Studies on the activation mechanism of the HIV-1 membrane fusion glycoprotein, gp41

Ashraf Khasawneh

MBBS, MBiomedSc (Part 1)

Supervisor: Dr. Andy Poumbourios
Co-supervisor: Associate Professor Heidi E Drummer

Thesis submitted for the degree of

Doctor of Philosophy

Accepted on recommendation of

Assistant Professor Michael J. Root
Dr. Robert Center

July 2011

Copyright Notices

Notice 1

Notice 2

Department of Microbiology
Monash University Victoria 3800 AUSTRALIA

Fusion Laboratory, Centre for Virology
Macfarlane Burnet Institute for Medical Research
and Public Health Victoria 3004 AUSTRALIA

Table of Contents

List of Figures

List of Tables

Abstract

Retrovirus entry follows receptor binding by the surface-exposed glycoprotein, which triggers the transmembrane glycoprotein (TM) to refold into a membrane fusion-active state. The trimer of hairpins is the best-characterized TM protein conformer and is composed of a central trimeric coiled coil associated with an outer antiparallel layer of 3 C-terminal segments. The trimer of hairpins represents the end point of a refolding pathway that apposes and catalyses fusion of the virus and cell membranes. The structure of pretriggered or 'prefusion' TM is unknown.

Cysteine replacement mutagenesis was applied to the human T cell leukemia virus type 1 TM, gp21, to ask if it resembles the prefusion form of its structural homologue, Ebola virus GP2, which is maintained as a trimer via a short coiled coil. Structural homology predicted that Ala-375 of gp21 is at the N-terminus of a corresponding short coiled coil and is oriented towards the 3-fold symmetry axis. The A375C substitution enabled interprotomer disulfide bonding, indicating that gp21 protomers are in close proximity in the Ala-375 region. I propose that prefusion gp21 resembles GP2 with a short coiled coil maintaining trimerization and functioning as a structural scaffold for the transition to fusion-active conformations.

The fusion activation of the human immunodeficiency virus type-1 TM, gp41, involves its transition to the trimer of hairpins via a prehairpin intermediate. Cysteine-replacement mutagenesis was used to examine the structures of prefusion gp41 and recombinant models of the prehairpin and trimer of hairpins. The T569C mutation, in the C-terminal segment of the predicted coiled coil enabled quantitative interprotomer disulfide bonding in prefusion gp41, whereas T538C in the N-terminal region did not. Thus the coiled coil appears to be also present in prefusion gp41. By contrast, S538C led to efficient interprotomer disulfide formation in the prehairpin and trimer of hairpins. Threonine-538 is within the N-terminal "polar segment" which mediates functionally

important hydrophobic interactions with the membrane-proximal ectodomain region (MPER) in the trimer of hairpins. A model for gp41 conformational activation is proposed, whereby a trimeric coiled coil in prefusion gp41 becomes extended into the polar segment in the prehairpin, providing a packing surface for C-terminal sequences to form a hairpin structure that extends to the terminal membrane-interactive sequences.

The disulfide-bonded region (DSR) of gp41 mediates association with the receptor-binding glycoprotein, gp120, and transmission of the activation signal from receptor-bound gp120. Forced evolution of the W596L.K601D DSR mutant that lacks gp120-gp41 association, was used to identify functional determinants that are linked to the DSR and therefore may be involved in the activation process. A D601H pseudoreversion in the DSR restored gp120-gp41 association but required an additional D674E mutation in the MPER for optimised replication competence. In an independent culture, D601H emerged together with D674N and deletion of Thr-389-Trp-390 in variable region 4 of gp120. Conservative substitutions at Asp-674 modulated virus entry in the context of W596L.K601H, indicating for the first time that the MPER is functionally linked to the association/activation synapse of gp120-gp41. My results reveal new information about the fusion-activation mechanism employed by retroviruses.

Declaration

I hereby declare that the experiments in this thesis constitute work carried out by the candidate except for Materials and Methods section 2.2.1.7 and Results section 4.3.3.2 in Chapter 4 which were performed in collaboration with Mr. Chan-Sien Lay. I also declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis. The content of this thesis is the result of work that was carried out since the 20th of January 2009, which was the official commencement date of the research program.

Ashraf Ibrahim Khasawneh

Acknowledgments

I would like to express my sincere respect and deepest gratitude to my principal supervisor, Dr Pantelis Poumbourios for giving me the opportunity to undertake my PhD project in the Fusion lab. His shrewd insights, and valued comments and suggestions, have been of such great benefit for improvement of my work, and in writing the thesis. I would also like to thank my co-supervisor, Heidi Drummer; your guidance support and advice has been invaluable.

A very warm thanks go to everyone in the Fusion Lab especially Annamarie Laumaea for her help with making the cysteine constructs and her help with the entry essays. Thank you also to Chan-Sien Lay for his help with making some of the MBP chimera and to Dr. David Harrison for his help and advice in the lab. I would like to thank Yousef Al-Hammad for helping in organising and putting the final touches on the thesis.

I would like also to acknowledge my sponsor the Hashemite University for giving me the opportunity to pursue my studies in such a wonderful country.

My family has always been an important source of encouragement and support especially my mother who showed extra care in the past 2 months when we were blessed with the birth of our baby girl, Farah. I'm indebted to my father who would have been really proud to see this moment. Last but by no means least it gives me immense pleasure to thank my wife, Wasan, for your love, understanding, and infinite support, and for the sacrifices which you made in order to make this happen. Without you this would not have been possible.

Abbreviations

Ab	Antibody
AIDS	Acquired immune deficiency syndrome
ARV	AIDS-associated retrovirus
ASLV	Avian sarcoma leukosis virus
BL21 (DE3)	Chemically competent E. coli cells
βME	β-mercaptoethanol
BMH	bis(maleimido)hexane cross linker
BMOE	bis(maleimido)ethane cross linker
C1-C5	Constant regions of HIV Env
C8	Mouse monoclonal antibody that binds to cytoplasmic tail of gp41
CA (p24)	Capsid
CCR5	Chemokine (CC motif) receptor 5
CD4	Cluster of differentiation 4
CMV	Cytomegalovirus
cpm	Count per minute
CT	Cytoplasmic tail of gp41
CX3CR1	Chemokine (C-X3-C motif) receptor 1
CXCR4	Chemokine (CXC motif) receptor 4
Cys	Cysteine
DE81	Diethylaminoethyl cellulose chromatography paper
DMEM	Dulbecco modified eagle medium
DMEM-10	DMEM with 10 % fetal calf serum
DMEM-15	DMEM with 15 % fetal calf serum
DNA	Deoxyribonucleic acid
dsDNA	Double-stranded DNA

DSR	Disulfide region (chain reversal region) of gp41
DTT	Dithiothreitol
DV012	Sheep polyclonal antibody raised against gp120
EBOV	Ebola virus
ECL	Extracellular loop
EM	Electron microscopy
Env	HIV-1 Envelope glycoprotein
ER	Endoplasmic Reticulum
ERVs	Endogenous retroviruses
ESCRT	Endosomal Sorting Complex Required for Transport
FP	Fusion peptide of gp41
FPLC	Fast protein liquid chromatography
GP	Surface glycoprotein of Ebola virus
gp120	Surface (SU) Envelope glycoprotein of HIV
gp21	Transmembrane (TM) envelope glycoprotein of HTLV
gp41	Transmembrane (TM) envelope glycoprotein of HIV
Gpr1	G protein-coupled receptor 1
h	Hour
HA	Influenza hemagglutinin
HI-FCS	Heat inactivated foetal calf serum
HIV-1	Human immunodeficiency virus type 1
HFV	Human foamy virus
HR1	N-terminal heptad repeat region of HIV-1 gp41
HR2	C-terminal heptad repeat region of HIV-1 gp41
HTLV	Human T-lymphotropic virus type I
IAPE	Intra-cisternal A-type particles with an envelope

IFL	Internal fusion loop
IN (p32)	Integrase
IPTG	isopropyl-β-D-thiogalactopyranoside
IRES	Internal ribosome entry site
kb	Kilobases
kDa	Kilodaltons
LAV	lymphadenopathy associated virus
LB	Lauria Bertani medium
LINEs	Long interspersed nuclear elements
LTR	Long terminal repeat
MA (p17)	Matrix
mAb	Monoclonal antibody
MBP	Maltose binding protein
MES	2-(N-morpholino)ethanesulfonic acid
MHC-I	Major histocompatibility complex type 1
Min	Minute
ml	Millilitre
µl	Microlitre
MLV	Murine leukemia virus
MMTV	Mouse mammary tumor virus
MPER	Membrane proximal region of gp41
mRNA	Messenger RNA
NC (p9 and p6)	Nucleocapsid
Nef	Negative factor
OD_{600}	Optical density of a sample measured at a wavelength of 600 nm
ORF	Open reading frame

PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PR (p10)	Protease
R5	CCR5-using strain of HIV
R5X4	CCR5 and CXCR4-using strain of HIV
Rev	Regulator of expression of viral proteins in HIV
RLU	Relative light unit
RNA	Ribonucleic acid
RRE	Rev responsive element
RSV	Rous sarcoma virus
RT (p66/51)	Reverse transcriptase
sCD4	Soluble CD4
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
sec	Seconds
SINEs	Short interspersed nuclear elements
SIV	Simian immunodeficiency virus
ssRNA	Single stranded RNA
T20	Enfuvirtide (fusion inhibitor)
Tat	Trans-activator of transcription
TB	Terrific broth medium
TEMED	N,N,N',N'-tetra-methyl ethylenediamine
Tm	Melting temperature of a primer
TMD	Trans membrane domain
U3	Untranslated 3’ region of LTR
U5	Untranslated 5’ region of LTR
UNAIDS	Joint United Nations programme on HIV-1/AIDS

V1/V2	First and second variable loop of HIV-1 gp120
Vif	Virion infectivity factor
v/v	Volume to volume
Vpr	Viral protein R
Vpu	Viral protein U
VSV G	Vesicular stomatitis virus glycoprotein G
WHO	World health organisation
WT	Wild type
w/v	Weight to volume
X4	CXCR4-using strain of HIV
2F5	Broadly neutralizing antibody, target MPER of gp41
3D	Three dimensional
4E10	Broadly neutralizing antibody, target MPER of gp41
6-HB	Six-helix bundle

Single letter code for amino acids

Amino acid	Abbreviation	Single letter code
Alanine	Ala	A
Arginine	Arg	R
Aspargine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Chapter 1: General Introduction

1.1 Introduction

Retroviruses were discovered in the early 1900's with the isolation of avian sarcoma/leukosis viruses (ASLV) from a chicken tumor (Ellerman and Bang, 1908; Rous, 1911). Twenty five years later John Bittner linked mammary carcinoma in mice to a filterable agent which was subsequently isolated by Ledwick Gross as the first murine leukemia virus (MLV) (Gross, 1951; Bittner, 1936). The isolation of retroviruses from other mammalian species, including cats, sheep and monkeys, in the mid twentieth century led to the search for human cancer viruses. In 1977, an epidemiological study showed a clustering of adult T-cell leukemia in specific parts of Japan (Takatsuki *et al.*, 1978; Uchiyama *et al.*, 1977), which was indicative of a transmissible agent. Three years later, Poiesz *et al* isolated the first human retrovirus, human T-cell leukemia virus 1 (HTLV-1) as the causative agent (Poiesz *et al.*, 1980). At around the same time, an epidemic of AIDS arose in several developed countries with human immunodeficiency virus (HIV) being isolated as the causative agent (Barre-Sinoussi *et al.*, 1983; Gallo *et al.*, 1984).

1.1.1 The Retroviridae: Evolution and Classification

Retroviruses are 80-100 nm-diameter enveloped RNA viruses. Their outer lipid envelope, derived from host cell membrane, displays the viral glycoproteins encoded by the *env* gene. The virion RNA is 7–12 kb in size, and it is linear, single-stranded, nonsegmented, and of positive polarity. Two copies of genomic RNA are packaged into each retroviral particle. Retroviruses have a unique distinguishing feature, which is the process of reverse transcription. The retrovirus genome exists in two major forms: single stranded RNA in the virion; and double stranded DNA integrated into the host cell genome. The conversion of viral genomic RNA to DNA is mediated by the virally

encoded enzyme reverse transcriptase (RT), whereas the transcription of integrated DNA to RNA is carried out by the cellular transcription machinery.

Retroviral classification schemes were initially based on virion morphology. In this system, B type particles were characterized by an eccentric core and prominent glycoprotein spikes, C type particles by a centric core, D type particles by a cone shaped core, and lentiviruses by a truncated cone shaped core (Coffin, 1992). With the availability of nucleotide sequence information, classification schemes based on molecular phylogeny could be produced. A scheme based on amino acid similarities of the RT (Llorens *et al.*, 2010) is shown in Figure 1.1.

The genomic organisation of retroviruses is characterised by 5' and 3' long terminal repeat (LTR) sequences, which flank the coding region comprised of the *gag*, *pol* and *env* genes. *gag* encodes the core structural proteins (matrix, capsid, nucleocapsid), *pol* encodes the viral replicative enzymes (RT/RNaseH, protease, integrase), while *env* encodes the envelope glycoproteins (Env). This basic genome organisation has undergone considerable diversification during the evolution of the Retroviridae.

Retroelements represent the most basic examples of the Retroviridae. Retroelements are mobile genetic elements that have the capacity to "jump" from one chromosomal location to another via an RNA intermediate that is reverse-transcribed to DNA by its encoded RT, and then integrated into a new location within the host genome by the integrase enzyme. The location of the insertion in the genome determines whether or not the transposition results in genetic alteration in the cell. Retroelements constitute a significant part of the genome, 60-70% in plants and ~ 50% in mammals (Sanmiguel and Bennetzen, 1998; Li *et al.*, 2004). The human genome content of retroelements reaches around 42% (Lander *et al.*, 2001). Retroelements are divided into two groups: the LTR and non-LTR elements. The non-LTR group includes the long interspersed repetitive elements

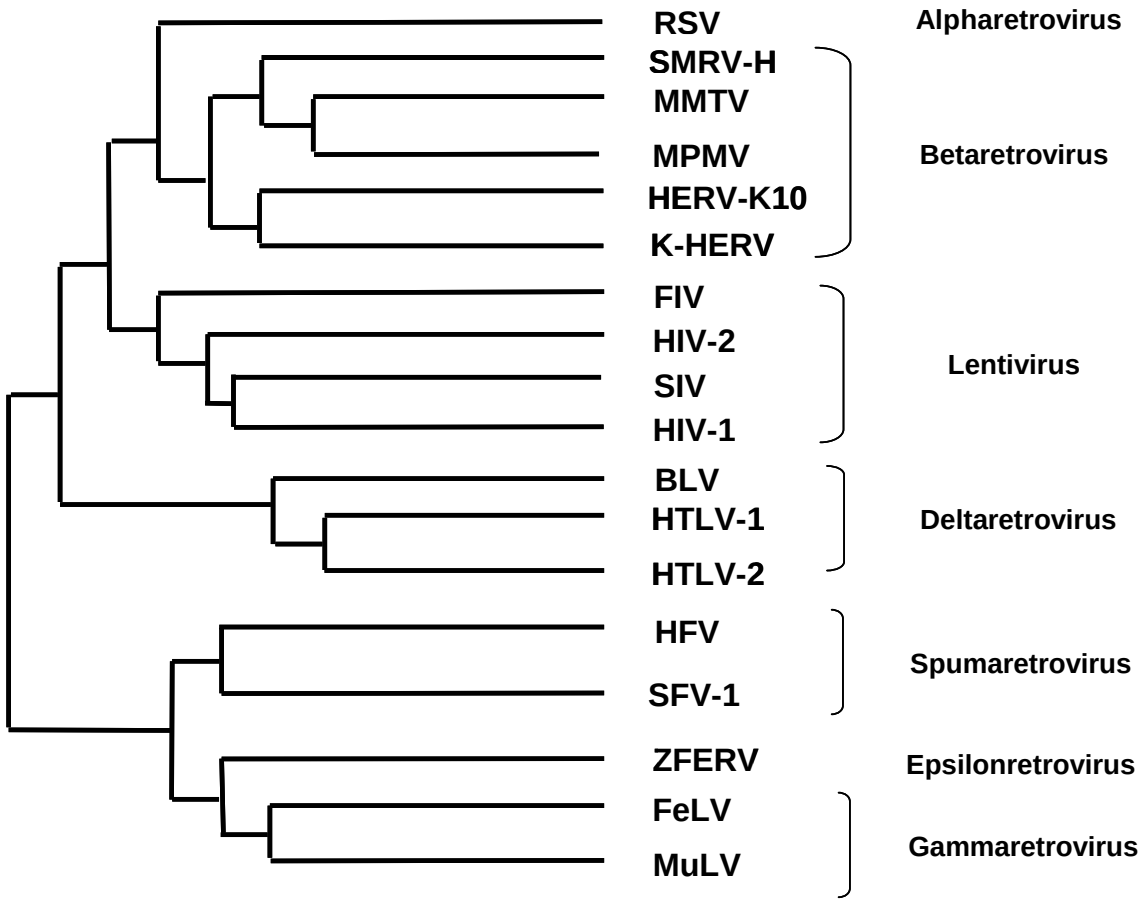


Figure 1.1: Classification of the retroviruses

This classification was based on RT amino acid sequence similarity. Abbreviations are: RSV, Rous sarcoma virus; SMRV, squirrel monkey retrovirus; MMTV, mouse mammary tumor virus; MPMV, Mason-Pfizer monkey virus; HERV, human endogenous retrovirus; FIV, feline immunodeficiency virus; HIV, human immunodeficiency virus; SIV, simian immunodeficiency virus; BLV, bovine leukemia virus; HTLV, human T-cell leukemia virus; HFV, human foamy virus; SFV, simian foamy virus; ZFERV, zebrafish endogenous retrovirus; FeLV, feline leukemia virus; MuLV, murine leukemia virus. This figure was reproduced with modifications from Gypsy Database (GyDB).

(**LINEs**, e.g. the Alu and MIR repeats), and the short interspersed repetitive elements (**SINEs;** e.g. the L1 and L2 sequences) (Figure 1.2) (Medstrand *et al.*, 2002). SINEs do not have protein coding capacity and depend on LINEs for their amplification. The LTR class include the retrotransposons and endogenous retroviruses (ERVs) (Figure 1.2). Because the LTR elements resemble the terminal regions of infectious retroviral genomes, Temin proposed that infectious retroviruses evolved from retroelements (Temin, 1980). It is now assumed that ERVs originated from a retroviral infection that led to integration of the viral genome into the germ line of the host and subsequent transmission of proviral elements to following generations. ERVs are generally degenerate and dormant with disrupted ORFs, except for a few examples, which retain their genes and express the corresponding proteins as intra-cisternal A-type particles with an envelope (IAPE) (Ribet *et al.*, 2008; Ribet *et al.*, 2004; Dewannieux *et al.*, 2004). MLVs include both endogenous and exogenous (infectious) viruses and have evolved into 4 classes: the ecotropic MLVs, which are capable of infecting murine cells in culture, polytropic and modified polytropic MLVs which infect murine and nonmurine cells, and xenotropic MLVs, which infect nonmurine cells only (Stoye and Coffin, 1987).

The human immunodeficiency viruses are members of the *Lentivirinae* genus. All known lentiviruses are exogenous, infect a wide variety of vertebrate hosts (primates, horses, cattle, goat, sheep and cats) and characteristically cause overt symptomatic disease after long incubation periods. In addition to the canonical 5'-LTR-*gag-pol-env*-3'-LTR genomic organisation, lentiviruses also comprise regulatory and accessory genes. Such genes can be involved in the synthesis and processing of viral RNA and in the modulation of host cell processes that favour the viral replicative cycle.

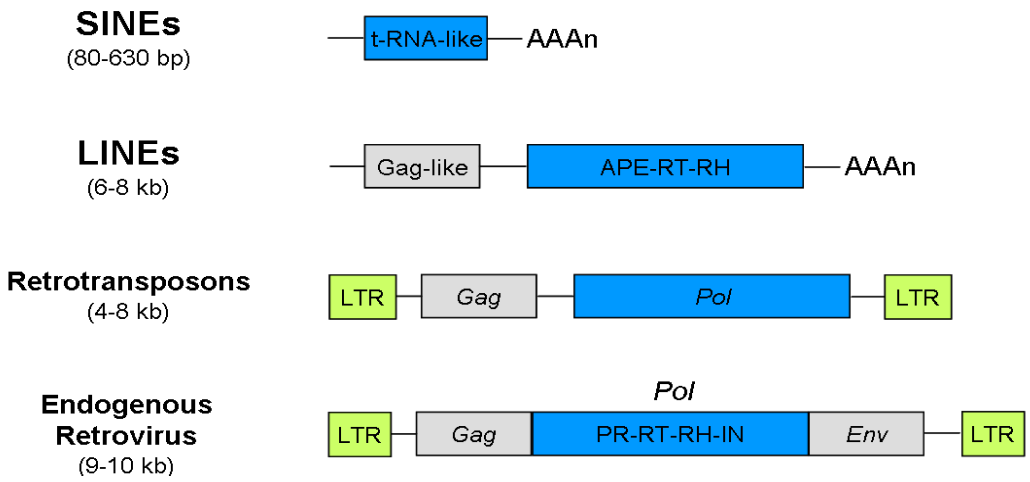


Figure 1.2 Structural features of important retroelements

The SINEs and the LINEs represent the non-LTR elements whereas the LTR class include retrotransposons and endogenous retroviruses. Abbreviations are: APE, apurinic-apyrimidinic endonuclease; RT, reverse transcriptase; RH, RNase H; IN, integrase. The usual length of the retroelements is shown in parenthesis. Modified from Bannert and Kurth (2004).

1.1.2 The Retroviridae: Genome Organisation

All Retroviridae share a similar genomic organization with a number of functionally homologous open reading frames. The viral genes are flanked by LTR sequences at both the 5' and 3' ends of the genome. The LTRs are involved in several steps in the retrovirus life cycle including reverse transcription, integration of the viral genome into the host cell genome, viral gene expression and generation of viral genomes for virus assembly (Garcia *et al.*, 1987; Coffin, 1990; Bushman, 2003). Retroviral LTRs are comprised of unique 3' (U3), flanking R and unique 5' (U5) regions and they are positioned at either end of the proviral DNA in the reverse transcription phase (Starcich *et al.*, 1985; Coffin, 1990; Bushman, 2003). The ends of the LTRs subsequently participate in the integration of the provirus into the host genome (Bushman, 2003). After virus integration, the 5' end of the LTR serves as a promoter for the entire retroviral genome, while the 3' end provides for viral RNA polyadenylation.

The coding regions of the retroviral genome include *gag, pol* and *env* genes, which encode the structural and enzymatic polyprotein precursors. The *gag* gene encodes the capsid (CA), matrix (MA) and nucleocapsid (NC) structural proteins, which form the virion core structure, while the *pol* gene encodes the viral enzymes: RT/RNaseH, protease (PR) and integrase (IN), the latter of which is responsible for proviral integration into the host genome. The *env* gene encodes the envelope glycoprotein precursor, which will be discussed in greater detail below.

With the exception of Gammaretroviruses such as MLV, all other retroviruses have additional genes that encode regulatory and/or accessory proteins. For example, HIV-1 and HTLV-1 encode regulatory proteins that modulate transcription (HIV-1 *tat*, HTLV-1 *tax*), and efficient cytoplasmic expression of unspliced and incompletely spliced viral RNA

transcripts encoding the viral structural and enzymatic proteins (HIV-1 *rev*, HTLV-1 *rex*). In addition to these regulatory genes, HIV-1 encodes 4 accessory genes: *vif, vpu, vpr and nef*. The accessory proteins translated from these genes are not absolutely required for infectivity in cell culture but play significant roles in HIV-1 infection *in vivo*. For example, host antiviral defense mechanisms such as the APOBEC 3G and 3F cytidine deaminases, which promote A-to-G hypermutation of the viral genome during reverse transcription, and tetherin, which blocks HIV-1 egress from infected cells, are counteracted by Vif and Vpu, respectively. Nef plays a role in immune evasion by mediating CD4 and MHC class-I down-regulation, while Vpr is involved in G2/M cell cycle arrest and nuclear targeting of the viral preintegration complex (Re and Luban, 1997; Emerman and Malim, 1998; Hout *et al.*, 2004; Levesque *et al.*, 2004; Le Rouzic and Benichou, 2005). Another example of a retroviral accessory protein is v-src (viral-sarcoma) of the alpharetrovirus Rous sarcoma virus (RSV), which is essential for sarcoma formation in chickens. v-src is a constitutively active homologue of the c-src tyrosine kinase proto-oncogene, which regulates cellular proliferation and movement. v-src is believed to have been acquired by the virus in the course of its interaction with the host genome (Stehelin *et al.*, 1977; Stehelin *et al.*, 1976). This discovery indicated that proto-oncogenes in the normal cellular genome could have potent transforming ability when inappropriately activated or over-expressed. This was an important landmark, which transformed the understanding of cancer.

1.2 Human Immunodeficiency Virus

Human immunodeficiency virus was discovered in 1983 by Luc Montagnier's group (Barre-Sinoussi *et al.*, 1983), and was later verified by the groups of Robert Gallo and Jay Levy (Gallo *et al.*, 1984; Popovic *et al.*, 1984; Levy *et al.*, 1984). It was initially given three different names (lymphadenopathy associated virus (LAV), human T-

lymphtropic virus type III (HTLV-III), AIDS-associated retrovirus (ARV)), but was renamed human immunodeficiency virus (HIV) in 1986 (Coffin *et al.*, 1986). It is the causative agent of acquired immunodeficiency syndrome (AIDS), and comprises two distinct species: HIV-1 and HIV-2. HIV-1 is considerably more pathogenic than HIV-2. Although the predominant mode of transmission for both HIV-1 and HIV-2 is through sexual contact, the higher viral load of HIV-1 contributes to a greater frequency of transmission and, as such, contributes to the majority of global infections (Reeves and Doms, 2002). HIV-1 strains are divided into three major groups, M, O, and N with group M being further divided into 9 subtypes (Hu *et al.*, 1996). Subtype B is the most prevalent subtype in western countries, while subtype C is predominant in Asia and sub Saharan Africa.

1.2.1 Epidemiology of HIV-1

The World Health Organization (WHO) estimates 33 million people worldwide are infected with HIV with slightly more than half being female; children less than 15 years of age represent 2 million cases (UNAIDS/WHO, 2010). Over the past few years the rate of new infections has declined due to the expansion of prevention programs, which focused on condom use, raising awareness about the risks of early sexual debut and multiple partners especially in developing countries. The introduction and expansion in the use of antiretroviral drugs has led to a drop in the rate of AIDS-related deaths from 2.1 million in 2004 to 1.8 million in 2009, and increased survival with a better quality life (UNAIDS/WHO, 2010). These figures highlight the importance of putting forward new strategies to prevent the spread of HIV-1.

There is great geographical variation in the distribution of HIV/AIDS (Figure 1.3). The highest prevalence rate (11 %) is seen in Sub-Saharan Africa with 22.5 million

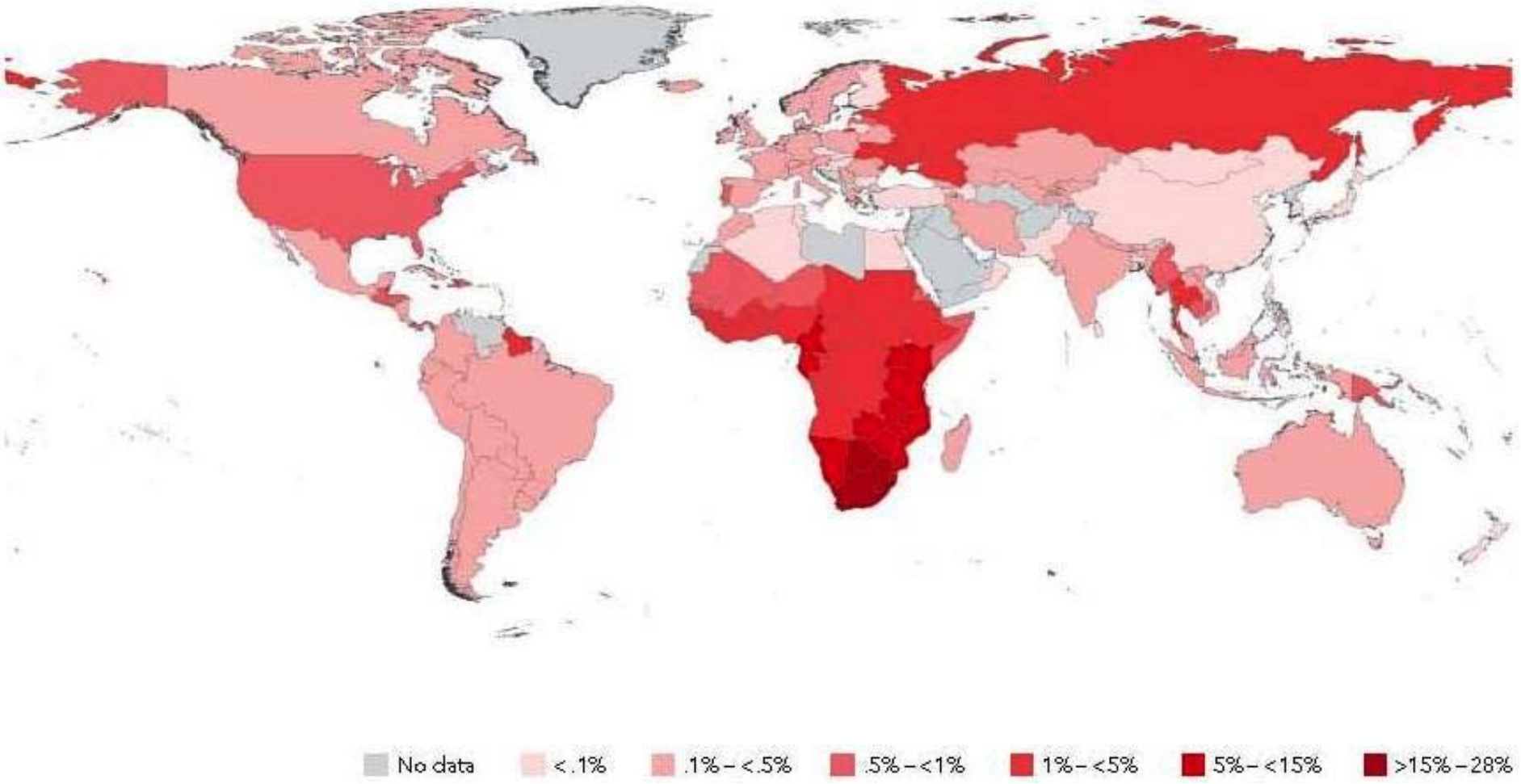


Figure 1.3 Global distribution of HIV-1 in 2009

Countries are shaded according to their HIV-1 prevalence as indicated in the bottom of the figure. Sub-Saharan Africa has the highest prevalence rates in the world. At the end of 2009, there were an estimated 33.3 million people living with HIV/ AIDS globally. Adapted from UNAIDS/WHO (2010).

people living with HIV-1. Poor resources, limited access to treatment and the absence of public education in developing countries represent a major challenge to the control and prevention of the disease. In Australia there are 17,000 people living with HIV-1, representing less than 0.1% of the population (UNAIDS/WHO, 2010). In Australia, the epidemic is mainly confined to the homosexual community.

