followed by 25 cycles of 98 °C for 15 sec, 55 °C for 30 sec, 72 °C for 90 sec and a final annealing step at 72 °C for 2 min. The presence of the ~1 kb PCR product (10 µL) was confirmed using a 0.8% agarose gel. Once positive clones were identified, replica colonies were inoculated and cultured as described above (Section 2.7.2). Plasmid DNA for transformation into commercial electro-competent *E. coli* cells (New England BioLabs®) was extracted from yeast cells using the Zymoprep™ Yeast Plasmid Miniprep Kit (Zymo Research Corporation®) according to the manufacturer's instructions. IMC chimeras were confirmed by sequencing as before (Section 2.2.5).

2.7.4 Transformation of electro-competent *E. coli* cells

Following the isolation of full length chimeric IMC DNA from yeast cells (Section 2.7.4), the DNA was then transformed into electro-competent *E. coli* cells (New England BioLabs®). Cells stored at -80°C were thawed on ice for 10 min and mixed gently by inversion. 25 µL of cells were transferred into pre-chilled microcentrifuge tubes and 1-2 µL of the desired DNA added to the cells. The cell/DNA mix was then transferred into pre-chilled 1 mm electroporation cuvettes (Biorad®) and electroporated using 2 kilo Volts (kV), 200 ohms (Ω) and 25 micro Farad (µF). 975 µL of pre-warmed SOC outgrowth medium (2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$ in distilled water) was added immediately with gentle mixing. Cells were allowed to recover for 1 hr at 37 °C. 100 µL cells were spread on Luria agar (LB, 1.5% w/v agar and 100 µg/mL carbenicillin disodium salt) plates and grown overnight at 37 °C[205]. Transformation efficiency controls were included at 10 pg of pUC19 (Invitrogen®). Colonies were picked, cultured and plasmid DNA extracted as before (Section 2.2.2) for screening.

2.7.5 Testing envelope functionality of chimeric IMCs

The functionality of the chimeric IMC constructs (pCMV-PBS-LTR-NL4-3Δgp160/URA + *env*) (Section 2.7.4) was tested using the TZM-bl cell assay as outlined above (Section 2.6) with modifications. Firstly, the chimeric IMC constructs were transfected (Section 2.6.1) with the following modifications: HEK 293T cells were seeded at a density of 10^4 cells/well in a 96 well plate (NEST®) in 100 µL of growth medium and incubated overnight at 37 ^{o}C in a water-jacket incubator (90% humidity and 5% CO_2). On the day of transfection the cell monolayers were 50-80% confluent. For each well of a 96 well plate (NEST®), 50 µL of DMEM and 2 µL of PEI were mixed with 0.225 µg of chimeric IMC DNA and 0.225 µg of pCMV_NL4-3LTR-Gag4 helper plasmid DNA (donated by Dr Manish Sagar, Brigham and Women's Hospital, Harvard Medical School, USA). HIV-$1_{pNL4.3}$ full-length proviral backbone and pCMV-PBS-LTR-NL4-3Δgp160/URA were used as positive and negative controls, respectively. This was done in duplicate for each chimeric IMC construct and included a no virus control to measure background. The DNA-PEI mixture was incubated at room temperature for 15 min. All growth medium was removed from HEK 293T cells and replaced with 100 µL fresh growth medium. 50 µL of the DNA-PEI

mix was transferred to each well of the HEK 239T cells plate and incubated for 48 hrs. After 24 hrs, TZM-bl cells were seeded at 10^4 cells per well in a 96 well plate (NEST®) in 200 µL growth medium and incubated overnight. A volume of 100 µL growth medium was removed and HEK 293T cells supernatants (100 µL) were then transferred to the TZM-bl cells. Infection was carried out for 48 hrs at 37 °C and virus input was not washed out during this time. Luciferase activity was measured as outlined before (Section 2.6) and IMC constructs that gave luminescence 2.5-fold above background were selected for replication in donor PBMCs as previously suggested by Li *et al.* (2005).

2.8 Replication assay

2.8.1 Isolation of peripheral blood monocytic cells (PBMCs)

Buffy packs were obtained from healthy donors who were negative for HIV-1, syphilis and hepatitis B and C from the Western Province Blood Transfusion Services. Ethics was approved by the Faculty of Health Sciences Human Research Ethics Committee at The University of Cape Town (approval number: 544/2017). Peripheral blood mononuclear cells (PBMCs) were isolated using Histopaque (Sigma-Aldrich®) gradient density centrifugation with Leucosep tubes (Greiner Bio-One®) according to the manufacturer's instructions (Thorsby & Bratlie, 1970). Fifteen mL Histopaque was loaded onto a Leucosep tube and centrifuged for 1 min at 2500 rpm at room temperature. Whole blood was diluted 1 in 3 with 50% Roswell Park Memorial Institute medium (RPMI) 1640 (Lonza®), with no supplementation, and 50% phosphate buffered saline (PBS; Sigma-Aldrich®) and 30 mL of the diluted blood was added onto the 50 mL Leucosep tube with Histopaque followed by centrifugation at 2500 rpm for 15 min at room temperature. After centrifugation, three layers were observed: PBMCs separate into a clean white buffy coat with red blood cells remaining in the Histopaque. The PBMCs were removed with a Pasteur pipette, and washed twice with PBS supplemented with 1% (v/v) FBS (Biocom Biotech®). After each wash the samples were centrifuged for 5 min at 1200 rpm. During the second wash, 10 µL of the PBMC sample in PBS was added to 90 µL trypan blue (Lonza®) and counted on a haemocytometer to determine cell number and viability. If PBMCs were still contaminated with red blood cells after the second wash an additional wash was performed to remove thrombocytes before counting. From one buffy-coat ± 300-500 million (Ml) cells can be obtained. Isolated PBMCs were cultured in high

glucose (4.5 g/mL) RPMI 1640 (Lonza®) with 10% (v/v) FBS (Biocom Biotech®), 2 mM L-glutamine (Sigma-Aldrich®), 25 mM HEPES (Lonza®) and 100 mg/mL penicillin and 100 mg/mL streptomycin (Sigma-Aldrich®) at 37 °C in a water-jacket incubator (90% humidity and 5% CO_2) for 72 hrs. To make cell stocks, cells were stained for cell viability (Section 2.4.1) and counted as before and centrifuged at 1500 rpm for 8 min and 20-50 MI cells were resuspended in a final volume of 1 mL in 10% DMSO (Sigma-Aldrich®) in FBS (Biocom Biotech®) and incubated at -80 °C overnight and stored in liquid N_2.

2.8.2 Transfection to produce infectious virus

To produce chimeric IMCs, HEK293T cells were transfected as above (Section 2.4.1) with the following modifications: PEI (6 µg) was mixed with 1 µg of each chimeric IMC backbone (pCMV-PBS-LTR-NL4-3Δgp160/URA + *env*) and 1 µg of the helper vector pCMV_NL4-3LTR-Gag4. After 48hrs, growth medium containing the viruses was collected and clarified through a 0.22 µm pore size filter. Virus stocks were stored at -80 °C in 500 µL growth medium supplemented with 20% FBS and an aliquot was inactivated with 1% Empigen® (Sigma Aldrich®) in TBS to determine p24 concentration (Section 2.5.2) and tissue culture infectious dose (TCID50).

2.8.3 Tissue Culture Infectious Dose (TCID50) for titration of infectious virus

In order to determine the virus titres of the chimeric IMCs (Section 2.8.2), the TZM-bl assay was used as described Edmonds *et al.* (2010) with a few modifications[216]. TZM-bl cells were seeded as before (Section 2.6) and infected with 4-fold serial dilutions of virus stocks, including a DMEM (growth medium only) control. Each dilution was performed in quadruplet and luminescence read after 48 hrs as previously described (Section 2.6). The titre of the virus stock was determined using the Reed-Muench method and expressed as log infectious units (IU)/mL[212]. TCID50 was used in order to calculate the multiplicity of infection (MOI), i.e. the number of infectious particles per cell infected.

2.8.4 Testing PBMC donor selection

To troubleshoot for differences in infectivity based on PBMC donor variation, PBMCs from three different HIV-negative donors were isolated as before (Section 2.8.1) and some cells were activated, and the remaining cells cryopreserved in liquid N_2. Proliferation was stimulated with 200 U/mL natural human Interleukin-2 (IL-2) (Gentaur®) and 0.5 µg/mL phytohemagglutinin-P lectin (PHA-P) (ThermoScientific™) for 72 hrs in high glucose (4.5 g/mL) RPMI 1640 (Lonza®) containing 10% (v/v) FBS (Biocom Biotech®), 2 mM L-glutamine (Sigma-Aldrich®), 25 mM HEPES (Lonza®) and 100 mg/mL penicillin and 100 mg/mL streptomycin (Sigma-Aldrich®). Activated PBMCs were seeded at a density of 10^6 cells/well in a 96 well plate (NEST®) in 100 µL and 300 TCID50 units of replication competent HIV-$1_{pNL4.3}$ diluted in RPMI (with no supplementation) to a total volume of 200 µL used to infect in triplicate. 50 µL culture supernatant was collected for p24 ELISA (Section 2.5.2) and 50 µL fresh RPMI (no supplementation) was added to the plate for further culturing. Viral replication in the three donors was compared by plotting the p24 ng/mL over days post infection. Donors which showed the highest replication for all viruses tested were taken forward for viral expansion of chimeric IMCs and all subsequent replication experiments.

2.8.5 Viral expansion in activated donor PBMCs

In order to produce high virus titre, IMCs produced in HEK 293T cells were expanded in PBMCs. Activated PBMCs were stained for cell viability as described before (Section 2.4.1) and plated in 6 well plates at a concentration of 5 X 10^6 cells per 2 mL of Full RPMI medium supplemented with IL-2 only. Virus stocks were thawed and 500 µL was added to the cells. Spinoculation was performed by centrifugation at 1500 *x g* for 2 hrs at room temperature to enhance cell infection. The infected cells were then transferred to T25-tissue culture flasks (NEST®) in a total volume of 5 mL Full RPMI supplemented with IL-2 so that cells were at a final concentration of 10^6 cells per mL. Cells were then incubated at 37 ^oC in a water-jacket incubator (90% humidity and 5% CO_2) for 14 days. Infected PBMCs were refreshed with 5 million activated PBMCs in fresh Full RPMI supplemented with IL-2 on day 7 and 500 µL virus supernatant was collected on days 7 and 10 post infection. Virus supernatant was aliquoted in cryovials and stored at -80 °C.

2.8.6 Concentrating IMCs by ultra-centrifugation

To concentrate chimeric IMCs expanded in PBMCs (Section 2.8.5), virus was layered onto a 20% glycerol cushion in a total volume of 5 mL and ultra-centrifuged in the Optima L-80 XP® ultracentrifuge at 26000 rpm for 2 hrs at 4 °C using the SW 55Ti rotor (Beckman Coulter). The virus pallets were resuspended in 50 µL of PBS and stored at -80 °C until required.

2.8.7 Infection of activated PBMCs for replication kinetics

Stored stocks of PBMCs identified to best support replication (Section 2.8.4) were thawed in supplemented high glucose (4.5 g/mL) RPMI 1640 (Lonza®) containing 10% (v/v) FBS (Biocom Biotech®), 2 mM L-glutamine (Sigma-Aldrich®), 25 mM HEPES (Lonza®) and 100 mg/mL penicillin and 100 mg/mL streptomycin (Sigma-Aldrich®). PBMCs were activated and seeded as before (Section 2.8.4) and 100 TCID50 units of each chimeric virus diluted in RPMI (with no supplementation) to a total volume of 200 µL used to infect in triplicate. HIV-$1_{pNL4.3}$ virus was used as a positive control and cells without infection was included as a negative control. To enhance cell infection, spinoculation by centrifugation at 1 500 x *g* for 2 hrs at room temperature was performed. Infected cells were then incubated at 37 ^{o}C in a water-jacket incubator (90% humidity and 5% CO_2) as before (Section 2.8.4). 50 µL virus supernatant was collected on day 0, 7, 10 and 14 post infection and stored at -80 ^{o}C for p24 ELISA (Section 2.5.2). 50 µL fresh RPMI (no supplementation) was added to the plate for further culturing. Replication kinetics was plotted as p24 ng/mL over days post infection, average of the slope of each curves was calculated and plotted as a bar graph using GraphPad Prism™ 5.

2.9 Statistical analyses

To determine the fold changes in the expression and secretion of the Env after the incorporation of the mutations (Section 2.2.1) compared to the wild type, a one-way ANOVA with Tukey post-tests was performed using the GraphPad Prism™ 5. Histograms of the mean with standard error bars from the three biological repeats were generated. This statistical test was also used to compare the level of demannosylation after endoglycosidase treatment as well as entry efficiencies of pseudoviruses and chimeric IMCs. For the replication assay, the mean slopes of the curves of p24 concentration over time were calculated and plotted as a bar graph with error bars indicating the standard deviation of the mean slopes for the two

donors. GraphPad Prism™ 5 was used to perform a two-way ANOVA with Tukey post-test. Statistical significance was indicated as *, ** and *** for p-values less than 0.05, 0.01 and 0.001.

CHAPTER 3: IMPACT OF SUBTYPE B TRANSMISSION MOTIFS ON SUBTYPE C ENVELOPE FUNCTION

3.1 Introduction

Seventy percent of HIV-1 infections worldwide are due to heterosexual transmission with the remainder owing to homosexual contact, maternal-infant infection, and injection drug use[217]. According to the 2017 UNAIDS Data Book, new infections among young women in 2016 were 44% higher than they were among men in the same age group[3], thus indicating that females at higher risk of infection than men[218]. With the female genital mucosa the primary site of HIV-1 transmission[14], studies have focussed on how mucosal surfaces function as a barrier to the incoming virus. The low transmission rate per heterosexual coital act suggests that the genital mucosa is a formidable barrier[219]. Furthermore, the frequency of multiple variant infections were found to increase dramatically in the presence of chronic ulcerative disease or inflammatory genital infection in the recipient, suggesting that disruption of the barrier might abrogate mucosal protection[179]. Sexually transmitted infections (STIs) that cause inflammation have also been associated with increased susceptibility to HIV infection[220] particularly in young women[221,222]. Elevated genital concentrations of chemokines and inflammatory cytokines have been suggested to contribute to the high risk of HIV acquisition in African women[223].

There has been a suggestion that the barrier function of the mucosa results in a genetic bottleneck during HIV transmission. The single variant, termed the transmitted founder (TF), is able to overcome the mucosal barrier and establish infection due to a viral phenotype that provides an advantage during transmission[15,177,178]. TFs lack diversity within the *env* gene[178] and are R5-tropic, supporting the hypothesis that the viral Envelope (Env) may be responsible for the selection of specific variants in the female genital tract (FGT)[14,179,182]. Currently prophylactic HIV-1 vaccine trials target either: a) Env[224] which is found on the surface of the virus and is responsible for binding to receptors on the surface of target cells and mediating the first step in viral replication[15,177,178], or b) Gag, the capsid protein which encloses the viral RNA[225]. It has been hypothesised that TFs carry motifs within Env that facilitate

transmission[226] and previous studies have identified phenotypic and genotypic characteristics of TFs that may play a role in their selection at the mucosal barrier[74,227–230].

Up to now, most research has focussed on HIV-1 subtype B which predominates in developed countries while subtype C contributes to 50% of infections globally[2,231] and is the dominant subtype in Southern Africa[17,232]. The identification of common phenotypic traits essential for HIV-1 transmission may help in the development of a universal preventative vaccine by targeting these motifs across subtypes. However, thus far, Env phenotypes have differed across subtypes except that all TFs utilise CCR5 co-receptor for viral entry[7] and are T-tropic[227]. Analysis of TF *env* sequences have revealed that subtype A and C TF Envs have shorter variable loop lengths and thus less potential N-glycosylation sites (PNGS)[15,65] whilst utilising CCR5 co-receptor only[227]. However, this trend for fewer PNG sites has not been consistent across all studies[186] and not always observed for subtype B analyses[8,185,190]. Subtype B TF Envs were found to be enriched for a Histidine at position 12 of the signal peptide (His12)[8]. This polymorphism was found to increase Env expression leading to more infectious virus[13]. Gnanakaran *et al*. (2011) also showed that TF Envs lacked a PNG at position 413-415 whereas variants from chronic infection tended to carry a PNG at this site. Although the PNG at 413-415 was proximal to receptor binding domains and could play a potential role in immune escape[8] the motif was not characterised.

The characterisation of TF Envs is important as the identification of potential transmission motifs can provide targets for the development of preventative vaccines. The two robust transmission motifs previously identified in TF *env* sequences of subtype B[8], may provide insight into conserved motifs that are important in the transmission of all other subtypes. Therefore, the overall aim of this study was to investigate whether a subtype C TF variant, CAP210 E8, was phenotypically advantaged by His12 and PNG413 transmission motifs. Unlike subtype B *env*, subtype C TFs carried a Glutamine at position 12 (Gln12) within the signal peptide (Env-SP) and the PNG at position 413 was present. Thus, in this study CAP210 E8 was mutated to carry His12 and PNG413 was deleted and the mutants were characterised phenotypically to determine their potential influence on Env expression, secretion, N-glycosylation and HIV-1 infectivity.

3.2 Results

3.2.1 Site-directed mutagenesis of the TF *env* signal peptide

In order to investigate whether the robust transmission motifs identified in subtype B TFs[8] influenced phenotypic properties of a subtype C TF variant lacking these motifs, these signatures were introduced by site-directed mutagenesis (SDM). The subtype C TF *env* clone, CAP210 E8 carried a Gln12 in its Env-SP that was mutated to either a His12 or an Ala12 residue (Figure 3.1). The Gln was mutated to an Ala to confirm that the His residue itself was important for Env function and not merely a change in amino acid identity. The PNG413 recognition site was deleted by replacing a Threonine (Thr) at position 415 with an Isoleucine (Ile). The Ile residue was selected as it was the most common amino acid at that position of subtype C Envs of variants from chronic infection. These mutations were introduced in both the full-length (gp160) and soluble (gp140) forms of this *env* clone (Riley Traviss, IBMS, UCT) (Figure 3.1). The soluble *env* clone (generated by Evelyn Ngwa, MCB, UCT) has a stop codon introduced at position 686 (HXB2 numbering) upstream of the gp41 membrane spanning domain so that only 20kDa of gp41 is included after translation[233,234]. Although gp120 monomers are typically used in studies where secreted Env is required, gp140 derivatives were used as it has been shown to readily form native Env trimers[146,235–237].

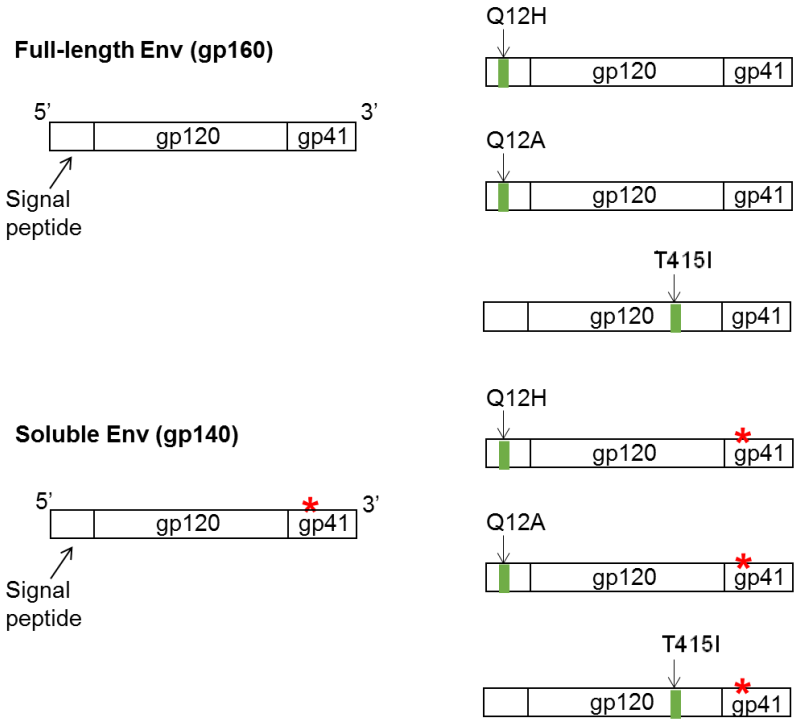

Figure 3.1: Graphical representation of the site-directed mutagenesis of the CAP210 E8 TF *env* clone. Soluble (gp140) and full-length (gp160) forms of the Env were mutated in their signal peptides and at a potential N-glycosylation site (PNGS), indicated in green, generating six different constructs. Position 12 of the signal peptide is indicated by an arrow with mutations at this position from Glutamine (Q) to Histidine (H) or Alanine (A) in the respective *env* constructs. The Threonine (T) at position 415 of the PNG was mutated to an Isoleucine (also indicated in green). The stop codon present in the soluble Env is indicated by a red asterisk.

3.2.1.1 Confirmation of successful site-directed mutagenesis PCR

The primers used for SDM PCR were designed to introduce a silent mutation 6 base pairs (bp) downstream of position 12 to include an additional *Nde*I restriction enzyme recognition site to screen for mutants.

SDM of gp160 and gp140 was successful as a PCR product of approximately 8.1 kilobases (kb) was observed, corresponding to the size of the CAP210 E8 TF *env* clone (data not shown). The transformation of the *Dpn*I-digested PCR products yielded ampicillin resistant colonies, and these were screened using restriction enzyme digestion by *Nde*I. A restriction map, indicating the position of the *Nde*I recognition sites and expected DNA fragments is shown in Figure 3.2A.

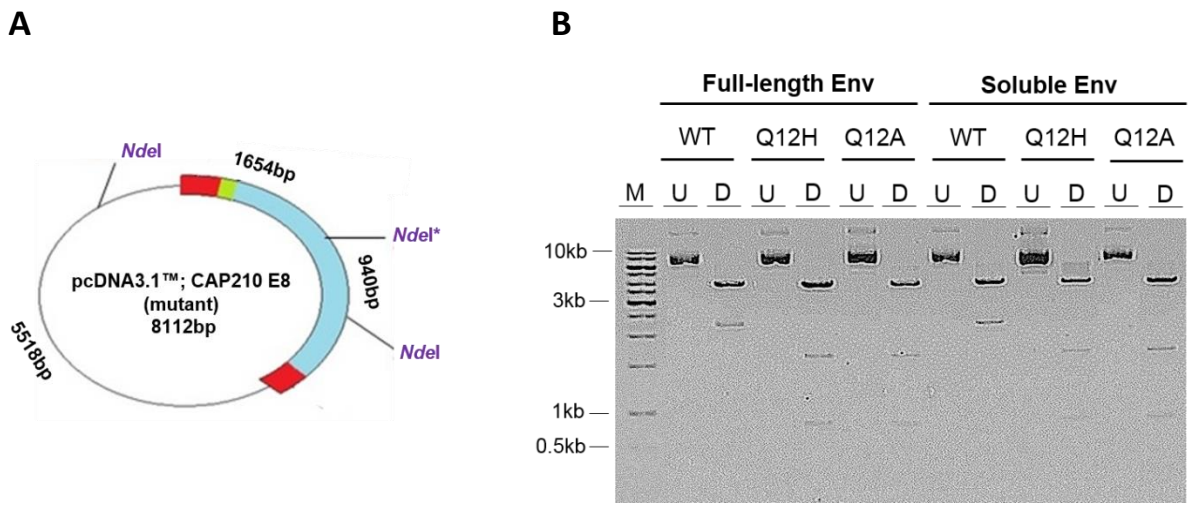

Figure 3.2: Restriction enzyme digestion to screen for successful site-directed mutagenesis of the CAP210 E8 TF *env* constructs. (A) A restriction enzyme map of CAP210 E8 *env* cloned into the pCDNA3.1 vector indicating the *Nde*I restriction enzyme sites with the site introduced *via* site directed mutagenesis indicated in purple by an asterisk. The signal peptide is indicated in green and the flanking regions of the *env* insert (in blue) are indicated in red. The sizes of the DNA fragments generated by *Nde*I digestion are shown. **(B)** Wild type (WT) and mutant (Q12H and Q12A) CAP210 E8 plasmid DNA was extracted from ampicillin resistant colonies after transformation and digested with *Nde*I before DNA fragments were separated on a 1% agarose gel. Gel electrophoresis of undigested (UD) and digested (D) WT and mutated clones are indicated.