1.2.2 HIV-1 virion structure

The virion is 80-120 nm in diameter, with approximately 14 envelope spikes embedded in an outer lipid bilayer envelope (Zanetti *et al.*, 2006; Zhu *et al.*, 2006). It contains two copies of positive sense genomic RNA strands, matrix, capsid, nucleocapsid proteins, and the virally encoded enzymes (Figure 1.4). The envelope spikes comprise the viral envelope glycoproteins (Env), gp120 and gp41, which exist as a trimer of heterodimers that are embedded in the lipid membrane. The transmembrane subunit, gp41, anchors the Env trimer into the viral plasma membrane and is non-covalently associated with the surface subunit, gp120. The Env trimer determines viral tropism and initiates viral entry into the cell via gp120-receptor interactions, while gp41, as a class I fusion glycoprotein, facilitates the fusion of the viral and target cell membranes. The MA protein (p17) lies immediately beneath the lipid bilayer, while the capsid protein (p24) forms the cone-shape structure seen under electron microscopy that contains the nucleocapsid comprised of NC (p7), two strands of genomic RNA, RT, and IN.

The RT enzyme lacks proof reading activity and therefore plays a major role in viral diversification through error-prone reverse transcription and retroviral recombination. It also contains RNase H activity, allowing it to degrade RNA from RNA:DNA hybrids. The NC protein surrounds the viral genomic RNA and plays an important role in virus packaging. Also packaged into the virion are the viral accessory proteins, Vif (Virion Infectivity Factor), Vpr (Viral Protein R), and Nef (Negative Regulatory Factor).

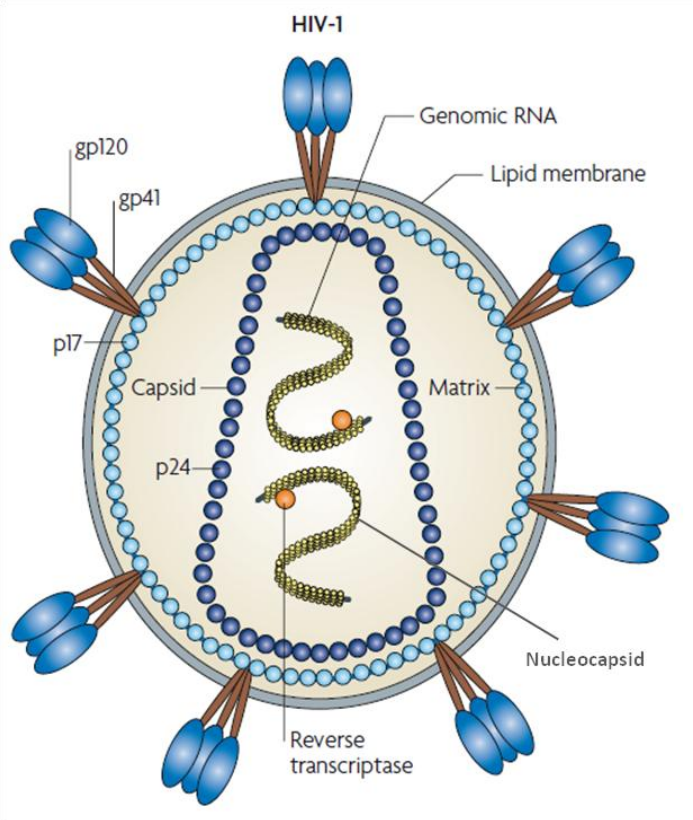


Figure 1.4 Schematic of the structure and composition of the mature HIV-1 virion

Adapted from Karlsson Hedestam *et al* (2008).

1.2.3 HIV-1 replication cycle

The primary cellular targets of HIV-1 are the cluster of differentiation 4 (CD4)-expressing cells of the immune system: T-cells, monocytes, macrophages and dendritic cells. HIV-1 replication, illustrated in figure 1.5, begins with the binding of the virion surface protein, gp120, to the host cell, followed by gp41-mediated virus-cell membrane fusion and entry of the viral capsid into the cytoplasm. A conserved region in gp120 binds to CD4 on the target cell inducing conformational changes in gp120 that expose a binding site for an additional receptor, i.e. the chemokine receptors (Kwong *et al.*, 1998). The G-protein coupled receptors, CXCR4 and CCR5, are the main co-receptors that mediate HIV-1 entry *in vivo*, but other receptors such as CCR3, apj, CX3CR1, Gpr1, Gpr15 and STRL33 can be used by HIV-1 as a coreceptor *in vitro* (Choe *et al.*, 1996; Choe *et al.*, 1998; Edinger *et al.*, 1998; Combadiere *et al.*, 1998; Alkhatib, 2009; Ray and Doms, 2006). The binding of gp120 to CD4 and co-receptor induces further conformational changes that result in the refolding of gp41 (Furuta *et al.*, 1998; Koshiba and Chan, 2003; Si *et al.*, 2004), which exposes the fusion peptide and leads to its insertion into the target cell membrane and eventual fusion with the virion envelope (Eckert and Kim, 2001; Colman and Lawrence, 2003). Release of the capsid into the cytoplasm and its uncoating are the following steps in the HIV-1 life cycle. The mechanism of envelope mediated HIV-1 entry is discussed in more details in section 1.3.5.

Retroviruses have acquired their name from the reverse transcription step where linear single stranded RNA (ssRNA) genome is converted into double stranded DNA (dsDNA). This conversion takes place in the cytoplasm of the newly infected host-cell by the RT enzyme in a process called reverse transcription. The double stranded DNA intermediate migrates to the nucleus in a process called nuclear translocation where the viral IN catalyses the integration of the viral genome into the host's genome, forming the

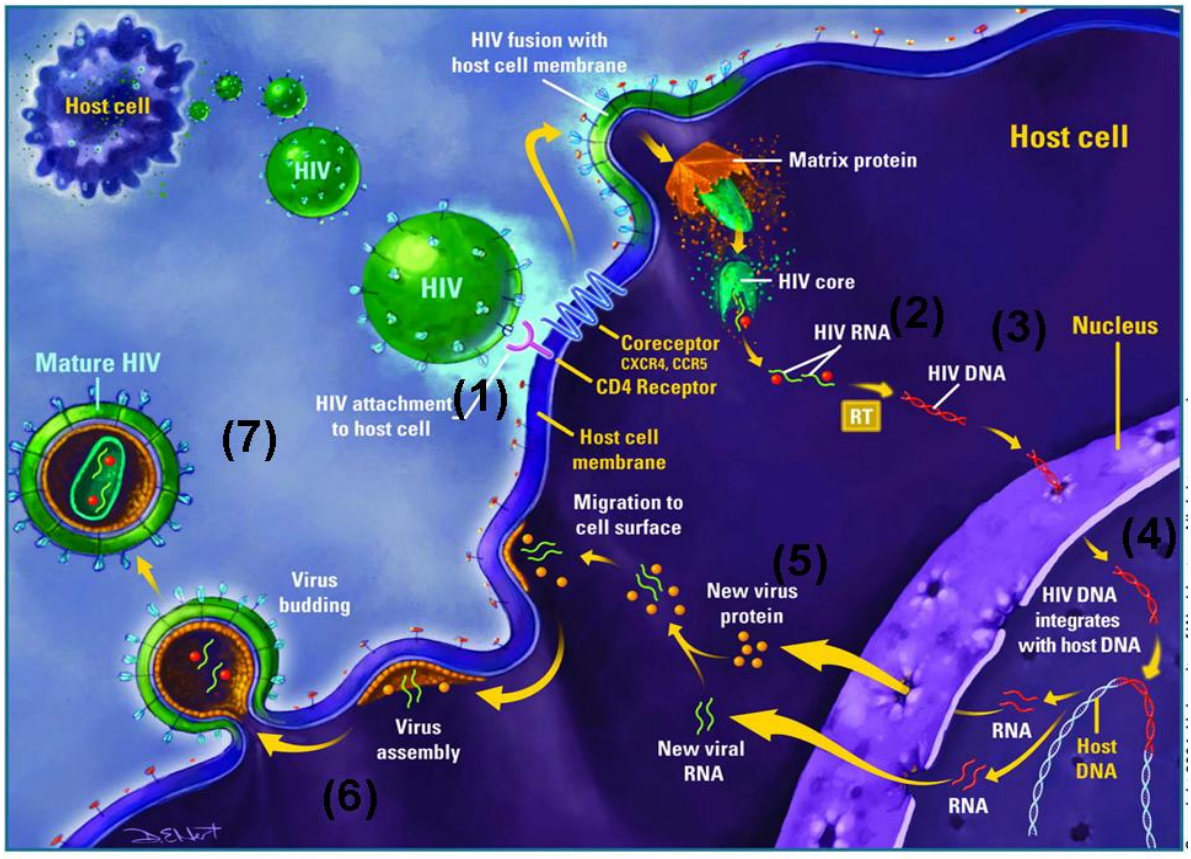


Figure 1.5 HIV-1 replication cycle

The virus particles bind to the cell-surface receptor (CD4) and coreceptors (CXCR4 and CCR5) **(1)** and enter the cell following viral and cellular membrane fusion releasing the viral RNA **(2)**. The RNA is reverse transcribed into double stranded DNA **(3)**, which is translocated into the nucleus and integrates into the host DNA **(4)** via the HIV-1 integrase. The proviral DNA is transcribed by host cell and viral factors and then exported into the cytoplasm for translation into new viral protein **(5)** which together with new viral RNA migrate to the cell surface for assembly **(6)** and shed from the cell as progeny HIV-1 virions **(7)**. Adapted from nwabr.org, URL:http://www.nwabr.org/education/pdfs/hiv_lifecycle.jpg.

provirus. After integration the viral genome is transcribed into mRNA from the promoter sequences in the LTR under the control of HIV-1 *tat*. The full-length and partially spliced RNA transcripts are transported to the cytoplasm in a *rev* controlled process via the *rev*-responsive element (RRE), an assembly of stem-loop structures present in the gp41 coding region (Felber *et al.*, 1989). Translation of the *gagpol* region gives rise predominantly to the Gag precursor protein. However, a rare frameshifting event, due to the presence of a stem-loop structure on the 3' side of NC, enables read-through of the *pol* region to yield GagPol precursor. The ratio of Gag to GagPol is approximately 20:1. The Gag and GagPol precursors assemble into arrays at the cell membrane predominantly via an N-terminal myristoyl group and a basic domain within MA. The p6 domain at the C-terminus of the Gag precursor recruits the host cell Endosomal Sorting Complex Required for Transport (ESCRT) machinery that facilitates the severing of nascent immature virions from the cell membrane. Activation of the viral PR leads to the cleavage of the Gag precursor to the virion structural proteins MA, CA, NC, the spacer peptides SP1 and SP2 and the C-terminal p6 protein (Dorfman *et al.*, 1994). GagPol is cleaved to the viral replicative enzymes, RT, RNaseH, PR and IN. These cleavage events enable virions to mature into infectious particles with the characteristic cone-shaped core.

1.3 The HIV-1 envelope glycoproteins

The HIV-1 Env glycoprotein is synthesized in the rough endoplasmic reticulum (ER) to generate the Env precursor protein, gp160 (Freed and Martin, 1995). In the ER, gp160 acquires high-mannose oligosaccharides, forms intramolecular disulfide bonds and undergoes oligomerization (Earl *et al.*, 1991). gp160 is then transported to the *Golgi* where it is cleaved by the cellular enzyme furin to generate the mature surface unit (SU) glycoprotein, gp120, and a transmembrane unit (TM) glycoprotein, gp41.

After gp160 cleavage, the gp41 anchors the Env complex in the membrane while gp120 remains associated with it through a non-covalent interaction. The Env glycoprotein complexes are then transported to the cell surface where they are either incorporated into a budding virus particle or internalized (Freed, 2001). The relatively weak gp120-gp41 interaction results in significant amounts of gp120 being spontaneously shed from the surface of Env-expressing cells and virions in a process thought to play a role in viral evasion from the host immune responses (Schneider *et al.*, 1986; Moore *et al.*, 2006). Biophysical and cryoelectron tomography approaches have revealed that the mature Env complex comprises a trimer of gp120 molecules in association with a trimer of gp41 molecules (Center *et al.*, 2002; Liu *et al.*, 2008).

1.3.1 Structure of the gp120 surface glycoprotein

The surface glycoprotein gp120 comprises 5 conserved regions (C1-C5) that alternate with 5 hypervariable regions (V1-V5) (Kuiken *et al.*, 2000). The conserved regions (C1-C5) are heavily glycosylated and form core structures, which are important for interaction with receptors and gp41 association (Kwong *et al.*, 2000; Wyatt *et al.*, 1998). The sequence variability and extensive glycosylation of the variable regions (V1-V5), which are exposed on the surface of the protein, act as a screen against the host immune response (Kwong *et al.*, 1998; Wyatt *et al.*, 1998). Kwong *et al.* was the first to solve the crystal structure of gp120 bound to domains 1 and 2 of CD4 and a Fab fragment derived from monoclonal antibody (mAb) 17b, which is directed to a portion of the chemokine receptor binding site (Kwong *et al.*, 2000; Kwong *et al.*, 1998). It consisted mainly of the gp120 core with the highly flexible V1/V2 and V3 variable loops deleted. The gp120 core structure can be divided into a conserved inner domain, a less conserved outer domain and the bridging sheet, which comprises 4 antiparallel β-strands (Figure 1.6A). The outer domain is exposed on the surface of the Env trimer, forming a heavily glycosylated surface

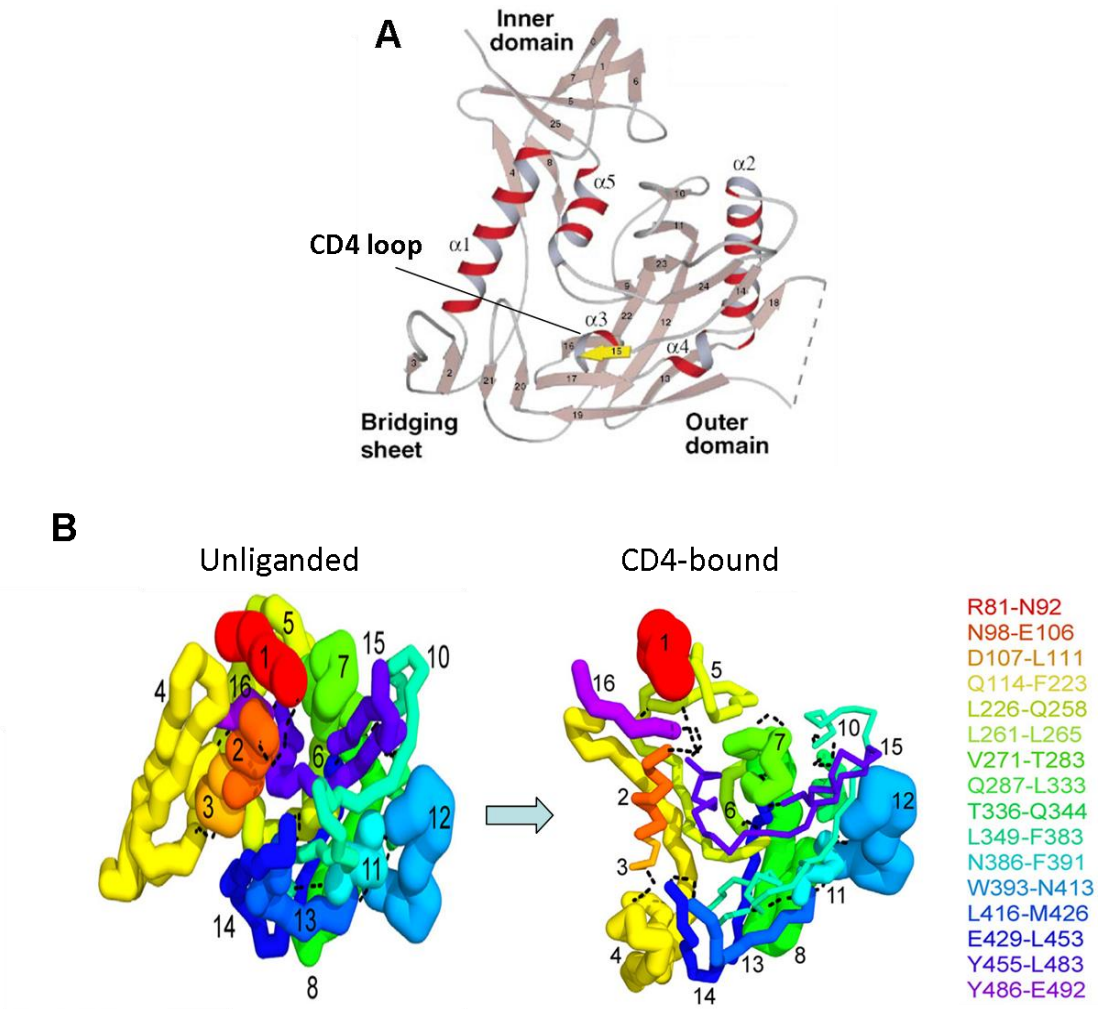


Figure 1.6 HIV-1 gp120 structure

(A) Ribbon diagram of HXBc2 gp120, core is comprised of helices in red and β-strands in salmon. Three structures constitute the gp120 core; the inner domain, bridging sheet and outer domain. The CD4 loop is also shown. Adapted from Kwong *et al.* (2000). **(B)** Homology model representing unliganded gp120 and the CD4-bound crystal structures of the HIV-1_{YU2} core gp120. Fragments are colored and numbered according to their position in sequence. In the unliganded gp120 structure, the V1/V2 loop stem (fragment 4) from the inner domain and fragment 14 from the outer domain are far apart from each other but upon CD4 binding the two fragments move closer to each other to form the bridging sheet. Adapted from Kong *et al.* (2010).

which plays an important role in evading cellular and humoral immunity (Kwong *et al.*, 2002). The inner domain faces the trimer axis and is likely to interact with gp41 (Liu *et al.*, 2008). From distal strands of the bridging sheet, the V1/V2 loops arise and are proposed to function as a conformational mask of the CD4 and co-receptor binding sites (Kwong *et al.*, 2000; Kwong *et al.*, 1998). The anti-parallel bridging sheet is highly conserved and plays an important role in co-receptor binding in the presence of CD4 (Kwong *et al.*, 2000; Trkola *et al.*, 1996). The crystal structure of CD4-bound gp120 containing the V3 loop in complex with an Fab fragment derived from a sulfated antibody mimic of the N-terminus of CCR5 was solved by Huang *et al.* (Huang *et al.*, 2007; Huang *et al.*, 2005). This structure showed that the V3 loop extends from the gp120 core by ~7 Å towards the target cell membrane, facilitating an interaction with the co-receptor.

More recent studies comparing the un-liganded gp120 (Chen *et al.*, 2005; Kong *et al.*, 2010) structure to the CD4-bound structure have shown extensive conformational changes upon CD4 binding, especially in the inner domain. The most significant conformational change is the formation of the bridging sheet. In the unliganded gp120 structure, the V1/V2 loop stem ("fragment 4" in Figure 1.6B) from the inner domain and "fragment" 14 from the outer domain are not adjacent to one another. Upon CD4 binding the two fragments become juxtaposed, forming the bridging sheet (Kong *et al.*, 2010; Chen *et al.*, 2005). The V1/V2 loop is believed to shield the V3 loop and the CD4 and co-receptor binding sites in the unliganded state, while in the CD4-bound state the V3 loop is pointed towards the cell membrane. The V3 loop mediates important contacts with the co-receptor (Huang *et al.*, 2007; Huang *et al.*, 2005), and determines the co-receptor preference of HIV-1 strains, which largely determines its tropism. Therefore the V3 loop represents an important target for neutralizing antibodies (Pantophlet and Burton, 2006).

In 2010, Pancera *et al* revealed the structure of CD4-bound gp120 with intact gp41 interactive regions where more than 70 amino acids were added to the previously determined core structure (Chen *et al.*, 2005; Chen *et al.*, 2009a; Zhou *et al.*, 2007; Pancera *et al.*, 2010). The new gp120 structure shows the gp41-interactive N- and C- termini of the mature protein clustered together and forming an ~ 35-Å projection at the envelope-proximal base of the molecule. The gp41-interactive termini are connected to the inner domain via a 7-stranded β-sandwich, which links together 3 structurally plastic layers that form excursions from the β-sandwich itself (layer 1), V1V2 (layer 2) and the outer domain (layer 3) (Figure 1.7A). CD4 engagement by residues in the outer domain and bridging sheet is believed to cause remodelling of the layers and transmission of this structural change to the gp41 association site via the β-sandwich, thereby releasing gp41 from the metastable state (Finzi *et al.*, 2010) (Figure 1.7B).

1.3.2 gp120 determinants of cellular tropism and disease progression

The main targets for HIV-1 *in vivo* are CD4-positive T-cells, macrophages and dendritic cells. Tropism within these cellular subsets is strongly determined by the co-receptor preference of HIV-1 strains. HIV-1 strains that use CXCR4 as a co-receptor are called X4 viruses, while those that use CCR5 are called R5 viruses (Alkhatib *et al.*, 1996; Dragic *et al.*, 1996; Feng *et al.*, 1996). Some strains are capable of infecting cells bearing either co-receptor and are known as R5X4 viruses (Doranz *et al.*, 1996). X4 strains can infect CXCR4+/CD4+ naïve T-lymphocytes and CCR5+/CXCR4+/CD4+ memory T-lymphocytes (Bjorndal *et al.*, 1997). On the other hand R5 strains preferentially infect CCR5+/CXCR4+/CD4+ memory T-lymphocytes, monocytes, macrophages, dendritic cells and microglia in the CNS (Alkhatib *et al.*, 1996; Deng *et al.*, 1996; Gartner *et al.*, 1986). While co-receptor usage strongly predicts HIV-1 tropism, it is not the case at all times. Some recent studies have shown that primary X4 viruses demonstrate macrophage

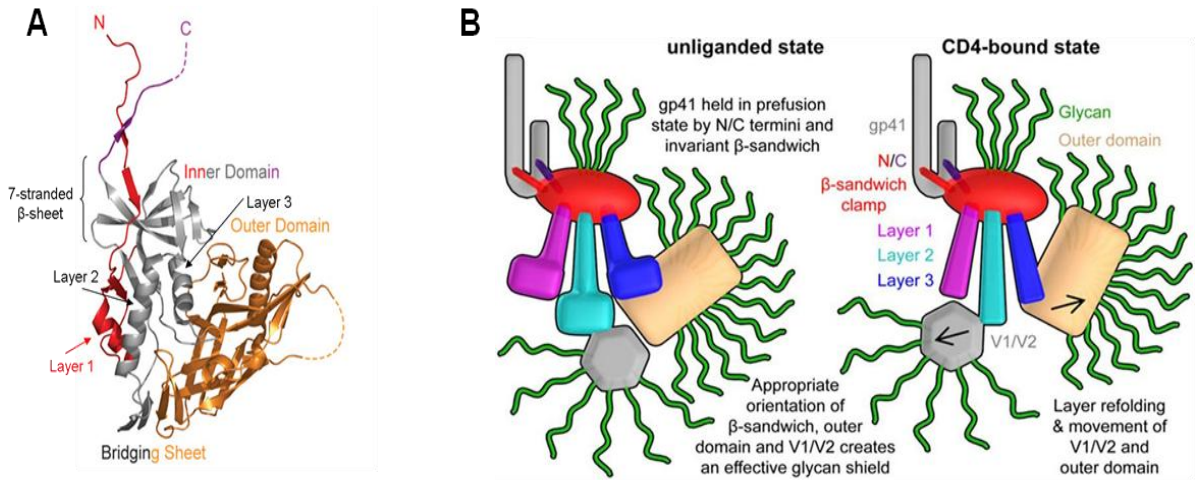


Figure 1.7 Structure of an HIV-1 gp120 core with intact gp41-interactive region

(A) Ribbon diagram showing HXBc2 gp120 core with intact N and C-termini, the invariant β-sandwich and the structurally plastic layers (1-3). **(B)** Model showing the role of gp120 layers in mobility and immune evasion. The β-sandwich holds gp41 in a metastable state via the N and C termini, while the layers are flexible enough to reposition the three highly glycosylated components: β-sandwich, V1/ V2 loops, and outer domain in response to CD4 binding. In the unliganded state (Left), the glycosylated structures arrange to form a protective shield against most antibodies. In the CD4-bound state (Right), the layers organize with outer domain to form the high-affinity binding sites for CD4 and coreceptor. Adapted from Pancera *et al.* (2010).

tropism, while some R5 primary viruses were incapable of infecting macrophages (Li *et al.*, 1999; Gorry *et al.*, 2001). The V3 loop in gp120 plays a major role in determining the co-receptor preference, where the absence of basic amino acids at position 11 and 25 favour CCR5 usage while their presence switch the preference to CXCR4 (Milich *et al.*, 1997; Hoffman *et al.*, 2002; Fouchier *et al.*, 1995). Other sites in gp120 such as V1/V2, V4 and V5 can also contribute to efficient utilization of a particular co-receptor (Cho *et al.*, 1998). The cytoplasmic tail of g41 has been shown to have an indirect effect on determining co-receptor usage through inducing allosteric changes in the gp120 (Taylor *et al.*, 2008).

The clinical course of HIV-1 infection can be divided into three stages. The acute phase is characterised by a peak in viral load, subsequent drop in the CD4+ cell number and the production of HIV-1 specific antibodies. The predominance of the R5 strains in this phase could be explained by the findings of Zhou *et al.* where they have shown that the exposure of resting CD4+ T-cells to X4 HIV-1 strains led to rapid cytolysis in cells that required virus entry, while this was not the case in activated CD4+ T-cells which could be seen late in the clinical course (Zhou *et al.*, 2008). This means that X4 strains are cytotoxic at this initial stage of infection. Neutralizing antibodies and cytotoxic T cells control virus replication to a great extent, and lead to a rebound in the CD4+ T-cell numbers in addition to viral load drop. The next stage is the asymptomatic phase where the viral replication is maintained at the minimal level and the CD4+ T-cell counts are stable or slowly declining. The length of this phase varies greatly between individuals depending on various viral and host factors. The third stage of the disease is characterised by a rebound in viral replication and a very fast drop in the T-cell counts that leads to the progression of the disease from HIV infection to the full blown picture of AIDS. In this final stage the immune system is very weak leading to the development of opportunistic infections and cancer.

During the late stages of the disease the viral co-receptor preference in 40-50% of patients was seen to switch from R5 into X4 and R5X4 (Connor *et al.*, 1997; Scarlatti *et al.*, 1997). The rapid progression of disease seen after co-receptor switch was thought to be due to higher cytotoxicity of the X4 strains (Picchio *et al.*, 1998), however other studies has shown that both R5 and X4 are equally cytopathic (Grivel and Margolis, 1999). It is the infection of the larger population of T-cells bearing the CXCR4 receptor which cause this rapid destruction of CD4+ T-cells (Penn *et al.*, 1999). On the other hand the majority of patients progress to AIDS while maintaining the R5 viral strains, even though at a slower pace than patients with X4 and R5X4 viruses. Disease progression in individuals infected with R5 strains is not only associated with enhanced viral tropism for the monocyte/macrophage lineage (Li *et al.*, 1999; Tuttle *et al.*, 2002), but also increased entry efficiency via CCR5 receptors (Jansson *et al.*, 1999; Low *et al.*, 2008; Coetzer *et al.*, 2008). Certain functional changes in R5 Env proteins have been determined in "late stage" individuals. These include the loss of N-linked glycosylation site in V2 that diminishes the ability of virus to use DC-SIGN for infectivity (Borggren *et al.*, 2008), while in other cases, an increase in positively-charged residues in V1/V2 and V4/V5 resulted in increased viral fitness and reduced sensitivity to entry inhibitors (Repits *et al.*, 2008). Enhanced fusogenicity due to the addition of an N-linked glycosylation site (N362) near the CD4 binding site has also been shown to be associated with disease progression in individuals exclusively harbouring R5 viruses (Sterjovski *et al.*, 2007). These studies has emphasised the important role of Env in directing the course of the disease.

1.3.3 Structure of the gp41 fusion glycoprotein

The gp41 is a class I fusion glycoprotein and consists of an N-terminal ectodomain, and a C-terminal transmembrane domain (TMD) and cytoplasmic tail. The ectodomain comprises an N-terminal fusion peptide (FP) connected via a polar segment to an α-helical coiled coil forming segment (helical region 1, HR1) (Chan *et al.*, 1997; Poumbourios *et al.*, 1997; Farzan *et al.*, 1998). The disulfide loop region (DSR) connects HR1 to a second α-helical segment (helical region 2, HR2), which is followed by the membrane proximal external region (MPER); the TMD and cytoplasmic tail (CT) are at the C-terminus (Figure 1.8A). The HR1 has a pocket-forming domain which is assembled in the coiled coil and enables an interaction with the pocket binding domain of HR2, largely through hydrophobic and aromatic amino acids (Chan *et al.*, 1998; Jiang *et al.*, 2002; Liu *et al.*, 2005). The 3-dimensional structure of gp41 in the prefusion state is unknown. A major focus of this project is to model the prefusion structure of gp41 on the fusion activation mechanism of the Ebola virus (EBOV); prefusion and postfusion structures of the EBOV glycoproteins are well characterised.

Prefusogenic gp41 is believed to be trapped in a metastable state via interactions with gp120. CD4 and co-receptor binding leads to conformational changes in gp120 that release gp41 and enable it to refold into low energy conformers. In one such conformer, the prehairpin intermediate, it is believed through structural homology with the hemagglutinin (HA) of influenza virus, that HR1 from 3 gp41 protomers form an extended trimeric coiled coil structure with the N-terminal FP being inserted into target cell membrane (Wild *et al.*, 1994; Furuta *et al.*, 1998; Si *et al.*, 2004). The post-fusion configuration is characterised by a well-defined, highly stable structure known as the six-helix bundle (6-HB), in which 3 HR2 segments pack into hydrophobic grooves on the exterior of the coiled coil in an antiparallel orientation (Figure 1.8B). Interactions between

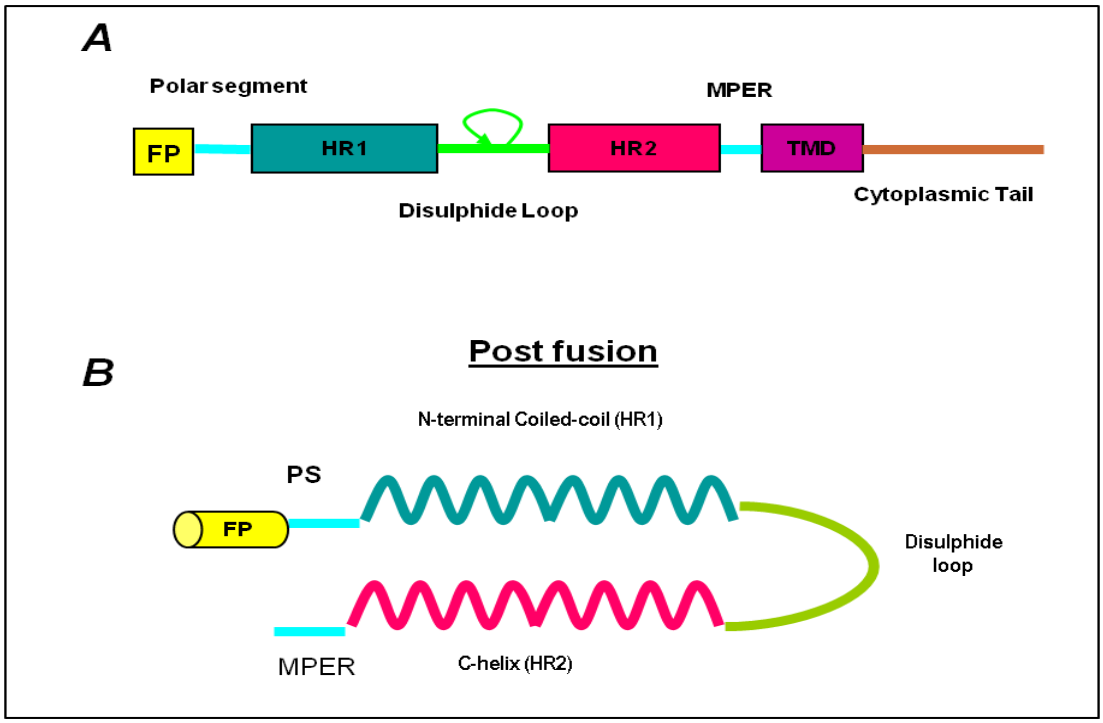


Figure 1.8 Schematic representation of HIV-1 gp41

(A) Linear schematic of gp41 (class I fusion protein) showing the key structural features from N to C-terminus: a fusion peptide (FP) at the N-terminus, an N-heptad repeat (HR1), a C-heptad repeat (HR2), a transmembrane domain (TMD), and a cytoplasmic tail. **(B)** Postfusion conformation of gp41 where the fusion peptide insert into the target cell membrane and the HR2 pack into hydrophobic grooves on HR1 forming the 6-HB.

pocket forming and binding domains contribute to 6-HB stability. The 6-HB conformation would appose the N- and C-terminal membrane-inserted ends of gp41, and the associated viral and cellular membranes for merger (Weissenhorn *et al.*, 1997a; Tan *et al.*, 1997; Chan *et al.*, 1997; Caffrey *et al.*, 1998; Yang *et al.*, 1999). The energy liberated as a result of the transition from the unstable prehairpin intermediate into the stable 6-HB is thought to be necessary for the creation of the fusion pore (Weissenhorn *et al.*, 1996; Hughson, 1997; Chan and Kim, 1998).

1.3.4 Structure of the Env trimeric spike

Biophysical approaches indicated that the Env spike on the surface of HIV-1 is trimeric (Chan and Kim, 1998). These data have been confirmed in cryo-electron tomography microscopy studies of the three dimensional (3D) structure of the virion-associated Env spike and its key features (Zhu *et al.*, 2006; Zanetti *et al.*, 2006; Liu *et al.*, 2008).

The 3D averaging of the first model released by Zhu *et al.* showed that the gp120 SU is composed of three lobes: a main lobe with two secondary lobes (lateral and proximal). Emanating from each proximal lobe was a short leg that meets with the viral membrane forming a foot. This structure was believed to be gp41, with the feet being the MPER (Zhu *et al.*, 2006). The model released months later by Zanetti *et al* showed great variation to Zhu's model despite using the same SIV mutant, but this could be due to the use of different data collection methods and analysis approaches. Zanetti's model showed three distinct gp120 lobes sitting on top of a narrow stalk (gp41) with a cavity in the gp120-gp41 interface (Zanetti *et al.*, 2006). The most recent model by Liu et al (Figure 1.9) showed the HIV-1 spike in the unliganded and CD4 bound states (Liu *et al.*, 2008). The CD4 moiety enabled the authors to dock the crystallographically determined gp120 monomers into the cryotomogram in a reliable manner. The unliganded HIV-1 Env spike

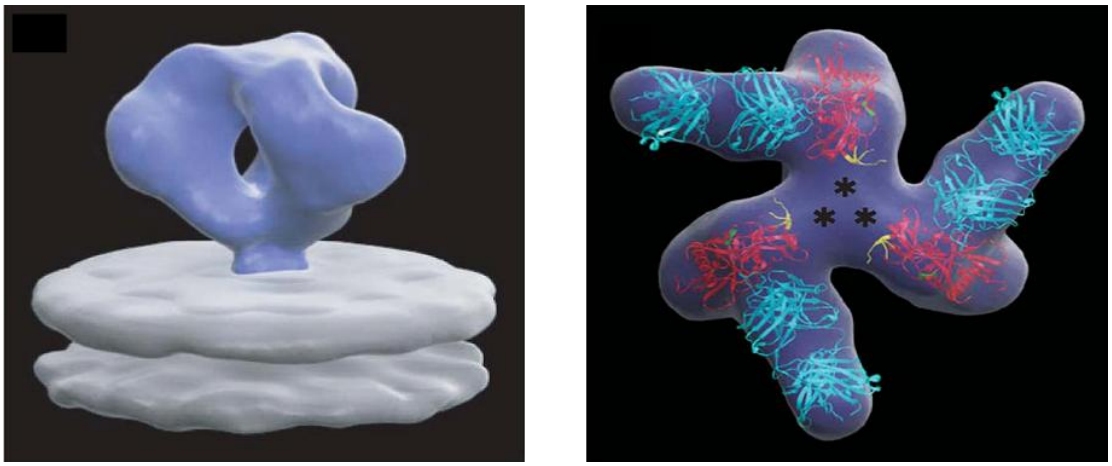


Figure 1.9 Electron tomographic structure of trimeric HIV-1 Env

Left panel: Side view of the trimeric Env spike**. Right panel:** gp120 (red ribbon) bound to domain 1 and 2 of CD4 (blue ribbon) was fitted into the Env spike and is shown from top. The asterisks show the predicted location of gp41 and V1/V2 loop of gp120. Stumps of the V1/V2 and V3 loops are shown in yellow and green, respectively. Adapted from Liu *et al.* (2008).

showed greater similarity with Zanetti's SIV model for both the SU and TM, while the CD4 bound structure showed an opening up of the gp120 lobes coupled with conformational changes in gp41 allowing it to move closer to the target cell membrane (Liu *et al.*, 2008). A precise 3D structure of HIV-1 Env would be very helpful in understanding the viral entry mechanism and in designing a neutralizing antibody eliciting vaccine.

1.3.5 Mechanism of Env-mediated HIV-1 entry

HIV-1 entry is initiated by the high affinity binding of gp120 to the CD4 receptor via the Phe-43 cavity, which is created by residues in the inner domain, bridging sheet and outer domain of gp120 (Kwong *et al.*, 1998; Wyatt *et al.*, 1998). This binding event leads to conformational changes in gp120 associated with the movement of V1/V2 loops to expose the coreceptor binding site comprised of the V3 loop and bridging sheet (Kwong *et al.*, 2000; Kwong *et al.*, 1998). This step is followed by the engagement of CCR5 at both the N terminus and ECL2 region triggering additional conformational changes (converts the flexible V3 stem into a rigid β-hairpin) leading to HIV-1 entry (Huang *et al.*, 2007). Receptor binding leads to conformational changes in gp41 exposing the N-terminal FP followed by its insertion into the target cell membrane (Freed *et al.*, 1990; Furuta *et al.*, 1998; Si *et al.*, 2004).

Membrane fusion begins with the insertion of gp41 FP into the outer leaflet of the target membrane and the gp41 adopts a prehairpin intermediate conformation bridging the viral and cellular membranes (Figure 1.10B) (Furuta *et al.*, 1998; Eckert and Kim, 2001). In the prehairpin conformation the trimeric coiled coil (HR1) is exposed and accessible to HR2-derived peptides like C34 which are capable of blocking further steps in the fusion cascade (Chan *et al.*, 1998). The prehairpin intermediate form is followed by the antiparallel packing of HR2 segments into hydrophobic grooves on the exterior of the

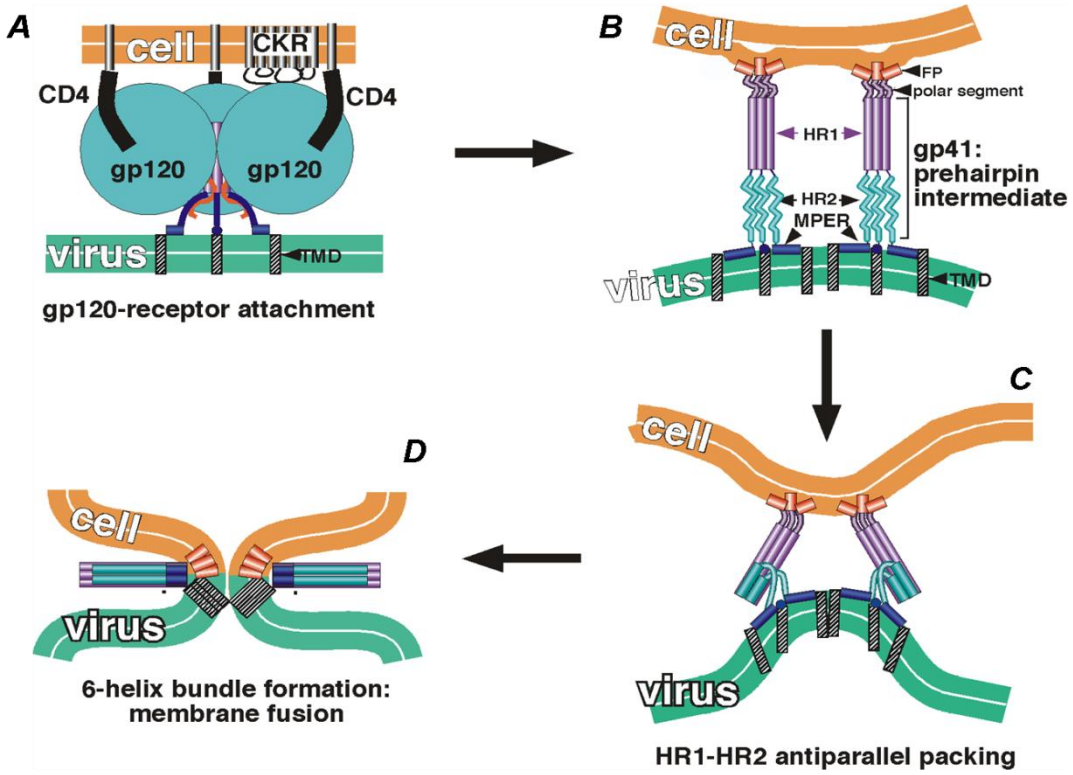


Figure 1.10 The HIV-1 fusion cascade

(A) The prefusion gp120-gp41 trimer, which is activated by the interactions of gp120 to CD4 and chemokine receptor. **(B)** Prehairpin intermediate formation and FP insertion into target cell membrane, followed by the packing of HR2 on to the N-terminal coiled coil in an antiparallel manner forming the 6-HB **(C, D)**. Those steps are followed by hemifusion where the outer leaflets of the viral and cellular membrane starts to mix, then pore formation and expansion allowing the viral core to enter the target cell. Modified from Bellamy-McIntyre *et al.* (2007).

coiled coil forming the 6-HB (Figure 1.10C and D) (Kilgore *et al.*, 2003). This process is believed to juxtapose the cellular and viral membranes, located at the N- and C- terminal ends of the gp41, for merger (Weissenhorn *et al.*, 1997b; Chan *et al.*, 1997). The mixing of the outer leaflets of the viral and cellular membranes is called lipid mixing (hemifusion), which is followed by small pore formation and pore expansion allowing for the viral nucleocapsid to be expelled into the cytoplasm (Eckert and Kim, 2001; Melikyan *et al.*, 2000).

1.4 Activation mechanisms of class I viral membrane fusion proteins

Class I fusion proteins share a number of key structural features (Figure 1.8A). The FP is found at the N-terminus but there are some exceptions like the α-retroviruses (ASLV) and filoviruses (EBOV), which have internal FPs located near the N-terminus (Adam *et al.*, 2004; Cheng *et al.*, 2004). The next structural feature is the N-terminal amphipathic α-helical region (HR1), followed by the C-terminal amphipathic α-helical region (HR2), the TMD and a CT of varying length. For HIV and SIV, the CT is an ~ 150-residue domain. Connecting these structures together are variable length linkers, their structure and role in the fusion mechanism will be discussed for each individual virus.

The fusion mechanism of enveloped viruses is mediated by one (e.g. retroviruses) or more (e.g. herpesviridae, poxviridae) of the viral glycoproteins. Viral glycoprotein attachment to cellular receptors can initiate the fusion process for some viruses (retroviruses), while a minority of viruses such as ASLV require an acidic environment in addition to the cellular receptors to trigger the fusion process (Chu *et al.*, 2006; Delboy *et al.*, 2006). Low pH, in the endocytic compartment, serves as the only trigger for fusion process activation in a wide range of viruses (*Orthomyxo-*, *Flavi-* and *Alphaviridae*) through protonation of certain pH-sensing residues like histidine (Skehel and Wiley, 2000; Chanel-Vos and Kielian, 2004; Kampmann *et al.*, 2006; Thoennes *et al.*, 2008). The

triggering step is followed by structural rearrangements which lead to the partial separation of the receptor binding subunit, allowing the metastable fusion subunit to move closer and eventually insert into the target cell, bridging the viral and cellular membranes.

This intermediate structure which adopts a homotrimer conformation is called a prehairpin; most of the time this conformation is short lived, except for the ASLV where it represents a stable state until further triggering (Narayan *et al.*, 2003). Following prehairpin formation the three C-terminal regions pack against the trimeric N-terminal region, pulling the N- and C-terminal ends together forming the trimer of hairpins (6-HB of HIV-1 gp41). The 6-HB represents a lower energy conformation than the metastable prefusion structure. 6-HB formation liberates energy, which is exploited in the catalysis of lipid mixing (hemifusion), small pore formation and pore expansion. The number of viral glycoproteins required to overcome the energy barrier to pore formation and expansion is a matter of debate. In the case of HIV-1 gp41, Yang *et al* suggest that one glycoprotein complex is enough while Herrera *et al* propose that 4-5 complexes are needed (Yang *et al.*, 2005; Herrera *et al.*, 2005).

1.4.1 Influenza virus HA – the prototype

The influenza virus HA is one of the best characterised class I fusion glycoproteins. HA is produced as the trimeric HA0 precursor protein, which is then cleaved in the *trans*-Golgi network, or by proteases present in the mucosa of the respiratory tract in released viruses, into two subunits; the receptor-binding subunit (HA1) and the fusion subunit (HA2), which remain connected to each other via a disulfide bond. The prefusion structure, which exists as a trimer, was solved 30 years ago (Wilson *et al.*, 1981; Wiley *et al.*, 1981). Figure 1.11A shows the prefusion conformation of an HA1-HA2 monomer which consists of a globular head domain (HA1) and a fibrous stalk (HA2). A prominent feature of the HA2 is the 54-residue amphipathic α-helical trimerization

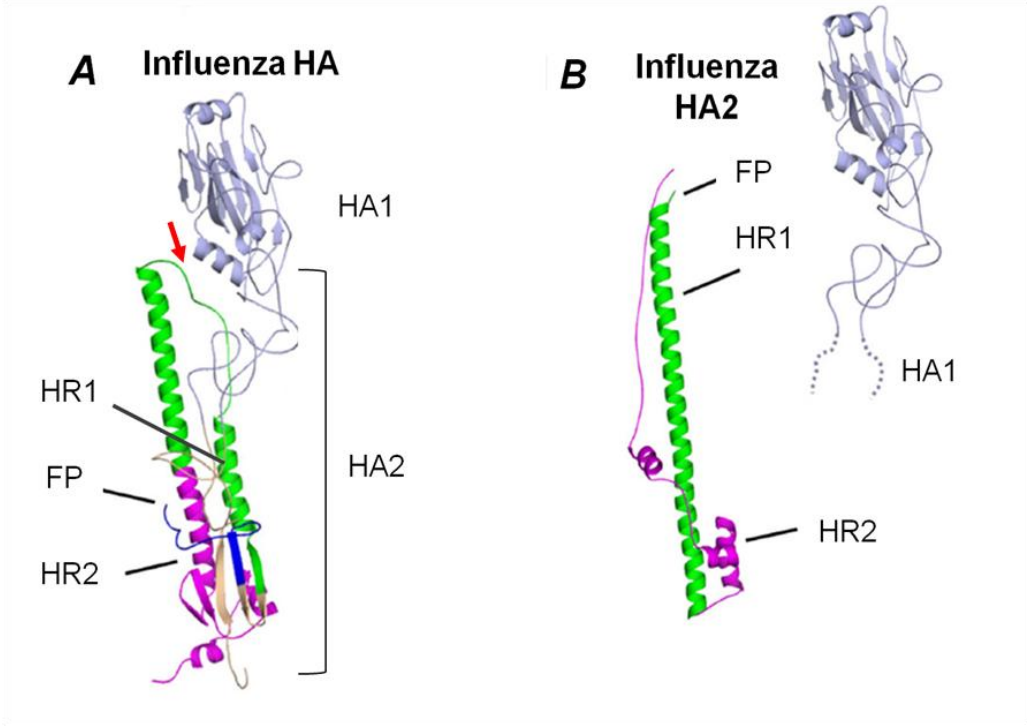


Figure 1.11 structure of influenza HA subunits in the pre and postfusion conformations

(A) HA1-HA2 prefusion monomer consists of a globular head domain (HA1) and a fibrous stalk (HA2). Low endosomal pH induce conformational changes in which the binding subunit (HA1) opens up to expose the fusion subunit (HA2). In the postfusion conformation, the green loop (red arrow in *A*) has undergone a coil-to-helix transition leading to the formation of HR1. The C-terminal part of the coiled coil melts to form HR2 that packs antiparallel to the HR1 **(B)**. Adapted from Lamb and Jardetzky (2007).

domain. This α-helix stabilizes the HA1-HA2 trimer by forming a central triple-stranded coiled coil.

The N-terminus, as in most other class I fusion proteins, is occupied by the FP that is sheltered in a pocket along the three fold axis; it is connected to an outer layer of short α-helices via two β-strands. The transition from the prefusion conformation to the postfusion follows HA binding to sialic acid on glycoproteins or glycolipids leading to internalization into the endosomes (Weis *et al.*, 1988). The acidic environment of the endosomes leads to protonation of three conserved residues (two aspartates and a histidine) in the FP leading to a conformational change in which the binding subunit (HA1) opens up to expose the fusion subunit (HA2) (Bullough *et al.*, 1994; Skehel *et al.*, 1982). The following step is a helical extension of the loop connecting the FP and coiled coil leading to a longer coiled coil (HR1) and FP insertion into the target membrane in a prehairpin conformation. Subsequently, the C-terminal segment of the coiled coil (HR2) becomes unstructured and inverts against the N-terminal portion of the central coiled coil. The C-terminal portion of the HA2 ectodomain continues to refold and pack against the central coiled coil in an antiparallel manner, forming the postfusion trimer of hairpins (a monomer is shown in Figure 1.11B). The trimer of hairpins is terminated and stabilised by an annular cap that directs the FP away from the 3-fold axis of symmetry. These refolding events bring both the C- and N-terminal ends of the HA2 into closer proximity and allow for merger between both membranes (Bullough *et al.*, 1994). In this trimer of hairpins structure the packing of the C-terminal "leash" into the grooves of the coiled coil is necessary for lipid mixing (Park *et al.*, 2003). The annular N-cap structure mediates interactions with membrane proximal residues and is important for pore formation and expansion (Chen *et al.*, 1999; Borrego-Diaz *et al.*, 2003).

It is generally agreed that a successful fusion event requires cooperativity between 3-6 HA trimers (Blumenthal *et al.*, 1996; Danieli *et al.*, 1996). Membrane fusion may also be promoted by the action of triggered HA trimers which are not at the fusion site but close to it in a phenomenon called the bystander effect (Kozlov and Chernomordik, 2002).

1.4.2 Ebola virus GP

The EBOV is an enveloped, negative stranded RNA virus of the *filoviridae* family. It causes severe hemorrhagic fever associated with up to 90% mortality. The EBOV genome has 7 genes which lead to the synthesis of 8 proteins; transcription of the 4th gene results in the expression of the envelope glycoprotein, GP (Sanchez *et al.*, 1996). GP is further cleaved by furin to yield two disulfide-linked subunits, GP1 and GP2. GP1 has a receptor-binding domain (Kuhn *et al.*, 2006), while the GP2 subunit is a class I fusion glycoprotein which comprises an internal FP in the form of a loop that is close to the N-terminus. As for retroviruses, HR1 is linked to HR2 via a DSR and a C-terminal TMD is present. The mechanism of entry into the host cell is still not well characterised but it is thought to proceed by endocytosis through clathrin coated pits and caveolae (Sanchez, 2007), or by macropinocytosis (Saeed *et al.*, 2010). The acidic environment of the endosomes is not only important for EBOV entry and infectivity (Takada *et al.*, 1997; Chandran *et al.*, 2005), but also for providing optimal conditions for the activity of the endosomal proteases, cathepsin B and L, which are essential for viral entry (Schornberg *et al.*, 2006; Lee *et al.*, 2008). Cleavage of GP by cathepsin B and L leads to the removal of the glycan cap and the mucin-like domain, and exposes the conserved residues involved in receptor binding triggering fusion (Chandran *et al.*, 2005; Hood *et al.*, 2009).

Exposure to low pH and cathepsin cleavage leads to conformational changes that allow binding to the TIM-1 receptor (Kondratowicz *et al.*, 2011) and subsequent conformational changes in GP1 that expose the fusion subunit GP2. The prefusion form of

EBOV GP2 was solved three years ago (Lee *et al.*, 2008) and it shows a structured HR1 domain; sequences that are C-terminal to HR1, including the retrovirus-like CX_6CC DSR motif and HR2 are disordered. The HR1 region is divided into 4 segments: HR1A and HR1B helices lie at ~ 90˚to the HR1D helix and are connected together via the HR1C loop (Figure 1.12). Three HR1D segments come together forming a trimer, which represents the nucleus from which the extended central trimeric coiled coil of the postfusion conformation originates. The disulfide bond and the β strands near the internal FP seem to play a role in stabilising the HR1A-D assembly, indicating a conformational signalling pathway between HR1 and the internal FP (Lee *et al.*, 2008). In the post fusion conformation, a loop-to-α-helix transition is seen in HR1C, which extends the central coiled coil in an N-terminal direction along the trimer axis (Figure 1.12). This extension of HR1 allows for fusion loop insertion into the target membrane in a putative prehairpin conformation. Further C-terminal, the CX_6CC DSR motif forms a chain reversal region at the base of the coiled coil, which is followed by the antiparallel packing of the C-terminal "leash" against the grooves of the coiled coil, in a trimer of hairpins conformation (Weissenhorn *et al.*, 1998b).

1.4.3 Retroviruses: HIV, HTLV-I and ASLV

gp120-CD4 binding is sufficient to induce a detectable gp41 prehairpin, however further interactions between CD4-gp120 and chemokine receptor appear to be required to drive the refolding of gp41 beyond this point (Furuta *et al.*, 1998). The post fusion trimer of hairpins structure of gp41 is well characterised (Weissenhorn *et al.*, 1997b), however the prefusion conformation is still a mystery. The overall organisation of structural elements, including the FP, HR1, DSR and HR2, and the postfusion trimer of hairpins conformation are conserved for gp41, gp21 and GP2 (Figure 1.13). This level of functional and structural conservation suggests that GP2 represents an attractive model for understanding retroviral

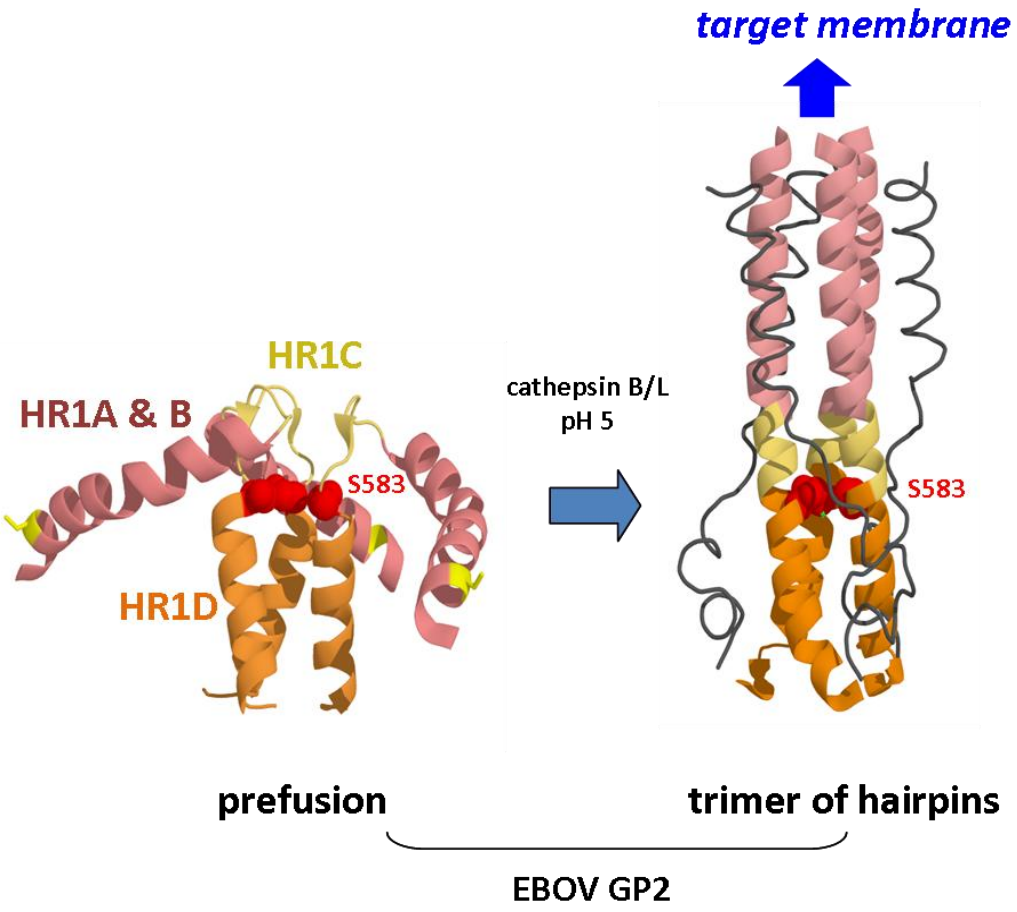


Figure 1.12 Pre and postfusion conformational differences in the Ebola virus GP2

Low pH and Cathepsin induces the transformation into the postfusion conformation where segments A-C of the HR1 helix extend to form the coiled coil and insert into the target cell membrane. Adapted from Lee *et al.* (2008).

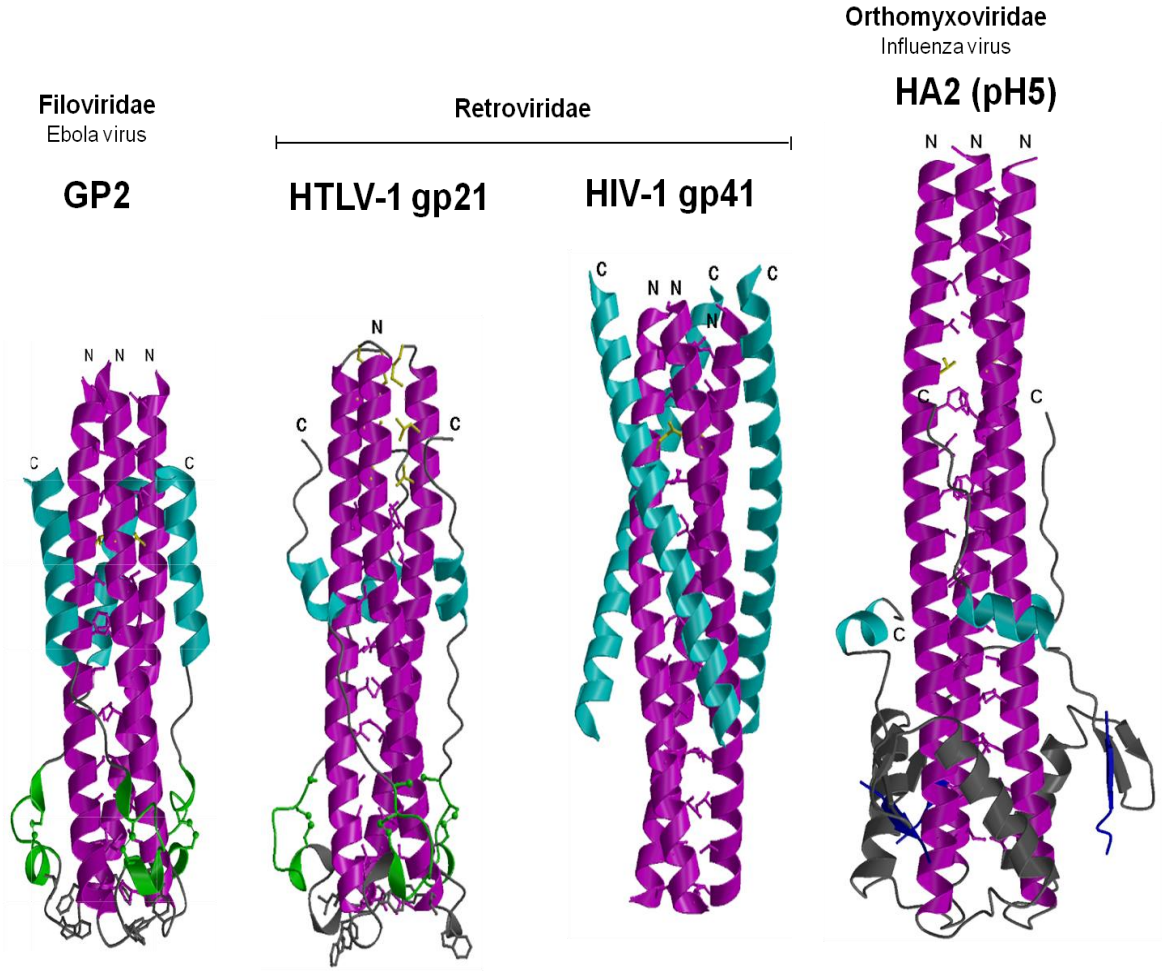


Figure 1.13 Conservation of the postfusion trimer of hairpins structure

Diagram showing the structural similarities of the post fusion conformation of the TM from two retroviruses (HIV-1 and HTLV-1), the Influenza virus HA2 and the EBOV GP2 suggesting a common mechanism of membrane fusion among these viruses.

TM proteins, as the pre- and postfusion forms of GP2 have been structurally characterised. In the prefusion form of GP2, we have seen that the C-terminal one-third of HR1 forms the trimeric coiled coil oligomerization domain, and upon triggering the unstructured regions of HR1 extend the coiled coil in the direction of the target cell. In chapters 3 and 4 of this thesis, I will describe the outcomes of using cysteine insertion mutagenesis to test the hypothesis that the pre-to-postfusion structural transition of gp21 and gp41 is analogous to GP2.

The mechanism whereby conformational signals evoked in retroviral SU glycoproteins by receptor(s) are transmitted to their partner TM glycoproteins, to trigger their refolding into fusion-active forms is being elucidated. Evidence is accumulating to suggest that the association site formed by the DSR of gp41 and the terminal C1 and C5 of gp120 could be a synapse for SU-to-TM conformational signalling (Poumbourios *et al.*, 2003; Maerz *et al.*, 2001; Binley *et al.*, 2003). For example, the simultaneous introduction of a Cys residue to the DSR and to C5 covalently links the gp41-gp120 heterodimer. This disulfide-linked glycoprotein complex is trapped in a fusion-inactive state and reduction of the intersubunit disulfide is required to activate membrane fusion (Binley *et al.*, 2003). Furthermore, in 2010 Bellamy-McIntyre *et al* have shown that W596L and W610F mutations in the DSR block formation of the gp41 prehairpin intermediate that is induced by gp120-CD4 interactions and the initial lipid mixing (hemifusion) phase of the membrane fusion cascade (Bellamy-McIntyre *et al.*, 2010). These workers proposed a model whereby the C1 and C5 regions form a clamp around the DSR to trap gp41 in a metastable conformation.