The expected banding pattern of wildtype (WT) CAP210 E8 full-length and soluble clones digested with *Nde*I included DNA fragments with molecular weights of 5518 bp and 2594 bp. The presence of the silent mutation that introduced an additional *Nde*I site digested the 2594 bp band into 1654 bp and 940 bp DNA fragments. This 940 bp band was used to distinguish WT constructs from Env-SP mutants. After digestion with *Nde*I, the banding pattern of the plasmid extracted from these ampicillin resistant colonies had DNA fragments of molecular weights corresponding to 5518 bp, 1654 bp and 940 bp indicating that all four clones tested had been successfully mutated by SDM. Furthermore, when the clones were sequenced the introduction of the relevant base changes was confirmed (Figure 3.3) and no additional PCR-derived mutations were identified in full-length *env* (data not shown).

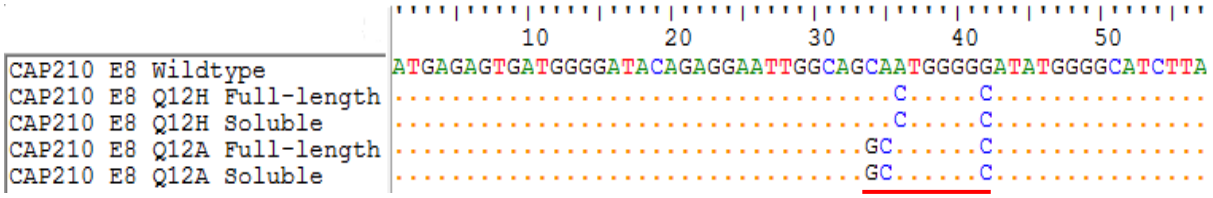

Figure 3.3: Confirmation of mutations at position 12 of Env-SP. WT CAP210 E8 TF *env* sequence was aligned to the four constructs generated by site-directed mutagenesis. Dots in the alignment indicate identical nucleotides to the CAP210 E8 WT sequence. The Glutamine to Histidine or Alanine mutation is shown as CAA to CAC or GCA, respectively. Nucleotide changes to introduce these mutations and the additional *Nde*I restriction enzyme site (underlined in red) in CAP210 E8 SDM constructs are shown. Only the relevant region of *env* is shown although the entire *env* was sequenced.

3.2.2 Charge distribution of signal peptides

The number of positively charged residues present in the N-terminal portion of Env-SP has previously been identified as determinants of the level of expression and secretion of Env[107]. Figure 3.4 shows the overall charge distribution (at physiological pH) within the first 14 amino acids of the CAP210 E8 TF *env* constructs versus a highly expressed subtype B TF *env*, AA01[13] carrying the transmission motif in the signal peptide. WT and signal peptide mutant Q12A appeared to have the same overall charge of +2 whilst Q12H had an overall charge of +3 due to the Gln to His substitution. The subtype B *env* however had the most positively charged N-terminal region with a net charge of +4.

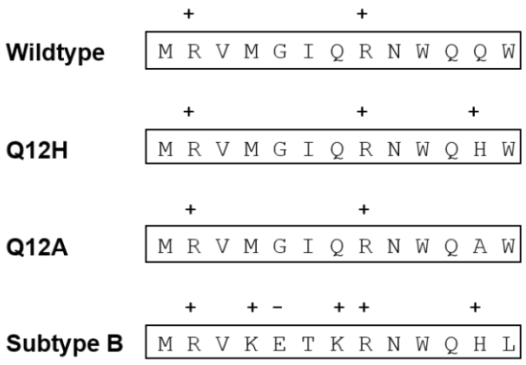

Figure 3.4: Schematic diagram of the charge distribution of the N-terminal region. The signal sequences of the CAP210 E8 TF *env* constructs (WT and signal peptide mutants, Q12H and Q12A) were aligned to that of a subtype B TF *env*[12] to compare differences in overall charge. Plus (+) and minus (-) signs denote positively and negatively charged residues at physiological pH, respectively.

3.2.3 Envelope expression

Signal peptide cleavage is intrinsically linked with Env translation, folding and processing ultimately influencing expression levels[238]. Due to the association between signal peptide

cleavage and protein expression levels HEK 293T cells were transiently transfected to express WT and mutant gp160 and gp140 and cells were lysed for Western blotting. As transfection efficiency could influence apparent expression levels, the firefly luciferase expression plasmid, pGL4-luc, was co-transfected (Figure 3.5). A representative graph of one independent experiment is shown to indicate the lack of variation in transfection efficiency between WT and mutants. This suggested that any variation in levels of cell-associated gp160, gp140 and gp120 detected by SDS-PAGE and Western blotting would not be due to transfection efficiency (Figure 3.6).

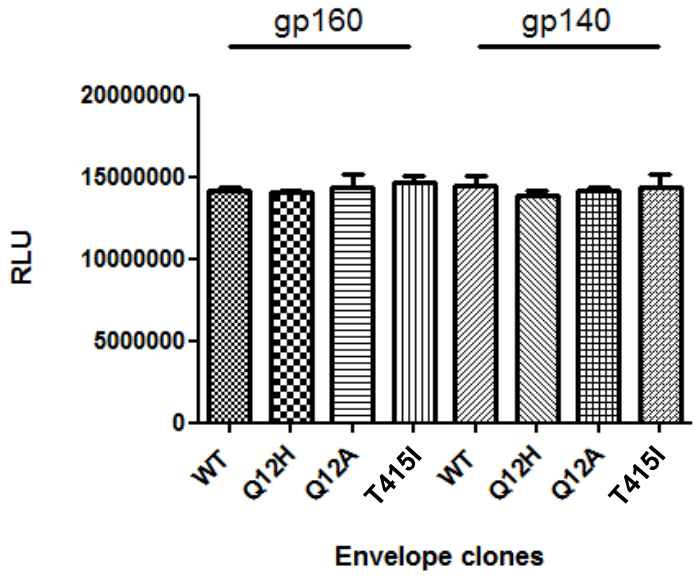

Figure 3.5: Transfection efficiency of wild type and mutated *env* constructs. HEK 293T cells were transfected with *env* and pGL4-luc plasmid DNA (3.25 µg and 0.5 µg, respectively). Cell lysates were used to measure luminescence and luciferase readings were expressed as relative light units (RLU) in duplicate. Bar graph showing the average RLUs was plotted using GraphPad Prism® with error bars indicating standard deviation.

The soluble recombinant gp120 used as a positive control migrated to approximately 129kDa which is 9 kDa greater than what was expected and gp160 and gp140 migrated at 164kDa and 145kDa, respectively (Figure 3.6). The observed molecular weights of gp160 and gp140 were very similar to the expected sizes (160kDa and 140kDa, respectively) and the slight difference could be due to variation in N-linked glycosylation[146,239]. All gp160 and gp140 constructs expressed Env as a single dense band. As Western blotting was able to resolve bands representing gp160 and gp140 of 164kDa and 145kDa, respectively, it was unclear why gp120 bands were not detected. It is possible that precursor Env, in this case uncleaved gp160 and

gp140, was more abundant relative to gp120 and thus easier to detect at this protein concentration. It has been reported that only a small fraction of gp160 is cleaved into gp120 and this is dependent on cell type[102]. Asmal *et al*. (2011) used Jurkat cells to compare Env expression levels making it difficult to make a direct comparison between the two sets of data. However, gp120 was not detected for some clones analysed by Asmal *et al*. (2011) and there seemed to be an association between increased expression of gp160 and detection of gp120[13]. It was thus possible that by increasing the concentration of total protein per sample analysed by SDS PAGE might have improved detection of gp120, however even after loading 150 µg of total protein, a single band representative of gp160 was observed (data not shown). However, we could not compare protein concentrations as these were not indicated by Asmal and colleagues.

Furthermore, the bands corresponding to gp160 and gp140 lacked the smear expected of heavily glycosylated proteins[235] and instead appeared as clear, sharp bands. Land *et al.* (2003) showed that most gp160 was found in the ER due to slow folding and delayed removal of Env-SP[126] thus most cell-associated Env would be carrying homogeneous high mannose N-glycans ($Glc_3Man_9GlcNac_2$). Therefore, the sharpness of the bands corresponding to gp160 and gp140 (Figure 3.5) could be due to homogeneous N-glycosylation.

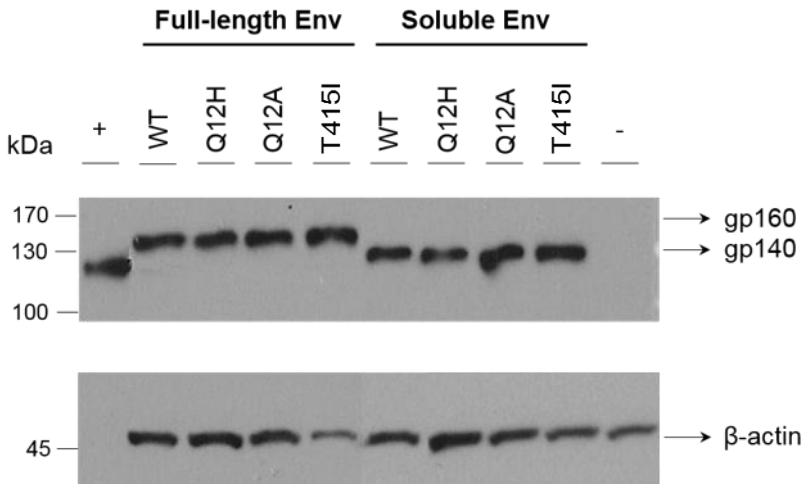

Figure 3.6: Expression of wild type and mutated *env* constructs. HEK 293T cells were transfected with *env* and pGL4-luc plasmid DNA (3.25 µg and 0.5 µg, respectively) and 100 µg of total protein was analysed by SDS-PAGE and Western blotting. Anti-gp120 and anti-actin antibodies were used to detect gp120 and gp160 and the loading control protein β-actin, respectively. The wild type (WT) and mutated (Q12H, Q12A and T415I) full-length and soluble Envs are indicated. Expected positions of the bands representing gp160, gp140 and beta-actin, indicated with an arrow, were based on the standard curve of the molecular weight marker. The plus sign (+) represents 30 ng of recombinant purified HIV-1 gp120 (IIB-CHO from the AIDS reagent programme) and the minus sign (-) represents cells transfected with an empty vector (pcDNA3.1 V5-His Topo® [Life technologies™]), used as a positive and negative control, respectively. The PageRuler Prestained Protein Ladder (Thermo Scientific) molecular weight ladder is indicated in kilo-Daltons (kDa). A representative of three independent experiments is shown.

Densitometry analyses normalised Env band intensities to the loading control, beta-actin (gp160/actin; gp140/actin) as well as transfection efficiency (Figure 3.5). Normalised densities of the mutants (Q12H, Q12A and T415I) were plotted as fold change in expression compared to WT (Figure 3.7). Signal peptide mutants, Q12H and Q12A, showed no significant difference in gp160 and gp140 expression compared to WT. Thus, unlike subtype B Env-SP, the His12 within the signal peptide did not influence gp160 nor gp140 expression[13,203]. However, as Western blotting might not be sensitive enough to detect slight changes in expression, flow cytometry might have improved detection limits.

In contrast, the expression of the T417I gp160 mutant was significantly 3.5-fold higher than WT ($p=0.0034$). The apparent lack of increased expression of the T415I gp140 mutant was likely due to the absence of a TM domain anchoring the mutant Env within the plasma membrane. Levels of cell-associated gp140 would not increase despite increased expression levels because, as soon as gp140 was processed and transported to the plasma membrane, it

was released into the medium. Therefore, if there was an increase in gp140 expression, then there should also be an increase in secreted gp140.

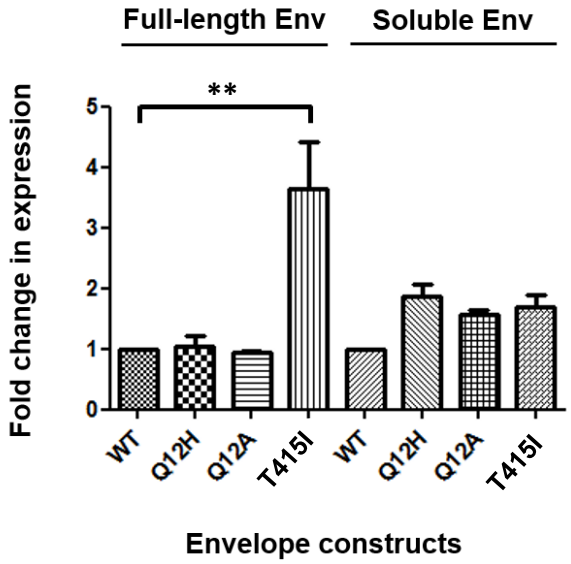

Figure 3.7: Fold changes in expression of WT and mutated Env constructs. Densitometry analysis of three independent Western blots comparing gp160 and gp140 mutants to wild type (WT) were used to determine the fold-change in expression relative to WT. Fold-changes in expression of the WT full-length (gp160) and soluble Env (gp140) versus those containing the Q12H, Q12A or T417I mutations are indicated by each bar. An asterisk indicates statistical significance (** <0.01). Error bars indicate standard deviation and data was compared using GraphPad Prism™ 5 with a one-way ANOVA with the Tukey's post-test.

3.2.4 Envelope secretion

As the His12 enhanced secretion of subtype B TF Env[13], and gp140 T415I cell-associated levels were not increased relative to WT despite increased gp160 expression, we wanted to determine whether CAP210 E8 mutants were secreted differently. To compare the secretion of gp140 carrying the subtype B transmission motifs to WT, the soluble *env* constructs were transfected into HEK 293T cells, the culture medium was collected and secreted Env was bound to N-*Galanthus nivalis* lectin-conjugated agarose beads in a pull-down assay (Section 2.4.4) to concentrate the glycoproteins. It was highly likely that the amount of Env bound to the agarose beads varied between samples, making it difficult to compare mutants to WT. Therefore, we normalised Env to the total amount of glycoproteins bound to the N-*Galanthus nivalis* lectin beads using Coomassie staining[209]. One half of the total volume protein eluted from the N-*Galanthus nivalis* lectin-conjugated agarose beads was used for detection of

gp120 *via* Western blotting (Figure 3.8A) and the other for quantification of total glycoprotein loaded on SDS-PAGE *via* Coomassie blue stain (Figure 3.8B).

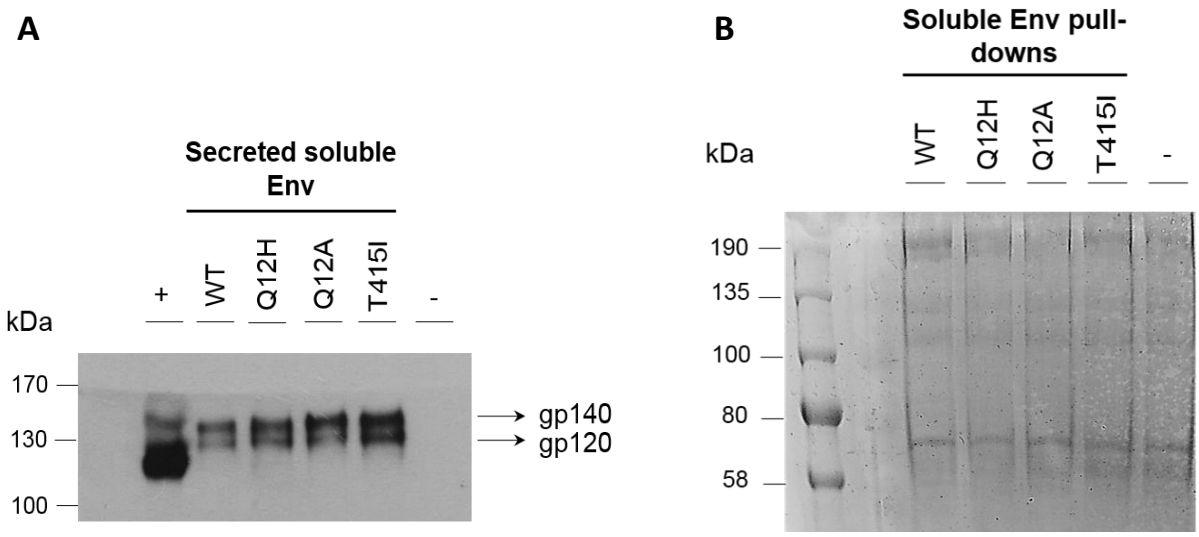

Figure 3.8: Secretion of wild type and mutated gp140. HEK 293T cells were transfected with 3.25 µg soluble *env* (gp140) plasmid DNA and Env secreted into the culture medium partially purified using N-*Galanthus nivalis* lectin agarose beads (Vectorlabs®). Bound protein was eluted and half run on SDS-PAGE and analysed by Western blotting (A) and the other half separated by SDS-PAGE and stained with Coomassie (B). Anti-gp120 antibody was used to detect gp120 and gp140 and the Coomassie stain was used to control for loading. The WT (WT) and mutated (Q12H, Q12A and T415I) soluble Envs are shown and densitometry analysis was used to calculate the amount of total gp120 gp14 relative to the total protein per lane of the Coomassie stained SDS PAGE. Expected positions of the bands representing gp140 and gp120, indicated with an arrow, were based on the standard curve of the molecular weight marker. The plus sign (+) represents 30 ng of recombinant purified HIV-1 gp120 (IIB-CHO from the AIDS reagent programme) and the minus sign (-) represents cells transfected with an empty vector (pcDNA3.1 V5-His Topo® [Invitrogen®]), used as a positive and negative control, respectively. The PageRuler Prestained Protein Ladder (Thermo Scientific) molecular weight ladder is indicated in kilo-Daltons (kDa). A representative of three independent experiments is shown.

Gp140 may be secreted in two forms: uncleaved gp140 and gp120. As expected, two bands were observed for all soluble Env pull-down samples and had an apparent molecular weight of 145kDa and 125kDa, respectively (Figure 3.8A). Gp140 carrying the signal peptide mutations (Q12H or Q12A) showed no significant difference in the amount of total gp140 and gp120 secreted compared to WT ($p > 0.05$) (Figure 3.9), once again, contrary to the findings of Asmal *et al.* (2011). However, less Q12A gp120 was detected in the culture medium, suggesting that gp140 Q12A was not cleaved as efficienctly. On the other hand, the T415I gp140 and gp120 mutants were secreted 3-fold and 4-fold higher than WT, respectively ($p < 0.05$), very similar to the fold change in gp160 expression. This strongly suggested that T415I

gp140 was expressed to higher levels compared to WT, similar to gp160, which resulted in increased levels of secretion into the culture medium (Figure 3.9).

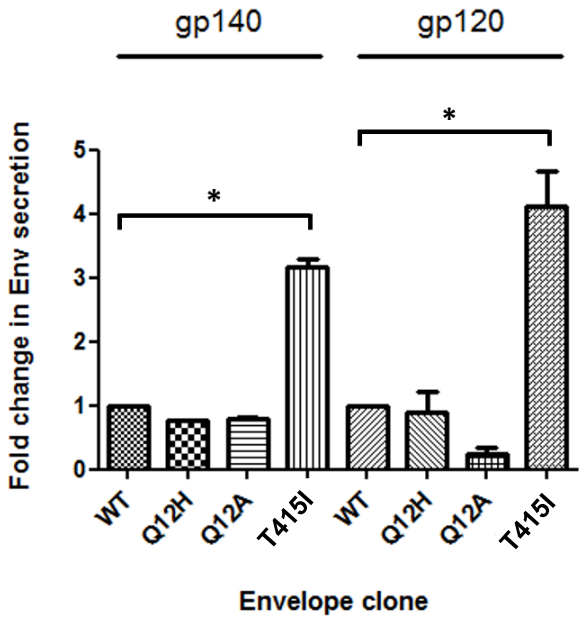

Figure 3.9: Secretion of WT and mutated gp140 and gp120. The level of wild type and mutated gp140 secreted into the culture medium was determined by normalising the band intensities of uncleaved gp140 and gp120 by the total band intensity of the corresponding lane of the Coomassie stained gel. Levels of secretion are shown as mean fold changes relative to wild type with error bars representing standard deviation. Statistical analysis was carried out using GraphPad Prism™ 5 one-way ANOVA with the Tukey's post-test. An asterisk indicates statistical significance (* < 0.05). This histogram represents the mean fold changes in Env expression of three independent experiments.

Gp160 and gp140 T415I, had significantly enhanced levels of both expression and secretion compared to wild type. The high secreted levels indicated rapid release of gp140 into the culture medium thus preventing cell-associated gp140 levels from increasing. Therefore, increased expression of Env did not result in greater accumulation of unprocessed precursor in the ER, but instead, led to an increased proportion of Env trafficked to the plasma membrane.

3.2.5 Envelope cleavage

As viral infectivity depends on the incorporation of functional trimers[92], gp160 cleavage into gp120 could be a very important determinant of pseudovirus fitness. His12 and PNG413 might influence gp160 cleavage so that fewer trimers are produced. However, cell-associated gp120

was not detected by Western blot and thus the cleavage efficiency of gp160 into gp120 could not be calculated as previously outlined[240]. Despite the absence of cell-associated gp120 (Figure 3.6), gp120 was detected in the culture medium for gp140 WT and all mutants (Figure 3.8A). This suggested that after gp140 cleavage, gp120 monomers or trimers were rapidly secreted into the culture medium leaving undetectable levels in the cell lysates.

When gp140 was expressed by HEK 293T cells, uncleaved gp140 was detected in the culture medium indicating that a subset of protein was not cleaved into gp120 in transit to the plasma membrane (Figure 3.8A). The efficiency at which gp140 was cleaved may be influenced by the levels of Env expression and the abundance of the protein inside the ER[84,107,124]. However, a fraction of gp140 was converted to gp120 (Figure 3.10) and we hypothesised that the relative level of gp120 as a percentage of total Env, i.e. cell associated-gp140 and gp140 + gp120 in the culture medium would allow comparison of cleavage efficiency between gp140 mutants and WT[240].

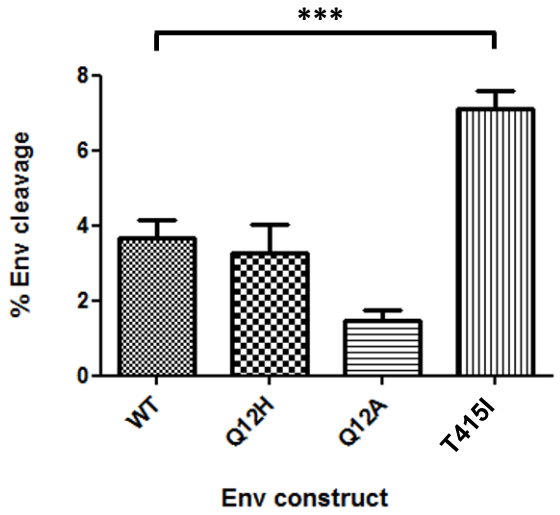

Figure 3.10: Cleavage of WT and mutated gp140 into gp120. To determine the efficiency at which gp140 was cleaved into gp120, the band intensity of gp120 was divided by the sum of the band intensities of gp120 and gp140 i.e. total Env expressed. Mean percentages of gp120 levels of three independent experiments were compared using GraphPad Prism™ 5 one-way ANOVA with the Tukey's post-test. Wild type (WT), Q12H, Q12A and T415I Envs are indicated by patterned bars. Error bars represent standard deviation. An asterisk indicates statistical significance (*** < 0.001).

Contrary to Q12H, the Q12A signal peptide gp140 mutant was cleaved approximately 2.3-fold less efficiently than WT. This suggested that when Alanine was at position 12, gp160 cleavage

was reduced without any concomitant change in expression whereas the presence of a Histidine had no effect. Q12A gp120 levels in the culture medium were lower than WT possibly due to decreased cleavage efficiency. However, as our method of determining cleavage was dependent on gp120 secretion, it is not possible to distinguish whether 1) more gp140 was cleaved and thus more gp120 secreted or 2) gp140 cleavage remained the same as WT but more gp120 was trafficked to the cell surface and directly released into the culture medium. Furthermore, as the relationship between an Alanine at position 12 and gp120 secretion levels and gp140 cleavage did not reach significance, the apparent association might be due to experimental error and needs to be confirmed with cell-associated gp120 expression studies. On the other hand, the gp140 T415I mutant was cleaved 1.5-fold significantly more efficiently than WT ($p = 0.0005$) (Figure 3.10). The significantly increased T415I gp160 expression (3-fold), increased gp140 and gp120 secretion (3-fold and 4-fold, respectively) and gp140 cleavage (1.5-fold) suggests that the loss of a PNG and/or Thr415 could enhance the production of functional trimers, increased Env incorporation and improved viral infectivity.