Mutational approaches have indicated that the 7-stranded β-sandwich, to which the terminal C1 and C5 protrusions are connected, also plays an important role in mediating association with gp41 (Sen *et al.*, 2008; Wang *et al.*, 2008; Yang *et al.*, 2003)

and in regulating its activation state (Finzi *et al.*, 2010). Evidence for the latter idea was provided by Finzi *et al.* (2010) who found that mutations in the layer 1 and 2 excursions from the β-sandwich increased resistance to sCD4-induced gp120 shedding and neutralization. The layer 1 mutations stabilized the sCD4-activated HIV-1 Env intermediate by slowing its reversion into the inactive form indicating its role in modulating conformational transition in gp41 (Finzi *et al.*, 2010). A model based on the involvement of layer 1 and 2 residues and the C1 and C5 regions in the CD4-induced activation of gp41 is shown in figure 1.14.

The results of a previous study by Bellamy-McIntyre *et al.* (2007) provided evidence that CD4-induced conformational activation of gp41 also involves a cooperative interaction between the polar segment (which connects the fusion peptide to the coiled coil) and the MPER at the C-terminus of the ectodomain. In this case, simultaneous mutations in these 2 regions ablated the responsiveness of the gp120-gp41 complex to CD4 ligation and the initiation of membrane fusion. In chapter 5 of this thesis, I will explore potential functional linkages between the DSR and MPER and how these relate to membrane fusion function.

Significant progress has been made in our understanding of the activation mechanisms for 2 other retroviruses: HTLV-1 (deltaretrovirus) and ASLV (alpharetrovirus). In the case of HTLV-1, entry involves the cell surface molecules glucose transporter protein I and neurophilin-1 (Manel *et al.*, 2003; Ghez *et al.*, 2006). In HTLV-1 the presence of a CXXC thiol-disulfide exchange motif in the C-terminal region of the SU (gp46) forms a labile disulfide bond with the CX_6CC motif within the DSR of the TM (gp21) subunit (Pinter *et al.*, 1997). Upon receptor binding, Ca++ release activates the CXXC motif, leading to reduction of the inter-subunit disulfide and enabling the conversion of gp21 into a fusion-active state (Li *et al.*, 2008; Wallin *et al.*, 2004). gp21

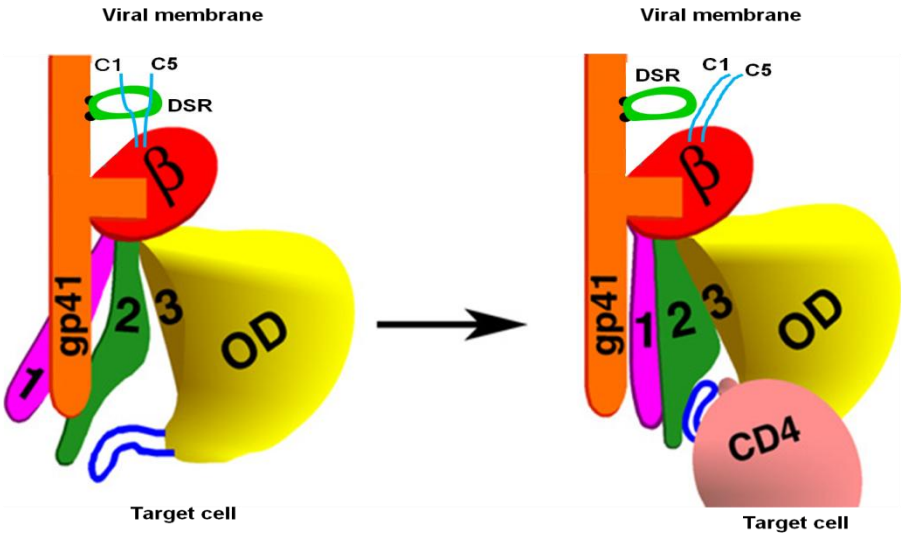


Figure 1.14 Activation model of HIV-1 Env glycoproteins triggered by CD4 binding

In the unliganded state (left), layer 1 and layer 2 assume a conformation that allows them to interact with the gp41 ectodomain. Upon CD4 binding, apposition of layer 1 and 2 occur leading to gp120 inner domain rearrangement, slowing CD4 off rate and allowing the gp41 ectodomain to undergo additional conformational changes necessary for HIV-1 entry. Note that the terminal C1 and C5 protrusions (light blue) are clamped to the DSR (forming the association site) in the unliganded state, but are released after CD4 binding. The β-sandwich is coloured red; layer 1, magenta; layer 2, green; layer 3, yellow; outer domain (OD), yellow; and the β 20-β 21 loop, blue. Modified from Finzi *et al.* (2010).

activation appears to involve the expulsion of a Phe residue, that is immediately adjacent to the CX_6CC DSR motif, from the core of the prefusion SU-TM complex (Maerz *et al.*, 2000).

Recent work has now revealed that the DSRs of HTLV-1, ASLV and EBOV have conserved roles in the activation of the fusion glycoprotein (Delos *et al.*, 2010). For ASLV, fusion activation requires 2 steps: SU-Tva receptor binding, which triggers the prehairpin intermediate and hemifusion, and low-pH-induced protonation of His-428 of SU within the endosome, which enables progression of the fusion cascade (Damico *et al.*, 1998; Hernandez and White, 1998; Narayan *et al.*, 2003; Delos *et al.*, 2010). A consensus functional model for the DSR of these viruses has now emerged based on the original work of Maerz *et al.* (2000). The expulsion of the hydrophobic residue following the CX_6CC motif from the prefusion SU-TM core is necessary for early conformational changes in the fusion glycoprotein including receptor-triggered membrane binding (Delos *et al.*, 2010). The Gly-Gly diad preceding the CX_6CC motif provides flexibility and enables the DSR to pack against the base of the central coiled coil to form a chain reversal. This chain reversal allows the C-terminal "leash" to pack against the surface of the coiled coil in an antiparallel orientation to form the trimer of hairpins (Maerz *et al.*, 2000; Delos *et al.*, 2010). The base of the coiled coil of N-helices contains a stretch of hydrophobic residues which functions in all three viruses to stabilise the postfusion hairpin (Delos *et al.*, 2010).

1.5 Aims and hypothesis

The overall aim of this thesis is to gain a better understanding of the activation mechanism of HIV-1 gp41 and the role of its structural elements in the conformational signalling/fusion activation pathway. We hypothesised that a portion of the central coiled coil of gp41 is formed in the prefusion conformation of gp120-gp41, which is induced to extend toward the membrane-inserted end of gp41 following receptor activation. The second aim is to identify regions within gp120-gp41 that interact functionally with the gp41 DSR. I hypothesise that the characterisation of such linkages will reveal structural determinants of the conformational signalling pathway in gp120-gp41. The outcomes of these studies will advance our understanding of HIV-1 Env structure and assist in the design of novel fusion inhibitors that target a fusion intermediate and/or a structural element involved in conformational signalling between gp120 and gp41.

Chapter 2: Materials and Methods

2.1 Studies with HTLV-1 Env

2.1.1 Cells

Human Embryonic Kidney 293T cells were cultured in DMEM-10 [Dulbeco Modified Eagle Medium (DMEM) (Sigma-Aldrich, USA), 10 % (v/v) heat inactivated foetal calf serum (HI-FCS) (Invitrogen), 2 mM L-glutamine (GE Healthcare, UK), 15 mM HEPES buffer solution (Invitrogen)] and supplemented with 40 μg/ml gentamycin (GE Healthcare) and 2 μg/ml minomycin-hydrochloride salt (Sigma, USA).

HeLa cells, which are an immortalized cell line derived from cervical carcinoma from Henrietta Lacks, were cultured in DMEM-10 and supplemented with 40 μg/ml gentamycin (GE Healthcare) and 2 μg/ml minomycin-hydrochloride salt (Sigma-Aldrich). Cells were incubated at 37 ˚C humidified with 5 % CO_2 and passaged twice weekly.

2.1.2 HTLV-1 Expression vectors

pCELT.1 is an HTLV-1 Env expression vector under the control of a CMV promoter. The HTLV-1 C-terminal tail is tagged with the mAb C8 epitope (Abacioglu *et al.*, 1994; Center *et al.*, 1998). pMBP/gp21(338-425) drives expression of the MBP-gp21 core domain chimera from a Tac promoter in *E.coli* (Kobe *et al.*, 1999; Center *et al.*, 1998).

2.1.3 Antibodies

The HTLV-1 gp46-specific mAb 67/5.5.13.1 is a mouse mAb that binds to an epitope located between residues 210-306 of the SU, gp46 (Abcam, USA). The anti-gp41 mAb C8 binds to the Pro-Asp-Arg-Pro-Asp-Gly sequence within the cytoplasmic tail. It was obtained through the AIDS Research and Reference Reagent Program from G. Lewis (Abacioglu *et al.*, 1994).

2.1.4 In vitro mutagenesis of pCELT.1

The QuikChange XL site-directed mutagenesis kit was used to introduce the A375C mutation into pCELT.1 and pMBP/gp21(338-425) using the primers: 5'-CTACTCAAAATTGCGCAGTATTGTGCCCAGAACAGACGAGGCC (HTLV-1 A375CqcF, forward primer) and 5'-GGCCTCGTCTGTTCTGGGCACAATACTGCGCAATTTTGAGTAG (HTLV-1 A375CqcR, reverse primer). Alanine-375 was chosen as the mutagenic target because it is the only hydrophilic residue that occupies an *a* or *d* position within the gp21 coiled coil. The sequences of mutants were confirmed by the ABI Prism BigDye terminator system (Applied Biosystems, USA).

2.1.5 HTLV-1 western blot

293T cells were transfected with 1 μg of wild type (WT) and A375C-mutated pCELT.1 using FuGENE HD (Roche, Switzerland), according to the manufacturer's instructions. The transfected cells were incubated at 37 °C for 48 h, washed twice in PBS and resuspended in 50 μl of ice-cold western lysis buffer (1 % Triton X-100, PBS, pH 7.4, 1 mM EDTA, 0.02 % sodium azide) for 20 minutes. The lysates were clarified by centrifugation at 10,000 x *g* at 4 °C for 10 min to remove cell debris. For sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE), 5 μl of 5-times sample buffer [10% (w/v) SDS, 250 mM Tris-HCl pH 6.8, 0.5% (w/v) bromophenol blue, 25 % (v/v) glycerol] was added to 25 μl of lysate in the presence of 5 mM iodoacetamide, and then boiled for 5 minutes. Samples were then electrophoresed in 12 % polyacrylamide gels in SDS-PAGE buffer [25 mM Tris, 190 mM glycine, 0.1% (w/v) SDS] under reducing (100 mM dithiothreitol (DTT)) (Thermo Fisher Scientific, USA) and non-reducing conditions. In some experiments, the stability of the A375C gp21 dimer was determined by including 0, 0.5, 5, 25 and 100 mM of DTT in the sample buffer.

The proteins were electrophoretically transferred to Hybond-C nitrocellulose membrane (Amersham, England) in transfer buffer [25 mM Tris, 190 mM glycine, 20 % (v/v) methanol] at 100 V for 1 h. The membranes were blocked with 5% (w/v) skim milk powder in PBS to prevent non-specific binding for 16 h at 4°C, washed seven times with PBST [PBS containing 0.2% (v/v) Tween 20], then probed with mAbs C8 or 67/5.5.13.1 for 16 h in 2.5 % skim milk powder-PBST at 4°C. The membranes were washed again seven times, and then probed for 1 h with goat anti-mouse Alexa Fluor 680 antibody (Invitrogen, USA) in 2.5 % skim milk powder-PBST at room temperature. After washing seven times with PBST, the membranes were visualised using the Odyssey infrared imaging system (LI-COR Biosciences, USA).

2.1.6 HTLV-1 cell-cell fusion assay

293T effector cells (500,000 cells per well of a 12-well culture plate [TPP, Switzerland]) were cotransfected with pCELT.1 and pT4*luc* (Maerz *et al.*, 2001). HeLa target cells (250,000 cells per well of a 12-well culture plate) were transfected with pCAG-T7, which directs the expression of T7 RNA polymerase from the CMV immediate-early enhancer and chicken β-actin promoter (Takikawa *et al.*, 2000) . At 24 h post-transfection, targets and effectors were each resuspended in 1 ml of complete medium; targets (100 µl) were co-cultured with effectors (100 µl) in a 96-well plate for a further 18 h at 37 °C. The cells were assayed for luciferase activity using the Promega SteadyGlo reagent (Promega, USA) according to the manufacturer's instructions.

2.1.7 Expression and purification of MBP/gp21 chimera

An *E. coli* strain BL21(DE3) colony representing each of the maltose-binding protein (MBP)/gp21 chimeras was grown in 100 ml terrific broth [TB, 1.2 % (w/v) bacto-tryptone, 2.4 % (w/v) bacto-yeast extract, 72 mM K_2HPO_4, 17 mM KH_2PO_4 and 0.4% glycerol] at 37 °C for 24 h. The 100 ml culture was then used to inoculate another 900 ml

of TB and agitated for another 4 h at 37 °C until the OD_{600} (optical density measured at a wavelength of 600 nm) was ≥ 1.8. The MBP/gp21 chimeras were induced at room temperature using 0.2 mM isopropyl-β-D-thiogalactopyranoside (IPTG, Sigma-Aldrich) for 18 h. The cells were pelleted at 18,000 x *g* in a 4°C cold centrifuge and lysed by sonication in 100ml ice-cold S-buffer [100 mM Tris-HCl pH 8, 300 mM NaCl, and 1 mM EDTA] and clarified by centrifugation (1 h, 20,000 x *g*, 4°C). MBP/gp21 was affinity-purified using amylose–agarose according to the manufacturer's instructions (New England Biolabs, USA). Briefly, the MBP/gp21 chimera present in a lysate obtained from 1 L of bacterial culture was adsorbed to 10 ml amylose-agarose. Following a wash step with 10 column volumes of S-buffer, the protein was eluted with 10 mM maltose in S-buffer. The MBP/gp21 oligomers were further purified using the AKTA-FPLC (Fast Protein Liquid Chromatography) and a Superdex 200 (HiLoad 26/60) gel filtration column equilibrated in S-buffer. Peak trimeric fractions were pooled together and concentrated by centrifugation at 4,800 x *g* in Ultracel 30 centrifugal columns (Amicon Ultra, Millipore Corporation, USA). MBP/gp21 protein concentration was quantified at 280 nm using a spectrophotometer ND-1000 (NanoDrop, Thermo Fisher Scientific, USA).

2.1.8 Coommassie staining of protein gels

A 10-16 % polyacrylamide step gradient tricine gel was made as follows: 16 % acrylamide solution [1 M Tris-HCl (pH 8.45), 0.1 % SDS, 16 % (w/v) acrylamide, 0.532 % (w/v) bisacrylamide, 0.05 % (w/v) ammonium persulphate, 0.05 % (v/v) TEMED (Invitrogen)] was overlayed with 10 % acrylamide solution [1 M Tris-HCl (pH 8.45), 0.1 % SDS, 10 % (w/v) acrylamide, 0.332 % (w/v) bisacrylamide, 0.05 % (w/v) ammonium persulphate, 0.05 % (v/v) TEMED (Invitrogen)] and left to set for 2 h at room temperature. The separating gel was overlayed with a 4 % polyacrylamide stacking gel [0.75 M Tris-HCl (pH 8.45), 0.075 % SDS, 4 % (w/v) acrylamide, 0.133 % (w/v) bisacrylamide, 0.05 %

(w/v) ammonium persulphate, 0.075 % (v/v) TEMED (Invitrogen)], which was left to set for 1 h prior to electrophoresis. Samples were prepared by adding 7 μg of each of the MBP/g21 chimeras to double-strength tricine sample buffer [150 mM Tris-HCl pH 7, 3% (w/v) SDS, 30 % (v/v) glycerol, 0.05 % (w/v) Coommasie brilliant blue G-250 (Bio-Rad Laboratories, USA)] and electrophoresed for 2.5 h using 100 mM Tris-HCl pH 8.9 as the anode buffer and 100 mM Tris-HCl pH 8.25, 100 mM tricine, 1 % SDS as the cathode buffer. The gels were stained in Coomassie brilliant blue solution [7.5 % (v/v) acetic acid, 50 % (v/v) methanol, 0.025 % (w/v) Coomassie brilliant blue R-250 (Bio-Rad Laboratories)] for 16 h, and then destained using 10% (v/v) acetic acid, 10 % (v/v) methanol for 24 h. The gels were then scanned in the LI-COR Odyssey infrared imager (LI-COR Biosciences).

2.2 Studies with HIV-1 Env

2.2.1 Cysteine mutagenesis

2.2.1.1 Cells

293T cells were cultured as described, above. BHK21 cells, which are derived from baby hamster kidney fibroblasts, were cultured in DMEM-10 and supplemented with 40 μg/ml gentamycin (GE Healthcare) and 2 μg/ml minomycin-hydrochloride salt (Sigma). Cells were incubated at 37 ˚C humidified with 5 % CO_2 and passaged twice weekly.

2.2.1.2 HIV-1 and HIV-2 Env Expression vectors

The pTMenv.2 Env expression vector (Figure 2.1) contains the full-length *env* gene [BH8 clone of HIV-1_{LA1} (Ratner *et al.*, 1985) under the control of the bacteriophage T7 DNA-dependent RNA polymerase promoter (Poumbourios *et al.*, 1997)]. The expression of Env from the T7 promoter requires cotransfection with pCAG-T7. The original plasmid pTM.1 (Moss *et al.*, 1990) was used as a negative control. The pMBP/gp41(528-L-677) *E. coli* expression vector comprises MBP linked to gp41 residues 528-677 of HIV-2_{ST}

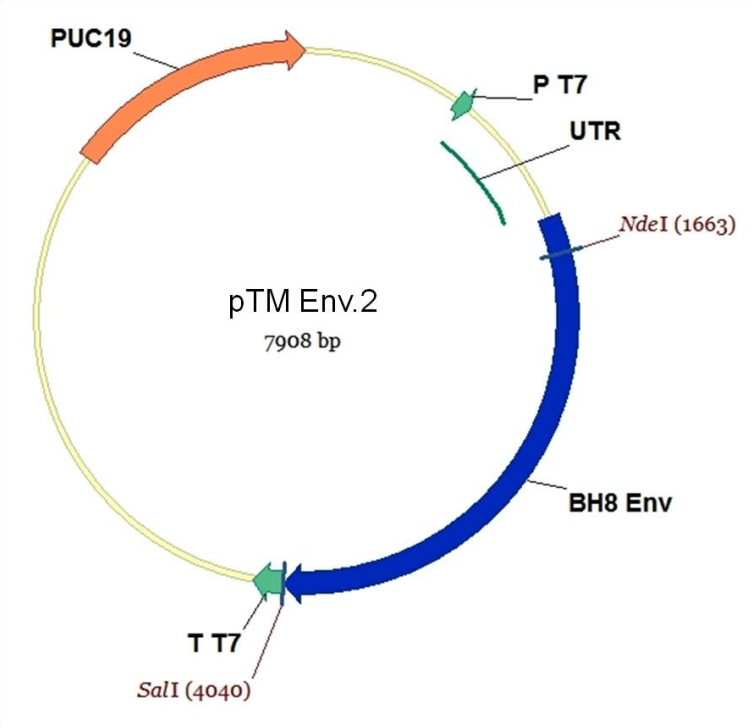


Figure 2.1 pTMenv.2 vector

This vector expressed the Env of the BH8 clone of HIV-1_{LA1}. Cysteine mutations were introduced into the BH8 *env* gene using this vector as a template in Quick Change PCR. The Env of the different cysteine mutants were expressed from this vector by co-transfecting 1 µg of pTMenv.2 and 1 µg of pCAG-T7 into 293T cell line for 48 hrs.

through an Asn-Ala linker incorporating a unique *Not*I site (Lay *et al.*, 2004; Bellamy-McIntyre *et al.*, 2007). In this construct, the native DSR sequence has been replaced with an artificial Ser-Gly-Gly-Arg-Gly-Gly linker. The expression of MBP/gp41(528-L-677) in *E. coli* is controlled by an inducible Tac promoter.

In the pT4*luc* bicistronic vector, human CD4 is constitutively expressed from the CMV immediate early promoter, while firefly luciferase is inducibly expressed from the bacteriophage T7 promoter by T7 DNA-dependent RNA polymerase (Maerz *et al.*, 2001). The pc.FUSIN (CXCR4) vector expresses the CXCR4 co-receptor from a CMV promoter, and was obtained through the AIDS Research and Reference Reagent Program from N. Landau (Deng *et al.*, 1996).

2.2.1.3 Antibodies

The details of the anti-gp41 mAb C8 are described above. DV-012 is a sheep polyclonal antibody specific for gp120 and was obtained through the AIDS Research and Reference Reagent Program from M. Phelan (Hatch *et al.*, 1992).

2.2.1.4 Construction of pTMenv.2 cysteine mutants

The pTMenv.2 plasmid was used as the backbone vector for introducing the cysteine mutations into HIV-1_{BH8} *env*. The cysteine substitution mutations were introduced at six different locations within the coiled coil and polar segment of gp41 at *a* or *d* positions (facing inwards towards the 3-fold axis of symmetry of the coiled coil). The hydrophilic residues mutated were Ser-534 (S534C), Thr-538 (T538C), Ala-541 (A541C), Gln-552 (Q552C), Gln-562 (Q562C) and Thr-569 (T569C) (Figure 2.2). A control cysteine mutant was made at a *b* position (facing outwards from the coiled coil) at Gln-567 (Q567C). The forward primers used for each mutant are listed in Table 2.1. The reverse primers are complementary to the forward primers (not shown). The QuikChange

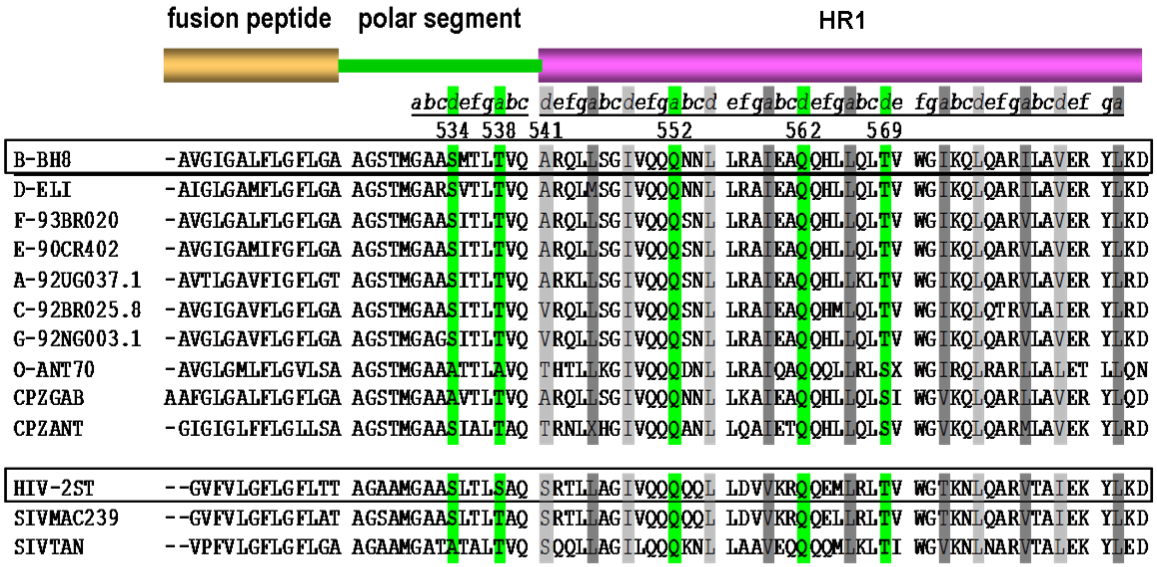


Figure 2.2 Alignment of gp41 amino acid sequences

The HIV-1_{BH8} and HIV-2_{ST} sequences represent the strains of interest in this study. The yellow segment represents the fusion peptide while the green and magenta represent the polar segment and HR1, respectively. The numbers show the sites where the cysteine mutations were introduced.

Primer name	Primer sequence 5'- 3'	Length mer	Tm ˚C
BH8-S534CqcF	GCACTATGGGCGCAGCGTGTATGACGCTGACGGTACAG	36	83.1
BH8-T538CqcF	GGCGCAGCGTCAATGACGCTGTGTGTACAGGCCAGACAATTATTG	42	81.3
BH8-A541CqcF	GCGTCAATGACGCTGACGGTACAGTGTAGACAATTATTGTCTGGTATAGTG	51	79.2
BH8-Q552CqcF	GTCTGGTATAGTGCAGCAGTGCAACAATTTGCTGAGGGCTATTGAG	46	80.3
BH8-Q567CqcF	GAGGCGCAACAGCATCTGTTGTGTCTCACAGTCTGGGGCATCAAGCAG	48	83.7

Table 2.1 HIV-1_{BH8} mutagenesis primers

XL site-directed mutagenesis kit was used according to the manufacturer's instructions (Stratagene, USA). In brief, the reaction mixture had the pTMenv.2 template, the forward and reverse primers with the single site mutation, the reaction buffer and the *Pfu* turbo DNA polymerase enzyme. The PCR cycling conditions were as follows: initial denaturation at 95 ˚C for 1 min; 18 cycles of 95 ˚C for 50 sec, 60 ˚C for 50 sec, 68 °C for 8 min (1 min/kb of plasmid length); a final extension step at 68 ˚C for 7 min. PCR tubes were left on ice for 2 minutes, digested with D*pn*I restriction enzyme at 37˚C for 1hr. The XL10-Gold Ultracompetent Cells were used for transformation. Plasmid DNA was extracted from five colonies per mutant (MOBIO Laboratories, USA), and the sequences of Env mutants were confirmed by the ABI Prism BigDye terminator system (Applied Biosystems).

2.2.1.5 Construction of MBP/gp41 chimera mutants

The cysteine substitution mutations were introduced into pMBP/gp41(528-L-677) at six different locations within the coiled coil and polar segment of gp41 at an *a* or *d* position. The hydrophilic residues mutated were Ser-534 (S534C), Ser-538 (S538C), Ser-541 (S541C), Gln-552 (Q552C), Gln-562 (Q562C) and Thr-569 (T569C) (Figure 2.2), four residues are conserved among the HIV-1_{BH8} and HIV-2_{ST} sequences. A control mutant wherein cysteine was introduced at an *f* position (i.e. facing outwards from the coiled coil), was made by replacing threonine-536 with Cys. The forward primers used for each mutant are listed in Table 2.2. The reverse primers are complementary to the forward primers (not shown). The mutagenesis was carried out using the QuikChange XL site-directed mutagenesis kit, and the sequences of cysteine mutants confirmed by the ABI Prism BigDye terminator system (Applied Biosystems).

Primer name	Primer sequence 5'-3'	Length mer	Tm ˚C
HIV-2ST-S534CqcF	GCCGCGATGGGCGCGGCGTGTTTGACGCTGTCGGCTCAGTC	41	91
HIV-2ST-T536CqcF	GGCGCGGCGTCCTTGTGTCTGTCGGCTCAGTCT	33	79.6
HIV-2ST-S538CqcF	GCGGCGTCCTTGACGCTGTGTGCTCAGTCTCGGACTTTATTG	39	80.4
HIV-2ST-S541CqcF	GACGCTGTCGGCTCAGTGTCGGACTTTATTGGCCGGG	37	80.3
HIV-2ST-Q552CqcF	GCCGGGATAGTGCAGCAATGTCAACAGCTGTTGGACGTGGTC	42	81
HIV-2ST-Q562CqcF	CTGTTGGACGTGGTCAAGAGATGTCAAGAAATGTTGCGACTGACC	45	80.3
HIV-2ST-T569CqcF	CAACAAGAAATGTTGCGACTGTGTGTCTGGGGAACAAAAAATCTCCAG	48	80.8

Table 2.2 HIV-2_{ST} mutagenesis primers

2.2.1.6 Construction of HIV-1_{AD8}-HIV-2_{ST} chimeric MBP/gp41 (528-L-677)$^{AD.ST}$ expression vectors

The N-to-C-terminal organization of the MBP/gp41(528-L-677)$^{AD.ST}$ chimera is as follows: MBP-(Asn-Ala-Ala-Ala)-HIV-1_{AD8} polar segment (res. 528-540)-HIV-2_{ST} HR1 (res. 541-596)-(Ser-Gly-Gly-Arg-Gly-Gly)-HIV-2_{ST} HR2 (res. 611-665)-HIV-1_{AD8} MPER (res. 666-677) and was obtained from Mr. Chan-Sien Lay (Burnet Institute). The T538C and I535A/T538C/V539G mutations were introduced to MBP/gp41(528-L-677)$^{AD.ST}$ using the QuikChange XL site-directed mutagenesis kit. The primer pairs used for T538C were: 5'-GGCGCCGCGTCAATAACGCTGTGTGTACAGTCTCGGACTTTATTG (STMG-Cys538f, forward primer) and 5'-CAATAAAGTCCGAGACTGTAC ACACAGCGTTATTGACGCGGCGCC (STMG-Cys538r, reverse primer), whereas for I535A/T538C/V539G, 5'-GCGTCAGCGACGCTGTGTGGGCAGTCTCGGAC (STMG-IAVG-Cys538f) was the forward primer and 5'-GTCCGAGACTGCCCACACAGCGTCGCTGACGC (STMG-IAVG-Cys538r) was the reverse primer. DNA sequences were confirmed using ABI Prism BigDye terminator system (Applied Biosystems).

2.2.1.7 Construction of a chimeric HIV-1_{AD8}-HIV-2_{ST} MBP-polar segment-HR1 chimera, MBP/gp41(528-591)$^{AD.ST}$

Wild type, T538C- and I535A/T538C/V539G-mutated MBP/gp41(528-L-677)$^{AD.ST}$ expression vectors were used as templates in PCR to produce corresponding MBP/gp41(528-589)$^{AD.ST}$ vectors, which comprise the polar segment of HIV-1_{AD8} (res. 528-540) fused to the HR1 of HIV-2_{ST} (res. 541-591). The forward primer was MBP-A3(f): 5'-GCCTGTCGTACTGCGGTGATCAAC, while the reverse primer was HIV-2_{ST} NheI-SalI(r): 5'-CGCCGCGTCGACTTAGTCCTTTAAGTATTTCTCGATAGC. The PCR products were ligated into the pMBP/gp41(528-L-677) vector through unique *Not*I-

*Sal*I restriction sites. DNA sequences were confirmed using ABI Prism BigDye terminator system (Applied Biosystems).

2.2.1.8 Construction of MBP/gp41(528-L-665)$^{AD.ST}$ chimera (tr665)

The N-to-C-terminal organization of the MBP/gp41(528-665)$^{AD.ST}$ chimera is as follows: MBP-(Asn-Ala-Ala-Ala)-HIV-1_{AD8} polar segment (res. 528-540)-HIV-2_{ST} HR1 (res. 541-596)-(Ser-Gly-Gly-Arg-Gly-Gly)-HIV-2_{ST} HR2 (res. 611-665). The vector encoding this chimera was obtained from Mr. Chan-Sien Lay (Burnet Institute). The T538C and I535A/T538C/V539G mutations were introduced to MBP/gp41(528-665)$^{AD.ST}$ by using the corresponding mutants of MBP/gp41(528-L-665)$^{AD.ST}$ in a PCR reaction with the primers: MBP-Ala3f as the forward primer and HIV-2_{ST} S665tr-TAA-SalI (r), 5'-G GCGGTCGACTTAGCTATTTAATTTTTGTAGTTCATACATG, as the reverse primer. The PCR products were ligated into the pMBP/gp41(528-L-677) vector through *Not*I-*Sal*I sites. DNA sequences were confirmed using ABI Prism BigDye terminator system (Applied Biosystems).

2.2.1.9 HIV-1 cysteine mutants western blot

293T cells were plated at 500,000 cells per well of a 12-well culture plate (TPP, Switzerland) and then co-transfected with 1 μg of pTMenv.2 and 1 μg of pCAG-T7 2 h later using FuGENE HD (Roche), according to the manufacturer's instructions. In order to promote cleavage of gp160 to gp120 and gp41, the transfected cells were incubated at 34 °C for 48 h prior to 2 phosphate-buffered saline (PBS) washes and lysis in 50 μl PBS containing 1 % Triton X-100, 0.02 % sodium azide and 1 mM EDTA for 20 minutes on ice. The lysates were clarified by centrifugation at 10,000 x *g* at 4 °C for 10 minutes to pellet cell debris. The samples were prepared for SDS-PAGE by adding 7 μl of 5-times sample buffer to 35 μl of lysate in the presence of 5 mM iodoacetamide (to alkylate free cysteines) and boiled for 5 minutes. Samples were then run in 10 % polyacrylamide gels

and electrophoresed in SDS-PAGE buffer under reducing (100 mM DTT) and nonreducing conditions. In some experiments, the samples were reduced with 0, 0.5, 5, 25 and 100 mM of DTT and 2 M Urea in the sample buffer. In this case, the samples were boiled for 1 h, prior to SDS-PAGE in a 6 M Urea-10 % polyacrylamide gel.

Proteins were transferred to Hybond-C nitrocellulose membrane (Amersham) and probed with C8 or DV-012 primary antibodies. Goat anti-mouse-Alexa Fluor 680 antibody (Invitrogen) or donkey anti-sheep Alexa Fluor 680 antibody (Invitrogen) were used as secondary antibodies, respectively, then the membranes were visualised using the Odyssey infrared imaging system (LI-COR Biosciences).

2.2.1.10 Luciferase reporter assay of cell-to-cell fusion for HIV-1 constructs

293T effector cells (500,000 cells per well of a 12-well culture plate) were cotransfected with WT and mutated pTM*env* and pCAG-T7 plasmids. BHK21 target cells (250,000 cells per well of a 12-well culture plate) were cotransfected with pT4*luc* (Maerz *et al.*, 2001) and pc.CXCR4 (Deng *et al.*, 1996). At 24 h post-transfection, targets and effectors were each resuspended in 1 ml of complete medium; targets (100 µl) were co-cultured with effectors (100 µl) in a 96-well flat-bottomed plate (BD Bioscience, USA) for a further 18 h at 37 °C. The cells were assayed for luciferase activity using the Promega SteadyGlo reagent (Promega) according to the manufacturer's instructions. 293T cells cotransfected with pTM.1 and pCAG-T7 were used as negative controls to determine the background level of the luciferase activity.

2.2.1.11 Expression and purification of MBP/gp41 chimera

An *E. coli* strain BL21(DE3) colony transformed with each of the MBP/gp41 expression plasmids was grown in 100 ml TB at 37 °C for 24 h with agitation. The 100 ml cultures were then used to inoculate 900 ml of TB, which was incubated at 37 °C until the OD_{600} reached between 1.8-2. The MBP/gp41 chimeras were induced at room temperature using 0.2 mM IPTG (Sigma-Aldrich) and incubated for a further 18 hrs. The cells were pelleted by centrifugation (18,000 x *g* for 1 h, 4 °C), lysed by sonication in 100ml ice-cold MES buffer [100 mM 2-N-morpholino ethanesulfonic acid (MES), pH 5, 300 mM NaCl, and 1 mM EDTA], and finally clarified by centrifugation (1 h, 20,000 x *g*, 4°C). The MBP/gp41 chimeras were affinity-purified using amylose–agarose according to the manufacturer's instructions (New England Biolabs) and as detailed in section 2.1.7. The protein-positive fractions were pooled and filtered through a 22 µm filter and then further purified by gel filtration using a Superdex 200 (HiLoad 26/60) column equilibrated in MES buffer and attached to an AKTA-FPLC system (Lay *et al.*, 2004). Fractions corresponding to trimeric protein were pooled and concentrated by centrifugation at 4,800 x *g* in Ultracel 30 centrifugal columns (Amicon Ultra, Millipore Corporation). The concentrations of MBP/gp41 samples were quantified at 280 nm using a spectrophotometer ND-1000 (NanoDrop). Protein samples were subjected to SDS-PAGE in 10-16 % tricine gels and stained with Coomassie brilliant blue following the same procedures in section 2.1.8.

2.2.1.12 Crosslinking of MBP/gp41 oligomers using homobifunctional Cys-reactive crosslinkers (BMOE and BMH)

The thiol-reactive crosslinkers used in this study were **BMOE,** bis(maleimido)ethane (8 Å spacer between thiol-reactive groups) and **BMH,** bis(maleimido)hexane (13 Å spacer) (Thermo Scientific, USA). Chimeric MBP/gp41 proteins in PBS containing 200 mM

imidazole pH 7.2 were treated with 4 mM BMOE or BMH for 1 h at room temperature. The reaction was quenched with 30 mM DTT for 15 minutes at room temperature prior to SDS-PAGE in 10–16% polyacrylamide gradient gels as mentioned in section 2.1.8. The protein bands were visualized following staining with Coomassie brilliant blue and scanning in a LI-COR Odyssey infrared imager (LI-COR Biosciences).

2.2.2 Forced sequence evolution of pAD8 mutants

2.2.2.1 Cells

293T and BHK21 cells were cultured as described, above. U87.CD4.CCR5 cells were obtained from the NIH AIDS Research and Reference Reagent Program from H. Deng and D. Littman (Bjorndal *et al.*, 1997). These cells were maintained in DMEM-15 (15 % (v/v) HI-FCS) and supplemented with 300 μg/ml G418 and 100 μg/ml puromycin (Sigma-Aldrich). The cells were incubated at 37 ˚C in a humidified incubator with 5 % CO_2.

2.2.2.2 Plasmids

pCDNA3.1-AD8*env* is a CMV promoter driven HIV-1_{AD8} Env expression vector (Poumbourios *et al.*, 2003). pΔKAD8$_{env}$ was derived by religation of the end-filled *Hind*III and *Eco*RI sites of pCDNA3.1-AD8*env* to remove a *Kpn*I site from the vector multicloning site (Figure 2.3). The pAD8 proviral clone was obtained from K. Peden (Theodore *et al.*, 1996). The pNL4.3.Luc.R^-E^- reporter vector was obtained from the AIDS Research and Reference Reagent Program from N. Landau (Deng *et al.*, 1996). pc.CCR5 vector expresses the CCR5 co-receptor from a CMV promoter, and was obtained through the AIDS Research and Reference Reagent Program from N. Landau (Deng *et al.*, 1996). The plasmids pCAG-T7, p*T4luc* and pCDNA3.1 were also used in this part of the study.

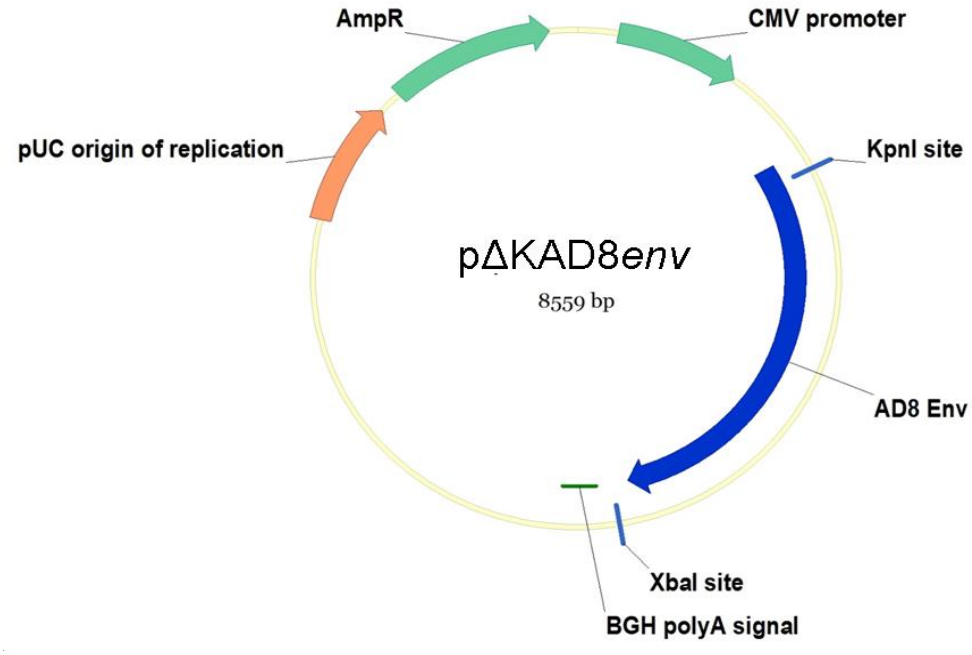


Figure 2.3 pΔKAD8env expression vector

This vector was used as a backbone for the ligation of Full-length *env* gene fragments which were generated by PCR amplification of genomic DNA extracted from the infected cells. The PCR products generated were subjected to electrophoresis and then purified using QIAEX II Gel Extraction Kit. The eluted DNA was then restricted with *Kpn*I and *Xba*I and ligated into the corresponding restriction sites in pΔKAD8$_{env.}$

2.2.2.3 Antibodies

Immunoglobulin G number 14 (IgG 14) was purified from the plasma of a HIV-1 positive individual using protein A-Sepharose (Amersham Biosciences). Monoclonal antibody C8 (against gp41) and polyclonal DV-012 antibody (against gp120) were used also.

2.2.2.4 Preparation of W596L.K601D HIV-1_{AD8} Env mutants

The pcDNA3.1-AD8*env* and pAD8 W596L.K601D mutants were obtained from Dr. P. Poumbourios. Briefly, the W596L.K601D double mutation was initially introduced to pcDNA3.1-AD8*env* by overlap-extension PCR using the overlapping forward primers: 5'-CTCGGTTGCTCTGGAGATCTCATCTGCACCAC (AD8.18/1A overlap1) and 5'-CAGCTCCTGGGGATTCTCGGTTGCTCTGGAGATC (AD8.18/1A overlap2), and the reverse primer, 5'-AATCCCCAGGAGCTGTTG (BH8.15). The sequences of mutant clones were verified by the ABI Prism BigDye terminator DNA sequencing system. The mutagenized DNA fragment bounded by *Eco*RI and *Bfu*A1 restriction sites was then ligated to the corresponding restriction sites within pAD8 to produce the W596L.K601D doubly mutated proviral clone. The veracity of proviral clones was confirmed by restriction analysis (*Hind*III, *Eco*RI plus *Bgl*I) and DNA sequencing.

2.2.2.5 Protein composition of virions

293T cells were plated at 350,000 cells per well of a 6-well culture plate (NUNC, Denmark). The cells were then transfected with 1 μg of pAD8 per well (6 wells per construct) using FuGENE HD (Roche, Switzerland). At 72 h post transfection, culture supernatants were filtered through 0.45 μm nitrocellulose filters (Sartorius Stedim Biotech, Germany) and virions pelleted through a 1.5 ml 25% (w/v) sucrose/PBS cushion in a Beckman ultracentrifuge (SW41 Ti rotor, 2.5 h, 25,000 rpm, 4 ˚C). Virus pellets were then

resuspended in 50 µl SDS-PAGE sample buffer [10% (w/v) SDS, 250 mM Tris-HCl pH 6.8, 0.5 % (w/v) bromophenol blue, 25 % (v/v) glycerol, 3 % (v/v) β-mercaptoethanol] and subjected to reducing SDS-PAGE in 7.5-15% polyacrylamide gradient gels. The proteins were transferred to Hybond-C nitrocellulose membrane (Amersham, England) and probed with IgG 14 primary antibody. The immunoblots were developed with IRDye 800CW-conjugated rabbit anti human Ig (Rockland, USA) and scanned using Odyssey infrared imaging system (LI-COR Biosciences).

2.2.2.6 Generation of the VSV G pseudotyped HIV-1_{AD8} virions:

293T cells were plated at 350,000 cells per well of a 6-well culture plate (NUNC, Denmark). The cells were then transfected with 0.75 µg of pAD8 and 0.25 µg of pHEF-VSVG (obtained from the NIH AIDS Research and Reference Reagent Program (Chang *et al.*, 1999)) per well (6 wells per construct) using FuGENE HD (Roche, Switzerland). At 72 h post transfection, culture supernatants were filtered through 0.45 µm nitrocellulose filters (Sartorius Stedim Biotech, Germany). Virus produced in 293T cells was normalised for RT activity and used to infect U87.CD4.CCR5 cells.

2.2.2.7 RT assay

The amount of virus present in cell culture supernatants was determined by measuring virion associated RT activity using a modified RT assay (Willey *et al.*, 1988). Briefly, 10 µl of culture supernatant was mixed with an equal volume of 0.3 % NP-40 (Nonyl phenoxypolyethoxylethanol) and combined with 40 µl of RT reaction mix [5 mM $MgCl_2$, 2 mM DTT, 50 mM Tris-HCl (pH 7.8), 7.5 mM KCl, 5 µg/ml poly(rA)-(dT) (Amersham Biosciences, UK) and 0.5 µCi [^{33}P]-dTTP (Perkin Elmer, USA)], in a 96 well round bottom microtitre plate (Sarstedt, USA). Following incubation for 3 h at 37 °C, 6 µl of the reaction mixture was spotted onto DE81 ion-exchange paper (Whatman, UK) and air -dried. The paper was then washed 5 times in double strength sodium chloride-sodium

citrate (SSC) buffer [0.3 M sodium citrate (pH 7), 0.03 M NaCl] (BIO-RAD, USA) and twice in 95 % (v/v) ethanol to remove unincorporated [^{33}P]-dTTP. Meltilex scintilant (Wallac, Finland) was applied to the paper and RT activity was determined by the level of [^{33}P]-dTTP bound to the DE81 filter, using a 1450 Microbeta Plus liquid scintillation counter (Wallac, Finland). The data were expressed as the average counts per minute (cpm) of three replicates.

2.2.2.8 14-Day Replication assay

U87.CD4.CCR5 cells (1 x 10^5 per well of a 96-well flat bottomed plate [BD Bioscience, USA]) were infected with equivalent amounts of WT and HIV-1_{AD8} mutants (according to RT activity) starting with neat virus stock (400,000 cpm) in the top row (A) of the plate. Serial 10-fold dilutions were performed until a 10^{-7} dilution was reached in the bottom row (H). Half of the supernatant volume was replaced with fresh DMEM-15 (half-media change) at 3, 7 and 10 days post infection, and RT assay was done for all viral supernatants according to the protocol in section 2.2.2.7.

2.2.2.9 Sequential passage of cell-free W596L/K601D virus in U87.CD4.CCR5

U87.CD4.CCR5 cells (1 x 10^6) in a 25 cm^2 culture flasks (BD Bioscience, USA) were infected with equivalent amounts of WT and W596L/601D-mutated HIV-1_{AD8} mutants (according to RT activity) in parallel and maintained in culture for 10 days. Half-media changes were done at days 3, 7 and 10, and virus-containing supernatants were stored at -80 ˚C. At day 10, the culture supernatants were filtered (0.45 μm pore size) and normalized according to RT activity prior to the next passage (5 passages in total). At the end of each passage (day 10) the cells were dislodged using trypsin-versene solution [1% (w/v) trypsin, 0.2 g/L EDTA], and then pelleted at 10,000 x *g* for 1 minute prior to storage at -80 ˚C.

2.2.2.10 Construction of pΔKAD8 Env revertants

2.2.2.10.1 Isolation of genomic DNA from viral infected cells

Genomic DNA was isolated from U87.CD4.CCR5 cells harvested at the end of each passage using DNeasy Blood & Tissue Kit (QIAGEN, Germany) according to the manufacturer's instructions.

2.2.2.10.2 PCR amplification of full-length HIV-1_{AD8} env from infected cells

Full-length *env* gene fragments were generated by PCR amplification of genomic DNA extracted from the infected cells, using Expand High Fidelity Enzyme (Roche, Switzerland). The primers used were 5'-AGAAAGAGCAGAAGACAGTGGCAATGA (Env1Bf, forward primer and 5'-CCCTTTCTAGAGGCCATCCAGCCATACTACGTTTTGACCACTTGC (AD8TAAXbaR, reverse primer). A 50 µl PCR reaction mixture contained 2 µl of template DNA, 5 µl 10x reaction buffer, 1 µl 10 mM dNTP mix, 1 µl of each primer at 10 µM, and 3.5 units Expand enzyme. The PCR cycling conditions were as follows: initial denaturation at 94 ˚C for 2 min; 34 cycles of 94 ˚C for 30 sec, 60 ˚C for 30 sec, 72 °C for 2 min; a final extension step at 72 ˚C for 7 min. PCR reactions were carried out using the Peltier PTC-200 Thermal Cycler (MJ Research, USA). PCR products were analysed by electrophoresis in 0.8% (w/v) agarose-TAE (40 mM Tris-acetate, 1 mM EDTA, pH 8.3) gels containing SYBR Safe (Invitrogen).

2.2.2.10.3 Cloning of PCR-amplified env DNA

The PCR products generated in section 2.2.2.10.2 were subjected to electrophoresis in 0.8 % agarose gels and then purified using QIAEX II Gel Extraction Kit (QIAGEN, Germany) according to the manufacturer's instructions. The eluted DNA was then restricted with *Kpn*I and *Xba*I and ligated into the corresponding restriction sites in pΔKAD8$_{env}$ (Figure 2.3). The entire *env* region was sequenced using the ABI BigDye

terminator system and the primers: Env1Bf, 5'-AGAAAGAGCAGAAGACAGTGGCAATGA (forward), Forwarduniseq, 5'-GTCAGCACAGTACAATGTACACATGG (forward), Tailseqf, 5'-GTCATTTCAGACCCACCTCCC and REKRf 5'-CCCACCAAGGCAAAGAGAAGAGTGGTGC (forward). The sequences were analysed using the Vector NTI software package (Invitrogen). Transfection-quality DNA corresponding to clones that were of interest was produced using a QIAGEN Plasmid Midi kit as recommended by the manufacturer.

2.2.2.11 Luciferase reporter assay of cell-cell fusion

These assays were performed as described in section 2.2.1.10 except that BHK targets were transfected with pc.CCR5 instead of pc.CXCR4.

2.2.2.12 Single cycle entry assay

Single cycle entry assays were conducted as described (Bellamy-McIntyre *et al.*, 2007). Briefly, Env-pseudotyped luciferase reporter viruses were produced by co-transfecting 293T cells with pΔKAD8$_{env}$ vectors plus the luciferase reporter virus vector, pNL4.3.Luc.R$^-$E$^-$ using FuGENE 6. At 48 h posttransfection, the virus containing supernatants were filtered through a 0.45 μm nitrocellulose filter prior to addition to U87.CD4.CCR5 target cells. At 48 h post infection, the cells were washed once with PBS prior to lysis and assay for luciferase using the Promega Luciferase Assay System.

Chapter 3: Probing the Structure of the gp21 Glycoprotein of HTLV-1 Using Cysteine Replacement Mutagenesis

3.1 Aim

To determine if structural changes occurring in gp21 of HTLV-1 during activation can be modelled on those observed for the Ebola virus (EBOV) homologue, GP2, using Cys replacement mutagenesis.

3.2 Introduction

Human T cell leukemia virus type 1 is a retrovirus with a wide geographic distribution associated with tropical spastic paraparesis (Osame *et al.*, 1986) and adult T cell leukemia (Poiesz *et al.*, 1980). Unlike other retroviruses, cell-free virions are non-infectious, therefore HTLV-1 transmission is believed to occur between infected and uninfected cells across a virological synapse (Fan *et al.*, 1992; Donegan *et al.*, 1994; Igakura *et al.*, 2003; Jolly and Sattentau, 2004). The HTLV-1 Env is initially synthesised as a polyprotein precursor, gp62, which is cleaved to yield a receptor-binding surface-exposed subunit (SU), gp46, and a noncovalently associated TM, gp21 (Pique *et al.*, 1992). The binding of gp46 to the cellular receptors, glucose transporter protein I and neurophilin-1 (Manel *et al.*, 2003; Ghez *et al.*, 2006), is believed to lead to conformational changes that trigger the reduction of a labile disulfide formed between SU and TM (Li *et al.*, 2008), and the refolding of gp21 into a fusion-active trimer of hairpins that involves the antiparallel packing of the C-terminal ectodomain segment onto the coiled coil (Kobe *et al.*, 1999).

The ectodomain of class I fusion glycoproteins, including that of retrovirus TMs, exhibits a common organization of structural/functional elements including the N-terminal FP or N-terminally-located fusion loop that is linked to HR1, a structurally variable segment or domain, the HR2, and the C-terminal TMD. X-ray crystallography of EBOV GP2 has revealed that HR1 of prefusion GP2 comprises 4 segments: a short trimeric coiled

coil formed by the C-terminal HR1D helix, which is connected to the N-terminal helical segments, HR1A and HR1B, via a random coil, HR1C (Figure 3.1A) (Lee *et al.*, 2008). In the fusion-activated or postfusion trimer of hairpins structure, HR1A, B and C form a continuous helical extension of the HR1D coiled coil (Malashkevich *et al.*, 1999; Weissenhorn *et al.*, 1998b), directing the fusion loop towards the target cellular membrane for insertion, and forming 3 hydrophobic grooves along the length of the coiled coil for the antiparallel packing of HR2 (Figure 3.1B).

The EBOV GP2 trimer of hairpins is strikingly similar to HTLV-1 gp21 (Figure 3.1C), the superpositioning of the 2 structures yielding a root mean square deviation of 1.8 Å for α-carbons within the coiled coil-CX_6CC-C-terminl region core (Kobe *et al.*, 1999; Malashkevich *et al.*, 1999; Poumbourios *et al.*, 1999; Weissenhorn *et al.*, 1998a). In these proteins, the CX_6CC motif forms a region of chain reversal at the base of the central coiled coil, while the C-terminal region packs against the coiled coil in an antiparallel orientation as a coil-helix-coil or "leash." By comparison, the structure of the gp41 trimer of hairpins has diverged from GP2 and gp21. While the structures of the GP2, gp21 and gp41 coiled coils are similar, the chain reversal region of gp41 is not observed in the available 3D structures due to disorder, and the antiparallel C-terminal region is completely α-helical (Yang *et al.*, 1999; Caffrey *et al.*, 1998; Weissenhorn *et al.*, 1997a; Chan *et al.*, 1997). The structural similarities between GP2 and gp21 and to a lesser degree gp41, led us to ask whether the structural activation mechanism of EBOV GP2 could be used to model retroviral TM protein activation.

The idea that the activation of retroviral TM proteins involves the N-terminal helical extension of a pre-existing trimeric coiled coil, as is the case for filoviruses, orthomyxoviruses and paramyxoviruses, is controversial due to the lack of 3D structural information for the prefusion form of TM. Current retroviral fusion activation models are

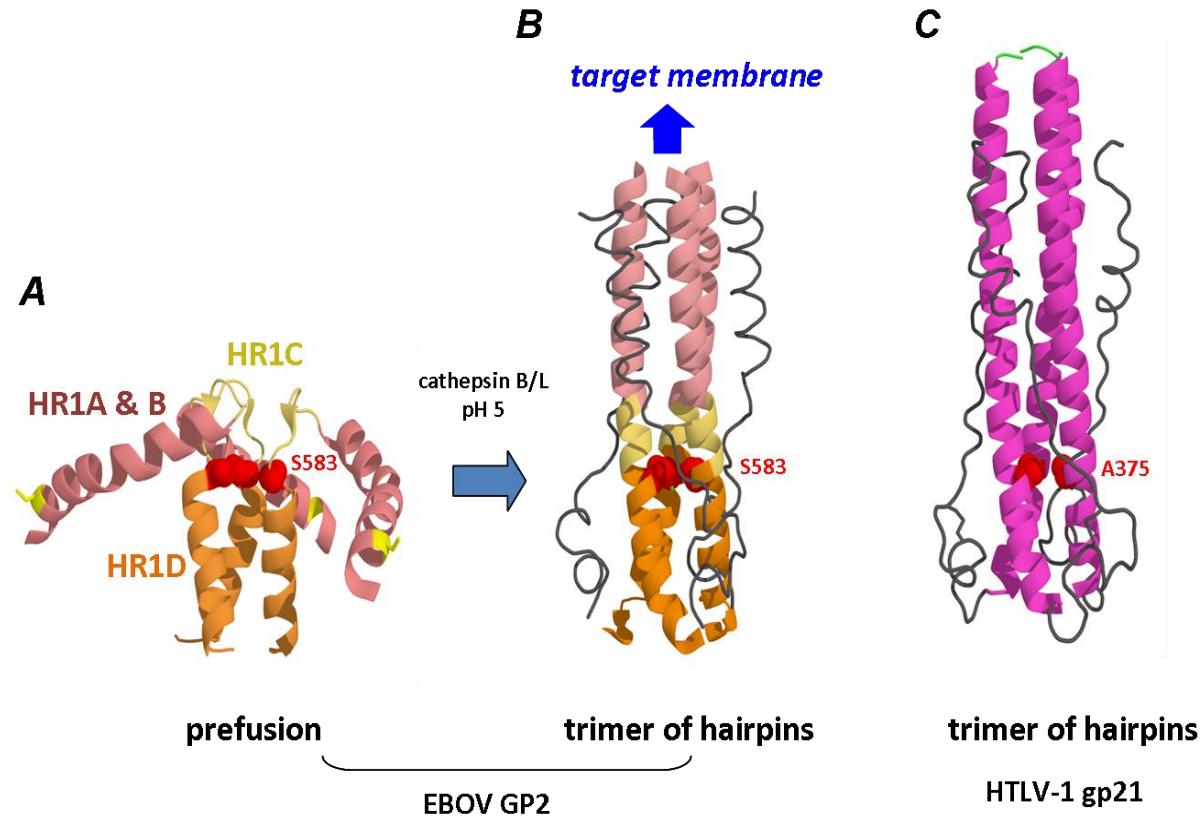


Figure 3.1 Conservation of the postfusion structure among the EBOV GP2 and gp21 of HTLV-1

(A) The GP2 prefusion structure of EBOV is comprised of four segments: three helical (HR1A, HR1B and HR1D) and the fourth segment has a loop structure (HR1C). Note that Serine 583 represents the most N-terminal residue in HR1D segment (PDB entry 3CSY). **(B)** Upon activation by low pH and cathepsins a series of conformational changes lead to the formation of the fusion-activated structure of GP2, the trimer of hairpins. A loop-to-helix transition by HR1C causes the central coiled coil to become extended in an N-terminal direction (PDB entry 1EBO). **(C)** The trimer of hairpins structures of GP2 and gp21 (PDB entry 1MG1) are very similar. Alanine 375 in gp21 is equivalent to Ser-583 in GP2. Figures drawn with Pymol.

based on early biochemical studies of HIV-1 gp41. The prevailing model is based on the initial studies of Wild *et al.* who proposed that the N-terminal HR1 sequences are not associated in the prefusion conformation but assemble into a coiled coil following receptor activation (Wild *et al.*, 1994). This model stems from the findings that the degree of helical structural disruption by substitutions in the central *a* position (Ile-573) of the HR1 correlated with the inhibition of fusion function, whereas precursor oligomerization was not affected by the same mutations. By contrast, alternative mutagenesis approaches indicated that the HR1 of gp41, and the hydrophobic heptad repeat therein, are critical determinants of Env oligomer formation (Poumbourios *et al.*, 1995; Poumbourios *et al.*, 1997; Center *et al.*, 1997). These data favour an alternative activation model for retroviruses that is analogous to that of EBOV with a preformed coiled coil within the prefusion TM that is extended towards the target membrane following activation.

In this chapter, I piloted a cysteine replacement mutagenesis approach to ask whether a coiled coil is present in the prefusion form of gp21, using the structurally homologous EBOV GP2 as a guide. We reasoned that the introduction of a cysteine to an *a* or *d* position within HR1 of gp21 should reveal the presence of a coiled coil in the prefusion trimer by forming an interchain disulfide, provided that the β-carbons of the introduced Cys residues are within 5.28 Å of each other (Reiter *et al.*, 1995; Sowdhamini *et al.*, 1989). Disulfide formation is indicated by the presence of a gp21 dimer in nonreducing SDS-PAGE.

A close inspection of prefusion EBOV GP2 showed that Ser-583 occupies an *a* position and represents the most N-terminal residue of the HR1D coiled coil (Figs. 3.1A; 3.2). Sequence and structural alignments of the GP2 and gp21 trimer of hairpins indicated that Ala-375 of gp21 is the structural equivalent of Ser-583 in a trimer of hairpins context. In a trimeric context, the β-carbon of Ala-375 is within 5.6 Å of the corresponding atom of

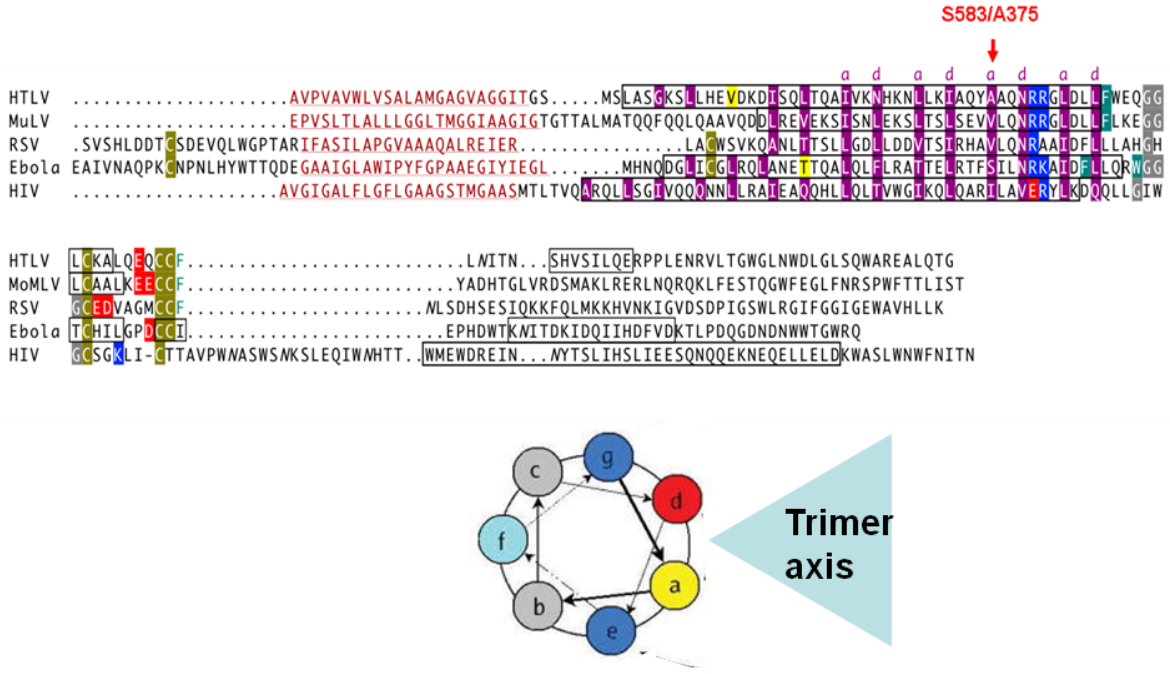


Figure 3.2 Alignment of retroviral TM protein sequences with EBOV GP2

The HR1 residues occupying *a* and *d* positions are highlighted with purple boxes. Serine-583 in EBOV GP2 and Ala-375 of HTLV-1 gp21 that was substituted with cysteine occupy analogous *a* positions. Fusion peptide or loop sequences are shown in red. Sequences predicted or shown to have α-helical structure are indicated by black boxes. Residues in the “a” and “d” positions of the heptad repeat are highlighted in magenta, and residues involved in irregular packing of the coiled coil are highlighted in yellow. In the chain-reversal region, acidic residues are highlighted in red, basic residues are highlighted in blue, hydrophobic residues involved in the formation of the hydrophobic core are highlighted in cyan, and glycines are highlighted in gray. Cysteines are highlighted in olive. The wheel below shows the heptad repeat in which the “a” and “d” positioned residues face inwards towards the trimeric axis. The figure was modified from (Poumbourios *et al.*, 1999).

an adjacent protomer and its small nonpolar side chain is associated with a small cavity in the core of the gp21 coiled coil. This residue therefore represents a good target for substitution with the small polar residue, Cys. By contrast, the remaining *a* and *d* positions of HR1 are occupied by the branched hydrophobic residues Leu and Ile, which form close knob-into-holes packing interactions, and the larger polar residues Asn-364 and Asn-378, which coordinate Cl^- ions in the coiled coil core (Kobe *et al.*, 1999). Molecular modelling suggests that Cys substitutions at these positions will disrupt such interactions due to cavity formation. The A375C mutation was therefore introduced into pCELT.1, which expresses fusion-competent HTLV-1 gp46.gp21, to determine whether an interchain disulfide could be mediated via the HR1 cysteine. The data indicate quantitative intermonomer disulfide formation for gp21-A375C consistent with the presence of a coiled coil in the prefusion TM.