3.2.6 Envelope N-glycosylation

Previous literature suggested that the loss of, or change, in a single PNG can lead to changes in the N-glycosylation profile of Env[241]. Changes in the distribution of N-glycans found on the fully processed Env have also been found to influence function by affecting binding to target cells and thus infectivity[150,157,172,241–245]. It has also been reported that an increase in Env expression can change the N-glycosylation profile of the protein such that predominantly high mannose N-glycans are added[137]. Upadhyay *et al.* (2018), also found that mutation of His12 altered the N-glycosylation profile. More recently, it was also shown that uncleaved gp140 trimers enter an "aberrant pathways" where they are hyperglycosylated[235].

Cell-associated and secreted gp140 had the same molecular weight (Figure 3.6 vs Figure 3.8A), suggesting that most secreted gp140 resembled unprocessed ER-associated Env. Therefore, to determine whether over-expression, cleavage and secretion of the T415I mutant resulted in gp140 carrying mainly high mannose, gp140 was treated with endoglycosidases after expression in HEK 293T cells[246]. Digestion of Env with Endo-*β*-*N*-acetylglucosaminidase H

(Endo H) and Peptide *N*-glycosidase F (PNGase F) (New England Biolabs®), removes high mannose and all N-glycans, respectively. The shifts in molecular weights of the secreted Env before and after treatment with these endoglycosidases were measured (Table 3.1) in order to determine the percentage of mannosylation (Figure 3.12). Previously, soluble gp120 constructs were utilised to characterise N-glycan structures on monomers[50,247], however more recently gp140 have been used to analyse N-glycosylation within the context of native trimeric Env[235].

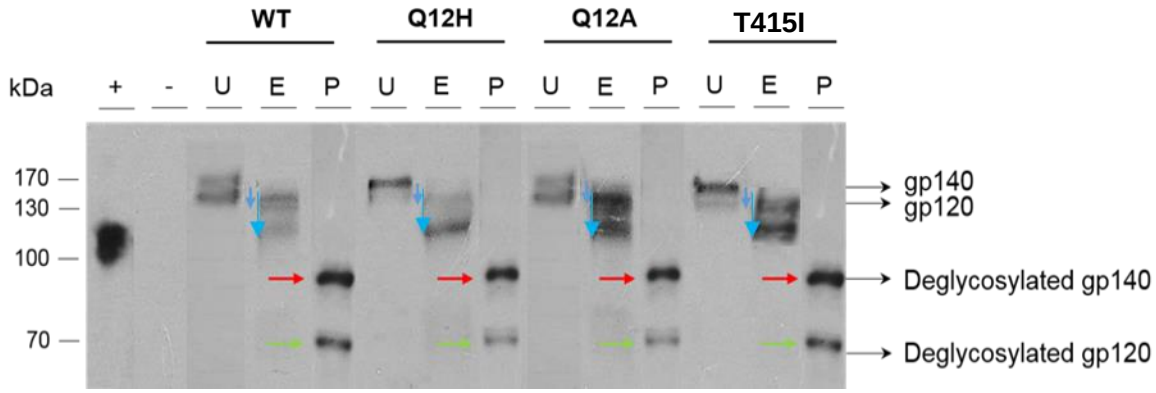

Figure 3.11: Digestion by endoglycosidases to determine the mannosylation of CAP210 E8 gp140 and gp120. Env was expressed in HEK 293T cells and soluble Env [Wild type (WT) gp140, Q12H gp140, Q12A gp140 and T415I gp140] was partially purified from the culture medium using N-*Galanthus nivalis* lectin agarose beads (Vectorlabs®) prior to digestion with Endo H (E), PNGase F (P) or left untreated (U) and analysed by SDS-PAGE and Western blotting. Red and green arrows indicate the position of deglycosylated gp140 and gp120 after PNGase F treatment (determined relative to the molecular weight marker), respectively. Blue arrows indicate the migration distance measured from the centre of the untreated Env band to the centre of the Endo H treated Env bands, either gp120 or uncleaved gp140. The plus sign (+) represents 30 ng of recombinant purified HIV-1 gp120 (IIB-CHO from the AIDS reagent programme). The minus sign (-) represents the negative control for the purification of the Env from the growth medium of HEK 293T cells expressing the empty vector, pcDNA3.1 V5-His Topo® (Invitrogen®). The PageRuler Prestained Protein Ladder (Thermo Scientific) molecular weight ladder is indicated in kilo Daltons (kDa) and used as a standard curve to determine the molecular weight of untreated, demannosylated and deglycosylated gp140 and gp120. A representative of three independent experiments is shown.

After PNGase F treatment of gp140 WT, 63kDa and 79kDa bands were observed that represented deglycosylated gp120 and gp140, respectively[242,246] (indicated by green and red arrows in Figure 3.10). These molecular weights are similar to those predicted for wild type CAP210 E8 gp120 (59kDa) and gp140 (77kDa) (Table 3.1), confirming the accuracy of the methodology. The molecular weight of total N-glycans, i.e. that contributed by high mannose and complex sugars was determined by measuring the difference in the apparent molecular

weight between untreated and PNGase F-treated gp140 and untreated and PNGase F-treated gp120. N-glycosylation contributed 77kDa and 70kDa to the molecular weight of WT gp140 and gp120, respectively. Similarly, to determine the percentage of high mannose residues present[246], the difference in apparent molecular weight between untreated gp140 and gp120 and corresponding Endo-H digested gp140 and gp120 were determined (Table 3.1).

Table 3.1: Predicted and apparent molecular weights of gp140 wild-type and mutants after endoglycosidase treatments

Env construct	Molecular weight (kDa) of Envs*							
	gp140				gp120			
	Predicted	Apparent (Untreated)	Apparent (Endo H)	Apparent (PNGase F)	Predicted	Apparent (Untreated)	Apparent (Endo H)	Apparent (PNGase F)
Wild type	77	153 (SD: 5.3)	138 (SD: 2.9)	87 (SD: 9.6)	58.6	139 (SD: 3.6)	120 (SD: 1.9)	66 (SD: 0.2)
Q12H	77	155 (SD: 2.1)	141 (SD: 0.7)	87 (SD: 13.6)	58.6	137 (SD: 5.7)	118 (SD: 0.2)	66 (SD: 2.6)
Q12A	77	150 (SD: 1.1)	134 (SD: 2.1)	88 (SD: 12.2)	58.5	134 (SD: 2.1)	118 (SD: 0.2)	65 (SD: 1.9)
T415I	77	152 (SD: 4.5)	148 (SD: 3.4)	85 (SD: 11.9)	58.6	136 (SD: 0.2)	119 (SD: 0.5)	67 (SD: 0.9)

*The molecular weight of unglycosylated Env (gp140 and gp120) for CAP210 E8 Env constructs was predicted using online tools (http://web.expasy.org/compute_pi and http://www.hiv.lanl.gov/cgi-bin/GENE_CUTTER/cutter.cgi). The apparent molecular weight of the Env proteins detected by SDS-PAGE corresponding to deglycosylated gp140 and gp120 were calculated using a standard curve from the PageRuler Prestained Protein Ladder (Thermo Scientific) molecular weight ladder. Apparent molecular weights tabulated are an average of three independent experiments.

After EndoH digestion, the apparent MW of WT gp140 and gp120 was reduced by ~16kDa and ~15kDa, respectively. When the percentage high mannose was calculated relative to the total molecular weight of N-glycans, mannosylation of gp140 and gp120 were 17% (SD: 5) and 28% (SD: 10), respectively (Figure 3.12). When the level of mannosylation of the soluble Env constructs containing the Q12H, Q12A or T415I mutations was compared to WT, there was

no significant difference (p = 0.2292). Therefore, changes in the signal peptide at position 12 and deletion of the PNG at position 413-415 did not influence Env-N-glycosylation.

The molecular weights of cell-associated gp140 and gp140 in the culture medium were similar and we proposed that uncleaved protein carried mostly high mannose residues. Previously, our lab showed that cell-associated gp140 was deglycosylated with Endo H, suggesting it only carried high mannose residues (data not shown). Furthermore, another study showed that gp140, if not cleaved, was hyperglycosylated due to non-native processing[235]. We showed that Endo H digestion of secreted gp140 did not generate deglycosylated protein, suggesting that uncleaved gp140 in the culture medium did not only carry high mannose residues but also processed complex N-glycans (Figures 3.11 and 3.12). Considering that gp120 was derived from cleaved gp140, we expected it to have similar levels of mannosylation. However, gp120 was more highly mannosylated than gp140 potentially due to four PNG sites within the 20kDa region of gp41 reported to carry complex N-glycans[248]. Therefore, the 11% difference in mannosylation between secreted gp140 and gp120 is likely due to the presence of additional complex N-glycans and not differences in processing.

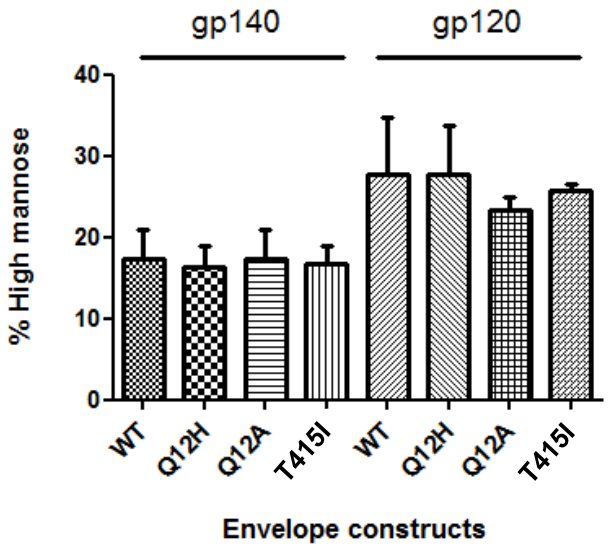

Figure 3.12: Percentage mannosylation of gp140 and gp120. Mannosylation, the level of high mannose residues relative to total N-glycans was calculated based on shifts in molecular weights of gp140 and gp120 after Endo H and PNGase F treatment. Shifts in molecular weight of gp140 or gp120 after Endo H treatment indicated the contribution of high mannose residues to the total molecular weight of the protein whereas shifts in molecular weight after PNGase F treatment indicated the molecular weight of total N-glycans. The ratio of molecular weight of high mannose relative to that of total N-glycans were calculated and expressed as a percentage mannosylation. WT (WT), Q12H, Q12A or T415I Envs (gp140 and gp120) are indicated by each bar. Each bar represents the mean high mannose N-glycan percentages representative of three independent experiments with standard deviations generated using GraphPad Prism™ 5 and analysed by a one-way ANOVA with Tukey's post-test.

3.2.7 Pseudovirion entry efficiency

As deletion of the PNG413 in subtype C CAP210 E8 TF Env increased expression 3.5-fold, we wanted to determine whether it enhanced Env entry efficiency. The signal peptide signature mutants were analysed to determine whether Env entry efficiency was affected, potentially *via* a pathway that did not include enhanced Env expression.

Pseudovirus (PSV) was generated by co-transfecting HEK 293T cells with CAP210 E8 constructs and the packaging vector, pSG3Δenv or pNL4.3-R-E-luc+. PSVs were generated using these two different packaging vectors to control for the potential influence of other viral proteins. PSVs, normalised to equal p24 concentration was added to TZM-bl cells[211]. PSVs generated using pSG3Δenv carrying His12 or Ala12 had significantly lower entry efficiency than WT (Figure 3.13A and B). When pNL4.3-R-E-luc+ was used to generate PSV, entry efficiency was only slightly lower and the difference between WT and mutants was no longer significant (Figure 3.14A and B). The loss of significance is likely due to the poor infectivity of the pNL4.3

backbone-generated PSVs (Figure 3.13C) coupled with high variability between biological repeats (Figure A2). Up to 2-fold difference in WT entry was observed between two independent experiments using this backbone when RLUs at 250 ng/mL p24 were compared (Figure A2.A). We did not add any reagents to improve infectivity as differences between clones might have been masked by enhancing agents. The introduction of either a Histidine or an Alanine at position 12 reduced Env entry efficiency, suggesting that the presence of a Glutamine at this position is important for Env function. The entry efficiency of T415I PSV generated using both pSG3ΔEnv and pNL4.3-R-E-luc+ was significantly higher (2-fold) than WT (Figure 3.13C). The increase in entry efficiency observed for the T415I mutant might be due to the 3.5-fold and 3-fold increase in Env expression and secretion, respectively.

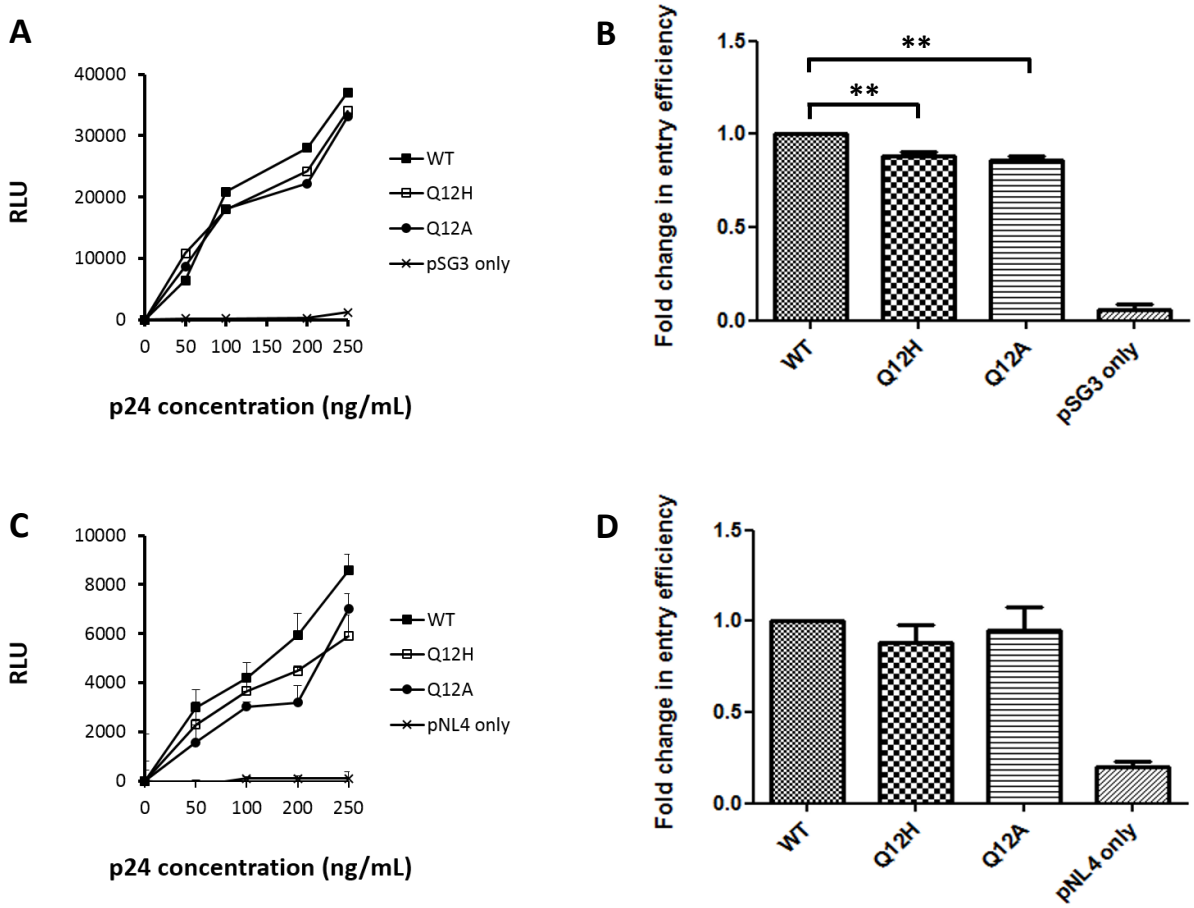

Figure 3.13: Entry efficiency of pseudovirus carrying the signal peptide subtype B Env transmission motif. Pseudovirions (PSV) were generated with A, B) pSG3Δenv or C, D) pNL4.3-R-E-luc+ HIV backbone and signature (Q12H) or non-signature (Q12A) signal peptide and WT (WT) Env constructs. A, C) PSV were diluted to 50, 100, 200 and 250 ng/mL p24 (x-axis) and used to infect TZM-bl luciferase reporter cells in triplicate in three independent experiments. A representative line graph is shown with error bars representing standard deviation of the three triplicate values. B, D) Relative light unit (RLU) values from three independent experiments were averaged at p24 concentration 250ng/ml with bar graphs showing change in entry efficiency relative to WT. Background RLU obtained from mock infection controls were subtracted from all values including WT, signal peptide mutants and the pNL4.3-R-E-luc+ HIV backbone only control. Graphs were drawn using GraphPad Prism™ 5 and the bar graph was analysed using a one-way ANOVA with Tukey's post-test. An asterisk indicates statistical significance (** < 0.01).

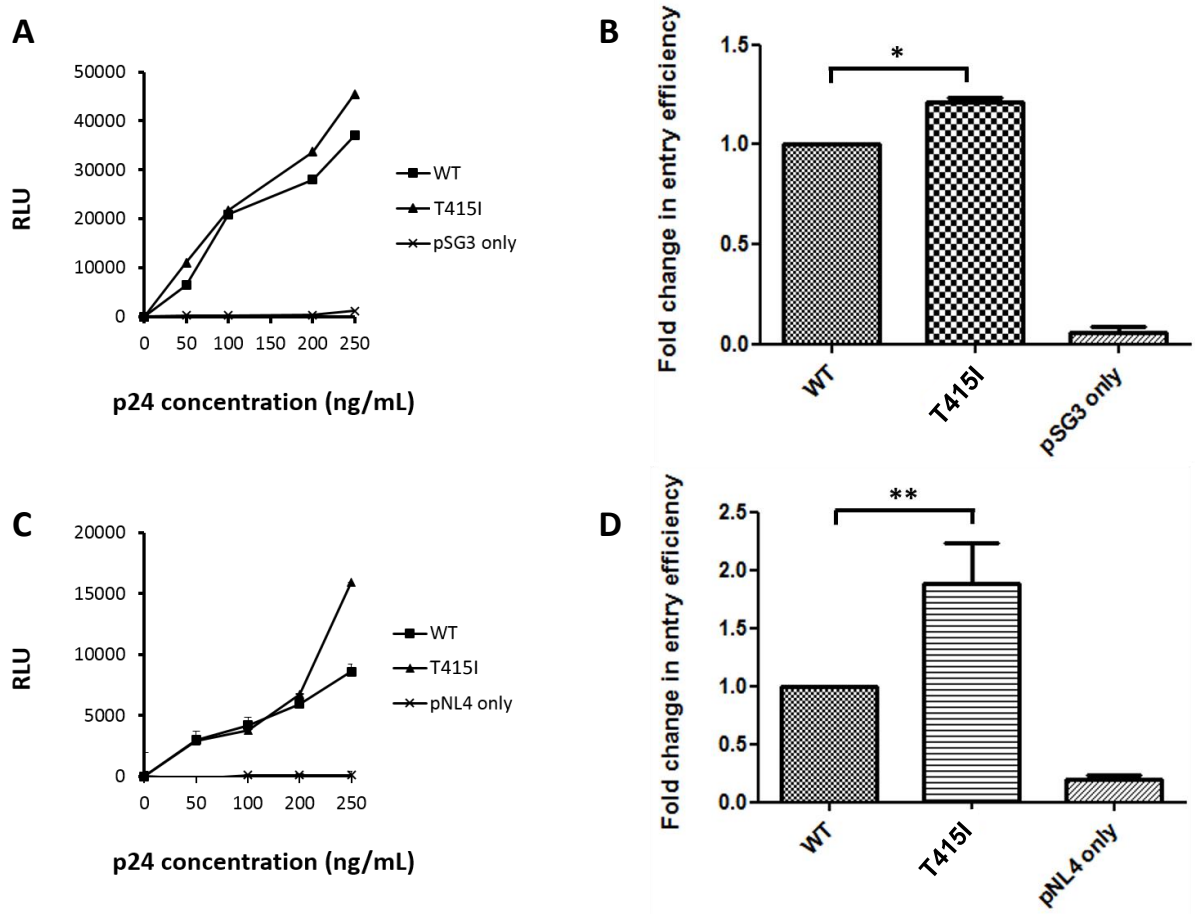

Figure 3.14: Entry efficiency of pseudovirus carrying the PNG415 subtype B Env transmission motif. Pseudovirions (PSV) were generated with A, B) pSG3Δenv or C, D) pNL4.3-R-E-luc+ HIV backbone with wild-type (WT) or Env lacking PNG413 (T415I). A, C) PSV were diluted to 50, 100, 200 and 250 ng/mL p24 (x-axis) and used to infect TZM-bl luciferase reporter cells in triplicate in three independent experiments. A representative line graph is shown with error bars representing standard deviation of the three triplicate values. B, D) Relative light unit (RLU) values from three independent experiments were averaged at p24 concentration 250ng/ml with bar graphs showing change in entry efficiency relative to WT. Background RLU obtained from mock infection controls were subtracted from all values including WT, T417I mutant and the pNL4.3-R-E-luc+ HIV backbone only control. Graphs were drawn using GraphPad Prism™ 5 and the bar graph was analysed using a one-way ANOVA with Tukey's post-test. An asterisk indicates statistical significance (* < 0.5; ** < 0.01).

3.2.8 Entry efficiency of chimeric Infectious Molecular Clones

In order to determine whether the subtype B transmission motifs influenced the entry efficiency of infectious molecular clones (IMCs), we compared the ability of chimeric IMCs expressing WT and mutant Envs (Section 2.8.2) to enter TZM-bl cells (Figure 3.15).

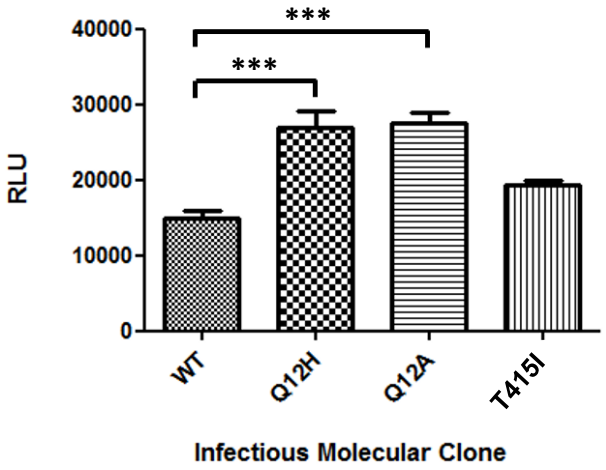

Figure 3.15: Entry efficiency of chimeric IMCs carrying the subtype B Env transmission motifs. Infectious molecular clones (IMCs) were produced *via* yeast homologous recombination to insert WT, Q12H, Q12A and T415I Env into pCMV-PBS-LTR-NL4-3Δgp160/URA. The chimeric IMCs (pCMV-PBS-LTR-NL4-3Δgp160/URA + *env*) were produced in HEK 293T cells and used to infect TZM-bl luciferase reporter cells at 50 ng/mL p24. Background relative light units (RLU) obtained from a mock infection control were subtracted from each experimental value. Bar graph showing the average of three independent experiments with standard deviation were drawn using GraphPad Prism™ 5 and analysed by one-way ANOVA with Tukey's post-test. An asterisk indicates statistical significance (*** < 0.001).

Interestingly, IMC entry efficiency of the signal peptide mutants showed opposing trends when compared to PSVs. Q12H and Q12A showed a significant increase in entry ($p = 0.0007$) compared to WT IMC (Figure 3.15). On the other hand, the entry efficiency of T415I IMCs (Figure 3.15) and PSVs (Figure 3.14) were similar although the difference did not reach significance at the lower IMC titre. In Figure 3.14C difference in entry between T415I and WT only became apparent at the highest p24 concentration, suggesting that more robust differences would have been clearer at both higher PSV and IMC titres. However, we cannot exclude the impact of intrinsic differences between PSVs and IMCs such as incorporation of Env due to differences in molar ratios of Env:Gag whereby PSV Env levels might be much higher than Gag due to co-transfection of separate vectors[249]. Inter assay variation was less pronounced for IMCs (Figure A2.B) compared to PSVs (Figure A2.A). However, since entry could only be tested at a single p24 concentration, i.e. 50 ng/mL, it is possible that this may have masked any variation between biological repeats that may only become apparent at higher titres as with PSVs (Figure A2.A).

Discussion

A genetic bottleneck occurs during HIV transmission when only a single variant, the TF overcomes the mucosal barrier in the genital tract to establish productive clinical infection[5,176,177]. Vaccines that elicit an immune response to TF Env might be able to prevent HIV-1 infection however the high diversity of Env makes it difficult to identify a target conserved in all variants. With the discovery that transmission involved a single variant, it was proposed that TFs might carry conserved motifs that provide them with an advantage during transmission. The identification of a conserved transmission motif might provide the perfect target for immunogen design[250].

In this study, the impact of His12 and PNG413[8] on a subtype C TF variant were investigated. Sequence analysis of over 500 subtype C acute infection *env* sequences from the Los Alamos HIV database revealed that 77% had a Glutamine at position 12 of the signal peptide and PNG413 was present in 95% subtype C TF sequences analysed. Therefore, CAP210 E8 was selected for this study as it represented the majority of subtype C *env* TF sequences. The apparent lack of subtype B transmission motifs in subtype C TFs would suggest that subtypes might have evolved alternative mechanisms for transmission. Furthermore, vaccines that target these motifs might be ineffective against subtype C viruses. Thus, the aim of this study was to characterise the subtype B transmission motifs in the context of a subtype C TF variant to determine whether they have the same effect on Env function as previously described.

Previous studies have suggested that positively charged amino acids found within Env-SP influences the rate at which it is cleaved from the N-terminal end of the protein[106,107]. Therefore, the more positive the SP, the slower Env-SP would be cleaved, leading to accumulation within the ER and reduced protein trafficking through the various compartments of the Golgi apparatus[107]. When the positively charged amino acids of the Env-SP were replaced with neutral ones, the expression and secretion of gp120 increased[107]. This was contrary to the findings of Asmal *et al.* (2011) who reported that when the positively charged Histidine was replaced with the neutral, Glutamine, expression decreased[13]. In our study the signal peptide of the subtype C TF Env had a net charge of +2, less than the charge of the subtype B Env-SP of the AA01 clone[13] (Figure 3.4). By substituting the Glutamine with

a Histidine, Env expression should have increased according to Asmal *et al.* (2011) and Updahyay *et al.* (2018) or decreased according to Li *et al.* (1994). However, when the Glutamine was substituted with both Alanine and Histidine there was no significant difference in Env expression compared to WT (Figure 3.7), suggesting that the positive charge of Histidine did not have an effect on Env-SP cleavage and thus did not alter cell-associated levels of Env.

When the natural signal sequence of gp120 was replaced, Herrera and colleagues (2000) found that by fusing the mature gp120 glycoprotein to the signal peptide of erythropoietin (EPO-SP) or that of the tissue plasminogen activator (t-PA-SP) higher amounts of gp120 were secreted. EPO-SP resulted in much higher secretion of gp120 than the t-PA-SP exchange[135], confirming that signal peptides can influence the secretion of the same protein differently. In our study, gp140 carrying the Q12H and Q12A polymorphism showed no significant difference in secretion compared to the WT (Figure 3.8). However, as gp120 was not detected in cell lysates, we were unable to calculate cleavage using cell–associated gp160 like previous studies[240]. The lack of a transmembrane (TM) region might influence gp140 processing and cleavage as the TM of gp160 has been shown to be important for processing and intracellular trafficking, suggesting that recombinant gp140 could be processed differently as it lacked gp41[251]. Based on N-glycosylation analysis of secreted Env, gp140 was processed similarly to that of gp120, with high mannose making up only, on average, 20% of total N-glycans. This could suggest cleaved and uncleaved gp140 underwent the same N-glycosylation and thus passed through the same processing pathway. The low level of mannosylation, however does suggest that gp140 did not undergo trimerisation as virus associated Env has been shown to carry mostly high mannose residues[137,248].

Previously it has been shown that an increase in expression can change the N-glycosylation profile of Envs from subtypes A, B and C, such that it comprises of predominantly oligomannose[137,153,252] and that mutating the Env-SP enriched for g160 monomers carrying high levels of high mannose N-glycans[131]. The authors suggest that gp160 retention due to slowed Env-SP cleavage may overwhelm the processing machinery in the ER leading to incorrectly processed protein[131]. No difference in the percentage of high mannose N-glycans was observed for the Env signal peptide mutants compared to the WT. This is in contrast to

previous findings when the substitution of positively charged Arginine residues at position 12 and 16 with uncharged Isoleucines resulted in subtype C viral like particle (VLP)-associated Env with predominantly high mannose N-glycans[131]. Even though CAP210 E8 was also subtype C, when the Glutamine residue at position 12 was substituted with either a charged Histidine or a non-polar Alanine, gp140 and gp120 were still poorly mannosylated. This suggested that position 12 in subtype C TF Env-SP is not important for N-glycosylation as it can accommodate positively charged, uncharged and non-polar residues. This was expected as the Env-SP mutants showed no significant difference in expression, or ER retention compared to WT, suggesting that ER processes were not altered by these mutations. However, cleavage of gp140 and secretion of gp120 was reduced for Q12A relative to WT suggesting that processing steps that occur later during trafficking of Env might be influenced by the presence of a small, non-polar amino acid residue at position 12 of the Env-SP. Overall, the presence of either Glutamine or Histidine at position 12 of Env-SP does not seem to affect the processing of Env. This conclusion supports the high frequency of Glutamine at this position in subtype C viruses. Our findings might differ from other studies because the impact of SP changes on Env processing and function might be dependent on the Env clone and/or cells used to express gp160 and gp140.

Despite lack of changes in Env processing, we wanted to confirm whether the signal peptide mutants had similar entry efficiency to WT. There was a significant decrease in entry efficiency of pseudovirus carrying the signal peptide mutations, Q12H and Q12A. However, this difference was not as apparent when PSV was produced with the pNL4.3-R-E-luc+ HIV backbone, suggesting that either interactions with other viral proteins might be important or variation could be due to experimental error. The low level RLU produced by the pNL4.3 backbone-generated PSVs (Figure 3.13C) coupled with the high variability between biological repeats, may have affected infectivity and thus statistical significance could not be reached. We did not add any reagents to improve infectivity as differences between clones might have been masked my enhancing agents. Furthermore, cleavage of Q12A gp140 was reduced with concomitant decrease in PSV entry efficiency however this relationship was not consistent as Q12H gp140 was cleaved similar to WT. Overall, these results suggest that Gln12 is important for Env function however the potential mechanism did not involve changed in Env expression, secretion or mannosylation.

Recently, another study mutated His12 to Glutamine, Arginine (R) and Tyrosine (Y) and reported that expression, incorporation into virus, and IMC infectivity were lowered compared to WT[203], supporting the findings of Asmal *et al.* (2011). The authors state that alteration in the signal peptide altered N-glycosylation and antibody sensitivity[203], supporting the hypothesis that the loss of His12 during chronic infection could facilitate immune escape[13]. When we compared, N-glycosylation between signal peptide mutants and WT, there were no changes in mannosylation. However, this could be due to sensitivity of the methodology as Upadhyay *et al*. (2018) only observed changes in N-glycosylation when using lectin-binding Western blotting and mass spectrometry analysis. Perhaps, using these more stringent methods might allow detection of changes in N-glycosylation of CAP210 E8 Q12H and Q12A.

The decrease in entry efficiency of the Q12H and Q12A mutants was in contrast to previous findings in which infectivity was enhanced for PSVs and IMCs carrying the His12 polymorphism[13,203]. Differences between the outcomes of the studies could be due to differences in methodology as Asmal *et al*. (2011) used Jurkat T-cell lines and we used HEK 293T cells for expression analysis although Upadhyay *et al*. (2018) also used HEK 293T cells. Cell type has previously been suggested to influence cleavage efficiency[240] and N-glycosylation patterns[253] of transiently expressed Env. Furthermore, although both studies used the pSG3Δenv backbone and TZM-bl cells for the PSV entry assay, Asmal *et al.* (2011) co-transfected a molar ratio of 4:1 of pSG3Δenv:Env whereas we used a molar ratio of 3:1. The use of differing ratios of plasmids during transfection could alter the amount of Env trimers incorporated into PSVs[249,254] thereby altering infectivity[89]. PSVs equivalent to 250 ng/mL, 200 ng/mL, 100 ng/mL and 50 ng/mL were used to infect TZM-bl cells and it was noted that apparent differences between mutants and WT were only detected at higher titers, suggesting that when PSV concentration was limiting, differences between Env clones are not detected. Asmal *et al.* (2011) also normalised to p24 concentration but did not report on the concentration and indicated change in entry efficiency relative to dilutions. As the authors detected significant differences across all dilutions, PSV titres were either much higher than ours or differences in trimer incorporation influenced PSV infectivity.

Based on the findings of Asmal *et al*. (2011), we expected an increase in PSV entry efficiency when we introduced His12 into our clone and hypothesised that this could translate into higher replicative fitness which would be advantageous during transmission. However recently it was suggested that low replicative capacity of transmitted variants could favour transmission as individuals progressed slower to disease and were asymptomatic for longer, increasing the probability of transmission[276,277]. The authors also suggested that their hypothesis would explain why subtype C variants that were less fit than other subtypes had high global prevalence. It is interesting to speculate that the absence of His12 plays a role in reducing the replicative fitness of subtype C transmitted variants whereas its presence enhanced the fitness of subtype B variants. However, when His12 was introduced within the signal peptide of our subtype C Env, PSV entry efficiency decreased suggesting that subtype C transmitted founders would have lower fitness and slower disease progression if it carried a His12 which would contribute to its increased probability of transmission. It is likely that subtype C viruses have evolved to lower fitness via evolutionary pathways distinct from that of subtype B so that His12 does not have the same effect on both subtypes. It is also important to consider that we used a subtype B backbone to generate pseudovirus and IMCs, so it is possible that interactions between subtype C Env and subtype B viral proteins influenced the effect of His12 on the viral fitness of our mutants.

As PSV entry efficiency might not reflect viral infectivity, we generated IMCs carrying the signal peptide mutations using the pCMV-PBS-LTR-NL4-3Δgp160/URA vector. We hypothesised that as Env is found on the IMC backbone, we might eliminate the influence of molar ratios and differences in incorporation of trimers. IMCs[214] generated from the signal peptide mutants showed a significant enhancement in entry of TZM-bl cells compared to WT ($p = 0.0007$). This was contrary to the decreased entry efficiency of Q12H and Q12A, irrespective of whether pSG3Δenv or pNL4.3-R-E-luc+ was used to generate the pseudovirus. Therefore, when PSVs and IMCs were generated with the same parent backbone, they still differed in entry efficiency: PSVs (decreased) vs IMCs (increased), suggested that HIV backbones were not influencing IMC entry efficiency. However, we were able to detect differences in IMC entry efficiency at 50 ng/mL p24, a titre where we could not detect variation in PSV entry efficiency. Thus, variation in entry efficiency could be due to intrinsic differences between IMCs and PSVs as suggested by Louder *et al.* (2005) and Provine *et al.*

(2009) where relative amounts of Env incorporated and efficiency of cleavage differed substantially between the PSV and IMC systems in subtype A and B viruses, respectively[249,255]. This hypothesis was supported when comparing our data to that of Upadhyay *et al*. (2018) as mutated IMCs behaved similarly in both studies although these authors could only detect differences at low titre contrary to our findings. There were also inconsistencies between Asmal *et al.* (2011) and Upadhyay *et al*., (2018) as the H12R mutation increased infectivity in the one and decreased IMC infectivity in the latter. The authors suggest that this could be due to electrostatic interactions with the SRP[203]. Furthermore, substitutions of Tyrosine at position 12 of JRFL with Arginine and Glutamine reduced IMC infectivity, suggesting that different amino acids at position 12 have different effects dependent on clone and, in our case, subtype.

In addition to Upadhyay and colleagues' inconsistencies between mutants, Asmal *et al*. (2011) also showed that PSV with a Glycine at position 12 had entry efficiency similar to clones with the His12 motif and the authors suggested that there were alternative determinants of infectivity. However, both reports did not consider the effect of PNG413 on PSV infectivity. At least one clone had both His12 and PNG413 (Appendix, Figure A1) and had higher Env incorporation than a clone where His12 and an adjacent amino acid residue were deleted. Also, the REJO isolate had a PNG at position 412 which could also influence infectivity. It is possible that PNG413 impacted the Env expression, secretion and incorporation experiments however we could not compare sequence data in detail (i.e. presence or absence of His12 and PNG413) with phenotype as Asmal and colleagues did not report on the accession numbers of their sequences. We thus went on to investigate whether PNG413 could modulate Env expression and processing as well as PSV infectivity.

Individual N-glycans have been identified to be important for Env folding[256,257] and can influence the conformation of neighbouring N-glycans[258]. These conformational changes due to N-glycosylation of particular sites can influence interactions between the Env on the incoming virus and binding to CD4 receptors on target cells[259]. The loss of, or change, in a specific PNG has also been shown to have differential effects on infectivity based on the location and proximity to other sites due to changes in binding interactions[241]. For example, the V3 loop, which remains stable during co-receptor binding[260], has been suggested to

determine the specificity of co-receptor binding based on the loss of a single PNG either on the loop or on a neighbouring region[241]. The way in which glycans proximal to CD4 binding sites are processed has also been proposed to influence how antibodies interact with this region[261]. Therefore, changes at a single PNG, can alter Env function.

The loss of PNG413 was suggested to be associated with receptor binding interactions based on the proximity of this PNG in relation to the CD4 and CCR5 co-receptor binding sites determined by protein modelling[8]. In this study, the expression, secretion and PSV entry efficiency of the CAP210 E8 T415I mutant was significantly higher compared to WT. It is possible that the 3-fold increase in secretion and 1.5-fold enhancement of gp140 cleavage was due to the 3.5-fold increase in Env expression. Perhaps the presence of PNG413 enhances interactions with chaperones leading to faster translocation into the ER and processing in the Golgi. Increased expression of Env did not lead to changes in N-glycosylation of CAP210 E8 T415I although as discussed before, this could be due to lack of assay sensitivity. Surprisingly, both WT and mutants were poorly mannosylated which was unlike the data reported by other studies[131,137] which reported Envs comprised of 98% oligomannose glycans. However, the authors highlighted that glycosylation patterns of virion-associated and recombinant gp120 differ which could have influenced our findings as gp140 might not be processed similarly to gp160 to form trimers.

Pseudovirus carrying the T415I mutation showed a 2-fold significant enhancement in entry efficiency compared to the WT (Figure 3.14). This observation was expected as a significant increase in Env expression (Figure 3.7) and secretion (Figure 3.9) was also observed. However, IMCs carrying the T415I mutation showed only a moderate increase (1.3-fold) in entry compared to its WT counterpart. This could be due to variation between PSVs and IMCs previously discussed (and illustrated in Figure A2). As this PNG occurs near the CCR5 co-receptor binding domain[8], it is possible that the enhanced infectivity observed for this mutant may be as a result of enhanced receptor affinity[262]. However, these findings need to be confirmed by measuring the incorporation of gp120 on the surface of pseudovirus as well as the measurement of binding affinities to receptors found on target cells, namely CD4, CCR5 and DC-SIGN. Since the densities of CD4 and CCR5 on TZM-bl cells are not the same as the virus would encounter on target cells, it is also important to compare this entry efficiency

data to entry into a different cell line such as 293-Affinofile cells in which the expression of these receptors on the cell surface can be stimulated by hormones[263,264] and thus the number of CD4 and CCR5 presented on the surface of these cells can be controlled.

Conclusion

Our study indicated that the subtype B transmission motifs were important for subtype C infectivity however the importance of the identity of the amino acid residue at positions 12 and 415 or presence and absence of PNG413 varies between variants and subtypes. Although, we cannot exclude that assay differences contributed to inconsistent findings the influence of Glutamine at position 12 clearly had a positive influence on CAP210 E8 fitness whereas, across two studies, this particular residue termed as “non-signature” negatively influenced viral infectivity. Furthermore, all studies, including ours, did not consider the combined effect of His12 and PNG413 and we should thus have included double mutants in order to investigate the potential effects of the presence both subtype B TF transmission motifs on our subtype C TF clone. Despite these considerations, subtype C fitness is not dependent on a His at position 12 nor a PNG at 413. It is interesting to speculate that PNG413, which seems to have a fitness cost, is found in subtype C TFs because it provides a specific advantage to transmission such as enhancing Env interactions with DC-SIGN or α4β7. During chronic infection, PNG413 is lost as it improves replicative capacity and might also facilitate immune escape.

CHAPTER 4: IMPACT OF SUBTYPE B TRANSMISSION MOTIFS ON SUBTYPE C REPLICATION

4.1 Introduction

Changes in the signal peptide and a potential N-glycan site within gp120 altered the entry efficiency of PSVs and IMCs (Chapter 3). Although, previous studies have consistently used pseudovirus-based assays to measure Env functionality, these assays are performed using cell lines and can only measure a single round of replication. For this reason, a system which supports and measures replication over multiple rounds of infection might better resemble *in vivo* viral replication. We therefore constructed chimeric infectious molecular clones (IMCs) carrying CAP210 E8 WT and mutants in the context of a full HIV-1 proviral backbone[65,249] and measured replication in peripheral blood mononuclear cells (PBMCs)[265,266].

The amplification of the *env* gene is the first step in the generation of chimeric infectious molecular clones (IMCs) using the yeast gap-repair system[213]. This involves the homologous recombination of PCR-amplified *env* DNA with a linearised proviral vector lacking the *env*, pCMV-PBS-LTR-NL4-3Δenv/URA3[214] *via* short sequences of homology in both DNA fragments. Homologous recombination should then result in the generation of a full-length viral genome to produce replication competent recombinant viruses.

Homologous recombination is a natural mechanism that is used by yeast cells to repair their chromosomal DNA using regions of homology between the damaged DNA and the chromosome in order to repair the gap[267]. The yeast gap-repair method offers several advantages over the bacterial cloning in that it is efficient and uses both positive and negative selection markers to ensure the insertion of the *env* gene in the correct orientation[268–270]. This method supersedes standard cloning methods which rely on restriction enzyme digestions, gel purifications, *in vitro* ligations, and screening of the inserts in order to produce infectious full-length proviral DNA vectors[214]. The subtype B backbone pCMV-PBS-LTR-NL4-3Δenv/URA3 used in these homologous recombination experiments was derived from the pREC_nfl_HIV-1Δenv/URA3 plasmid in a series of cloning steps[214]. To generate the pCMV-PBS-LTR-NL4-3Δenv/URA3 backbone, the 5'LTR was replaced with a cytomegalovirus (CMV)

promoter in order to prevent recombination of the entire HIV genome. Thus the expression of the HIV genes is under the control of this promoter. The *env* gene was replaced by the *ura3* gene, which encodes the orotidine-5'phosphate decarboxylase protein (URA3) which is involved in the biosynthesis of uracil. The plasmid also contains the beta-isopropylmalate dehydrogenase gene for the biosynthesis of leucine (LEU), an autonomously replicating sequence (ARSH4), and the yeast centromere sequence (CEN6), which all enables the recombinant plasmid to be cultured in yeast cells[213].

After successful amplification of the *env* gene of interest, the PCR product is co-transformed with the proviral backbone pCMV-PBS-LTR-NL4-3Δenv/URA3 for homologous recombination in yeast cells. Colonies are then grown on negative selection medium. This medium contains 5-fluoro-1,2,3,6-tetrahydro-2,6-dioxo-4-pyrimidine carboxylic acid (FOA) which inhibits the growth of yeast which have not replaced the *ura3* with *env* by homologous recombination. Yeast cells still containing the pCMV-PBS-LTR-NL4-3Δenv/URA3 with *ura3* will produce the URA3 enzyme which uses FOA to produce a toxic by-product 5'fluorouracil. Surviving colonies are then screened by PCR for the presence of the *env* gene as well as direct DNA sequencing for confirmation of full-length chimeric IMC vector for the production of replication competent viruses. Clones with the highest entry efficiency in TZM-bl cells are then selected to compare replicative capacity of these viruses in PBMCs.

4.3 Results

4.3.1 Amplification of the *env* gene from CAP210 E8 constructs

In order to generate chimeric IMCs carrying the *env* sequence of interest, the yeast gap-repair system was used[213,214]. The first step in this assay was the amplification of the CAP210 E8 TF *env* constructs using primers designed to include flanking regions homologous to the IMC backbone to facilitate recombination in yeast cells.

Optimisation of the PCR reaction included the use of various DNA Polymerases in order to get sufficient amplification of the *env* gene from the CAP210 E8 TF *env* constructs. The first of these optimisation steps was a comparison of the efficiency of the KAPA High Fidelity *Taq* Polymerase (KAPA Biosystems®) and *Phusion* (Thermo Scientific).

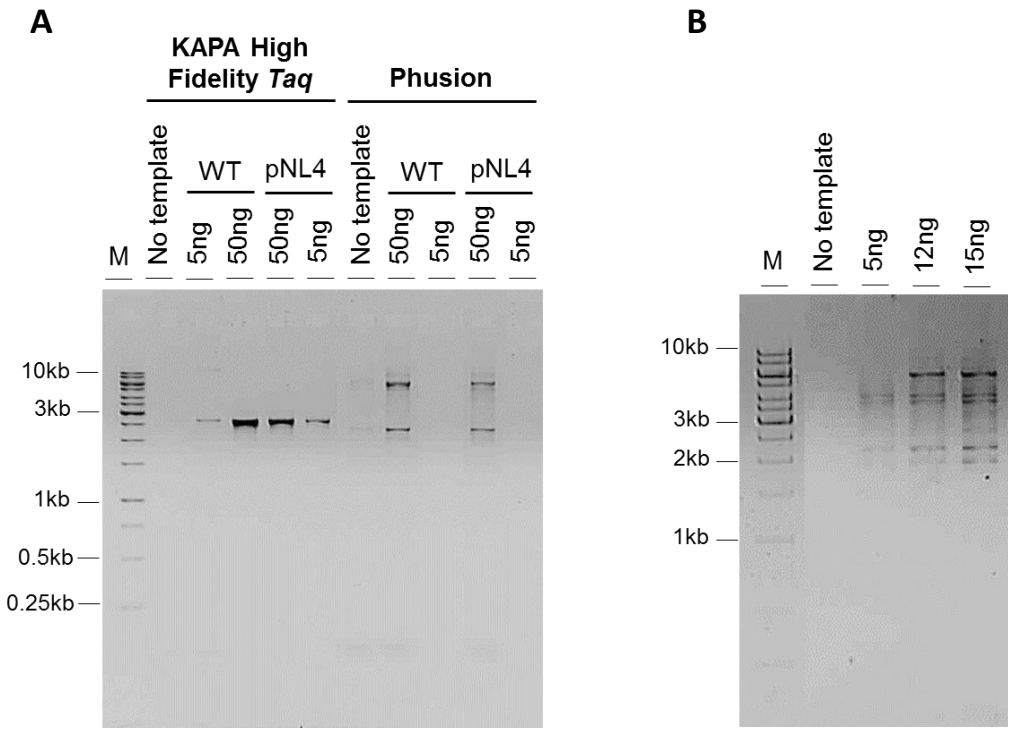

Figure 4.1: PCR amplification of *env* using KAPA High Fidelity *Taq* and *Phusion* polymerases. Complimentary primers flanked by short sequences of homology to the *pol* and 3'LTR, were used to amplify *env* from wildtype (WT) CAP210 E8 TF and the HIV-$1_{pNL4.3}$ full-length proviral backbone (pNL4). (A) The KAPA High Fidelity *Taq* and *Phusion* DNA polymerases were tested for their ability to amplify env from WT and pNL4 at concentrations of 5ng and 50ng. (B) A gradient of WT template concentrations (5, 12 and 15 ng) were tested using the *Phusion* polymerase to reduce non-specific bands during *env* amplification. Five microliters of each of the PCR products were visualised on a 0.8% ethidium bromide stained agarose gel. M indicates the DNA molecular weight ladder.