3.3 Results

3.3.1 Prefusion model of HTLV-1

3.3.1.1 Expression and processing of gp62 protein

The pCELT.1 plasmid expresses the precursor protein gp62 which is further cleaved by cellular proteases to the SU, gp46, and TM, gp21; gp21 contains a C-terminal epitope tag that is recognised by mAb C8 (Maerz *et al.*, 2001; Lee *et al.*, 1984). It was first necessary to establish that the cysteine mutant is expressed and processed at wild-type levels. The WT and A375C Env-expressing plasmids were transfected into 293T cells and the cell lysates were harvested according to the protocol in section 2.1.5; an empty vector (pCDNA3.1) was also included as a control. The mouse mAbs, 67/5.5.13.1 (directed to gp46) and C8 (directed to gp21), were used in a western blot under reducing conditions (100 mM DTT) to detect the gp62 precursor and the cleavage products gp46 and gp21 respectively. Figure 3.3A & B shows that that the presence of the cysteine did not affect gp62 expression and processing to gp46 and gp21, as wild-type-equivalent levels of these species were observed for A375C. The antibody specificity was demonstrated by the absence of protein bands in the mock (pCDNA3.1) transfection.

Having confirmed that A375C was successfully expressed in 293T cells, we next tested whether this construct was able to mediate membrane fusion. Wild type, A375C and the fusion incompetent mutant, G390P (Maerz *et al.*, 2000) were cotransfected with the luciferase-reporter construct pT4*luc* (Maerz *et al.*, 2001) into 293T cells, while HeLa targets were transfected with the bacteriophage T7 RNA polymerase expression vector, pCAG-T7. At 24 h posttransfection, effectors and targets were cocultured for a further 18 h to enable fusion and expression of the luciferase reporter. Reporter gene induction requires the formation and expansion of a fusion pore to enable the T7 polymerase to access the T7 promoter, which controls luciferase expression. Figure 3.3C shows

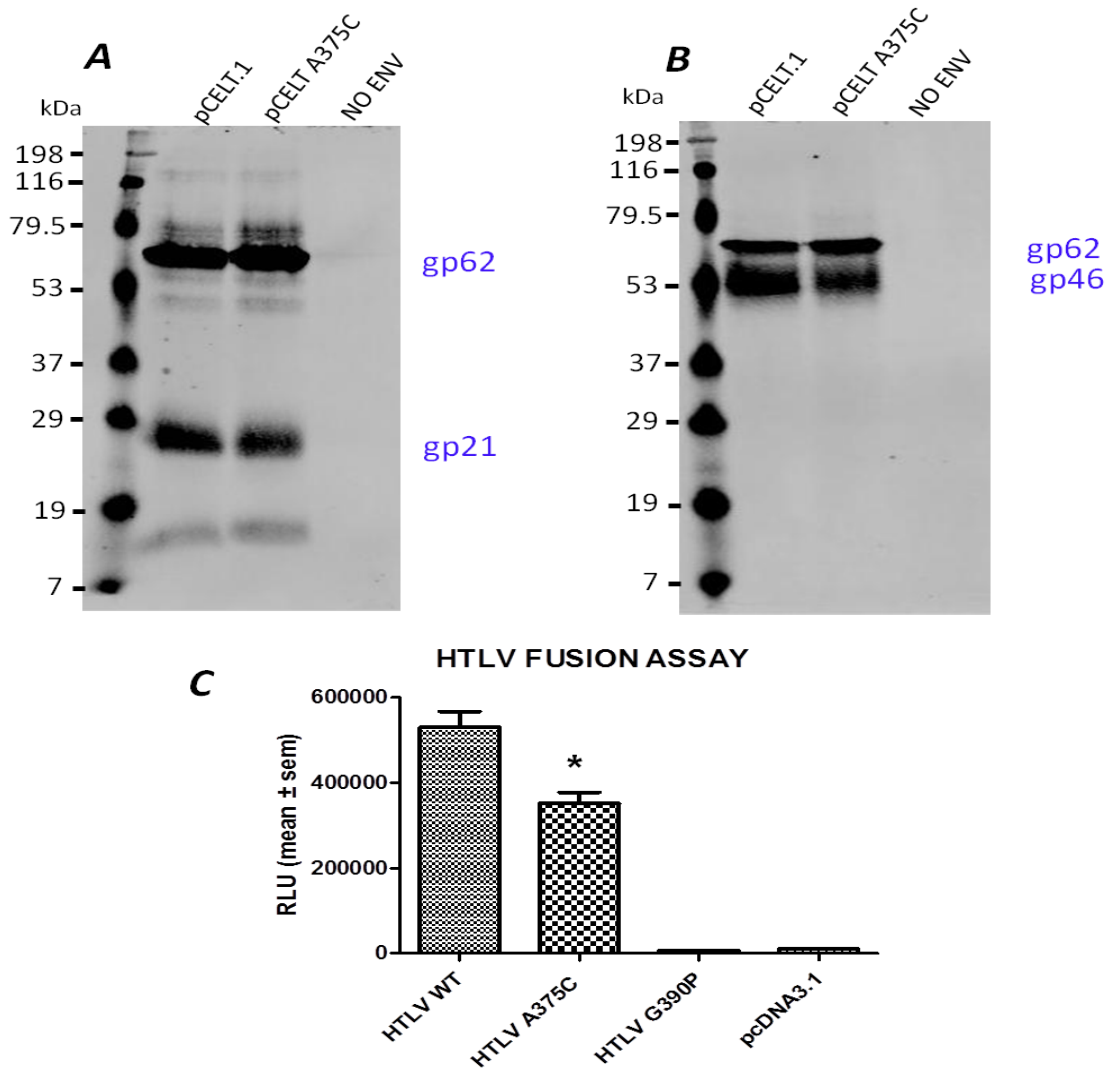


Figure 3.3 Expression, processing and fusion activity of A375C

293T cells were transfected with WT and A375C-mutated pCELT.1 vectors or the pCDNA3.1 vector control (No Env). At 48 h posttransfection, the cells were lysed subjected to SDS-PAGE and Western blot with **(A)** mAb C8, directed to the epitope tag at the C-terminus of gp21 or **(B)** mAb 67/5.5.13.1 directed to gp46. The data are representative of two independent experiments. **(C)** Cell-cell fusion assay. 293T effector cells were contransfected with pCAG-T7 and WT or mutated pCELT.1 vectors, while HeLa targets were transfected with the luciferase reporter plasmid, pT4luc. The effectors and targets were cocultured for 24 h prior to lysis and luciferase assay. The data represent the mean relative light units (RLU) ± standard error from 3 independent experiments. *, $P < 0.05$; 2-sample t-test assuming unequal variances.

substantial fusion activity for WT with an ~ 30 % reduction in activity for A375C ($P < 0.05$, 2 sample t-test assuming unequal variances). By contrast, G390P lacked detectable fusion activity above the empty vector (pCDNA3.1) control. Overall, these data indicate that the A375C mutant folds into a membrane fusion-competent structure.

3.3.1.2 DTT titration and quantification of dimer bands in pCELT$_{A375C}$

The A375C-mutated Env was expressed and processed successfully to gp46 and gp21 and able to mediate fusion. The next step was to determine whether Cys-375 could form an interchain disulfide bond. 293T cells expressing WT and A375C Envs were lysed in the presence of the alkylating agent, iodoacetamide to limit the possibility of postlysis thiol isomerization by free sulhydryls. The alkylated lysates were boiled in the presence of different concentrations of DTT and then electrophoresed in SDS-12 % polyacrylamide gels. gp21 monomers and dimers were detected with mAb C8 in a western blot (Center *et al.*, 1998). Two prominent bands corresponding to epitope-tagged gp62 and gp21 were observed for the WT at all DTT concentrations, whereas an additional ~ 48 kDa band was observed for A375C, a molecular weight that closely approximates that expected for the epitope-tagged gp21 dimer (~ 24,500 Da per monomer) (Figure 3.4). The intensity of the putative dimer band corresponded to ~ 55 % of total gp21 at 0 mM DTT. Step-wise decreases in dimer band intensity concomitant with increases in monomer band intensity were observed with increasing DTT, consistent with disulfide-mediated dimerization of gp21 via Cys-375 (Figure 3.4). These findings suggest that the cysteines at position 375 within the trimeric coiled coil are close enough to form a disulfide bond. The percentage dimer observed at 0 mM DTT was close to the expected value of 66% if in every trimer two out of three cysteines formed a disulfide while the third is left free. Thus disulfide formation is highly efficient in this context. The data are consistent with a coiled coil being present in the prefusion conformation of gp21.

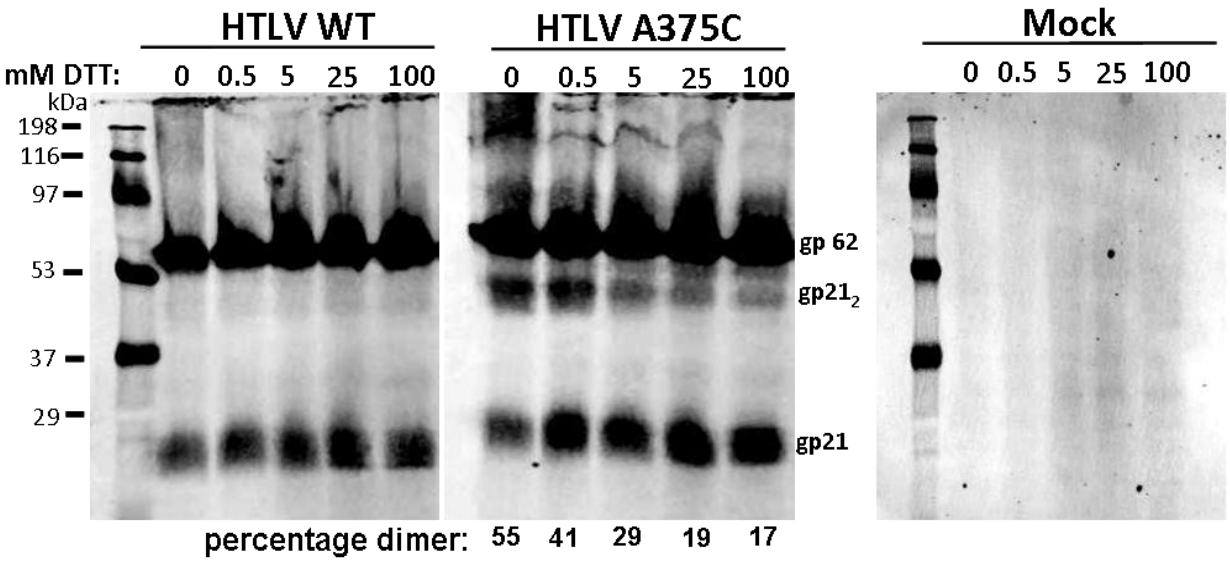


Figure 3.4 Disulfide-mediated dimerization of gp21 via Cys-375

293T cells were transfected with pCELT.1, pCELT$_{A375c}$ or pcDNA3.1 (mock) expression vectors. At 48 h post-transfection, the cells were lysed in the presence of 5 mM iodoacetamide, boiled in the presence of SDS and various concentrations of DTT prior to SDS-PAGE in 12 % polyacrylamide gels. The proteins were visualized by Western blotting with mAb C8 and Alexa Fluor 680-conjugated goat anti-mouse Ig. The A375C mutant shows a putative dimer band at ~48 kDa (gp21$_2$), which was reduced gradually with increasing concentration of DDT. The relative intensities of the putative dimer band (gp21$_2$) in relation to the total amount of gp21 at the various DTT concentrations are presented as percentages below the A375 blot.

3.3.2 Postfusion model of HTLV-1

3.3.2.1 The MBP/gp21 (338-425) chimera acquires trimeric structure

The 2.5 Å-resolution crystal structure of MBP/gp21(338-425) revealed the fusion-activated trimer of hairpins conformation of gp21 (Kobe *et al.*, 1999). We used this protein to examine the distance constraints on disulfide formation as the β-carbon of Ala-375 is 5.6 Å away from a corresponding atom of an adjacent protomer, which is 0.3 Å outside the optimal disulfide bonding distance. Wild type and A375C-mutated MBP/gp21(338-425) chimeras were expressed in *E. coli,* affinity-purified using amylose–agarose and the trimers isolated using the AKTA-FPLC Superdex 200 (HiLoad 26/60) column equilibrated in S-buffer. The resultant WT chimera is expected to be trimeric (Center *et al.*, 1998), while the effect of the cysteine substitution on the protein profile is not known. The purified MBP/gp21 chimeras were concentrated by centrifugation and then 100 µg of each protein was reanalysed by gel filtration using an analytical Superdex 200 HR (10/30) column. The gel filtration profiles shown in figure 3.5A show that the WT and A375C both eluted as trimers (elution time of 28 minutes), indicating that MBP/gp21(338-425) is a valid tool for studying the effect of the cysteine substitution on the postfusion conformation.

3.3.2.2 Absence of disulfide links in MBP/gp21 A375C trimer of hairpins

We next asked whether Cys-375 could form an interchain disulfide in the context of the fusion-activated trimer of hairpins conformation. SDS-PAGE analysis of the purified MBP/gp21 WT and A375C oligomers in the presence of 3 % β-mercaptoethanol showed a single band migrating at ~53 kDa (Figure 3.5B), the expected molecular weight of MBP/gp21(338-425). Predominantly monomeric protein was also observed in the absence of β-mercaptoethanol with a low level (< 1 %) of dimeric protein for both WT and A375C mutant. The low level of dimer could be attributed to the free sulfhydryl group

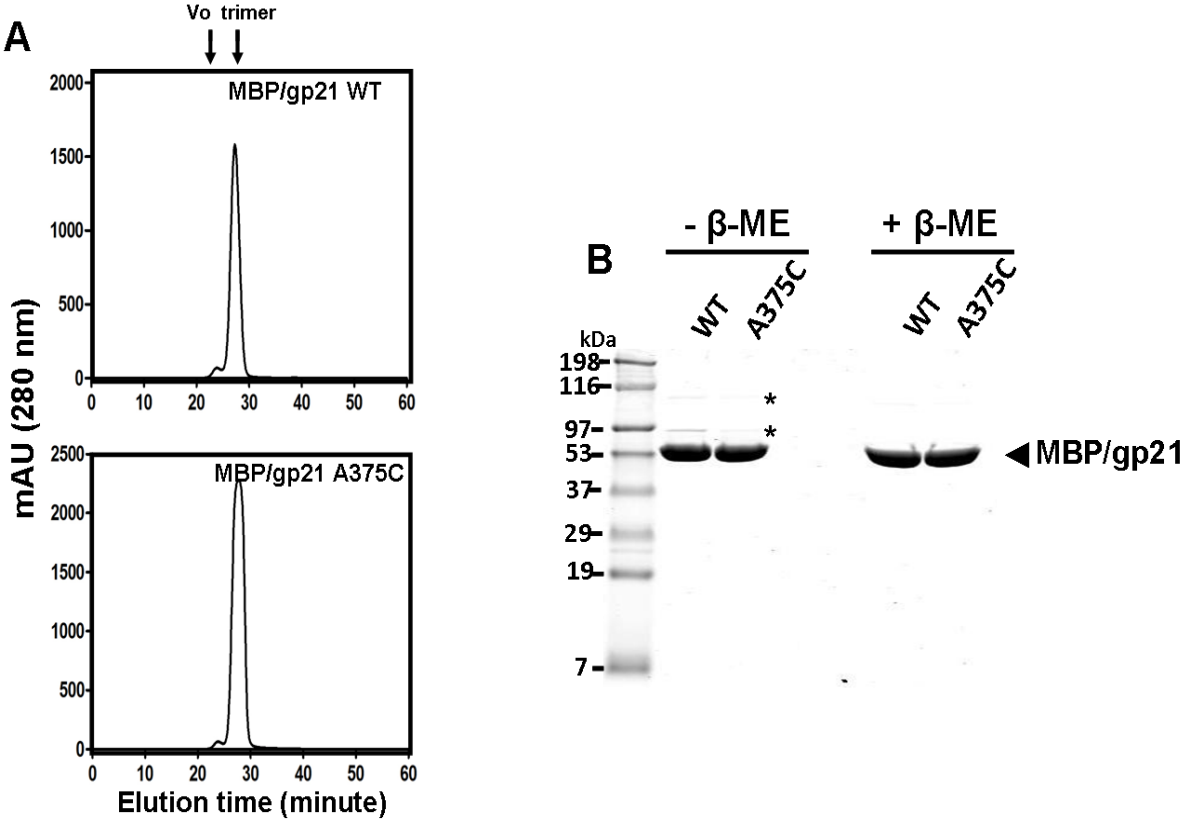


Figure 3.5 Gel filtration and SDS-PAGE profiles of MBP/gp21(338-425)-A375C

(A) Gel filtration profiles of MBP/gp21(338-425) (top) and MBP/gp21(338-425)-A375C (bottom). The protein samples were analysed by Superdex 200 HR (10/30) gel filtration at a flow rate of 0.5 ml/min. The peak elution time for WT and A375C is 28 min, which corresponds to trimer. Vo: void volume. **(B)** SDS-PAGE under reducing (+ βME) and non-reducing (- βME) conditions of gp21(338-425) and MBP/gp21(338-425)-A375C. A minor dimeric species observed under nonreducing conditions is indicated by an asterisk. Its presence does not correlate with the introduction of Cys-375 to the gp21 coiled coil. The lower asterisk marks a nonspecific protein that persists despite the purification procedure.

of Cys-401 or residual oligomer that resists disruption by the SDS (Center *et al.*, 1998; Kobe *et al.*, 1999). The presence of Cys-375 did not result in quantitative disulfide stabilization of dimer, in contrast to the results obtained with prefusion gp21. The inappropriate distance (more than 5.28 Å) and the structural rigidity of the coiled coil in the trimer of hairpins context (Kobe *et al.*, 1999) could be the reason for the lack of pairing of the cysteines in the MBP/gp21-A375C chimera.

3.4 Discussion

The obvious structural similarities between EBOV GP2 and retroviral TMs in the fusion-activated conformation (Kobe *et al.*, 1999; Caffrey *et al.*, 1998; Chan *et al.*, 1997; Fass *et al.*, 1996; Malashkevich *et al.*, 1998; Tan *et al.*, 1997; Weissenhorn *et al.*, 1997b) prompted us to ask whether the recently elucidated prefusion structure of GP2 (Lee *et al.*, 2008) can be used to model the prefusion state of its closest retroviral homologue, gp21. We piloted the use of Cys-substitution mutagenesis to probe the proximity of HR1 sequences in the gp21 trimers. Alanine-375 of gp21 was substituted with cysteine because a structure-based alignment indicated that it corresponds to Ser-583 of prefusion EBOV GP2, which is the most N-terminal *a*-position residue of the coiled coil formed by HR1D. We observed efficient Cys-375-dependent disulfide-linked stabilization of gp21 dimer, which strongly suggests that within the trimer, HR1 segments about residue 375 are in close proximity to each other ($\leq$ 5.28 Å for β-carbons). The gp21 was processed from gp62 in 293T cells that in our hands are not permissive for HTLV-1 Env-mediated fusion (Poumbourios *et al.*, unpublished data). Hence, we expect that this form of gp21 is in an untriggered or prefusion conformation. These data support a model whereby the C-terminal part of the coiled coil up to at least residue 375 is formed in the prefusion conformation, and represents the building block for receptor-triggered structural transitions that are associated with membrane fusion function. This model for retroviral TM protein

activation is therefore analogous to that formulated for other class I fusion proteins based on x-ray crystallographic analysis of pre- and fusion-activated conformations (Wilson *et al.*, 1981; Bullough *et al.*, 1994; Lee *et al.*, 2008; Weissenhorn *et al.*, 1998b).

Inter-monomer disulfide formation was not observed in the MBP/gp21(338-425) fusion-activated trimer of hairpins. The most likely reason for this observation is that the cysteines at position 375 are too far apart to form a disulfide bond. This idea is supported by the fact that the β-carbon of Ala-375 is 5.6 Å distant from the corresponding atom of an adjacent monomer. Also, being in a central location within the trimer of hairpins, which is stabilised by both knob-into-holes packing interactions in the coiled coil interface and interactions between the coiled coil and C-terminal leash, the backbone about the cysteines might not have sufficient flexibility to enable the thiols to come within range for disulfide bond formation (Caffrey *et al.*, 1998). Our finding that prefusion gp21 can form an intermonomer disulfide through Cys-375 suggests that the helices forming the coiled coil are closer together at this position when compared to the trimer of hairpins.

Unlike other retroviruses, cell-free HTLV-1 virions are poorly-infectious, therefore transmission is believed to occur between infected and uninfected cells across a virological synapse (Jolly and Sattentau, 2004; Fan *et al.*, 1992; Donegan *et al.*, 1994; Igakura *et al.*, 2003). The binding of gp46 to the HTLV-1 entry receptors, glucose transporter protein I and neurophilin-1 (Manel *et al.*, 2003; Ghez *et al.*, 2006), is believed to lead to conformational changes in SU that trigger the reduction of a labile disulfide formed between SU and TM (Li *et al.*, 2008), and the refolding of gp21 into a fusion-active form. The data presented here suggest that the conformational changes accompanying the fusion function of gp21 are analogous to those observed for EBOV GP2. These changes include the N-terminal helical extension of a coiled coil such that the fusion peptide approaches the target cellular membrane for insertion, and the antiparallel

packing of the C-terminal leash onto the coiled coil such that the virus and cell membranes are apposed for fusion (Kobe *et al.*, 1999). The implications of these findings for other retroviruses such as HIV-1 will be discussed in greater depth in Chapter 4.

Chapter 4: Probing the Conformational States of HIV gp41 Using Cysteine Replacement Mutagenesis

4.1 Aim

To determine if structural changes occurring in gp41 of HIV-1 during activation can be modelled on those observed for EBOV, GP2 and HTLV-1 gp21.

4.2 Introduction

The binding of gp120 to CD4 and a co-receptor, CCR5 and/or CXCR4, leads to a conformational change that both exposes and induces rearrangement of gp41 such that it moves closer to the target cell and adopts a pre-hairpin intermediate conformation. It is thought that the fusion peptide becomes inserted into the target cell membrane at this stage. The antiparallel packing of the HR2 onto the central coiled coil to form the 6-HB brings together the cellular and viral membranes and leads to their fusion. The fusion-activated or postfusion structure of gp41 is similar to that of HTLV-1 gp21 (Kobe *et al.*, 1999), EBOV GP2 (Weissenhorn *et al.*, 1998b) and influenza virus HA2 (Bullough *et al.*, 1994) with all sharing an N-terminal coiled coil, a region of chain reversal, and an antiparallel C-terminal segment. These structural similarities suggest a common membrane fusion mechanism.

In the previous chapter we modelled the activation mechanism of HLTV-1 gp21 on that of the GP2 of EBOV, and found evidence that the C-terminal segment of the gp21 coiled coil is formed in the prefusion conformation as is the case for the analogous HR1D segment in GP2. This evidence was obtained by substituting a Cys residue into an *a* position of the gp21 coiled coil and demonstrating the formation of a disulfide bond in the context of the prefusion conformation. The overall similarities between gp41, gp21 and GP2 fusion-activated trimer of hairpins structures and the positive results achieved with HTLV-1 encouraged us to use the same methodology to assess the various conformational states of HIV gp41 as they relate to membrane fusion function.

Superpositioning of the closely homologous gp21 and GP2 trimer of hairpins structures indicated that Ala-375 of gp21 was a good target for Cys replacement because it corresponds to Ser-583, the N-terminal *a*-position residue of the HR1D coiled coil in prefusion GP2. However, it was not possible to identify a Ser-583 homolog in gp41 for Cys replacement mutagenesis because the structure of the gp41 trimer of hairpins has diverged from gp21 and GP2. Cysteines were therefore substituted into 6 *a* or *d* positions occupied by hydrophilic residues (Ser-534, Thr/Ser-538, Ala-541, Gln-552, Gln-562 and Thr-569) along the length of the gp41 polar segment and HR1 (Figure 4.1). Interprotomer disulfide formation was examined in 2 expression systems: 1) A bacteriophage T7 promoter-driven HIV-1_{BH8} Env expression vector, pTMenv.2. The cotransfection of mammalian cells with pTMenv.2 together with a T7 polymerase expression vector (pCAG-T7) leads to the expression of functional gp120-gp41 complex. The outward-facing *b*-position hydrophilic residue (Gln-567) was substituted with Cys to control for *a*- and *d*-position-independent disulfide formation. 2) An *E. coli*-based system that expresses the HIV-2_{ST} gp41 ectodomain (minus the fusion peptide) in the low energy trimer of hairpins conformation as a chimera with MBP (MBP/gp41[528-L-677]). In this case, the outward-facing *f*-position residue, Thr-536, was substituted with Cys as the control.

The fusion competent mutant, T569C, was able to form an inter-protomer disulfide in gp160-derived gp41, indicating that the C-terminal segment of the coiled coil is likely to be formed prior to receptor-induced activation, as is the case with EBOV GP2 and HTLV-1 gp21. By contrast, the Cys-569-mediated disulfide was not formed in MBP/gp41(528-L-666) consistent with an inter-Cβ distance of 5.4 Å for this position, which is greater than the optimal separation required for disulfide bonding of 5.28 Å. Instead, interprotomer disulfide linkages were formed following a S538C mutation in the polar segment of the MBP/gp41(528-L-677) trimer of hairpins model and also in a model

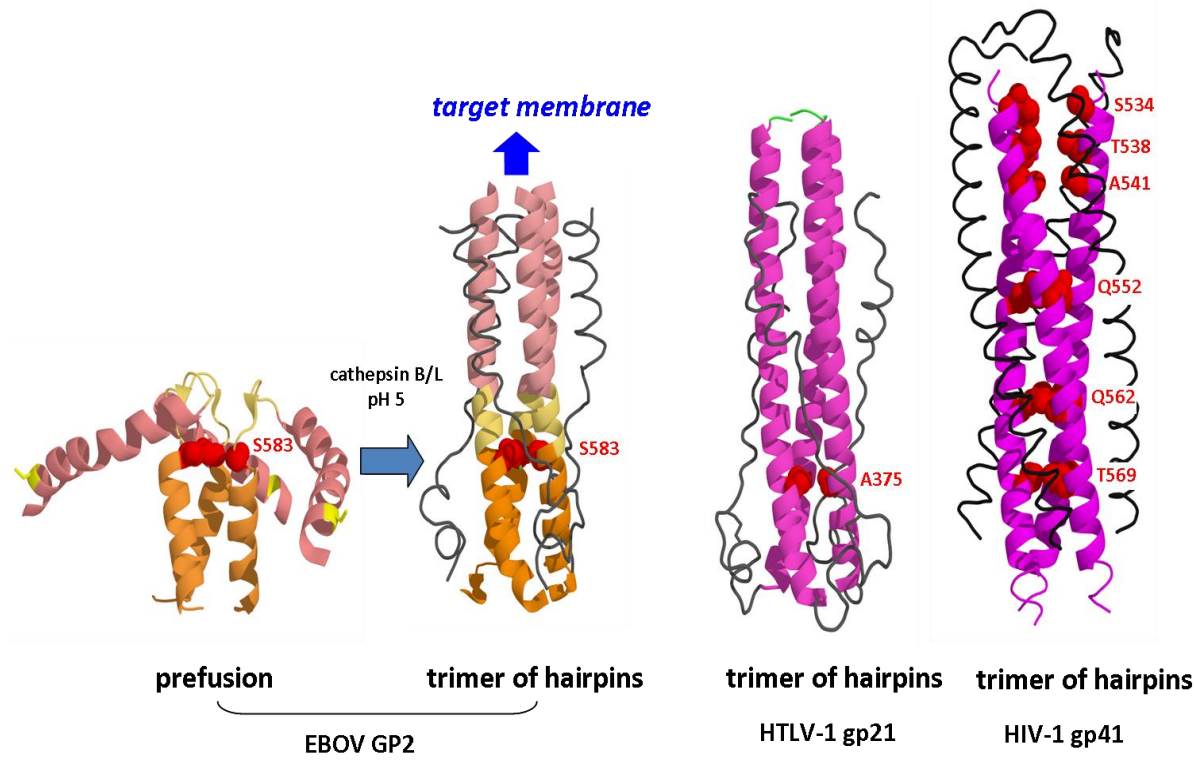


Figure 4.1 Conservation of the postfusion structure among the Ebola virus GP2, gp21 of HTLV-1 and gp41 of HIV-1

Ser-534, Thr-538, Ala-541, Gln-552, Gln-562 and Thr-569 represent the Cys substituted residues in gp41 coiled coil.

of the gp41 prehairpin intermediate, MBP/gp41(528-591)$^{AD.ST}$. The data suggest that the C-terminal segment of HR1 forms a coiled coil in prefusion gp41. Activation results in the propagation of helical structure in an N-terminal direction to include the polar segment, prior to formation of the 6-HB.