Wild-type (WT) CAP210 E8 *env* clone and the HIV-$1_{pNL4.3}$ proviral vector (PNL4) were used as positive controls for the PCR amplification of the *env* gene. As we were introducing the native pNL4 *en*v back into the pNL4.3 vector, we planned to compare this to the replication capacity of HIV-$1_{pNL4.3}$ provirus as these two constructs should be similar. Template DNA at 5ng and 50 ng were tested, and KAPA High Fidelity *Taq* polymerase appeared to yield a PCR product at both of these concentrations with the 50ng lane containing a more intense band than that of 5 ng. No product was observed for the 5 ng template controls using *Phusion* polymerase however non-specific bands were observed at 50 ng. In an attempt to minimise these non-specific bands, a gradient of template concentrations were used in order to amplify *env* from the WT. Again, non-specific bands were observed across all concentrations tested and this polymerase was thus excluded from all further optimisation steps (Figure 4.1).

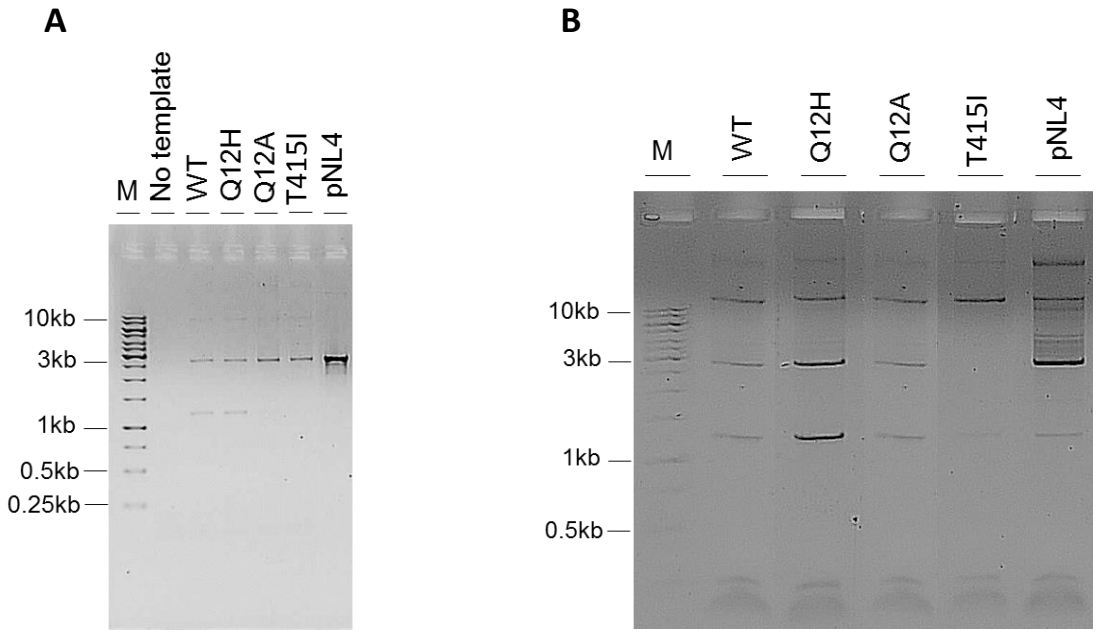

Figure 4.2: PCR amplification of *env* using the KAPA High Fidelity *Taq* polymerase. (A) The KAPA High Fidelity *Taq* polymerase (KAPA Biosystems®) was used to amplify *env* from wildtype (WT), signal peptide mutants (Q12H and Q12A), N-glycan mutant (T415I) and pNL4. (B) Multiple PCR products were pooled to purify the *env* band from the agarose gel. PCR products were visualised on a 0.8% ethidium bromide stained agarose gel for the presence of a ~3 kb band.

Since the KAPA High Fidelity DNA polymerase yielded no non-specific PCR products for the WT and pNL4 samples (at 5ng and 50ng starting template), it was used to amplify the *env* gene from the other CAP210 E8 constructs namely, the signal peptide mutants Q12H and Q12A, and the N-glycan mutant T415I. A single non-specific band was observed at ~1250 bp for the WT and Q12H samples (Figure 4.2A). No non-specific bands were observed for the other *env* PCR products however, when multiple reactions were pooled and run on an agarose gel, non-specific bands were observed consistently across all samples (Figure 4.2B). No PCR product representative of the *env* band was observed for the T417I mutant. Due to these findings we switched to Platinum High Fidelity *Taq* (Invitrogen®).

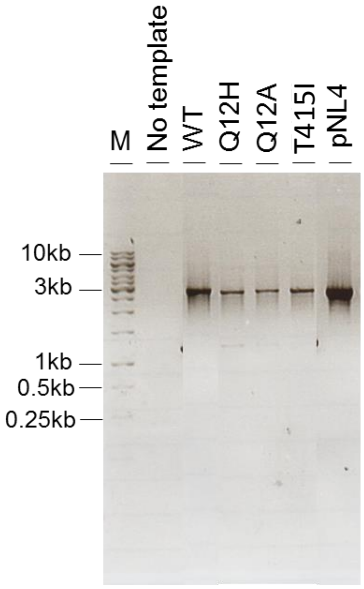

Figure 4.3: PCR amplification of *env* using Platinum High Fidelity *Taq* polymerase. The Platinum High Fidelity *Taq* polymerase (Invitrogen®) was used to amplify *env* from wildtype (WT), signal peptide mutants (Q12H and Q12A), N-glycan mutant (T415I) and pNL4. PCR products were visualised on a 0.8% ethidium bromide stained agarose gel for the presence of a ~3 kb band.

Platinum High Fidelity *Taq* polymerase PCR products yielded a single band for all *env* constructs that coincided with the expected size of *env* (Figure 4.3). However, direct DNA sequencing showed a number of PCR-derived insertions and deletions which would have rendered IMCs generated using these *env* amplicons, non-functional. Our final alternative DNA polymerase was the Q5 High Fidelity kit (New England Biolabs®).

Initial optimisation of the PCR using the Q5 polymerase was at the DNA template concentration level as previous PCR experiments yielded non-specific bands when template concentrations were too high (data not shown). The optimal DNA template concentration was found to be 1 ng, however the primer concentrations needed to be optimised based on the ratio of template DNA to primers available for *env* amplification (Figure 4.4).

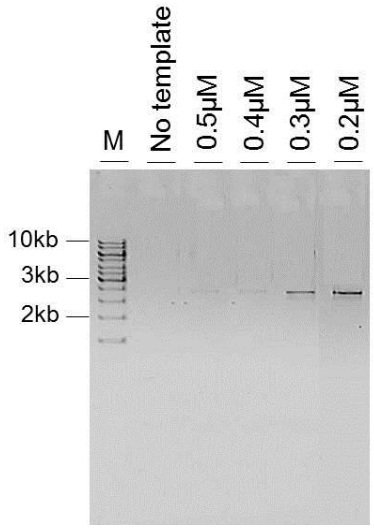

Figure 4.4: Q5 High Fidelity polymerase primer concentration optimisation. A gradient of primer concentrations (0.5, 0.4, 0.3 and 0.2 µM) were tested to amplify wildtype (WT) CAP210 E8 TF *env* using Q5 High Fidelity polymerase (New England Biolabs®). Five microliters of each of the PCR products were visualised on a 0.8% ethidium bromide stained agarose gel. M indicates the DNA molecular weight ladder.

Using WT DNA as the template, a gradient of primer concentrations were tested based on the recommended concentration of 0.5 µM. The optimal primer concentration was 0.2 µM as it yielded the most dense *env* band. This primer concentration was then used for all subsequent *env* PCR amplifications with 1 ng template DNA (Figure 4.5).

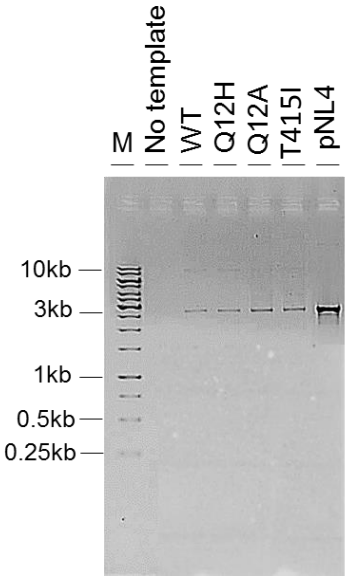

Figure 4.5: PCR amplification of *env* using Q5 High Fidelity polymerase PCR after primer concentration optimisation. Primer concentrations at 0.2 µM were used to amplify *env* from wildtype (WT), signal peptide mutants (Q12H and Q12A), N-glycan mutant (T415I) and pNL4 with the Q5 High Fidelity polymerase (New England Biolabs®). PCR products were visualised on a 0.8% ethidium bromide stained agarose gel.

The Q5 polymerase yielded PCR products for all *env* constructs. A band of ~3 kb was observed, with the *env* amplified from the pNL4 construct yielding the most intense band. Multiple PCR products were pooled and purified as described (Section 2.7.1) in preparation for transformation of yeast cells for homologous recombination with the linearised pCMV-PBS-LTR-NL4-3Δgp160 URA backbone.

4.3.2 Screening yeast for homologous recombination by colony PCR

In order to generate chimeric IMC backbones carrying the env gene of interest, the *env* amplicon was co-transformed with the shuttle vector pCMV-PBS-LTR-NL4-3Δenv/URA3 into competent *S. cerevisiae* cells and grown on negative selection minimal medium deficient of Leucine and in the presence of 5-fluoro-1,2,3,6-tetrahydro-2,6-dioxo-4-pyrimidine carboxylic acid (FOA). Cells which have replaced the *ura3* with *env* in the pCMV-PBS-LTR-NL4-3Δenv/URA3 vector formed colonies in the presence FOA. Colonies were screened by PCR for the presence of the *env* gene (Figure 4.6). Primers complimentary to a region spanning positions 237 to 585 (HXB2 numbering) within the *env* were used in colony PCR to generate a band of 1 kb. Of the twenty-five colonies tested (five per *env* construct), twenty carried the *env* gene insert (Figure 4.6). The sizes of the PCR products were higher than the expected 1 kb due to the presence of the 5' *vpu-rev* flanking region which is not always fully sequenced, therefore the exact length and identity is often unknown.

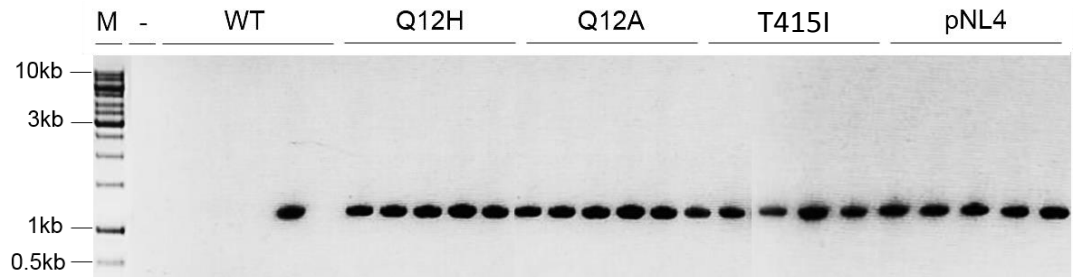

Figure 4.6: Screening yeast colonies for homologous recombination. Yeast cells were transformed with the *env* PCR product and linearised pCMV-PBS-LTR-NL4-3Δgp160/URA3 backbone. Yeast colonies were screened by colony PCR using FW-Env-Y and RV-Env-Y *env*-specific primers. Ten microliters of the reaction was separated on a 0.8% agarose gel. PCR products of a yeast colonies that have replaced the *ura3* gene with the *env* amplicon are indicated by a band at ~1 kb and the water control (contamination control) indicated by a minus sign, respectively. A line indicates the lanes with screened colonies. M indicates the DNA molecular weight marker.

4.3.3 Testing functionality of chimeric viruses using entry efficiencies into TZM-bl cells

Once colonies that carried chimeric IMC DNA were identified, plasmid was prepared and HEK293T cells were co-transfected with the IMC construct and the pCMV_NL4-3LTR-Gag4 (helper plasmid) to generate chimeric virus. Chimeric IMCs were used to infect TZM-bl cells and those that generated luminescence readings 2.5 times higher than background were regarded as carrying functional Env. Only clones with the highest fold change were selected for further analysis (Table 4.1). However, as chimeric viruses were not normalised prior to infection of TZM-bl cells, factors other than IMC fitness could be contributing to increased entry such as transfection efficiency and quality of plasmid DNA. Clones were then sent for full *env* sequencing (Section 2.2.5) to verify the presence of the *env* gene and that no PCR-derived mutations were present (data not shown).

Table 4.1 Infection of TZM-bl cells by chimeric IMCs to determine functionality.

	WT clone 2	Q12H clone 5	Q12A clone 4	T415I clone 5	pNL4 clone 2	HIV-$1_{pNL4.3}$[α]	pCMV-PBS-LTR-NL4Δgp160/URA[#]
Average RLU*	349663	475957	437668	352613	627138	7701954	15461
Fold-change[#]	22.61	30.78	28.30	22.81	40.56	498.17	1.00
Functionality[#]	Yes	Yes	Yes	Yes	Yes	Yes	No

* Functionality of chimeric IMCs was tested by infecting TZM-bl cells with HEK 293T cell culture supernatant containing virus in duplicate and calculating the average relative light units (RLU) determined using a luciferase reporter assay.

[#] To determine functionality, the RLU obtained from duplicate wells was divided by the RLU obtained from the mock infection (pCMV-PBS-NL4Δgp160/URA only) and indicated as fold-change above background. Functional (Yes) or non-functional (No) clones are indicated if the fold background was greater or less than 2.5, respectively.

[α] HIV-$1_{pNL4.3}$ was used as a positive control.

4.3.4 Testing donor PBMCs that can support replication of HIV-$1_{pNL4.3}$

After isolation of PBMCs from blood (Section 2.8.1), three different donors were tested for their ability to support the replication of HIV-$1_{pNL4.3}$ virus. We selected, HIV-$1_{pNL4.3}$ as despite it being X4-tropic as it was previously shown that CD4+ T cells infected with R5 variants produced more progeny than HIV-$1_{pNL4.3}$ when PHA and CD3/CD28 antibodies were used to stimulate PBMCs. We thus rationalised that if HIV-$1_{pNL4.3}$ replication was detectable under those conditions then our R5 chimeric Env IMCs would be too[275]. Virus was harvested 4, 7 and 10 days post-infection. p24 ELISA of day 10 harvests showed that donors 1 and 2 could support replication of HEK 293T cell-produced HIV-$1_{pNL4.3}$ virus as these gave the highest p24 values of 373 ng/mL (SD: 157.59) and 300 ng/mL (SD: 163.07), respectively (Figure 4.7).

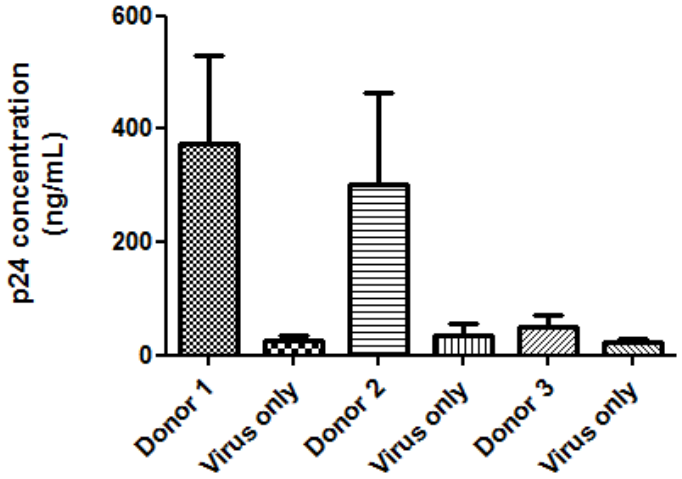

Figure 4.7: Pilot replication assay to test donor PBMCs using HIV-$1_{pNL4.3}$. Virus produced in HEK 293T cells was used to infect three different activated donor peripheral blood mononuclear cells (PBMCs) with HIV-$1_{pNL4.3}$ in triplicate. Virus supernatant was harvested 4, 7, 10 and 14 days post-infection. A bar graph showing the mean p24 ELISA concentrations (ng/mL) 10 days post infection with error bars indicating the standard deviations between triplicate values was generated. Virus only controls were included in order to show the initial viral titres.

4.3.5 Replication of expanded IMCs using donor PBMCs

Since HEK 293T cell-produced IMCs gave poor Tissue Culture Infectious Dose 50 (TCID50) values (Table 4.2), these virus stocks were expanded in the two donor PBMCs identified to support replication (Section 2.8.6; Section 4.2.4) and TCID50 assays[212] performed 7 and 10 days post-infection. . It was found that expansion either improved or decreased TCID50/mL values whereas there was no change for some. The variation depended on donor with most TCID50 values lower after expansion in donor 1 compared to donor 2. There was also no consistent increase in TCID50 values with increase in days post-infection (dpi) (Table 4.2). As expected, the positive control, HIV pNL4.3 provirus had higher TCID50 than the other variants. Due to the low IMC titre, all harvests were pooled, concentrated using ultracentrifugation (Section 2.8.5) and then used to infect activated donor PBMCs at 100 TCID50 units.

Table 4.2: TCID50 values using TZM-bl infectivity assay of virus before and after viral expansion in PBMCs

Chimeric virus ID	TCID50/mL*				
	Before expansion	After expansion			
		PBMC Donor 1		PBMC Donor 2	
		7 dpi	10 dpi	7 dpi	10 dpi
WT	0	0.0905	0	1.77	0
Q12H	2.15	1.77	0.0905	1.77	0
Q12A	3.16	1.77	0.0905	0.0905	0
T415I	0	1.77	0.0905	4.64	1.77
pNL4	3.16	0.0905	0	2.15	1.77
HIV-$1_{pNL4.3}$	3.16	1.77	2.15	3.16	1.77

*TCID50/mL values were calculated using the method described by Reed & Meunch (1983) after infection of TZM-bl cells with serially diluted virus stocks[205].
dpi: days post-infection

We expected pNL4 IMC (pCMV-PBS-LTR-NL4-3Δgp160/URA vector + pNL4 *env*) and the proviral HIV-$1_{pNL4.3}$ to replicate to similar levels in PBMCs. However, this was not the case as HIV-$1_{pNL4.3}$ yielded higher TCID50/mL values in both donors than the pNL4 IMC. Overall, IMCs appeared to have expanded/replicated better in PBMCs isolated from Donor 2 than from that of Donor 1 as TCID50 measurements were higher 7 dpi. After ultracentrifugation, the TCID50 measurements (Table 4.2) improved by around 5-fold (data not shown) however they were still quite low even after pooling and infections were therefore carried out using 100 TCID50 units of virus as opposed to the 300 TCID50 units used during the pilot study (Sections 2.8.7 and 4.2.4). Replication was monitored over 2 weeks and virus supernatants harvested 4, 7, 10 and 14 days post infection and virus titres quantified by p24 ELISA. However day 4 harvests showed no detectable levels of p24 protein (data not shown) and was therefore excluded. The average p24 concentrations (ng/mL) between the two donors for each virus were plotted over time (Figure 4.8A), since the overall trend between the two donors were similar (Figure A3). The mean slope of these curves was also plotted as a bar graph[271] (Figure 4.8B)

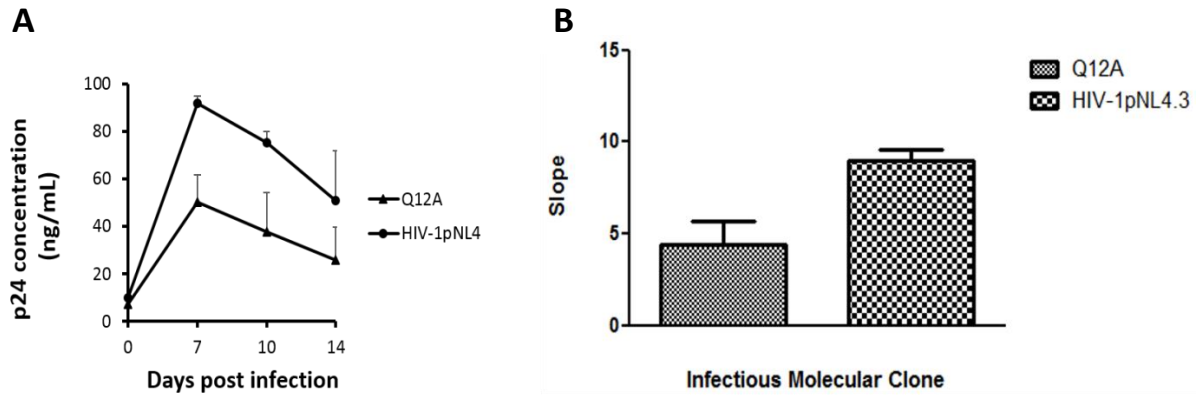

Figure 4.8: Replication of chimeric IMCs in PBMCs. PBMCs from two donors were infected with viruses with equal TCID50 generated from CAP210 E8. Viral replication was quantified by p24 ELISA on days 0, 7, 10, and 14 post infection. (A) The mean p24 concentration from the two independent assays from 2 donors were plotted over time. Error bars represent the standard error of the mean. HIV-1_{pNL4-3} virus was used as positive control. (B) Mean slopes of p24 curves were calculated using the slopes between RLU values at days 0 and 7, 0 and 10, and 0 and 14. Slope values for each virus was used to calculate the mean and standard deviation of two independent donors and used to generate a bar graph[269] in GraphPad Prism™ 5.

The signal peptide mutant, Q12A, and the positive control HIV-$1_{pNL4.3}$ were the only two IMCs to show detectable levels of p24 after ELISA (Figure 4.8) and thus a comparison between IMC entry efficiencies into TZM-bl cells versus replication of IMCs in PBMCs could not be made.

4.4 Discussion

Previous studies suggest that there are inconsistencies in the findings with regards to infectivity based on the viral expression system used even when the same producer cell line is used[249,255]. Discrepancies in results may also be due the comparison of entry efficiency data of pseudovirus which only undergoes a single round of replication versus that of infectious molecular clones (IMCs) which are replication competent. In order to address this issue, chimeric IMCs carrying the *env* of interest were generated using the yeast recombination assay[213]. Chimeric IMCs were then produced in HEK293T cells, expanded in peripheral blood mononuclear cells (PBMCs) and used to infect these cells in order to compare replicative fitness.

In order to measure replicative capacity, *env* was shuttled into the pCMV-PBS-LTR-NL4-3Δgp160/URA vector to generate recombinant viruses using the yeast gap repair system[213]. The first step in the generation of these full-length proviral vectors was the PCR amplification of the *env* gene from CAP210 E8. As part of a series of optimization steps in order to get PCR product, a range of DNA polymerases were tested including varying DNA template concentrations, primer concentrations and PCR run parameters. The Q5 High Fidelity *Taq* Polymerase was chosen as it yielded the desired ~3 kb product (Figure 4.5). Although larger non-specific bands were also present which could not be removed even after several troubleshooting runs, direct sequencing of the final IMC constructs showed that there were no additional PCR-derived mutations present (data not shown).

Clones able to enter TZM-bl cells (Table 4.1) were taken forward to be produced at a larger scale in HEK 293T cells. When the TCID50 was determined for chimeric IMCs produced in HEK 293T cells, they were found to be poor and thus virus could not be used to infect donor PBMCs. This may be due to the efficiency of transfection due to the length of the vector[272] which ultimately affects the amount of virus particles the cells can produce. For this reason, we then expanded these HEK 293T cell-produced virus into in two different donor PBMCs (identified to support replication) to improve TCID50 measurements (Table 4.2). Even after viral expansion in PBMCs, TCID50 values were still low and therefore virus was concentrated by ultracentrifugation and the used to infect activated donor PBMCs.