4.3 Results

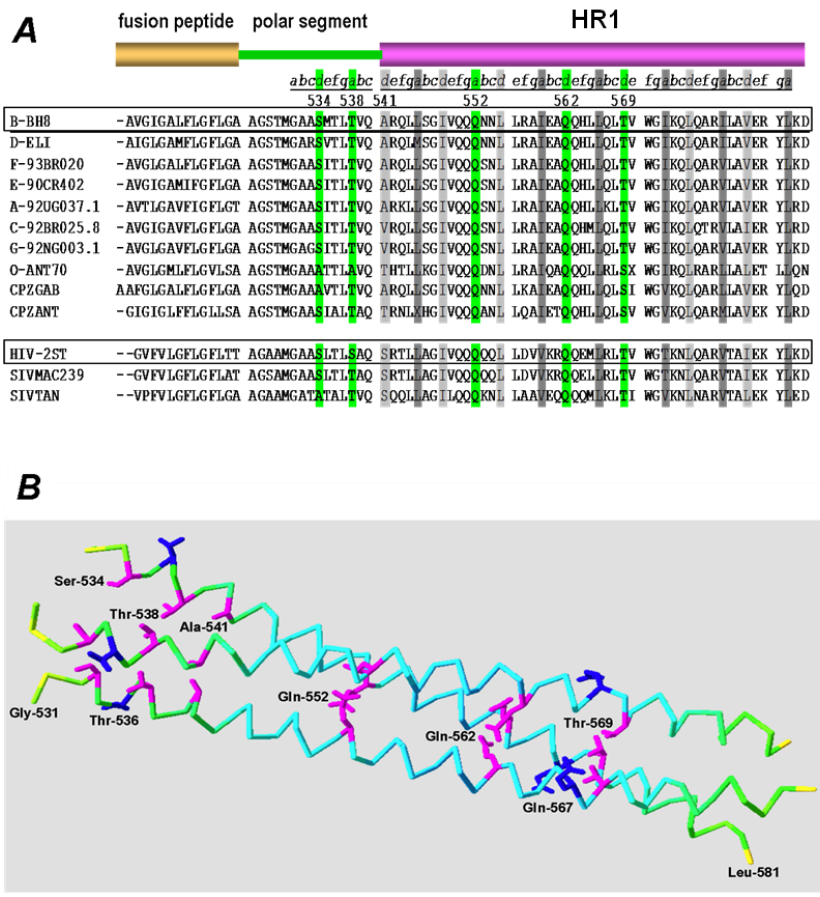
4.3.1 gp160-derived gp41: the prefusion conformation

4.3.1.1 Cysteine scanning mutagenesis of gp41

An alignment of gp41 polar segment-HR1 sequences derived from HIV-1, HIV-2 and selected nonhuman primate immunodeficiency viruses is shown in figure 4.2A. Conserved hydrophilic residues present in *a* or *d* positions of the heptad repeat that were selected for substitution with cysteine are highlighted in green. The small hydrophobic sidechain of Ala-541 was also targeted. We avoided changing *a*- and *d*-position hydrophobic residues as they are involved in knob-into-holes packing interactions and cavities introduced at these sites lead to loss of function (Poumbourios *et al.*, 1997). The structure of the coiled coil formed by the polar segment-HR1 helices present in the gp41 trimer of hairpins (PDB ID 2X7R), the *a*- and *d*-position residues targeted with Cys substitutions and interprotomer distances between β-carbons are shown in figure 4.2B. Note that in the gp41 trimer of hairpins, the interprotomer distance between Cβ atoms of residues selected for Cys substitution are larger than 5.28 Å, which is optimal for disulfide-bonding with 10.8 Å for Ser-534, 9.8 Å for Thr-538, 8.4 Å for Ala-541, 7.0 Å for Gln-552, 6.2 Å for Gln-562 and 5.4 Å for Thr-569 (Figure 4.2B).

4.3.1.2 Expression and processing of Cys-substituted Env glycoproteins

The mutations were initially introduced into the pTMenv.2 plasmid. This plasmid directs expression of the precursor protein gp160, which is transported to the *Golgi*



		Ser-534	Thr-536	Thr-538	Ala-541	Gln-552	Gln-562	Gln-567	Thr-569
Position in heptad		*d*	*f*	*a*	*d*	*a*	*d*	*b*	*d*
Cβ - Cβ distance (Å)		10.8	21.2	9.8	8.4	7.0	6.2	15.5	5.4
B-value	Cα	54.02	46.87	45.09	39.87	23.15	14.71	21.70	22.86
	Cβ	58.59	49.81	45.70	36.42	22.12	13.92	20.14	21.42
	γ	56.35	51.10 49.54	49.81(Oγ1) 40.47(Cγ2)		28.43	18.58	24.91	22.05(Oγ1) 26.55(Cγ2)

Figure 4.2 Cysteine substitution locations in the polar segment-HR1 coiled coil of gp41

(A) Alignment of gp41 polar segment-HR1 amino acid sequences. The BH8 and HIV-2_{ST} sequences are the main concern of this study. The yellow segment represents the fusion peptide while the green and magenta regions represent the polar segment and HR1 respectively. The numbers show the sites where the cysteine mutations were introduced. **(B)** The polar segment-HR1 coiled coil is derived from the Buzon *et al.* 2010 structure (PDB entry 2X7R). It is colorcoded according to *B*-value. The color gradient is low = blue➔green➔yellow➔orange➔red = high. Below is a table listing heptad position, Cβ-Cβ distances and *B*-values for the residues mutated. The *B*-values were obtained from the 2X7R coordinates.

complex where it is cleaved by furin into an extracellular glycosylated SU, gp120, and a TM, gp41 (Stein and Engleman, 1990; Earl *et al.*, 1991; Willey *et al.*, 1988). It was first necessary to establish that the cysteine mutants are expressed and processed to gp120 and gp41 at levels similar to the WT. The WT, S534C, T538C, A541C, Q552C, Q562C and T569C Env-expression plasmids were transfected into 293T cells and the Env-containing lysates harvested according to the protocol described in section 2.2.1.9; the pTM.1 empty vector (No Env) was included as a negative control. The mouse mAb C8, directed to the gp41 cytolasmic tail, and the sheep polyclonal antibody DV012 directed to gp120 were used in a western blot to detect the gp160 and the cleavage products gp41 and gp120, respectively. Figure 4.3 shows that the mutated gp160 precursors were expressed and processed to gp41 and gp120. The S534C, T538C and A541C mutants were processed in a similar manner to WT, whereas decreased processing was observed for Q552C, Q562C and T569C, consistent with slower translocation to the *Golgi* where processing occurs. The absence of protein bands in the mock transfection lanes (pTM.1) demonstrated the band specificity. We note that an ~ 80 kDa species, which corresponds in molecular weight to a gp41 dimer, was observed in the T569C lane (asterisk) (Figure 4.3).

4.3.1.3 Cell-cell fusion activity of Cys substitution mutants

The next step was to demonstrate that the cysteine mutants are capable of mediating membrane fusion. 293T cells coexpressing Env and T7 polymerase were co-cultured with BHK cells expressing both CD4 and CXCR4, and harbouring a T7 promoter driven luciferase reporter. Cell-cell fusion enables T7 polymerase to transcribe luciferase, which can be used to measure fusion activity, as detailed in section 2.2.1.10. The S534C and A541C mutants retained ~ 80 % fusion activity relative to WT, whereas T538C and T569C had 40% and 50% fusion activity, respectively (Figure 4.4). The other two

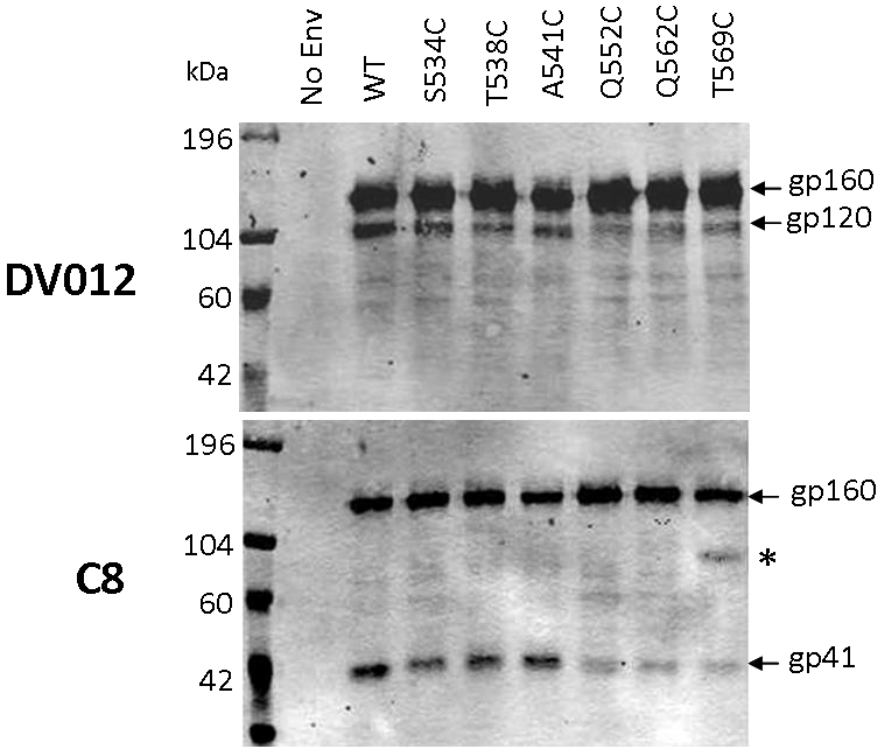


Figure 4.3 Western blot of HIV-1 Env cysteine mutants

Wild type and mutated pTMenv.2 vector-transfected 293T cells were lysed and subjected to reducing SDS-PAGE and electrophoretic transfer to nitrocellulose. The top panel was probed with DV012 (directed against gp120) and Alexa fluor 680-conjugated donkey anti-sheep Ig. The bottom panel was probed with mAb C8 (directed against the gp41 cytoplasmic tail) and Alexa fluor 680-conjugated goat anti-mouse Ig. The precursor protein gp160 expression is equal for WT and all cysteine mutants but processing to gp120 and gp41 is reduced for Q552C, Q562C and T569C. The pTM.1 empty vector was used as a negative control (No Env). The asterisk in the C8 blot may represent an incompletely reduced gp41 dimer.

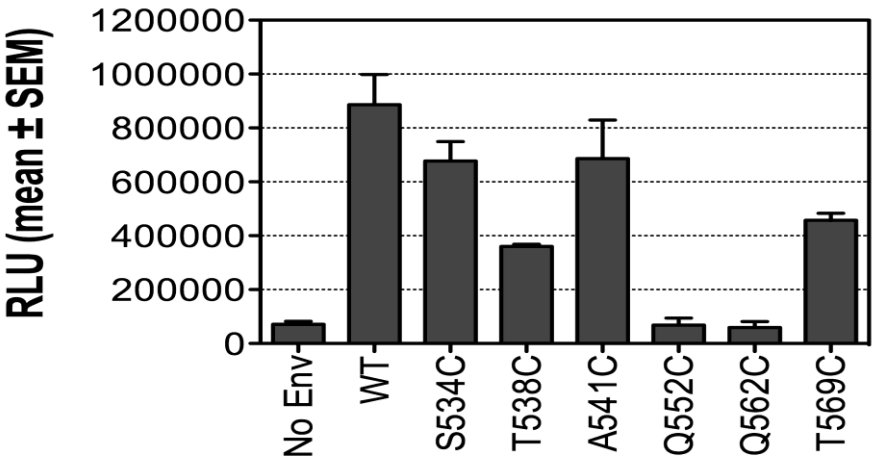


Figure 4.4 Cell-cell fusion activity of cysteine mutants

293T effector cells were cotransfected with the T7 polymerase expression vector, pCAG-T7, together with pTMenv.2 vectors. BHK21 target cells were cotransfected with pT4luc and pc.CXCR4. At 18 h posttransfection, the effectors and targets were mixed and co-cultured for 24 h prior to lysis and assay for luciferase activity. Mean relative light units (RLU) ± standard error is shown from 3 independent experiments.

mutants, Q552C and Q562C, showed a fusion level equivalent to that of the empty vector control. The Q552C and Q562C mutants were excluded from further analysis because their lack of fusion function suggests that they may form non-native structures. Having confirmed that S534C, A541C, T538C and T569C retained significant levels of fusion activity, we used these mutants to put our hypothesis to the test by assessing their ability to form inter-protomer disulfide bonds in prefusion gp41.

4.3.1.4 Assessment of interprotomer disulfide formation by gp41 Cys mutants

The WT and cysteine mutant Env glycoproteins were expressed in 293T cells. The clarified cell lysates were then subjected to SDS-PAGE in 10 % polyacrylamide gels in the presence and absence of DTT. The western blot was probed with both mAb C8 (directed to gp41) and DV012 (directed to gp120). The C8 antibody was used to detect any putative dimer formed by the cysteine mutants, whereas DV012 was used as a control to confirm that the dimer band (if present) is not simply due to degradation of gp160 or gp120. Under non-reducing conditions, prominent bands corresponding in molecular weight to gp160 and gp41 were observed for the WT and all mutants examined (Figure 4.5A). An additional ~ 80 kDa band, which corresponds to the predicted molecular weight of gp41 dimer, was detected in the T569C lane but not in the WT lane. The ~ 80 kDa band was also observed at trace levels in the S534C, T538C and A541C lanes. The T569C ~ 80 kDa band was largely reduced with 100 mM DTT, suggesting that it is a disulfide linked dimer of gp41 (Figure 4.5B). Under both nonreducing and reducing conditions, a low level of background staining was observed in the vicinity of the ~ 80 kDa band, however these species were also present for WT and the other mutants suggesting that they are break-down products of gp160. The use of DV012 in the western blot revealed strong bands corresponding to gp160 and gp120 and several faint bands corresponding to degradation products of these proteins (Figure 4.5C). Importantly, the DV012 antibody did not detect a

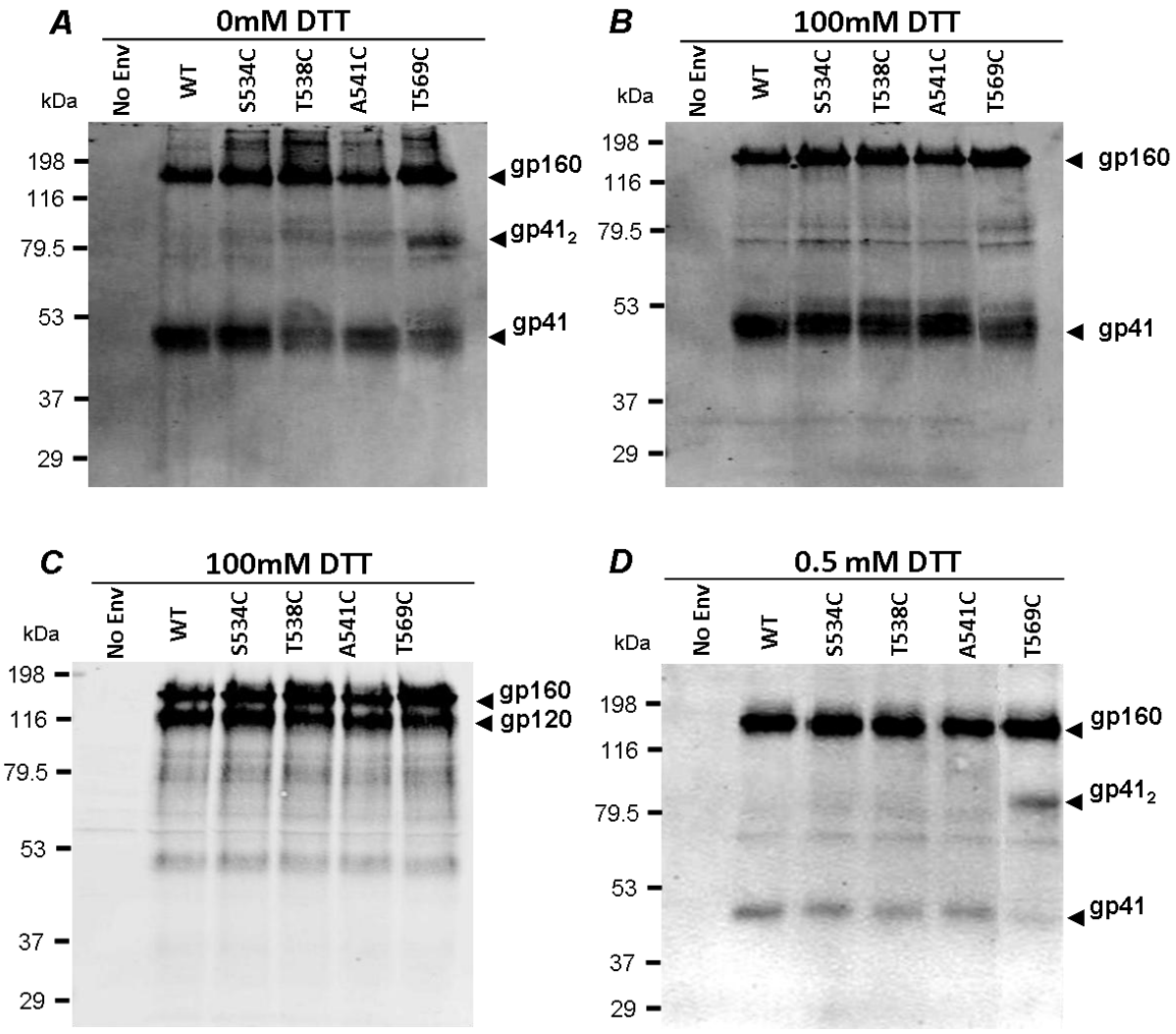


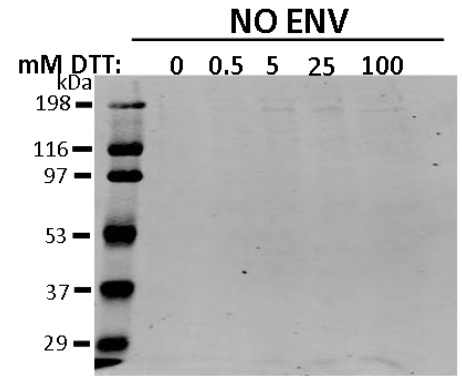
Figure 4.5 SDS-PAGE and western blot of cysteine mutants

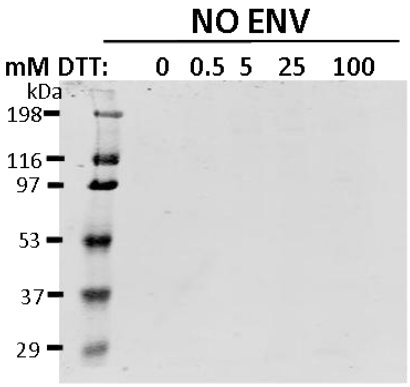
Lysates of 293T cells expressing functional cysteine mutants were boiled in the presence of the indicated concentrations of DTT and then subjected to SDS-PAGE (10 % polyacrylamide gels) and western blotting with mAb C8 (A, B, D) and DV012 (C). The blots are representative of three independent experiments.

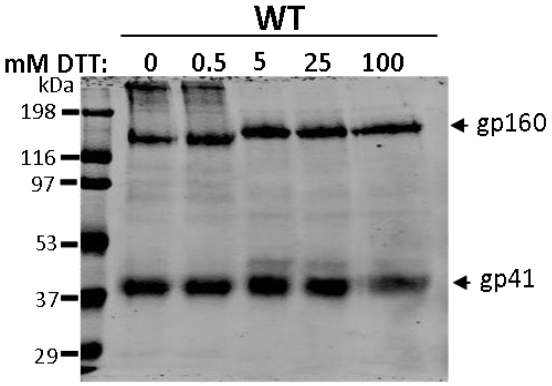
prominent T569C-specific ~ 80 kDa band, indicating that the ~ 80 kDa species detected by mAb C8 under non-reducing conditions is likely to be a disulfide linked dimer of gp41 mediated through Cys-569. Finally, we found that 0.5 mM DTT was sufficient to reduce the trace amounts of ~ 80 kDa species observed for S534C, T538C and A541C under non-reducing conditions, whereas the T569C-specific ~ 80 kDa species was not affected by the low concentration of reducing agent (Figure 4.5D). These data suggest that the ~ 80 kDa species observed for S534C, T538C and A541C is due to the inefficient formation of a tenuous disulfide that can be reduced by a low concentration of reducing agent. By contrast, disulfide formation by Cys-569 appears to be efficient and robust because it appears to be partially resistant to reduction even at 100 mM DTT.

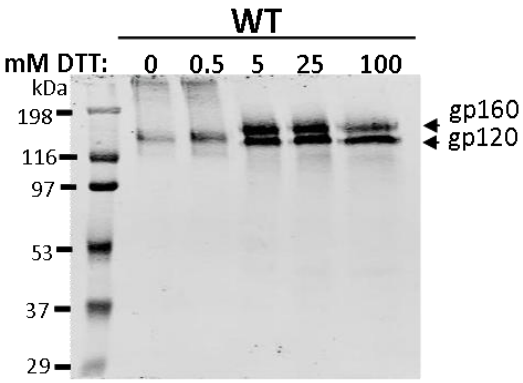
4.3.1.5. Reduction of the Cys-569-mediated interprotomer disulfide requires the presence of urea

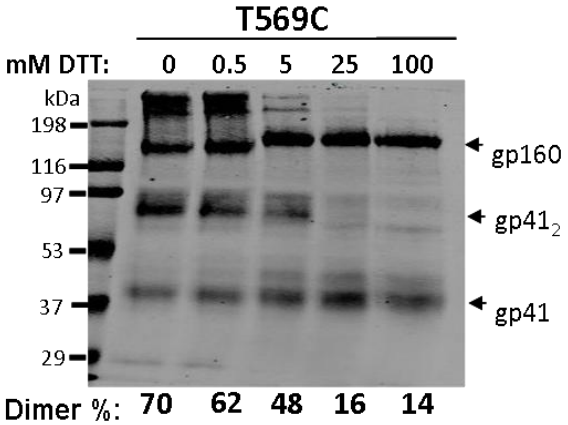
The 80 kDa band observed for T569C was not completely reduced following boiling in the presence of 100 mM DTT and 2 % SDS. We reasoned that the persistence of this species could be due to the maintenance of residual α-helical 2° structure and therefore quaternary structure in the vicinity of Cys-569 that may enable oscillation between reduced and oxidized states in the presence of DTT. To test this idea, we treated the WT-, T569C- and empty vector-transfected 293T cell lysates with increasing DTT concentrations (0, 0.5, 5, 25 and 100 mM) and 2 M urea, a strong protein denaturant, prior to electrophoresis in SDS-polyacrylamide gels containing 6 M urea. For these experiments, the transfected 293T cells were lysed in the presence of 5 mM iodoacetamide in order to inhibit post-lysis thiol-disulfide exchange reactions mediated by free sulfhydryl groups. For T569C, the ~ 80 kDa putative gp41 dimer band accounted for ~ 60 – 70 % of total gp41 with 0, and 0.5 mM DTT, whereas stepwise reductions in band intensity with 5 and 25 mM DTT and a virtual absence of this band with 100 mM DTT were observed (Figure 4.6). The stepwise

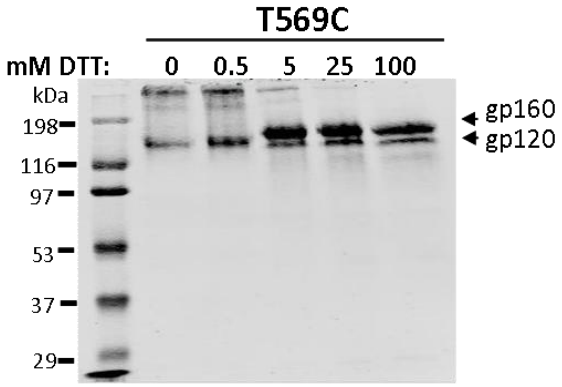

NO ENV
mM DTT: 0 0.5 5 25 100
kDa
198
116
97
53
37
29

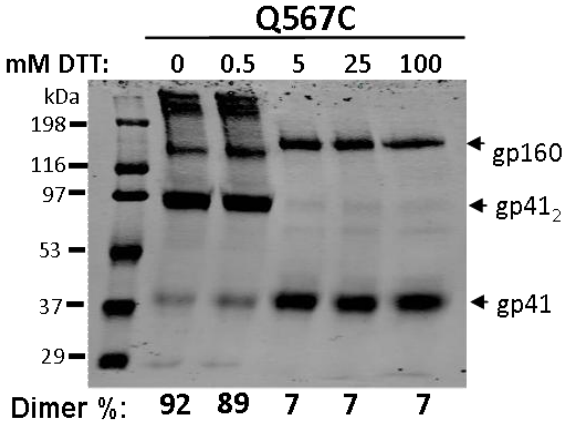

NO ENV
mM DTT: 0 0.5 5 25 100
kDa
198
116
97
53
37
29


WT
mM DTT: 0 0.5 5 25 100
kDa
198
116
97
53
37
29
gp160
gp41


WT
mM DTT: 0 0.5 5 25 100
kDa
198
116
97
53
37
29
gp160
gp120


T569C
mM DTT: 0 0.5 5 25 100
kDa
198
116
97
53
37
29
gp160
gp41$_2$
gp41
Dimer %: 70 62 48 16 14


T569C
mM DTT: 0 0.5 5 25 100
kDa
198
116
97
53
37
29
gp160
gp120


Q567C
mM DTT: 0 0.5 5 25 100
kDa
198
116
97
53
37
29
gp160
gp41$_2$
gp41
Dimer %: 92 89 7 7 7

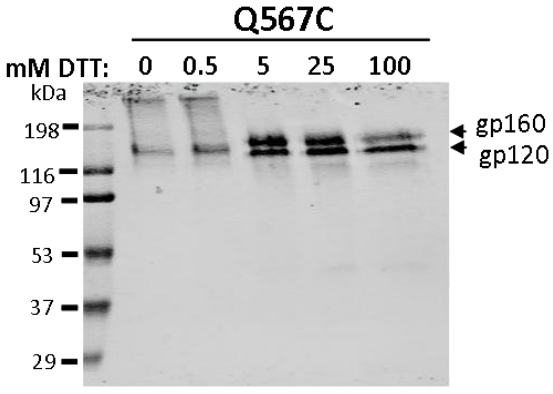

Q567C
mM DTT: 0 0.5 5 25 100
kDa
198
116
97
53
37
29
gp160
gp120

Figure 4.6 SDS-PAGE of cysteine mutants in the presence of urea

293T cells, cotransfected with pCAG-T7 and pTMenv.2 expression vectors, were lysed in the presence of 5 mM iodoacetamide at 48 h posttransfection. The lysates were boiled in the presence of 2 M urea, 2 % SDS and the indicated concentration of DTT prior to SDS-PAGE in 10 % polyacrylamide gels containing 6 M urea. The proteins were visualized by western blotting with mAb C8 (left-hand side panels), and DV012 (right-hand side panels). No Env: pTM.1 empty vector was used instead of pTMenv.2.

reductions in ~ 80 kDa band intensity with 5 and 25 mM DTT were associated with concomitant increases in gp41 monomer band intensity indicating that the 80 kDa band corresponds to a disulfide-linked gp41 dimer. The gp41 dimer species was not present in mAb C8 western blots of WT- and mock-transfected cells, or in blots where the gp120-specific antibody DV012 was used indicating that the introduced cysteines at position 569 form the inter-protomer disulfide in gp41. The gp41 dimer band represented ~ 60 - 70% of total gp41 at 0 and 0.5 mM DTT which is close to the expected value of 66 % if two out of three cysteines formed a disulfide bond in every trimer. Thus disulfide formation via Cys-569 is highly efficient in this context.

The data suggest that the *d*-position T569C mutation enables quantitative disulfide bond formation in the context of a functional gp41 trimer. In a trimeric coiled coil context, the sidechain of Cys-569 will be orientated towards the 3-fold symmetry axis and juxtaposed with Cys-569 sidechains of partner protomers, which may enable disulfide formation. We reasoned that substitution of an outward facing sidechain, for example the *b*-position residue Gln-567, should not enable quantitative disulfide formation in a trimeric coiled coil context because the distance between Cβ atoms is predicted to be > 5.28 Å (15.5 Å) by available x-ray structures of the coiled coil region. The Q567C mutants was expressed and processed in a similar manner to the WT (compare the 100 mM DTT lanes for WT and Q567C in figure 4.6), however, it lacked detectable cell-cell fusion function (886,793 ± 11,1671 RLU for WT *versus* 64,095 ± 16,550 RLU for T567C; n = 4), consistent with a compromised structure. Sodium dodecyl sulfate-PAGE in the presence of urea and western blot analysis indicated that ~90% of the total gp41 protein presented as a disulfide-linked dimer with 0 and 0.5 mM DTT was completely reduced with 5 mM DTT (Figure 4.6). This 9:1 dimer:monomer stoichiometry is not possible in a trimeric protein and is likely to be due to the formation of higher-order aggregates linked through

Cys-567.

From these data I infer that the C-terminal portion of the central coiled coil, at least up to Thr-569, is present in the prefusogenic form of gp41 (i.e. the form that exists prior to engagement of receptors by gp120) because substituting this *a*-position residue with Cys enables quantitative interprotomer disulfide bond formation. By contrast, the polar segment mutations, S534C and T538C, and A541C at the N-terminus of HR1 did not enable interprotomer disulfides to form (Figure 4.5), indicating that these regions are not closely apposed in the prefusion structure.

4.3.2 Interprotomer disulfide formation in a fusion-activated trimer of hairpins model protein: MBP/gp41(528-L-677)

4.3.2.1 Cysteine-substituted MBP/gp41(528-L-677) chimeras acquire trimeric structure

MBP/gp41(528-L-677) is a bacterially expressed chimeric protein that models the gp41 trimer of hairpins. The chimera comprises MBP joined through Asn-Ala to the HIV-2_{ST} polar segment-HR1 segment (residues 528-596), a Ser-Gly-Gly-Arg-Gly-Gly linker, and the HIV-2_{ST} HR2-MPER segment (residues 610-677) (Figure 4.7). HIV-2_{ST} sequences were used in this case because we found that analogous chimeras containing HIV-1 sequences derived from several clades resulted in the production of high molecular weight aggregates. MBP/gp41(528-L-677) forms stable, soluble trimers following expression in *E. coli* and limited chymotrypsin proteolysis of the purified trimer releases the constituent peptides of the trimer of hairpins, Thr-538-Trp-596 and Val-611-Asn-677 (Bellamy-McIntyre *et al.*, 2007). Wild type and cysteine substituted MBP/gp41 chimeras were expressed in *E. coli,* affinity-purified using amylose–agarose and further purified by gel filtration. An examination of the purified and concentrated chimeras by analytical gel filtration (Figure 4.8) indicated that the WT and cysteine mutants eluted at 28 min,

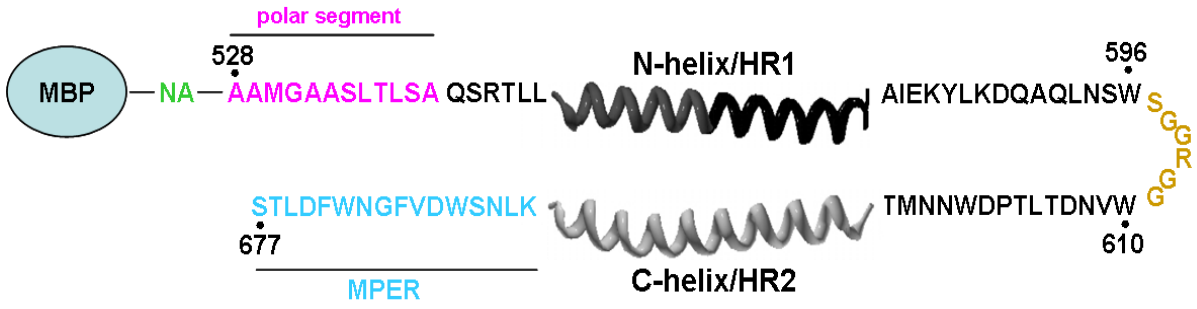


Figure 4.7 Schematic of MBP/gp41(528-L-677) chimera

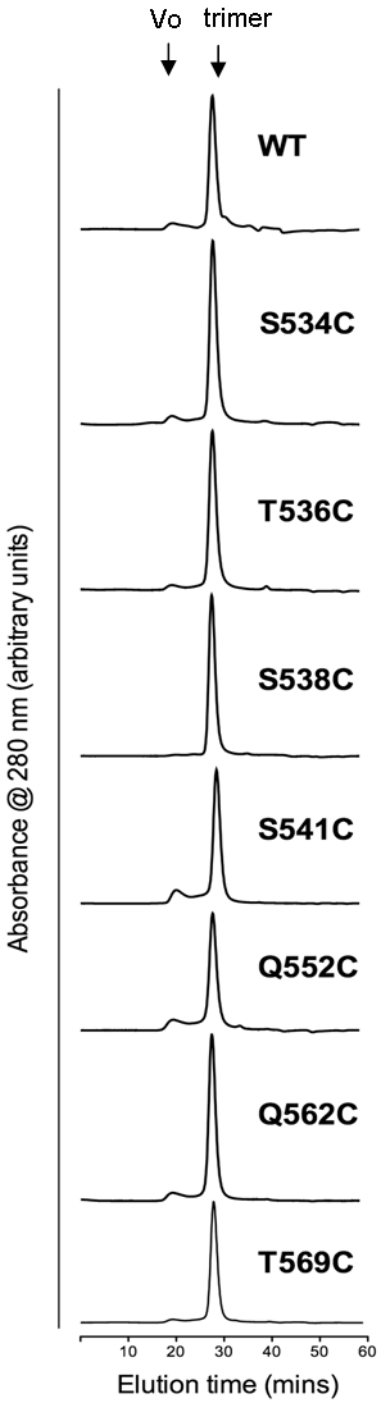


Figure 4.8 Gel filtration of purified MBP/gp41(528-L-677) cysteine mutants in a Superdex 200 (HR 10/30) column (0.5 ml/min flow rate)

The elution positions of blue dextran (denoting the void volume, Vo), and the MBP/gp41(529-665) reference trimer (Lay *et al.*, 2004) are indicated.

corresponding to the elution time of MBP/gp41(529-665) whose trimeric structure has been verified by sedimentation equilibrium (Lay *et al.*, 2004). These trimeric proteins were used to monitor near-neighbour interprotomer interactions via disulfide formation in the context of a fusion-activated trimer of hairpins structure.