When p24 ELISAs were performed on virus supernatants harvested over 14 days, detectable levels of p24 protein were only observed for the signal peptide mutant, Q12A, and the positive control HIV-$1_{pNL4.3}$ (Figure 4.8). This may be due to the sensitivity of the ELISA although we are routinely able to detect virus at concentrations as low as 500 pg/mL. Furthermore, replication of HIV-$1_{pNL4.3}$ provirus was not detected when it usually replicates to high titres (data not shown). Since the amount of TCID50 units used for the infection was reduced from 300 to 100, it may be possible that the multiplicity of infection (MOI) was too low for productive, detectable infection. As possible T lymphocyte quiescence was a concern and HIV preferentially targets activated CD4+ T lymphocytes[273], IL-2 was used to stimulate cells[274] and was replenished after each harvest[273]. However, the decrease in replication after 7 days suggests that activation might not have been optimal and that PBMCs were dying over time.

Cytopathic events have been shown to limit replication in PBMCs although HIV-$1_{pNL4.3}$ has been shown to be less pathogenic than other variants[275]. However as the Env might have impacted the cytopathic nature of HIV-$1_{pNL4.3}$ we should have tested cell viability during and after replication of chimeric IMCs to determine whether the PBMCs remained viable.

Since replication was not observed for the Q12H and T415I mutants, we could not compare the entry of chimeric IMCs into TZM-bl cells to the replicative capacity in PBMCs. Other clones of these chimeric IMCs which were identified during screening (Table 4.1) should be tested for their ability to replicate so that the effect of these motifs can be characterised based on infectivity in the context of replication competent virus.

CHAPTER 5: SUMMARY AND CONCLUSIONS

In this study, the robust transmission motifs identified in subtype B TF Envs[8] were investigated for their potential role in the expression, processing and infectivity of a subtype C TF Env from CAP210 E8. CAP210 E8 represented the consensus of subtype C TFs as it lacked the transmission motifs: it carried a Glutamine at position 12 (Gln12) and a PNG at position 413-415. Site-directed mutagenesis successfully introduced the transmission motifs into CAP210 E8 to generate signal peptide mutants, Q12H and Q12A as well as the PNG mutant T415I. Full-length (gp160) and soluble (gp140) forms of these Env constructs were generated and expressed in HEK 293T cells for expression, secretion, cleavage efficiency and N-glycosylation analyses as well as the production of PSV to test infectivity of TZM-bl cells and chimeric IMCs for replication in donor PBMCs.

Table 5.1 Overview of the phenotypes of signal peptide and N-glycan mutants relative to wild-type (WT)

Mutant ID	Expression	Secretion	Cleavage	Mannosylation	PSV entry efficiency		IMC entry efficiency
					pSG3	pNL4.3	pNL4.3
Q12H	0	0	0	0	<*	0	>*
Q12A	0	<	<	0	<*	0	>*
T415I	>*	>*	>*	0	>*	>*	>

* Difference was significant ($p < 0.05$)
0: no change; >: increase; <: decrease relative to WT

Referring to Table 5.1, it is evident that the loss of Threonine at position 415 and/or the loss of PNG413 enhanced CAP210 E8 infectivity most likely by altering expression. On the other hand, the analysis of the findings of the signal peptide mutants are more complicated as there were inconsistencies between mutants and between assays. Q12A had lower secretion and apparent cleavage than WT which correlated with the observed decrease in PSV entry efficiency. As secretion and apparent cleavage of Q12H did not change, this would suggest that it is the type of amino acid at position 12 that is important and not Histidine or charge. Gln12 is the preferred polymorphism for subtype C TF, and deletion of this amino acid results in loss of fitness. However, Q12H and Q12A IMCs could enter TZM-bl cells better than WT. This could suggest that in multiple rounds of infection, the presence of His12 is beneficial.

Furthermore, contrary to previous findings[13,203], His12 did not enhance Env expression, processing and PSV infectivity. The increase in IMC entry efficiency was the only consistent finding between this study and earlier reports[13,203]. As Asmal and colleagues tested PSVs generated by pNL4.3 and Upadhyay *et al.* (2018) compared IMCs of REJO and JRFL isolates, it is unlikely that the HIV backbone, (pSG3 vs pNL4.3) and the assay (PSV vs IMC) was directly responsible for the differences between our study and theirs. We would like to suggest that the reason for the inconsistencies within our study is multifactorial that include protein-protein interactions, cell type factors and assay conditions whereas differences between our study and earlier reports is due to subtype differences. It was very unfortunate that we were unable to measure and compare the replication of the chimeric IMCs in donor PBMCs as this might have shed light on the potential impact that different amino acids at position 12 might have on HIV-1 replication, a more physiologically relevant model. Therefore, future studies will include: optimising the replication assay, testing clones other than CAP210 E8, comparing incorporation of gp120 in PSVs and IMCs and determining N-glycosylation of PSV- and IMC-associated Env using mass spectrometry.

In conclusion, these findings suggest that position 12 of the Env signal peptide is important to subtype C and B HIV infectivity although the mechanism needs to be further investigated. However, the presence of Histidine at this site decreased CAP210 E8 fitness and that subtype C variants have evolved and selected for a Glutamine at this position most likely because it benefits viral survival. Similarly, PNG413 (or Thr415) is important for viral infectivity but as it is present in subtype C TFs and absent in subtype B variants, it also might have evolved to impact HIV-1 fitness in a subtype specific manner. This has serious implications for vaccine design as any selection pressure that drives the loss of His and/or PNG413 might enhance the viral fitness of diverse variants, resulting in increased disease progression.

REFERENCES

1. Barre-Sinoussi, F. *et al.* Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Sci.* **220,** 868–871 (1983).
2. Gilbert, M. T. P. *et al.* The emergence of HIV/AIDS in the Americas and beyond. *Proc. Natl. Acad. Sci. U. S. A.* **104,** 18566–70 (2007).
3. UNAIDS. *UNAIDS Data 2017*. 8 (2017).
4. Delwart, E. L., Sheppard, H. W., Walker, B. D., Goudsmit, J. & Mullins, J. I. Human immunodeficiency virus type 1 evolution in vivo tracked by DNA heteroduplex mobility assays. *J. Virol.* **68,** 6672–83 (1994).
5. Abrahams, M.-R. *et al.* Quantitating the multiplicity of infection with human immunodeficiency virus type 1 subtype C reveals a non-poisson distribution of transmitted variants. *J. Virol.* **83,** 3556–3567 (2009).
6. Chohan, B. *et al.* 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. *J. Virol.* **79,** 6528–6531 (2005).
7. Keele, B. F. *et al.* Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 7552–7557 (2008).
8. Gnanakaran, S. *et al.* Recurrent signature patterns in hiv-1 b clade envelope glycoproteins associated with either early or chronic infections. *PLoS Pathog.* **7,** e1002209 (2011).
9. Chen, B. *et al.* Structure of an unliganded simian immunodeficiency virus gp120 core. (2005).
10. Wilen, C. B., Tilton, J. C. & Doms, R. W. in *Advances in Experimental Medicine and Biology* 223–242 (2012). doi:10.1007/978-1-4614-0980-9
11. Wilen, C. B., Tilton, J. C. & Doms, R. W. HIV: cell binding and entry. *Cold Spring Harb. Perspect. Med.* **2,** (2012).
12. Arthos, J. *et al.* HIV-1 envelope protein binds to and signals through integrin

alpha4beta7, the gut mucosal homing receptor for peripheral T cells. *Nat. Immunol.* **9,** 301–9 (2008).

13. Asmal, M. *et al.* A signature in HIV-1 envelope leader peptide associated with transition from acute to chronic infection impacts envelope processing and infectivity. *PLoS One* **6,** (2011).
14. Haynes, B. F. & Shattock, R. J. Critical issues in mucosal immunity for HIV-1 vaccine development. **5,** 3–9 (2010).
15. Derdeyn, C. a *et al.* Envelope-constrained neutralization-sensitive HIV-1 after heterosexual transmission. *Science* **303,** 2019–2022 (2004).
16. Hemelaar, J., Gouws, E., Ghys, P. D. & Osmanov, S. Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. *AIDS* **20,** W13-23 (2006).
17. Hemelaar, J., Gouws, E., Ghys, P. D. & Osmanov, S. Global trends in molecular epidemiology of HIV-1 during 2000– 2007. *AIDS* **25,** 679–689 (2011).
18. Perelson, A. S., Neumann, A. U., Markowitz, M., Leonard, J. M. & Ho, D. D. HIV-1 Dynamics in Vivo : Virion Clearance Rate , Infected Cell Life-Span , and Viral Generation Time. 1–6 (2013).
19. Varmus, H. Regulation of HIV and HTLV gene expression. *Genes Dev.* **2,** 1055–1062 (1988).
20. Frankel, A. D. & Young, J. A. T. HIV-1: Fifteen Proteins and an RNA. *Annu. Rev. Biochem.* **67,** 1–25 (1998).
21. Fanales-Belasio, E. *HIV virology and pathogenetic mechanisms of infection : a brief overview*. (2010). doi:10.4415/ANN
22. Johnson, W. E. & Desrosiers, R. C. V IRAL P ERSISTENCE : HIV ' s Strategies of Immune System Evasion. *Annu. Rev. Med.* **53,** 499–518 (2002).
23. Wu, Y. HIV-1 gene expression: Lessons from provirus and non-integrated DNA. *Retrovirology* **1,** 1–10 (2004).
24. Mervis, R. J. *et al.* The gag gene products of human immunodeficiency virus type 1: alignment within the gag open reading frame, identification of post-transcriptionnal modifications and evidence for alternative gag precursors. *J. Virol.* **62,** 3993–4002 (1988).
25. Veronese, F. M., Copeland, T. D., Oroszlan, S., Gallo, R. C. & Sarngaharan, M. G. Biochemical and immunological analysis of human immunodeficiency virus gag gene

products p17 and p24. *J. Virol.* **62,** 795–801 (1988).

26. Ono, a, Orenstein, J. M. & Freed, E. O. Role of the Gag matrix domain in targeting human immunodeficiency virus type 1 assembly. *J. Virol.* **74,** 2855–2866 (2000).
27. Gottlinger, H. G., Sodroski, J. G. & Haseltine, W. A. Role of capsid precursor processing and myristoylation in morphogenesis and infectivity of human immunodeficiency virus type 1. *Proc. Natl. Acad. Sci.* **86,** 5781–5785 (1989).
28. Kondo, E. & Göttlinger, H. G. A conserved LXXLF sequence is the major determinant in p6gag required for the incorporation of human immunodeficiency virus type 1 Vpr. *J. Virol.* **70,** 159–164 (1996).
29. Poon, D. T., Wu, J. & Aldovini, A. Charged amino acid residues of human immunodeficiency virus type 1 nucleocapsid p7 protein involved in RNA packaging and infectivity. *J Virol* **70,** 6607–6616 (1996).
30. Levin, J. G., Guo, J., Rouzina, I. & Musier-Forsyth, K. Nucleic Acid Chaperone Activity of HIV-1 Nucleocapsid Protein: Critical Role in Reverse Transcription and Molecular Mechanism. *Prog. Nucleic Acid Res. Mol. Biol.* **80,** 217–286 (2005).
31. Robinson, H. L. Vaccines: New Hope for an AIDS Vaccine. *Nat. Rev. Immunol.* **2,** 239–250 (2002).
32. Rich, R. L. & Myszka, D. G. Spying on HIV with SPR. **1,** 124–133 (2003).
33. Staropoli, I., Chanel, C., Girard, M. & Altmeyer, R. Processing , Stability , and Receptor Binding Properties of Oligomeric Envelope Glycoprotein from a Primary HIV-1 Isolate *. *J. Biol. Chem.* **275,** 35137–35145 (2000).
34. Didigu, C. A. & Doms, R. W. Novel approaches to inhibit HIV entry. *Viruses* **4,** 309–324 (2012).
35. Kwong, P., Wyatt, R., Robinson, J. & Sweet, R. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* **393,** (1998).
36. Berger, E. A., Murphy, P. M. & Farber, J. M. Chemokine receptors as HIV-1 coreceptors: roles in viral entry, tropism, and disease. *Annu. Rev. Immunol.* **17,** 657–700 (1999).
37. Nora, T. *et al.* Functional diversity of HIV-1 envelope proteins expressed by contemporaneous plasma viruses. *Retrovirology* **5,** 1–16 (2008).
38. Yamamoto, T., Tsunetsugu-yokota, Y., Mitsuki, Y. & Mizukoshi, F. Selective

Transmission of R5 HIV-1 over X4 HIV-1 at the Dendritic Cell – T Cell Infectious Synapse Is Determined by the T Cell Activation State. (2009). doi:10.1371/journal.ppat.1000279

39. Haynes, W., Iii, G. R. & Network, H. I. V. Homozygous and Heterozygous CCR5 - [DELTA] 32 Genotypes Are Associated With Resistance to HIV Infection. *J. Acquir. Immune Defic. Syndr.* **27,** 472–481 (2001).
40. Kwon, D. S., Gregorio, G., Bitton, N., Hendrickson, W. a & Littman, D. R. DC-SIGN-mediated internalization of HIV is required for trans-enhancement of T cell infection. *Immunity* **16,** 135–44 (2002).
41. Geijtenbeek, T. B. H. *et al.* DC-SIGN , a Dendritic Cell – Specific HIV-1-Binding Protein that Enhances trans -Infection of T Cells. **100,** 587–597 (2000).
42. Turville, S. G. *et al.* HIV gp120 receptors on human dendritic cells. *Blood* **98,** 2482–2488 (2001).
43. Geijtenbeek, T. B. H. *et al.* DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells. *Cell* **100,** 587–597 (2000).
44. Coleman, C. M. & Wu, L. HIV interactions with monocytes and dendritic cells: Viral latency and reservoirs. *Retrovirology* **6,** 1–12 (2009).
45. Starcich, B. R. *et al.* Identification and characterization of conserved and variable regions in the envelope gene of HTLV-III/LAV, the retrovirus of AIDS. *Cell* **45,** 637–48 (1986).
46. Wyatt, R. *et al.* The antigenic structure of the HIV gp120 envelope glycoprotein. *Nature* **393,** 705–711 (1998).
47. Checkley, M. A., Luttge, B. G. & Freed, E. O. HIV-1 Envelope Glycoprotein Biosynthesis , Trafficking , and Incorporation. *J. Mol. Biol.* **410,** 582–608 (2011).
48. Pancera, M. *et al.* Structure of HIV-1 gp120 with gp41-interactive region reveals layered envelope architecture and basis of conformational mobility. *PNAS* **107,** 1166–1171 (2010).
49. Pancera, M. *et al.* Structure and immune recognition of trimeric pre-fusion HIV-1 Env. *Nature* **514,** 455–461 (2014).
50. Leonard, C., Spellman, M. & Riddle, L. Assignment of intrachain disulfide bonds and characterization of potential glycosylation sites of the type 1 recombinant human immunodeficiency virus envelope glycoprotein (gp120. *J. Biol. …* **265,** 10373–10382

(1990).

51. Bhattacharya, J., Peters, P. J. & Clapham, P. R. Human immunodeficiency virus type 1 envelope glycoproteins that lack cytoplasmic domain cysteines: impact on association with membrane lipid rafts and incorporation onto budding virus particles. *J. Virol.* **78,** 5500–6 (2004).
52. Do, Y. *et al.* Unliganded HIV-1 gp120 core structures assume the CD4-bound conformation with regulation by quaternary interactions and variable loops. 5663–5668 (2012). doi:10.1073/pnas.1112391109
53. Mclellan, J. S. & Pancera, M. Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9. *Nature* **480,** 336–343 (2012).
54. Pancera, M. *et al.* Structure of HIV-1 gp120 with gp41-interactive region reveals layered envelope architecture and basis of conformational mobility. *PNAS* **107,** 1–6 (2010).
55. Pejchal, R. *et al.* A Potent and Broad Neutralizing Antibody Recognizes and Penetrates the HIV Glycan Shield. *Science (80-.).* **334,** 1097–1103 (2011).
56. Zhou, T. *et al.* Structural definition of a conserved neutralization epitope on HIV-1 gp120. *Nature* **445,** 732–737 (2007).
57. Diskin, R. *et al.* Increasing thePotency and Breadth of an HIV Antibody by Using Structure-Based Rational Design. *Science (80-.).* **334,** 1289–1294 (2011).
58. Lasky, L. A. *et al.* Delineation of a Region of the Human lmmunodeficiency Virus Type 1 gp120 Glycoprotein Critical for Interaction with the CD4 Receptor. *Cell* **50,** 975–985 (1987).
59. Kowalski, M. Functional regions of the Envelope Glycoprotein of Human Immunodeficiency Virus type 1. *Science (80-.).* **237,** 1351–1355 (1987).
60. Olshevsky, U. D. Y. *et al.* Identification of Individual Human Immunodeficiency Virus Type gpl20 Amino Acids Important for CD4 Receptor Binding. *J. Virol.* **64,** 5701–5707 (1990).
61. Wyatt, R. *et al.* Involvement of the V1/V2 variable loop structure in the exposure of human immunodeficiency virus type 1 gp120 epitopes induced by receptor binding. *J. Virol.* **69,** 5723–33 (1995).
62. Guttman, M., Kahn, M., Garcia, N. K., Hu, S.-L. & Lee, K. K. Solution structure, conformational dynamics, and CD4-induced activation in full-length, glycosylated,

monomeric HIV gp120. *J. Virol.* **86,** 8750–64 (2012).

63. Chan, D. C., Fass, D., Berger, J. M. & Kim, P. S. Core structure of gp41 from the HIV envelope glycoprotein. *Cell* **89,** 263–73 (1997).
64. Weissenhorn, W., Dessen, A., Harrison, S. C., Skehel, J. J. & Wiley, D. C. Atomic structure of the ectodomain from HIV-1 gp41. *Nature* **387,** 426–430 (1997).
65. Etemad, B. *et al.* Human Immunodeficiency Virus Type 1 V1-to-V5 Envelope Variants from the Chronic Phase of Infection Use CCR5 and Fuse More Efficiently than Those from Early after Infection. *J. Virol.* **83,** 9694–9708 (2009).
66. Troyer, R. M. *et al.* Changes in Human Immunodeficiency Virus Type 1 Fitness and Genetic Diversity during Disease Progression †. **79,** 9006–9018 (2005).
67. Parrish, N. F. *et al.* 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 Pathog.* **8,** e1002686 (2012).
68. Freed, E. O., Myers, D. J. & Rissert, R. Identification of the Principal Neutralizing Determinant of Human Immunodeficiency Virus Type 1 as a Fusion Domain. *J. Virol.* **65,** 190–194 (1991).
69. Sok, D. *et al.* Article A Prominent Site of Antibody Vulnerability on HIV Envelope Incorporates a Motif Associated with CCR5 Binding and Its Camouflaging Glycans Article A Prominent Site of Antibody Vulnerability on HIV Envelope Incorporates a Motif Associated with CCR5 . *Immunity* **45,** 31–45 (2016).
70. Napier, K. B., Wang, Z., Peiper, S. C. & Trent, J. O. CCR5 interactions with the variable 3 loop of gp120. *J. Mol. Model.* **13,** 29–41 (2007).
71. Shioda, T. *et al.* A Naturally Occurring Single Basic Amino Acid Substitution in the V3 Region of the Human Immunodeficiency Virus Type 1 Env Protein Alters the Cellular Host Range and Antigenic Structure of the Virus. *J. Virol.* **68,** 7689–7696 (1994).
72. Fouchier, R. O. N. A. M. *et al.* Phenotype-Associated Sequence Variation in the Third Variable Domain of the Human Immunodeficiency Virus Type 1 gpl20 Molecule. *J. Virol.* **66,** 3183–3187 (1992).
73. Pollakis, G. *et al.* Phenotypic and Genotypic Comparisons of CCR5- and CXCR4-Tropic Human Immunodeficiency Virus Type 1 Biological Clones Isolated from Subtype C-Infected Individuals. *J. Virol.* **78,** 2841–2852 (2004).
74. Salazar-Gonzalez, J. F. *et al.* Genetic identity, biological phenotype, and evolutionary

pathways of transmitted/founder viruses in acute and early HIV-1 infection. *J. Exp. Med.* **206,** 1273–1289 (2009).

75. Ping, L.-H. *et al.* Comparison of viral Env proteins from acute and chronic infections with subtype C human immunodeficiency virus type 1 identifies differences in glycosylation and CCR5 utilization and suggests a new strategy for immunogen design. *J. Virol.* **87,** 7218–33 (2013).
76. Abrahams, M. *et al.* Rapid, complex adaptation of transmitted HIV-1 full-length genomes in subtype C-infected individuals with differing disease progression. *AIDS* **27,** 507–18 (2013).
77. Behrens, A. *et al.* Composition and Antigenic Effects of Individual Glycan Sites of a Trimeric HIV-1 Envelope Article Composition and Antigenic Effects of Individual Glycan Sites of a Trimeric HIV-1 Envelope Glycoprotein. *CellReports* **14,** 2695–2706 (2016).
78. Pancino, G., Ellerbrok, H., Sitbon, M. & Sonigo, P. in *Simian Immunodeficiency Virus SE - 5* (eds. Letvin, N. & Desrosiers, R.) **188,** 77–105 (Springer Berlin Heidelberg, 1994).
79. Allan, J. S. *et al.* Major glycoprotein antigens that induce antibodies in AIDS patients are encoded by HTLV-III. *Science* **228,** 1091–4 (1985).
80. Feizi, T. & Larkin, M. Inhibitors of the biosynthesis and processing og N-linked oligosaccharides. *Glycobiology* **1,** 236–236 (1991).
81. Leonard, K., Spellman, W., Harris, R. J. & Thomas, N. 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 *. (1990).
82. Bernstein, H. B., Tucker, S. P., Hunter, E., Schutzbach, J. S. & Compans, R. W. Human immunodeficiency virus type 1 envelope glycoprotein is modified by O-linked oligosaccharides. *J. Virol.* **68,** 463–8 (1994).
83. Silberstein, S. & Gilmore, R. Biochemistry, molecular biology, and genetics of the oligosaccharyltransferase. *FASEB J.* **10,** 849–858 (1996).
84. Earl, P. L., Moss, B. & Doms, R. W. Folding, interaction with GRP78-BiP, assembly, and transport of the human immunodeficiency virus type 1 envelope protein. *J. Virol.* **65,** 2047–55 (1991).
85. Stein, S. & Engleman, G. Intracellular Processing of the gp160 HIV- 1 Envelope

Precursor OF THE GOLGI COMPLEX *. *J. Biol. Chem.* **265,** 2640–2649 (1990).

86. Checkley, M. A., Luttge, B. G. & Freed, E. O. HIV-1 Envelope Glycoprotein Biosynthesis, Trafficking, and Incorporation. *J. Mol. Biol.* **410,** 582–608 (2011).
87. Murakami, T. & Freed, E. O. Genetic Evidence for an Interaction between Human Immunodeficiency Virus Type 1 Matrix and alpha -Helix 2 of the gp41 Cytoplasmic Tail. *J. Virol.* **74,** 3548–3554 (2000).
88. Zhu, P. *et al.* Electron tomography analysis of envelope glycoprotein trimers on HIV and simian immunodeficiency virus virions. *Proc. Natl. Acad. Sci.* **100,** 15812–15817 (2003).
89. Bachrach, E., Dreja, H., Lin, Y., Corbeau, P. & Piechaczyk, M. Effects of virion surface gp120 density on infection by HIV-1 and viral production by infected cells. *Virology* **332,** 418–429 (2002).
90. Nickel, W. & Seedorf, M. Unconventional mechanisms of protein transport to the cell surface of eukaryotic cells. *Annu. Rev. Cell Dev. Biol.* **24,** 287–308 (2008).
91. Caporale, M. *et al.* The Signal Peptide of a Simple Retrovirus Envelope Functions as a Posttranscriptional Regulator of Viral Gene Expression The Signal Peptide of a Simple Retrovirus Envelope Functions as a Posttranscriptional Regulator of Viral Gene Expression †. *J. Virol.* **83,** 4591–4604 (2009).
92. Murakami, T. Retroviral env glycoprotein trafficking and incorporation into virions. *Mol. Biol. Int.* **2012,** 1–11 (2012).
93. Chen, Y. *et al.* SPD--a web-based secreted protein database. *Nucleic Acids Res.* **33,** D169-73 (2005).
94. Braakman, I. & Bulleid, N. J. Protein Folding and Modification in the Mammalian Endoplasmic Reticulum. *Annu. Rev. Biochem.* **80,** 71–99 (2011).
95. Levine, C., Mitra, D. & Sharma, A. The efficiency of protein compartmentalization into the secretory pathway. *… Biol. cell* 279–291 (2005). doi:10.1091/mbc.E04
96. Evans, E. a, Gilmore, R. & Blobel, G. Purification of microsomal signal peptidase as a complex. *Proc. Natl. Acad. Sci. U. S. A.* **83,** 581–5 (1986).
97. Johnson, a E. Protein translocation at the ER membrane: A complex process becomes more so. *Trends Cell Biol.* **7,** 90–5 (1997).
98. Kapp, K., Schrempf, S., Lemberg, M. K. & Dobberstein, B. in *Protein Transport into the Endoplasmic Reticulum* (ed. Zimmermann, R.) 1–16 (Landes Bioscience, 2009).

99. Martoglio, B. & Dobberstein, B. Signal sequences : more than just greasy peptides. *Trends Cell Biol.* **8924,** 14119–14123 (1998).

100. von Heijne, G. Signal sequences. *J. Mol. Biol.* **184,** 99–105 (1985).

101. Hegde, R. S. & Bernstein, H. D. The surprising complexity of signal sequences. *TRENDS Biochem. Sci.* (2006). doi:10.1016/j.tibs.2006.08.004

102. Li, Y., Luo, L., Thomas, D. Y. & Kang, C. Y. The HIV-1 Env Protein Signal Sequence Retards Its Cleavage and Down-regulates the Glycoprotein Folding. *Virology* **272,** 417–428 (2000).

103. Blobel, G., Walter, P. & Gilmore, R. Intracellular protein topogenesis. *Prog. Clin. Biol. Res.* **168,** 3–10 (1984).

104. Goder, V. & Spiess, M. Topogenesis of membrane proteins: determinants and dynamics. *FEBS Lett.* **504,** 87–93 (2001).

105. Shaw, a S., Rottier, P. J. & Rose, J. K. Evidence for the loop model of signal-sequence insertion into the endoplasmic reticulum. *Proc. Natl. Acad. Sci. U. S. A.* **85,** 7592–6 (1988).

106. Li, Y. *et al.* Effects of inefficient cleavage of the signal sequence of HIV-1 gp 120 on its association with calnexin, folding, and intracellular transport. *Proc. Natl. Acad. Sci. U. S. A.* **93,** 9606–11 (1996).

107. Li, Y., Luo, L., Thomas, D. Y. & Kang, O. Y. Control of Expression, Glycosylation, and Secretion of HIV-1 gp120 by Homologous and Heterologous Signal Sequences. *Virology* **204,** 266–278 (1994).

108. Walter, P. & Johnson, A. E. Signal Sequence Recognition and Protein Targeting to the Endoplasmic Reticulum Membrane. *Annu. Rev. Cell Biol.* **10,** 87–119 (1994).

109. Walter, P. & Blobel, G. Disassembly and reconstitution of signal recognition particle. *Cell* **34,** 525–533 (1983).

110. Andrews, D. W., Walter, P. & Ottensmeyer, F. P. Evidence for an extended 7SL RNA structure in the signal recognition particle. *EMBO J.* **6,** 3471–7 (1987).

111. Zhang, L., Leng, Q. & Mixson, J. A. Alteration in the IL-2 signal peptide affects secretion of proteinsin vitro andin vivo Alteration in the IL-2 signal peptide affects secretion of proteins in vitro and in vivo. *J. Gene Med.* **7,** 354–365 (2005).

112. Zheng, N. & Gierasch, L. M. Signal Sequences: The Same Yet Different. *Cell* **86,** 849–852 (1996).

113. High, S. Pergamon PROTEIN AT THE MEMBRANE OF THE I . INTRODUCTION Eukaryotic cells are characterised by membrane enclosed subcellular compartments which perform , a variety of specialised functions (deDuve , 1991). These membranes act as selective barriers allow. 233–250 (1992).
114. Rapoport, T. a, Jungnickel, B. & Kutay, U. Protein transport across the eukaryotic endoplasmic reticulum and bacterial inner membranes. *Annu. Rev. Biochem.* **65,** 271–303 (1996).
115. Walter, P. & Blobel, G. Translocation of proteins across the endoplasmic reticulum III. Signal recognition protein (SRP) causes signal sequence-dependent and site-specific arrest of chain elongation that is released by microsomal membranes. *J. Cell Biol.* **91,** 557–61 (1981).
116. Ataide, S. F. *et al.* The crystal structure of the signal recognition particle in complex with its receptor. *Science* **331,** 881–886 (2011).
117. Voeltz, G. K., Rolls, M. M. & Rapoport, T. a. Structural organization of the endoplasmic reticulum. *EMBO Rep.* **3,** 944–50 (2002).
118. Walter, P., Gilmore, R. & Blobel, G. Protein translocation across the endoplasmic reticulum. *Cell* **38,** 5–8 (1984).
119. Millman, J. S. & Andrews, D. W. Switching the model: a concerted mechanism for GTPases in protein targeting. *Cell* **89,** 673–6 (1997).
120. Hegde, R. S. & Kang, S. The concept of translocational regulation. *J. Cell Biol.* **182,** 225–232 (2008).
121. Helenius, A Trombetta, E.S Hebert, D.N Simons, J. . Calnexin, calreticulin and the folding of glycoproteins. *Trends Cell Biol.* **7,** 193–200 (1997).
122. Helenius, A. Essay How N-linked Oligosaccharides Affect Glycoprotein Folding in the Endoplasmic Reticulum. *Mol. Biol. Cell* **5,** 253–265 (1992).
123. Frickel, E.-M. *et al.* ERp57 is a multifunctional thiol-disulfide oxidoreductase. *J. Biol. Chem.* **279,** 18277–87 (2004).
124. Land, A., Zonneveld, D. & Braakman, I. Folding of HIV-1 envelope glycoprotein involves extensive isomerization of disulfide bonds and conformation-dependent leader peptide cleavage. *FASEB J.* **17,** 1058–67 (2003).
125. Gething, M. J. Role and regulation of the ER chaperone BiP. *Semin. Cell Dev. Biol.* **10,** 465–72 (1999).

126. Land, A. & Braakman, I. Folding of the human immunodeficiency virus type 1 envelope glycoprotein in the endoplasmic reticulum. *Biochimie* **83,** 783–790 (2001).
127. Kornfeld, R. & Kornfeld, S. Assembly of asparagine-linked oligosaccharides. *Annu. Rev. Biochem.* **54,** 631–664 (1985).
128. McCune, J. M. *et al.* Endoproteolytic cleavage of gp160 is required for the activation of human immunodeficiency virus. *Cell* **53,** 55–67 (1988).
129. Freed, E. O., Myers, D. J. & Risser, R. Mutational analysis of the cleavage sequence of the human immunodeficiency virus type 1 envelope glycoprotein precursor gp160. *J. Virol.* **63,** 4670–5 (1989).
130. Moulard, M. & Decroly, E. Maturation of HIV envelope glycoprotein precursors by cellular endoproteases. *Analysis* **1469,** (2000).
131. Crooks, E. T., Tong, T., Osawa, K. & Binley, J. M. Enzyme digests eliminate nonfunctional Env from HIV-1 particle surfaces, leaving native Env trimers intact and viral infectivity unaffected. *J. Virol.* **85,** 5825–39 (2011).
132. von Heijne, G. Analysis of the distribution of charged residues in the N-terminal region of signal sequences: implications for protein export in prokaryotic and eukaryotic cells. *EMBO J.* **3,** 2315–2318 (1984).
133. Auriol, L. U. C. D., Vaquero, C. & Sitbon, M. Functional tolerance of the human peptide to mutations in the amino-terminal and Functional Tolerance of the Human Immunodeficiency Virus Type 1 Envelope Signal Peptide to Mutations in the Amino-Terminal and Hydrophobic Regions. *J. Virol.* **66,** 5114–5118 (1992).
134. Boyd, D. & Beckwith, J. The Role of Charged Amino Acids in the Localization of Secreted and Membrane Proteins Minireview. *Cell* **62,** 1031–1033 (1990).
135. Herrera, a M., Musacchio, a, Fernández, J. R. & Duarte, C. a. Efficiency of erythropoietin's signal peptide for HIV(MN)-1 gp 120 expression. *Biochem. Biophys. Res. Commun.* **273,** 557–9 (2000).
136. Pfeiffer, T., Pisch, T., Devitt, G., Holtkotte, D. & Bosch, V. Effects of signal peptide exchange on HIV-1 glycoprotein expression and viral infectivity in mammalian cells. *FEBS Lett.* **580,** 3775–3778 (2006).
137. Bonomelli, C. *et al.* The Glycan Shield of HIV Is Predominantly Oligomannose Independently of Production System or Viral Clade. 1–7 (2011). doi:10.1371/journal.pone.0023521

138. Crooks, E. T., Tong, T., Osawa, K. & Binley, J. M. Enzyme Digests Eliminate Nonfunctional Env from HIV-1 Particle Surfaces , Leaving Native Env Trimers Intact and Viral Infectivity Unaffected . *J. Virol.* **85,** 5825–5839 (2011).
139. Michalski, C. J., Li, Y. & Kang, C. Y. Induction of cytopathic effects and apoptosis in Spodoptera frugiperda cells by the HIV-1 Env glycoprotein signal peptide. *Virus Genes* **41,** 341–50 (2010).
140. Dell, A., Galadari, A., Sastre, F. & Hitchen, P. Similarities and differences in the glycosylation mechanisms in prokaryotes and eukaryotes. *Int. J. Microbiol.* **2010,** 148178 (2010).
141. Yan, A. & Lennarz, W. J. Unraveling the mechanism of protein N-glycosylation. *J. Biol. Chem.* **280,** 3121–3124 (2005).
142. Kelleher, D. J. & Gilmore, R. An evolving view of the eukaryotic oligosaccharyltransferase. 47–62 (2005). doi:10.1093/glycob/cwj066
143. Welply, J. K., Shenbagamurthi, P. & Lennarz, W. J. Substrate Recognition by Oligosaccharyltransferase. *J. Biol. Chem.* **258,** 11856–11863 (1983).
144. Gavel, Y. & Heyne, G. Von. Sequence differences between glycosylated and non-glycosylated Asn-X-Thr / Ser acceptor sites : implications for protein engineering. *Protein Eng.* **3,** 433–442 (1990).
145. Apweiler, R., Hermjakob, H. & Sharon, N. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim. Biophys. Acta - Gen. Subj.* **1473,** 4–8 (1999).
146. Go, E. P. *et al.* Characterization of Glycosylation Profiles of HIV-1 Transmitted / Founder Envelopes by. *J. Virol.* **85,** 8270–8284 (2011).
147. Rao, R. S. P., Buus, O. T. & Wollenweber, B. Distribution of N-glycosylation sequons in proteins: How apart are they? *Comput. Biol. Chem.* **35,** 57–61 (2011).
148. Petrescu, A. J., Milac, A. L., Petrescu, S. M., Dwek, R. A. & Wormald, M. R. Statistical analysis of the protein environment of N-glycosylation sites: Implications for occupancy, structure, and folding. *Glycobiology* **14,** 103–114 (2004).
149. Hettkamp, H., Legler, G. & Bause, E. Purification by affinity chromatography of glucosidase I, an endoplasmic reticulum hydrolase involved in the processing of asparagine???linked oligosaccharides. *Eur. J. Biochem.* **142,** 85–90 (1984).
150. Vigerust, D. J. & Shepherd, V. L. Virus glycosylation : role in virulence and immune

interactions. *Trends Microbiol.* **15,** (2007).

151. Stanley, P. Golgi glycosylation. *Cold Spring Harb. Perspect. Biol.* **3,** (2011).
152. Chackerian, B., Rudensey, L. M. & Overbaugh, J. Specific N-linked and O-linked glycosylation modifications in the envelope V1 domain of simian immunodeficiency virus variants that evolve in the host alter recognition by neutralizing antibodies . Specific N-Linked and O-Linked Glycosylation Modification. *J. Virol.* **71,** 7719–7727 (1997).
153. Doores, K. J. *et al.* Envelope glycans of immunodeficiency virions are almost entirely oligomannose antigens. *PNAS* **107,** 13800–13805 (2010).
154. Coss, K. P. *et al.* HIV-1 Glycan Density Drives the Persistence of the Mannose Patch within an Infected Individual. *J. Virol.* **90,** 11132–11144 (2016).
155. Zhu, X., Borchers, C., Bienstock, R. J. & Tomer, K. B. Mass Spectrometric Characterization of the Glycosylation Pattern of HIV-gp120 Expressed in CHO Cells. *Peptides* 11194–11204 (2000).
156. Go, E. P. *et al.* Characterization of Host-Cell Line Specific Glycosylation Profiles of Early Transmitted/Founder HIV‑1 gp120 Envelope Proteins. *J. Proteome Res.* **12,** 1223–1234 (2013).
157. Moore, P. L. *et al.* Evolution of an HIV glycan – dependent broadly neutralizing antibody epitope through immune escape. *Nat. Med.* 1–6 (2012). doi:10.1038/nm.2985
158. Yang, W. *et al.* Glycoform analysis of recombinant and human immunodeficiency virus envelope protein gp120 via higher energy collisional dissociation and spectral-aligning strategy. *Anal. Chem.* **86,** 6959–6967 (2014).
159. Go, E. P. *et al.* Glycosylation Benchmark Profile for HIV-1 Envelope Glycoprotein Production Based on Eleven Env Trimers. *J. Virol.* **91,** e02428-16 (2017).
160. Pritchard, L. K. *et al.* Structural Constraints Determine the Glycosylation of HIV-1 Envelope Trimers Article Structural Constraints Determine the Glycosylation of HIV-1 Envelope Trimers. *CellReports* **11,** 1604–1613 (2015).
161. Chung, N. P. Y. *et al.* Stable 293 T and CHO cell lines expressing cleaved , stable HIV-1 envelope glycoprotein trimers for structural and vaccine studies. **11,** 1–14 (2014).
162. Doores, K. J. The HIV glycan shield as a target for broadly neutralizing antibodies. *FEBS J.* **282,** 4679–4691 (2015).

163. Pritchard, L. K. *et al.* Glycan clustering stabilizes the mannose patch of HIV-1 and preserves vulnerability to broadly neutralizing antibodies. *Nat. Commun.* **6,** 1–11 (2015).

164. Raska, M. *et al.* Glycosylation Patterns of HIV-1 gp120 Depend on the Type of Expressing Cells and Affect Antibody Recognition * □. 20860–20869 (2010). doi:10.1074/jbc.M109.085472

165. Li, Y. A. N., Luo, L., Rasool, N. & Kangt, C. Y. Glycosylation Is Necessary for the Correct Folding of Human Immunodeficiency Virus gpl20 in CD4 Binding. *J. Virol.* **67,** 584–588 (1993).

166. Ng, D. T. W., Hiebert, S. W. & Lamb, R. A. Different Roles of Individual N-Linked Oligosaccharide Chains in Folding , Assembly , and Transport of the Simian Virus 5. *Mol. Cell Biol.* **10,** 1989–2001 (1989).

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

168. Heifetz, A., Keenan, R. W. & Elbein, A. D. Mechanism of Action of Tunicamycin on the Transferaset. 2186–2192 (1979).

169. Elbein, A. & Heath, E. Inhibitors of the Biosynthesis and Processing of N-Linked Oligosaccharide. *Crit. Rev. Biochem. …* **16,** 21–49 (1984).

170. Fenouillet, E. & Jones, I. M. The glycosylation of human immunodeficiency virus type 1 transmembrane glycoprotein (gp41) is important for the efficient intracellular transport of the envelope precursor gp160. *J. Gen. Virol.* **76 (Pt 6),** 1509–14 (1995).

171. Wang, W. *et al.* A systematic study of the N-glycosylation sites of HIV-1 envelope protein on infectivity and antibody- mediated neutralization. *Retrovirology* **10,** (2013).

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

173. Shen, R., Raska, M., Bimczok, D., Novak, J. & Smith, P. D. HIV-1 Envelope Glycan Moieties Modulate HIV-1 Transmission. *J. Virol.* **88,** 14258–14267 (2014).

174. Wolinsky, S. *et al.* Selective transmission of human immunodeficiency virus type-1 variants from mothers to infants. *Science (80-.).* **255,** 1134–1137 (1992).

175. Zhu, T. *et al.* Genotypic and Phenotypic Characterization Infection HIV-1 in Patients with Primary. *Science (80-.).* **261,** 1–4 (1993).

176. Derdeyn, C. A. *et al.* Envelope-Constrained Neutralization-Sensitive HIV-1 after

Heterosexual Transmission. *Science (80-.).* **303,** 2019–2022 (2004).

177. Salazar-Gonzalez, J. F. *et al.* Deciphering human immunodeficiency virus type 1 transmission and early envelope diversification by single-genome amplification and sequencing. *J. Virol.* **82,** 3952–70 (2008).
178. Abrahams, M. *et al.* Quantitating the multiplicity of infection with human immunodeficiency virus type 1 subtype C reveals a non-poisson distribution of transmitted variants. *J. Virol.* **83,** 3556–3567 (2009).
179. Haaland, R. E. *et al.* Inflammatory genital infections mitigate a severe genetic bottleneck in heterosexual transmission of subtype A and C HIV-1. *PLoS Pathog.* **5,** e1000274 (2009).
180. Nofemela, A. *et al.* Defining the human immunodeficiency virus type 1 transmission genetic bottleneck in a region with multiple circulating subtypes and recombinant forms. *Virology* **415,** 107–113 (2006).
181. Miller, W. C., Rosenberg, N. E., Rutstein, S. E. & Powers, K. a. Role of acute and early HIV infection in the sexual transmission of HIV. *Curr. Opin. HIV AIDS* **5,** 277–82 (2010).
182. Grivel, J.-C., Shattock, R. J. & Margolis, L. B. Selective transmission of R5 HIV-1 variants: where is the gatekeeper? *J. Transl. Med.* **9 Suppl 1,** S6 (2011).
183. Simmonds, P., Balfe, P., Ludlam, C. a, Bishop, J. O. & Brown, a J. Analysis of sequence diversity in hypervariable regions of the external glycoprotein of human immunodeficiency virus type 1. *J. Virol.* **64,** 5840–50 (1990).
184. Palmer, S. *et al.* Multiple , Linked Human Immunodeficiency Virus Type 1 Drug Resistance Mutations in Treatment-Experienced Patients Are Missed by Standard Genotype Analysis Multiple , Linked Human Immunodeficiency Virus Type 1 Drug Resistance Mutations in Treatment-Experie. *J. Clin. Microbiol.* **43,** 403 (2005).
185. Chohan, B. *et al.* 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. *Society* **79,** 6528–6531 (2005).
186. Sagar, M. *et al.* Selection of HIV Variants with Signature Genotypic Characteristics during Heterosexual Transmission. **199,** (2009).
187. Frost, S. D. W. *et al.* Characterization of Human Immunodeficiency Virus Type 1 (HIV-1) Envelope Variation and Neutralizing Antibody Responses during Transmission of

HIV-1 Subtype B Characterization of Human Immunodeficiency Virus Type 1 (HIV-1) Envelope Variation and Neutra. *J. Virol.* **79,** 6523–6527 (2005).

188. Salazar-Gonzalez, J. F. *et al.* Deciphering Human Immunodeficiency Virus Type 1 Transmission and Early Envelope Diversification by Single-Genome Amplification and Sequencing Deciphering Human Immunodeficiency Virus Type 1 Transmission and Early Envelope Diversification by Single-Genome. *J. Virol.* **82,** 3952–3970 (2008).
189. Alexander, M. *et al.* Donor and Recipient Envs from Heterosexual Human Immunodeficiency Virus Subtype C Transmission Pairs Require High Receptor Levels for Entry . 4100–4104 (2010). doi:10.1128/JVI.02068-09
190. Isaacman-Beck, J. *et al.* Heterosexual transmission of human immunodeficiency virus type 1 subtype C: Macrophage tropism, alternative coreceptor use, and the molecular anatomy of CCR5 utilization. *J. Virol.* **83,** 8208–8220 (2009).
191. Brenchley, J. M. *et al.* CD4+ T cell depletion during all stages of HIV disease occurs predominantly in the gastrointestinal tract. *J. Exp. Med.* **200,** 749–59 (2004).
192. McKinnon, L. R. *et al.* Characterization of a Human Cervical CD4+ T Cell Subset Coexpressing Multiple Markers of HIV Susceptibility. *J. Immunol.* **187,** 6032–6042 (2011).
193. Cicala, C. *et al.* The integrin 4 7 forms a complex with cell-surface CD4 and defines a T-cell subset that is highly susceptible to infection by HIV-1. 20877–20882 (2009).
194. Perciani, C. T. *et al.* αEβ7 , α4β7 and α4β1 integrin contributions to T cell distribution in blood , cervix and rectal tissues : Potential implications for HIV transmission. *PLoS One* **13,** e0192482 (2018).
195. Nawaz, F. *et al.* 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 Pathog.* **7,** 1–14 (2011).
196. Gonzalez, M. W., Devico, A. L., Lewis, G. K. & Spouge, J. L. Conserved Molecular Signatures in gp120 Are Associated with the Genetic Bottleneck during Simian Immunodeficiency Virus (SIV), SIV- Human Immunodeficiency Virus (SHIV), and HIV Type 1 (HIV-1). **89,** 3619–3629 (2015).
197. Rizzuto, C. & Sodroski, J. Fine definition of a conserved CCR5-binding region on the

human immunodeficiency virus type 1 glycoprotein 120. *AIDS Res Hum Retroviruses* **16,** 741–749 (2000).

198. Irungu, J. *et al.* Comparison of HPLC/ESI-FTICR MS Versus MALDI-TOF/TOF MS for Glycopeptide Analysis of a Highly Glycosylated HIV Envelope Glycoprotein. *J. Am. Soc. Mass Spectrom.* **19,** 1209–1220 (2008).
199. Sato, S. *et al.* Potent Antibody-Mediated Neutralization and Evolution of Antigenic Escape Variants of Simian Immunodeficiency Virus Strain SIVmac239 In Vivo. *J. Virol.* **82,** 9739–9752 (2008).
200. Wain-hobson, S., Sonigo, P., Danos, O., Cole, S. & Alizon, M. Nucleotide Sequence of the AIDS Virus , LAV. *Cell* **81,** 9–17 (1985).
201. Sanders, R. & Hsu, S. Evolution rescues folding of human immunodeficiency virus-1 envelope glycoprotein GP120 lacking a conserved disulfide bond. *Mol. Biol. …* 4707–4716 (2008). doi:10.1091/mbc.E08
202. Wang, B. *et al.* Incorporation of High Levels of Chimeric Human Immunodeficiency Virus Envelope Glycoproteins into Virus-Like Particles Incorporation of High Levels of Chimeric Human Immunodeficiency Virus Envelope Glycoproteins into Virus-Like Particles . *J. Virol.* **81,** 10869–10878 (2007).
203. Upadhyay, C. *et al.* Alterations of HIV-1 envelope phenotype and antibody-mediated neutralization by signal peptide mutations. *PLOS Pathog.* **14,** e1006812 (2018).
204. Kong, L. *et al.* Expression system-dependent Modulation of HIV-1 Envelope Glycoprotein Antigenecity and Immunogenicity. *J. Mol. Biol.* **403,** 131–147 (2010).
205. Green, M. R. & Sambrook, J. *Molecular cloning: A Laboratory Manual*. *Molecular cloning: A Laboratory Manual* **4,** (2012).
206. Harwood, A. J. The Rapid Boiling Method for Small-Scale Preparation of Plasmid DNA. *Basic DNA RNA Protoc.* **58,** 265–268 (1996).
207. Platt, E. J., Wehrly, K., Kuhmann, S. E., Chesebro, B. & Kabat, D. Effects of CCR5 and CD4 Cell Surface Concentrations on Infections by Macrophagetropic Isolates of Human Immunodeficiency Virus Type 1. *J. Virol.* **72,** 2855–2864 (1998).
208. Uphoff, C. C. & Drexler, H. G. Basic Cell Culture Protocols. *Methods Mol. Biol.* **946,** 1–13 (2013).
209. Welinder, C. & Ekblad, L. Coomassie Staining as Loading Control in Western Blot Analysis. *J. Proteome Res.* **10,** 1416–1419 (2011).

210. Lin, G. *et al.* Differential N-Linked Glycosylation of Human Immunodeficiency Virus and Ebola Virus Envelope Glycoproteins Modulates Interactions with DC-SIGN and DC-SIGNR. *Society* **77,** 1337–1346 (2003).
211. Wei, X., Decker, J. M., Wang, S. & Hui, H. Antibody neutralization and escape by HIV-1. *Nature* **837,** 835–837 (2003).
212. Montefiori, D. C. Evaluating neutralizing antibodies against HIV, SIV, and SHIV in luciferase reporter gene assays. *Curr. Protoc. Immunol.* **Chapter 12,** Unit 12.11 (2005).
213. Marozsan, A. J. & Arts, E. J. Development of a yeast-based recombination cloning/system for the analysis of gene products from diverse human immunodeficiency virus type 1 isolates. *J. Virol. Methods* **111,** 111–120 (2003).
214. Dudley, D. M., Henry, K. & Gibson, R. M. A novel yeast-based recombination method to clone and propagate diverse HIV-1 isolates. *Biotechniques* **46,** 458–467 (2009).
215. Gietz, R. D. & Schiestl, R. H. Frozen competent yeast cells that can be transformed with high efficiency using the LiAc/SS carrier DNA/PEG method. *Nat. Protoc.* **2,** 1–4 (2007).
216. Edmonds, T. G., Ding, H., Yuan, X., Wei, Q. & Smith, K. S. Replication Competent Molecular Clones of HIV-1 Expressing Renilla Luciferase Facilitate the Analysis of Antibody Inhibition in PBMC. *Virology* **408,** 1–13 (2010).
217. Shaw, G. M. & Hunter, E. HIV transmission. *Cold Spring Harb. Perspect. Med.* **2,** (2012).
218. Hladik, F. & Hope, T. J. HIV infection of the genital mucosa in women. *Curr. HIV/AIDS Rep.* **6,** 20–8 (2009).
219. Fox, J. & Fidler, S. Sexual transmission of HIV-1. *Antiviral Res.* **85,** 276–285 (2010).
220. Mlisana, K. *et al.* Symptomatic vaginal discharge is a poor predictor of sexually transmitted infections and genital tract inflammation in high-risk women in South Africa. *J. Infect. Dis.* **206,** 6–14 (2012).
221. Freeman, E. E. *et al.* Herpes simplex virus 2 infection increases HIV acquisition in men and women: systematic review and meta-analysis of longitudinal studies. *Aids* **20,** 73–83 (2006).
222. Boily, M. C. *et al.* Heterosexual risk of HIV-1 infection per sexual act: systematic review and meta-analysis of observational studies. *Lancet Infect. Dis.* **9,** 118–129 (2009).

223. Masson, L. *et al.* Genital Inflammation and the Risk of HIV Acquisition in Women. *Clin. Infect. Dis. HIV/AIDS* **61,** 260–269 (2015).

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

225. Safrit, J. T. *et al.* Status of vaccine research and development of vaccines for HIV-1 ☹. *Vaccine* **34,** 2921–2925 (2016).

226. Kariuki, S. M., Selhorst, P., Ariën, K. K. & Dorfman, J. R. The HIV‐1 transmission bottleneck. *Retrovirology* **14,** 1–19 (2017).

227. Parker, Z. F. *et al.* Transmitted/founder and chronic HIV-1 envelope proteins are distinguished by differential utilization of CCR5. *J. Virol.* **87,** 2401–11 (2013).

228. Baalwa, J. *et al.* Molecular identification , cloning and characterization of transmitted / founder HIV-1 subtype A , D and A / D infectious molecular clones. *Virology* **436,** 33–48 (2012).

229. Parrish, N. F. *et al.* Phenotypic properties of transmitted founder HIV-1. *PNAS* **1110,** 6626–6633 (2013).

230. Wilen, C. B. *et al.* Phenotypic and Immunologic Comparison of Clade B Transmitted / Founder and Chronic HIV-1 Envelope Glycoproteins ☐. 8514–8527 (2011). doi:10.1128/JVI.00736-11

231. Tebit, D.M.& Arts, E.J. Tracking a century of global expansion of HIV to drive understanding and to combat disease, *Lancet Infect. Dis*. **11,** 45-46 (2011).

232. Van Harmelen, J. H. *et al.* A predominantly HIV type 1 subtype C-restricted epidemic in South African urban populations. *AIDS Res. Hum. Retroviruses* **15,** 395–8 (1999).

233. Chakrabarti, B. K. *et al.* Modifications of the Human Immunodeficiency Virus Envelope Glycoprotein Enhance Immunogenicity for Genetic Immunization Modifications of the Human Immunodeficiency Virus Envelope Glycoprotein Enhance Immunogenicity for Genetic Immunization. *J. Virol.* **76,** 5357–5368 (2002).

234. Stamatatos, L., Lim, M. & Cheng-mayer, C. Generation and Structural Analysis of Soluble Oligomeric gp140 Envelope Proteins Derived from Primary HIV Type 1 Isolates. *AIDS Res. Hum. Retroviruses* **16,** 981–994 (2000).

235. AlSalmi, W. *et al.* A new approach to produce HIV-1 envelope trimers: Both cleavage and proper glycosylation are essential to generate authentic trimers. *J. Biol. Chem.*

290, 19780–19795 (2015).

236. Go, E. P. *et al.* Glycosylation site-specific analysis of HIV envelope proteins (JR-FL and CON-S) reveals major differences in glycosylation site occupancy, glycoform profiles, and antigenic epitopes' accessibility. *J. Proteome Res.* **7,** 1660–1674 (2008).
237. Go, E. P. *et al.* Glycosylation Site-Specific Analysis of Clade C HIV-1 Envelope Proteins research articles. 4231–4242 (2009).
238. Stern, B. & Olsen, L. Improving mammalian cell factories: The selection of signal peptide has a major impact on recombinant protein synthesis and secretion in mammalian cells. *Trends Cell …* (2007).
239. Fenouillet, E. & Gluckman, C. Role of N-linked glycans of envelope glycoproteins in infectivity of human Role of N-Linked Glycans of Envelope Glycoproteins in Infectivity of Human Immunodeficiency Virus Type 1. *J. Virol.* **64,** 2841–2848 (1990).
240. Herrera, C. *et al.* The impact of envelope glycoprotein cleavage on the antigenicity , infectivity , and neutralization sensitivity of Env-pseudotyped human immunodeficiency virus type 1 particles. **338,** 154–172 (2005).
241. Wang, W. *et al.* A systematic study of the N-glycosylation sites of HIV-1 envelope protein on infectivity and antibody-mediated neutralization. *Retrovirology* **10,** 1–14 (2013).
242. Botarelli, P. & Houlden, B. N-glycosylation of HIV-gp120 may constrain recognition by T lymphocytes. *J. …* **147,** 3128–3132 (1991).
243. Lee, W. Mutational analysis of conserved N-linked glycosylation sites of human immunodeficiency virus type 1 gp41 . Mutational Analysis of Conserved N-Linked Glycosylation Sites of Human Immunodeficiency Virus Type 1 gp4l. *J. Virol.* **66,** 1799–1805 (1992).
244. Reynard, F., Willkomm, N., Fatmi, A., Verrier, B. & Bedin, F. Characterization of the antibody response elicited by HIV-1 Env glycomutants in rabbits. *Electrophoresis* **25,** 535–546 (2007).
245. Liu, Y. *et al.* Env length and N-linked glycosylation following transmission of human immunodeficiency virus Type 1 subtype B viruses. **374,** 229–233 (2005).
246. Lin, G. *et al.* Identification of gp120 Binding Sites on CXCR4 by Using CD4-Independent Human Immunodeficiency Virus Type 2 Env Proteins Identification of gp120 Binding Sites on CXCR4 by Using CD4-Independent Human Immunodeficiency Virus Type 2

Env Proteins. *J. Virol.* **77,** 931–942 (2003).

247. Geyers, H. & Schneidern, J. Carbohydrates of Human Immunodeficieny Virus. *J. Biol. …* **263,** 11760–11767 (1988).

248. Pritchard, L. K., Harvey, D. J., Bonomelli, C., Crispin, M. & Doores, K. J. Cell- and Protein-Directed Glycosylation of Native Cleaved HIV-1 Envelope. *J. Virol.* **89,** 8932–8944 (2015).

249. Provine, N. M., Puryear, W. B., Wu, X., Overbaugh, J. & Haigwood, N. L. The Infectious Molecular Clone and Pseudotyped Virus Models of Human Immunodeficiency Virus Type 1 Exhibit Significant Differences in Virion Composition with Only Moderate Differences in Infectivity and Inhibition Sensitivity . *J. Virol.* **83,** 9002–9007 (2009).

250. Burton, D. R. A vaccine for HIV type 1: the antibody perspective. *Proc. Natl. Acad. Sci. U. S. A.* **94,** 10018–23 (1997).

251. Miyauchi, K. *et al.* The membrane-spanning domain of gp41 plays a critical role in intracellular trafficking of the HIV envelope protein. *Retrovirology* **7,** 95 (2011).

252. Kong, L. *et al.* Expression-System-Dependent Modulation of HIV-1 Envelope Glycoprotein Antigenicity and Immunogenicity. *J. Mol. Biol.* **403,** 131–147 (2010).

253. Raska, M. *et al.* Glycosylation patterns of HIV-1 gp120 depend on the type of expressing cells and affect antibody recognition. *J. Biol. Chem.* **285,** 20860–9 (2010).

254. Hammonds, J. *et al.* Gp120 stability on HIV-1 virions and Gag-Env pseudovirions is enhanced by an uncleaved Gag core. *Virology* **314,** 636–649 (2003).

255. Louder, M. K. *et al.* HIV-1 envelope pseudotyped viral vectors and infectious molecular clones expressing the same envelope glycoprotein have a similar neutralization phenotype , but culture in peripheral blood mononuclear cells is associated with decreased neutralization sens. *Virology* **339,** 226–238 (2005).

256. Moore, J. P., Willey, R. L., Lewis, G. K., Robinson, J. & Sodroski, J. Immunological evidence for interactions between the first, second, and fifth conserved domains of the gp120 surface glycoprotein of human immunodeficiency virus type 1. *J. Virol.* **68,** 6836–6847 (1994).

257. Mathys, L., François, K. O., Quandte, M., Braakman, I. & Balzarini, J. Deletion of the highly conserved N-Glycan at Asn260 of HIV-1 gp120 affects folding and lysosomal degradation of gp120, and results in loss of viral infectivity. *PLoS One* **9,** 1–11 (2014).

258. Kong, L. *et al.* Complete epitopes for vaccine design derived from a crystal structure

of the broadly neutralizing antibodies PGT128 and 8ANC195 in complex with an HIV-1 Env trimer. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **71,** 2099–2108 (2015).

259. Townsley, S., Li, Y., Kozyrev, Y., Cleveland, B. & Hu, S.-L. The conserved role of an N-linked glycan on the surface antigen of HIV-1 modulating virus sensitivity to broadly neutralizing antibodies against the receptor and coreceptor binding sites. *J. Virol.* JVI.02321-15 (2015). doi:10.1128/JVI.02321-15
260. Huang, C. C. *et al.* Structural biology: Structure of a V3-containing HIV-1 gp120 core. *Science (80-.).* **310,** 1025–1028 (2005).
261. Chen, M. *et al.* HIV gp120 V1/V2and C2-V3domains glycoprotein compatibility is required for viral replication. *Virus Res.* **79,** 91–101 (2001).
262. Platt, E. J., Durnin, J. P. & Kabat, D. Kinetic Factors Control Ef ciencies of Cell Entry, Ef cacies of Entry Inhibitors, and Mechanisms of Adaptation of Human Immunode ciency Virus†. *Society* **79,** 4347–4356 (2005).
263. Johnston, S. H. *et al.* A quantitative affinity-profiling system that reveals distinct CD4/CCR5 usage patterns among human immunodeficiency virus type 1 and simian immunodeficiency virus strains. *J. Virol.* **83,** 11016–11026 (2009).
264. Chikere, Kelechi C; Chou, Tom; Gorry, Paul R; Lee, B. Affinofile Profiling: How efficiency of CD4/CCR5 usage impacts the biological and pathogenic phenotype of HIV. *Virology* **435,** 81–91 (2013).
265. Arien, K. K. *et al.* The replicative fitness of primary human immunodeficiency virus type 1 (HIV-1) group M, HIV-1 group O, and HIV-2 isolates. *J. Virol.* **79,** 8979 (2005).
266. Ochsenbauer, C. *et al.* Generation of Transmitted / Founder HIV-1 Infectious Molecular Clones and Characterization of Their Replication Capacity in CD4 T Lymphocytes and Monocyte-Derived Macrophages. *J. Virol.* **86,** 2715–2728 (2012).
267. Aylon, Y. & Kupiec, M. New insights into the mechanism of homologous recombination in yeast. *Mutat. Res. - Rev. Mutat. Res.* **566,** 231–248 (2004).
268. Baudin, A., Ozier-kalogeropoulos, O., Denouel, A., Lacroute, F. & Cullin, C. A simple and efficient method for direct gene deletion in Saccharomyces cerevisiae. *Nucleic Acids Res.* **21,** 3329–3330 (1993).
269. Lorenz, M. C. *et al.* Gene disruption with PCR products in Saccharomyces cerevisiae. *Gene* **158,** 113–117 (1995).
270. Sikorski, R. S. & Hieter, P. A system of shuttle vectors and yeast host strains designed

for efficient manipulation of DNA in Saccharomyces cerevisiae. *Genetics* **122,** 19–27 (1989).

271. Weber, J. *et al.* Novel Method for Simultaneous Quantification of Phenotypic Resistance to Maturation, Protease, Reverse Transcriptase, and Integrase HIV Inhibitors Based on 3'Gag(p2/p7/p1/p6)/PR/RT/INT-Recombinant Viruses: a Useful Tool in the Multitarget Era of Antiretro. *Antimicrob. Agents Chemother.* **55,** 3729–3742 (2011).
272. Godbey, W. T., Wu, K. K. & Mikos, A. G. Size matters: Molecular weight affects the efficiency of poly(ethylenimine) as a gene delivery vehicle. *J. Biomed. Mater. Res.* **45,** 268–275 (1999).
273. Card, C. M. *et al.* Reduced Cellular Susceptibility to In Vitro HIV Infection Is Associated with CD4 + T Cell Quiescence. *PLoS One* **7,** e45911 (2012).
274. Kinter, a L. *et al.* HIV replication in CD4+ T cells of HIV-infected individuals is regulated by a balance between the viral suppressive effects of endogenous beta-chemokines and the viral inductive effects of other endogenous cytokines. *Proc. Natl. Acad. Sci. U. S. A.* **93,** 14076–14081 (1996).
275. Schweighardt, B. *et al.* R5 Human Immunodeficiency Virus Type 1 (HIV-1) Replicates More Efficiently in Primary CD4 T-Cell Cultures Than X4 HIV-1, *Journal of Virology*, Sept. 2004, p. 9164–9173 (2004)
276. Venner, C. M. *et al.* Infecting HIV-1 Subtype Predicts Disease Progression in Women of Sub-Saharan Africa, *EBioMedicine* **13,** 305–314 (2016)
277. Kiguoya M. W. *et al.* Subtype-Specific Differences in Gag- Protease-Driven Replication Capacity Are Consistent with Intersubtype Differences in HIV-1 Disease Progression, Journal of Virology **91,** e00253-17 (2017).

APPENDIX I - SUPPLEMENTARY FIGURES

Figure A1: Alignment of subtype B transmission motifs. CAP210 E8 sequence was aligned to subtype B sequences: HXB2 and five sequences analysed by Aslam *et al.* (2011) and REJO and JRFL analysed by Upadhyah *et al*. (2018). The putative transmission motifs, His12 and PNG413, are indicated by grey bars and the amino acids mutated in this study are indicated with arrows. The Histidine in the signal peptide is indicated as position 12 despite HXB2 numbering indicating that it is at position 9. Gnanakaran *et al*. (2011) indicated that the numbering of some positions was based on positions within alignments and not always HXB2 numbering. To avoid confusion, we selected to continue with numbering the His as position 12. The second transmission motif was identified as the Threonine at position 415 but as this amino acid determined the presence of a PNG at 413, the N-glycan site was considered to be important for transmission and thus we refer to PNG413 as the transmission motif and substituted T415 with an Isoleucine (I), the common amino acid found in this position, to delete the PNG[8]. HXB2 numbering was determined as explained at www.lanl.gov.

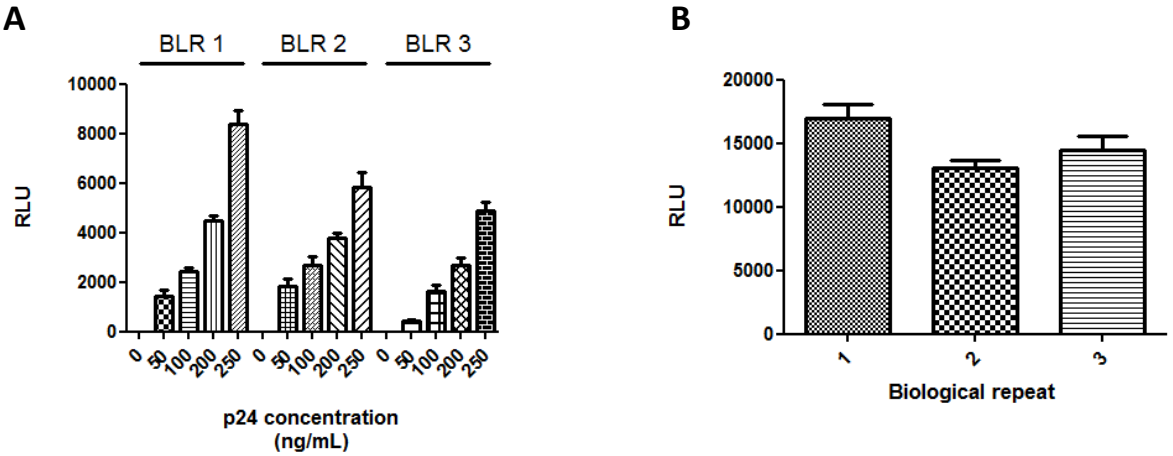

Figure A2: Entry efficiency of pseudovirus and IMCs shows high inter-assay variability. Using Wild type (WT) as an example, A) Pseudovirions (PSVs) generated with pNL4.3-R-E-luc+ HIV backbone were diluted to 0, 50, 100, 200 and 250 ng/mL p24 and used to infect TZM-bl luciferase reporter cells in triplicate. The bars indicate the relative light unit (RLU) values generated with error bars representing standard deviation of the three triplicate values. Three biological repeats (BLR) show the variation between assays using PSVs that were generated using the same HIV backbone. B. IMCs were diluted to 50 ng/mL p24 and used to infect TZM-bl cells in triplicate. Bars show RLUs of three independent experiments with error bars indicating the standard deviation of the mean.

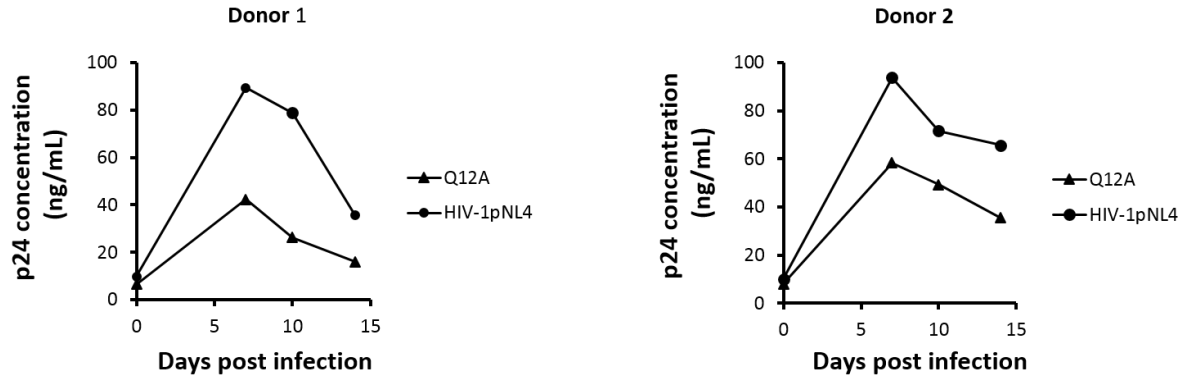

Figure A3: Replication of chimeric IMCs in PBMCs from two different donors show similar trend. PBMCs from two donors were infected with viruses with equal TCID50 generated from CAP210 E8. Viral replication was quantified by p24 ELISA 0, 7, 10, and 14 post infection. The p24 concentration in ng/mL from the two independent assays from 2 donors were plotted on separate line graphs over time.

APPENDIX II - SOLUTIONS

Plasmid DNA isolation and gel electrophoresis

10X TBE

108 g Tris-Cl

55 g Boric Acid

9.3 g EDTA

Top up to 1 L with dH_2O after adjusting pH to 8.5 with 10% citric acid

6X DNA loading dye

0.025 g Bromophenol Blue

0.025 g Xylene Cyanol

6 mL 50% glycerol

4 mL dH_2O

0.8% agarose gel

0.8 g agarose

100 mL dH_2O

Add 2 µL EtBr after heating

STET buffer

8 g Sucrose

5 g Triton X-100

10 mL 0.5 M EDTA, pH 8.0

5 mL 1 M Tris-Cl, pH 8.0

Top up to 100 mL with sterile dH_2O

SDS-PAGE and western blotting

RIPA buffer

10 mM Tris-Cl (pH 7.4)

150 mM NaCl

1% Sodium deoxycholate

0.1% SDS

1% Triton-X100

Make up to 100 mL with dH_20

10X running buffer

30.285 g Tris-Cl

144.13 g Glycine

10 g SDS

Make up to 1 L with dH_2O after adjusting pH to 8.3 with HCl

5X protein loading buffer

1 mL 1 M Tris-Cl (pH 6.8)

2.5 mL 20% SDS

4 mL 50% glycerol

0.01 g Bromophenol blue

Make up to 10 mL with dH_2O

Transfer buffer (TOWBIN)

3.025 g Tris-Cl

14.4 g Glycine

200 mL Methanol

Make up to 1 L with dH_2O

8% resolving gel

2 mL 30% Acrylamide/bis-Acrylamide

2.8 mL 1M Tris-Cl (pH 8.8)

37.5 µL 20% SDS

40 µL 20% Ammonium persulphate

20 µL TEMED

2.56 mL dH_2O

7% stacking gel

0.85 mL 30% Acrylamide/bis-Acrylamide

0.625 mL 1M Tris-Cl (pH 6.8)

25 µL 20% SDS

30 µL 20% Ammonium persulphate

10 µL TEMED

3.46 mL dH_2O

Partial protein purification

MES binding buffer

6.5 mL 1 M NaCl

5 mL 100 mM $CaCl_2$

10 mL 100 mM MES monohydrate

Make up to 50 mL with dH_2O after adjusting pH to 6.0 with HCl

1X MMP competitive binding buffer

9.709 g MMP

Make up to 50 mL with 1X PBS

1X PBS wash buffer

150 µL 100 mM $CaCl_2$

245 µL 100 mM $MgCl_2$

49.71 mL 1X PBS

Coomassie stain

50 mL Methanol

10 mL Glacial acetic acid

1 g Coomassie Brilliant Blue R-250

Make up to 100 mL with dH_2O

Destaining solution

125 mL Ethanol

50 mL Glacial acetic acid

Make up to 500 mL with dH_2O

<u>Yeast gap repair assay</u>

YEPD media

10 g Yeast Extract

20 g Peptone

20 g Dextrose

Make up to 1 L with dH_2O

Mix and autoclave.

Yeast Dropout media

6.7 g YNB (Yeast Nitrogen Base)

20 g Dextrose

Make up to 1 L with dH_2O

Stir until completely dissolved and autoclave

Yeast Amino Acid Dropout plates + FOA

Add 20 g agar

Add 6.7 g YNB (Yeast Nitrogen Base)

20 g Dextrose

0.69 g CSM-Leu

1 g FOA

Make up to 1 L with dH_2O

Stir until completely dissolved, autoclave and pour

10X TE Buffer

10 mL 1M Tris-Cl (pH 8.0)

2 mL 0.5M EDTA (pH 8.0)

Make up to 1 L with dH_2O

Replication assay

Full RPMI

Remove 60 mL RPMI from the bottle and add the following:

50 mL FBS

10 mL Pen/Strep

(Supplement with IL-2 and/or pHA, if necessary)