4.3.2.2 Disulfide formation in MBP/gp41(528-L-677) cysteine substitution mutants

Sodium dodecyl sulfate-PAGE analysis of the purified MBP/gp41 WT and cysteine-mutated oligomers in the presence of 3% β-mercaptoethanol showed a single band migrating at ~53 kDa (Figure 4.9), consistent with the expected size of MBP/gp41(528-L-677) monomer. Whereas monomeric protein was also observed for WT, Q552C, Q562C and T569C with SDS-PAGE in the absence of β-mercaptoethanol, an additional ~ 100 kDa band, consistent with an MBP/gp41 dimer, was observed for 3 chimeras with Cys substitutions in the N-terminal region: S534C and S538C in the polar segment, and S541C at the N-terminus of HR1. The dimer band was strongest for the polar segment mutant, S538C.

An underlying assumption of the current study is that the substitution of Cys residues into *a* or *d* positions of a trimeric coiled coil will favour disulfide formation because the inward-facing orientation of the sidechains will appose the reactive thiol groups. To test this idea, Cys was substituted for the outward facing *f*-position residue Thr-536. This mutation is close to Ser-538, which when substituted with Cys, enabled the most efficient disulfide formation in MBP/gp41(528-L-677). Figure 4.10 indicates a low level of disulfide formation via Cys-536 with dimer accounting for ~ 6 % of the total MBP/gp41. This low level of disulfide bonding via outward facing Cys-536 residues may correspond to a subset of MBP/gp41 conformers with disordered terminal segments that allow thiol groups to approach each other sufficiently closely to form the disulfide. Importantly, the dimer yield for S538C was 36 % of total MBP/gp41, indicating disulfide

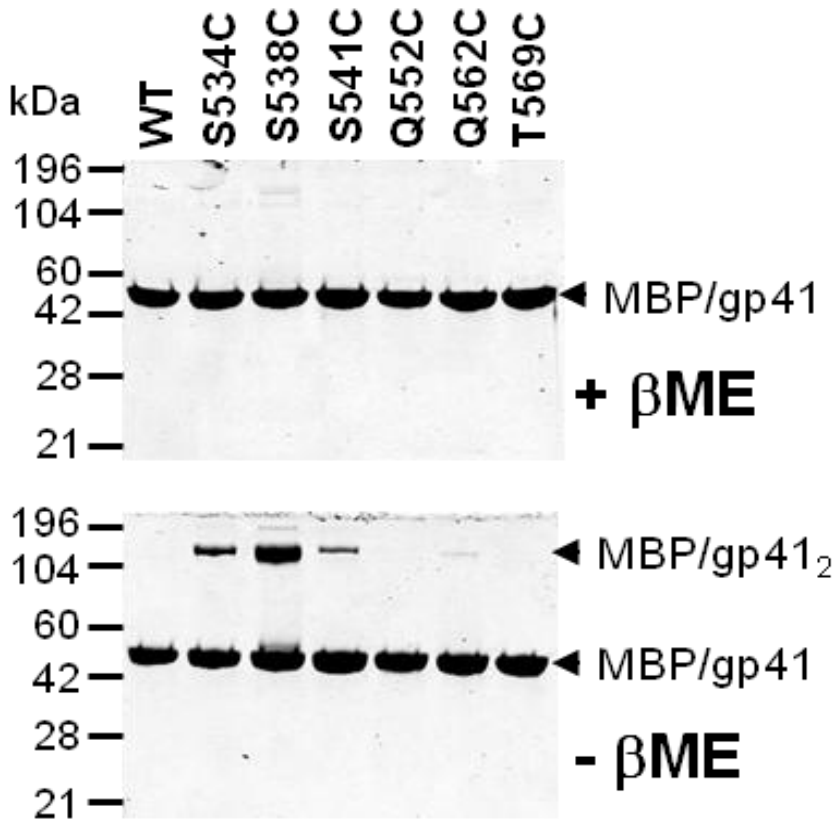


Figure 4.9 SDS-PAGE of MBP/gp41(528-L-677) cysteine mutants

The purified WT and Cys-mutated MBP/gp41(528-L-677) trimers were treated with 20 mM iodoacetamide and then subjected to SDS-PAGE under reducing (+ βME) and non-reducing (- βME) conditions in 10–16% polyacrylamide gradient gels. The protein bands were visualized following staining with Coomassie brilliant blue and scanning in a LI-COR Odyssey infrared imager. Dimers were observed for the polar segment mutants (S534C and S538C) and the N-terminal HR1 mutant (S541C).

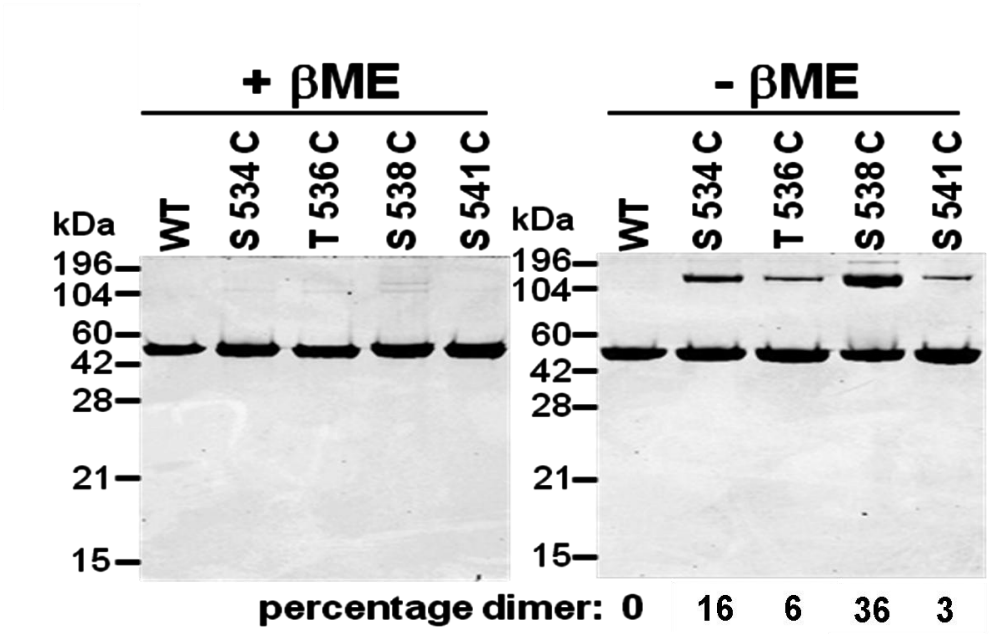


Figure 4.10 Position dependency of disulfide formation

The efficiencies of disulfide formation by *a*- (S538C) and *d*- (S534C, A541C) position mutants were compared to that of an *f*-position control mutant, T536C. The formula used for calculating values for percentage dimer was intensity (dimer band) ÷ intensity (monomer + dimer bands) x 100. The proteins were analysed as for Figure 4.9.

bonding via this *a*-position cysteine is significantly more efficient than bonding via the *f* position. By contrast, the dimer yield for S534C and S541C was 16 % and 3 % of total MBP/gp41, respectively (i.e. close to that of the T536C control) suggesting that disulfide bonding in this case may be independent of structure and sidechain orientation. Taken together, these data indicate that the *a*-position 538 sidechain within the polar segment becomes juxtaposed in the trimer of hairpins, consistent with the N-terminal helical extension of the coiled coil beyond Ser-541, which is the boundary previously identified by limited proteolysis and crystallography of the 6-HB (Chan *et al.*, 1997; Weissenhorn *et al.*, 1997a; Caffrey *et al.*, 1998; Yang *et al.*, 1999).

4.3.2.3 Thiol-specific chemical crosslinking of cysteine mutants with bis(maleimido)ethane (BMOE) and bis(maleimido)hexane (BMH)

The available crystal structures of the gp41 trimer of hairpins suggest that interprotomer Cβ – Cβ distances for cognate *a* and *d* position residues is greater than 5.28 Å, which is the optimal separation for disulfide formation. However, our data suggest that Cys-538 in the polar segment comes within range for disulfide bonding. The homobifunctional sulfhydryl-to-sulfhydryl crosslinkers, BMOE (8 Å spacer) and BMH (13 Å spacer), were therefore used to further examine the separation between introduced Cys residues in trimeric MBP/gp41. The data presented in figure 4.11, indicates that BMOE and BMH were unable to bridge cognate cysteines at 552, 562 and 569 within the 6-HB core domain. By contrast, both crosslinkers increased the efficiency of covalent stabilization for MBP/gp41(528-L-677) containing Cys-534, Cys-538 and Cys-541 in the N-terminal region. A decrease in intensity of the crosslinked S538C dimer band with DTT treatment was observed and can be attributed to reduction of the spontaneously-formed Cys-538 disulfide. Notably, BMH was able to efficiently form crosslinks with the outward facing Cys-536 for which the interprotomer Cβ-Cβ separation is 21.2 Å in the 2X7R

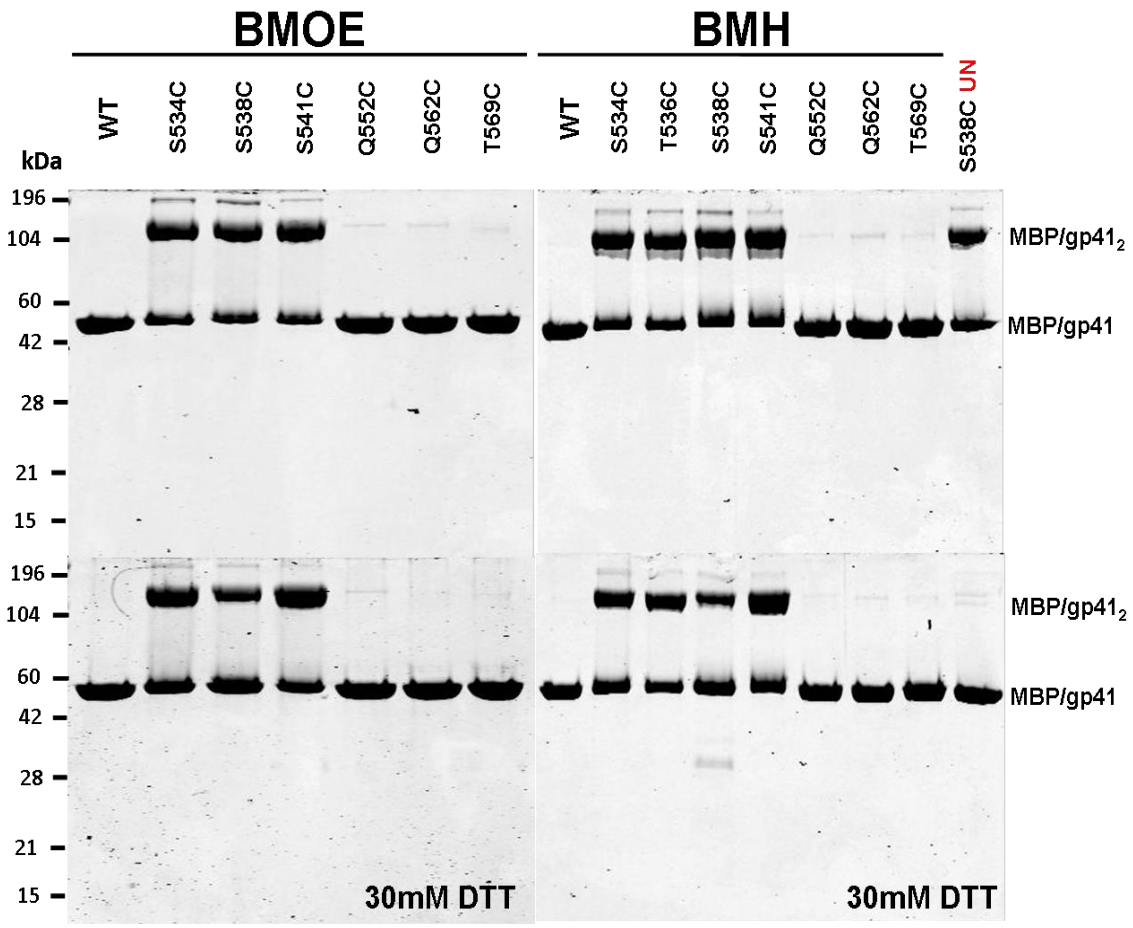


Figure 4.11 Thiol-specific chemical crosslinking of MBP/gp41 Cys mutants

Purified MBP/gp41 trimers were treated with 4 mM of BMOE (8 Å spacer, left panels) or 4 mM BMH (13.8 Å spacer, right panels) for 1 h. The reaction was quenched with 30 mM DTT (bottom panels only) and then subjected to SDS-PAGE in 10–16% polyacrylamide gradient gels. The protein bands were visualized following staining with Coomassie brilliant blue and scanning in a LI-COR Odyssey infrared imager. The untreated MBP/gp41(528-L-677) containing Cys-538 (S538C UN) was used to control reduction by 30mM DTT.

crystal structure. This separation exceeds the reach of the reactive maleimide groups by some 8 Å, indicating that the N-terminal region must be quite flexible for efficient crosslinking at Cys-536 to be allowed. The data in this section and in the preceding section 4.3.2.2 are consistent with both α-helical structure extending beyond the core region to include the polar segment, and with a high level of flexibility in this terminal region as is suggested by the 2X7R crystal structure to enable disulfide bonding. On the other hand, the absence of disulfide formation via the more closely spaced cysteines at 552, 562 and 569 is consistent with the structural rigidity of this central segment of the 6-HB. It follows that the inability of BMOE and BMH to crosslink through these cysteines at 552, 562 and 569 is due to steric occlusion by tightly packed coiled coil and antiparallel HR2 helices.

4.3.3 Probing the structure of the polar segment in gp41 conformers using mutation-directed disulfide formation

The results of previous studies have indicated an important structural and functional role for the polar segment in forming a clasp at the membrane interactive end of the trimer of hairpins through interactions with the MPER (Bellamy-McIntyre *et al.*, 2007; Lay *et al.*, 2004). x-ray crystallography revealed that the polar segment and MPER become juxtaposed and interact in the trimer of hairpins by forming helical extensions of the central trimeric coiled coil and HR2 helices, respectively (Buzon *et al.*, 2010) (Figure 4.12). Simultaneous I535A/V539G mutations in the polar segment of MBP/gp41(528-L-677) have been shown to increase the chymotrypsin sensitivity of the terminal clasp, indicative of destabilization. The clasp-destabilizing effects of I535A/V539G directly correlated with inhibition of cell-cell fusion and infectivity (Table 4.1), whereas early events such as precursor processing, gp120-gp41 surface-expression and association, and responsiveness to CD4 were not affected. These data led to the conclusion that the

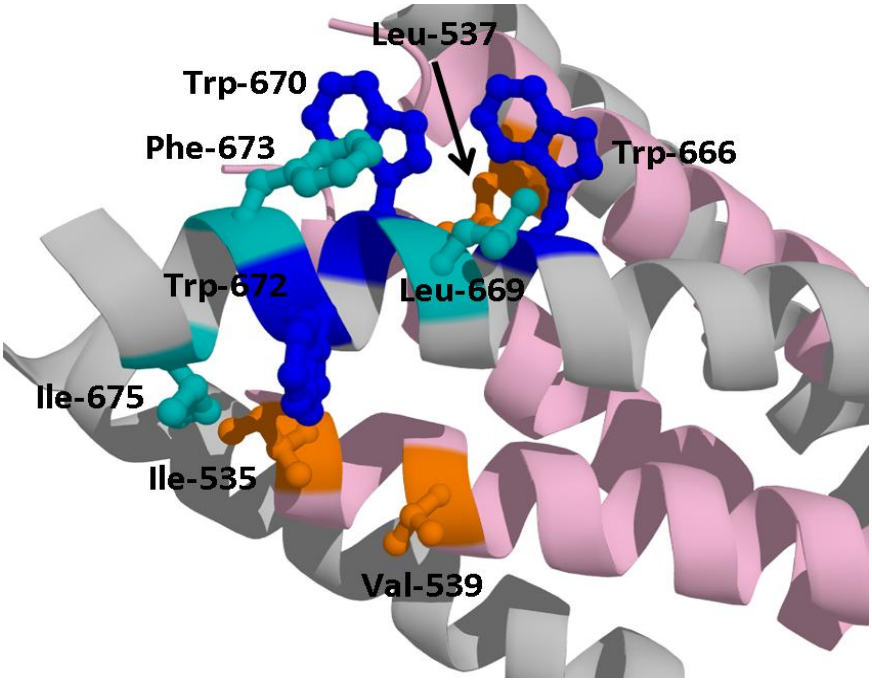


Figure 4.12 Hydrophobic interactions stabilize the terminal clasp

The terminal fusion clasp is stabilized by interactions between hydrophobic residues of the polar segment and MPER: W^{666}-L^{537}-W^{670} and W^{672}-I^{535}-I^{675}. Polar segment helices: pink, MPER helices: grey. Drawn from 2X7R using pymol.

	A Cleavage at Leu-537	*B Cell-cell Fusion*	*C Infectivity*
WT	15±1	100±10	12,970±1,569
I535A/V539G	27±3*	4±1**	5±0.3*
tr665	40±2*	Nd	Nd
No Env	Nd	4±1**	2±0.5**

Table 4.1 Effect of mutations on clasp stability, cell-cell fusion and infectivity

A) MBP/gp41(528-L-677)$^{AD.ST}$ was treated with limiting chymotrypsin (10:1 protein:protease). The protease sensitivity of the polar segment was calculated as described (Bellamy-McIntyre *et al.*, 2007). B) Cell-cell fusion activity was normalized to WT. C) Single cycle infectivity of gp120.gp41-pseudotyped luciferase reporter viruses (mean ± SEM relative light units x 10^{-3}). All assays: n=3; *, $P < 0.02$; **, $P < 0.01$, relative to WT; 2-sample t test assuming unequal variances. Summarized from Lay *et al.* Manuscript in Preparation.

assembly of hydrophobic clusters stabilizing the clasp provide energy for the destabilization and fusion of the virus and cell membranes (Lay *et al.*, Manuscript in Preparation). The finding that the introduction of Cys-538 to the polar segment in MBP/gp41(528-L-677) leads to disulfide formation provides an approach to probe for coiled coil formation within this N-terminal region in recombinant proteins designed to represent various gp41 fusion intermediates prior to and including terminal clasp formation.

Before embarking on these studies, a new set of MBP/gp41 chimeras were prepared in which the polar segment and MPER of HIV-2_{ST} were swapped with the corresponding sequences of HIV-1_{AD8}, the strain used for mammalian cell-based analyses of the terminal clasp in Lay *et al.* (The new chimeras are denoted with the AD.ST superscript). The swap was made because HIV-1_{AD8} and HIV-2_{ST} exhibit striking amino acid differences in the MPER, including A667D, S668V and W670G. Thus, MBP/gp41(528-L-677)$^{AD.ST}$ represents the entire trimer of hairpins including the terminal clasp, tr665 lacks the MPER and represents the 6-HB core with a "naked" polar segment extension, whereas MBP/gp41(528-591)$^{AD.ST}$ represents the polar segment-HR1 region prior to 6-HB formation (i.e. the prehairpin) (Figure 4.13). The terminal clasp destabilizing mutation, I535A/V539G, was used in conjunction with T538C (HIV-1 has Thr at position 538) to study how these changes affect polar segment structure in the various recombinant model proteins.

4.3.3.1 Cys-538-dependent disulfide bonding in MBP/gp41(528-L-677)$^{AD.ST}$ and tr665

The T538C and I535A/T538C/V539G mutations were introduced to MBP/gp41(528-L-677)$^{AD.ST}$ whereas T538C was introduced to tr665. The chimeras were expressed in *E. coli,* affinity-purified using amylose–agarose and then further purified with gel filtration using a Superdex 200 column linked to an AKTA-FPLC system. The purified

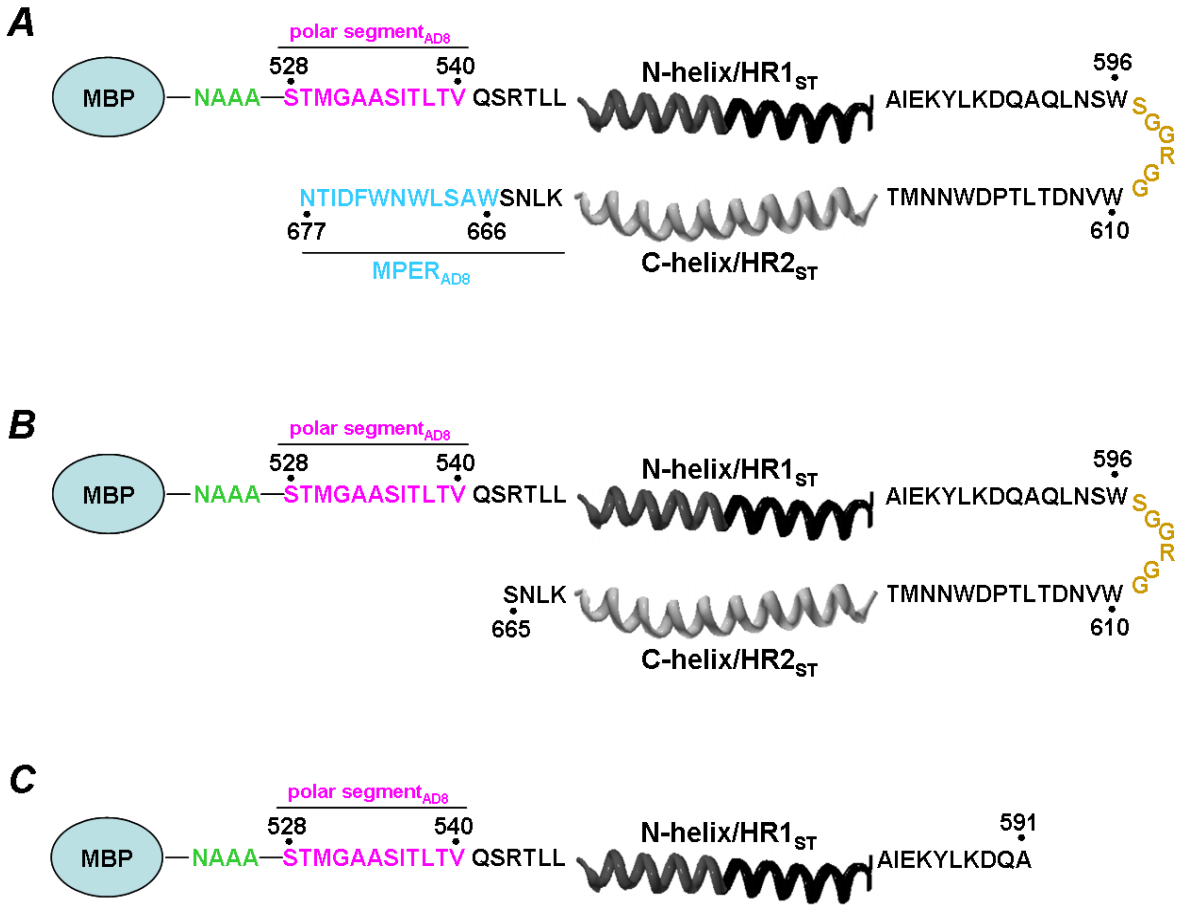


Figure 4.13 Schematic of MBP/gp41 chimera representing various gp41 fusion intermediates

A) MBP/gp41(528-L-677)$^{AD.ST}$

B) tr665

C) MBP/gp41(528-591)$^{AD.ST}$

chimeras exhibited a trimeric structure as seen in figure 4.14A. SDS-PAGE of chimeric MBP/gp41 proteins under non-reducing conditions showed dimer formation for T538C-containing chimeras, but not for the WT, with almost complete reduction to monomer in the presence of 3 % β-ME (Figure 4.14B). The efficiency of T538C-mediated disulfide bonding was decreased in the I535A/V539G mutant (compare T538C with I535A/T538C/V539G in figure 4.14B and figure 4.16), suggesting that the clasp destabilizing mutation affects the proximity of Cys-538 residues within gp41 trimers. By contrast, deletion of the MPER in chimera tr665 did not affect Cys-538-mediated disulfide bonding in comparison to MBP/gp41(528-L-677)$^{AD.ST}$ (Figure 4.14B and Figure 4.16). These data suggest that coiled coil extension into the polar segment is initially independent of the MPER. However, normal clasp formation appears to be required for an efficient disulfide to form as the clasp-destabilizing mutation I535A/V539G decreases Cys-538-mediated bonding efficiency.

4.3.3.2 Polar segment assembly occurs prior to trimer of hairpins formation

The data suggest that the terminal clasp forms in 2 stages with polar segments initially assembling into a trimeric coiled coil and providing a hydrophobic interaction surface for the MPER to allow subsequent terminal clasp formation. However, it is not clear whether polar segment assembly requires formation of the 6-HB core, comprising the trimeric HR1 coiled coil and the antiparallel HR2 outer layer. In order to clarify the order in which these conformational changes occur, polar segment-HR1-only MBP/gp41(528-591)$^{AD.ST}$ chimeras (Figure 4.13) containing T538C and I535A/T538C/V539G were expressed and characterised.

Superdex 200 gel filtration chromatography confirmed that the purified chimeras exhibited a trimeric structure (Figure 4.15A). SDS-PAGE under non-reducing conditions indicated that dimers were formed with equal efficiency for the T538C and

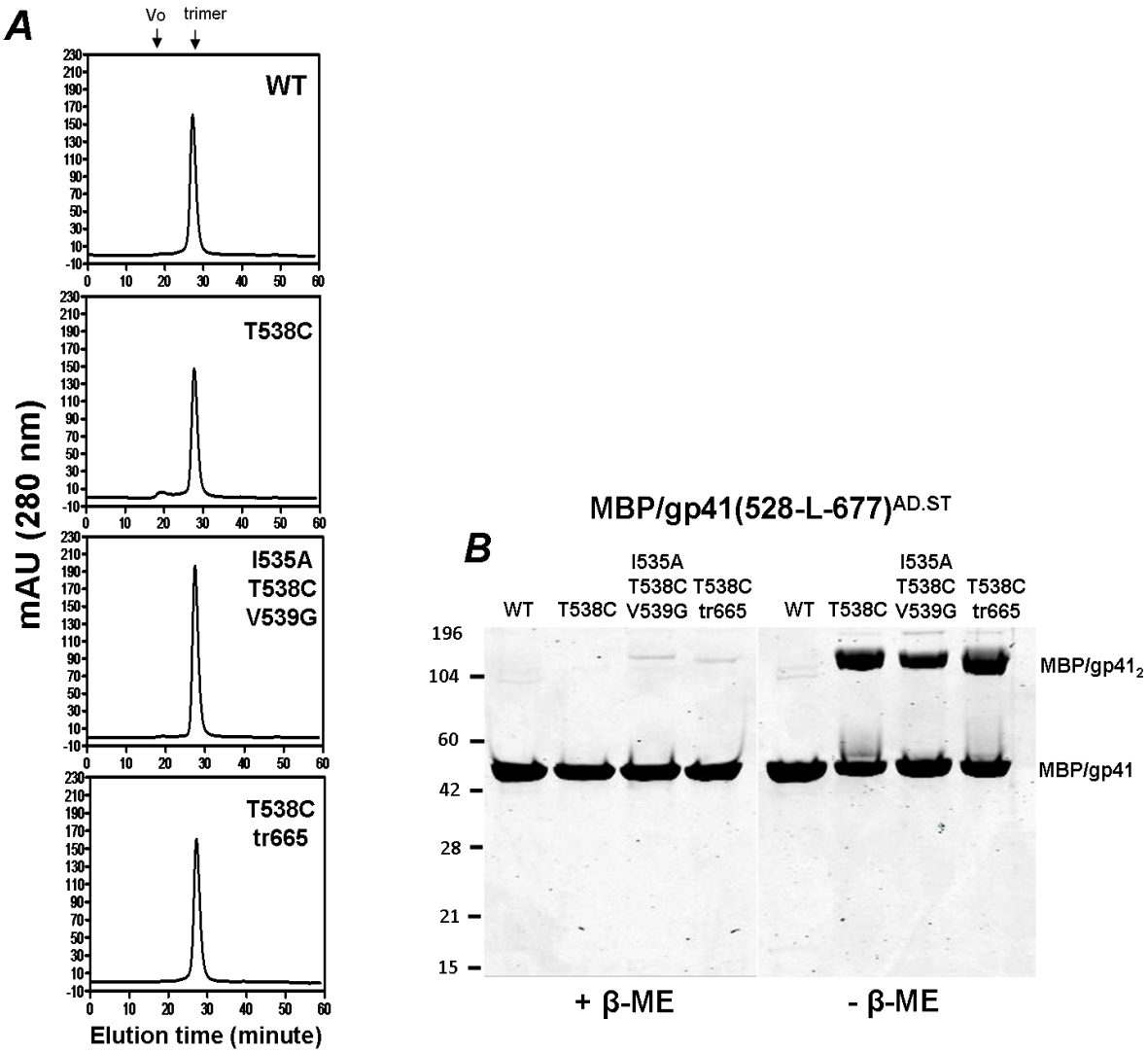


Figure 4.14 Cysteine-538-mediated disulfide bonding of polar segments is independent of the MPER

(A) Analytical Superdex 200 (HR 10/30) gel filtration of WT, T538C-, 1535A/T538C/V539G-mutated MBP/gp41(528-L-677)$^{AD.ST}$ and tr665-T538C (lacking the MPER) (0.5 ml/min flow rate). The elution positions of blue dextran (denoting the void volume, Vo), and the MBP/gp41(529-665) reference trimer (Lay *et al.* 2004) are indicated. **(B)** The MBP/gp41(528-L-677)$^{AD.ST}$ WT, T538C, I535A/T538C/V539G and tr665-T538C chimeras were subjected to SDS-PAGE under reducing (+ βME) and non-reducing (- βME) conditions. All three chimera containing the T538C mutation formed dimers but the presence of the I535A/V539G mutation caused a pronounced reduction (more than 40%) in the dimer band (refer for figure 4.16 for band intensity quantification). The absence of the MPER did not affect the dimer band intensity.

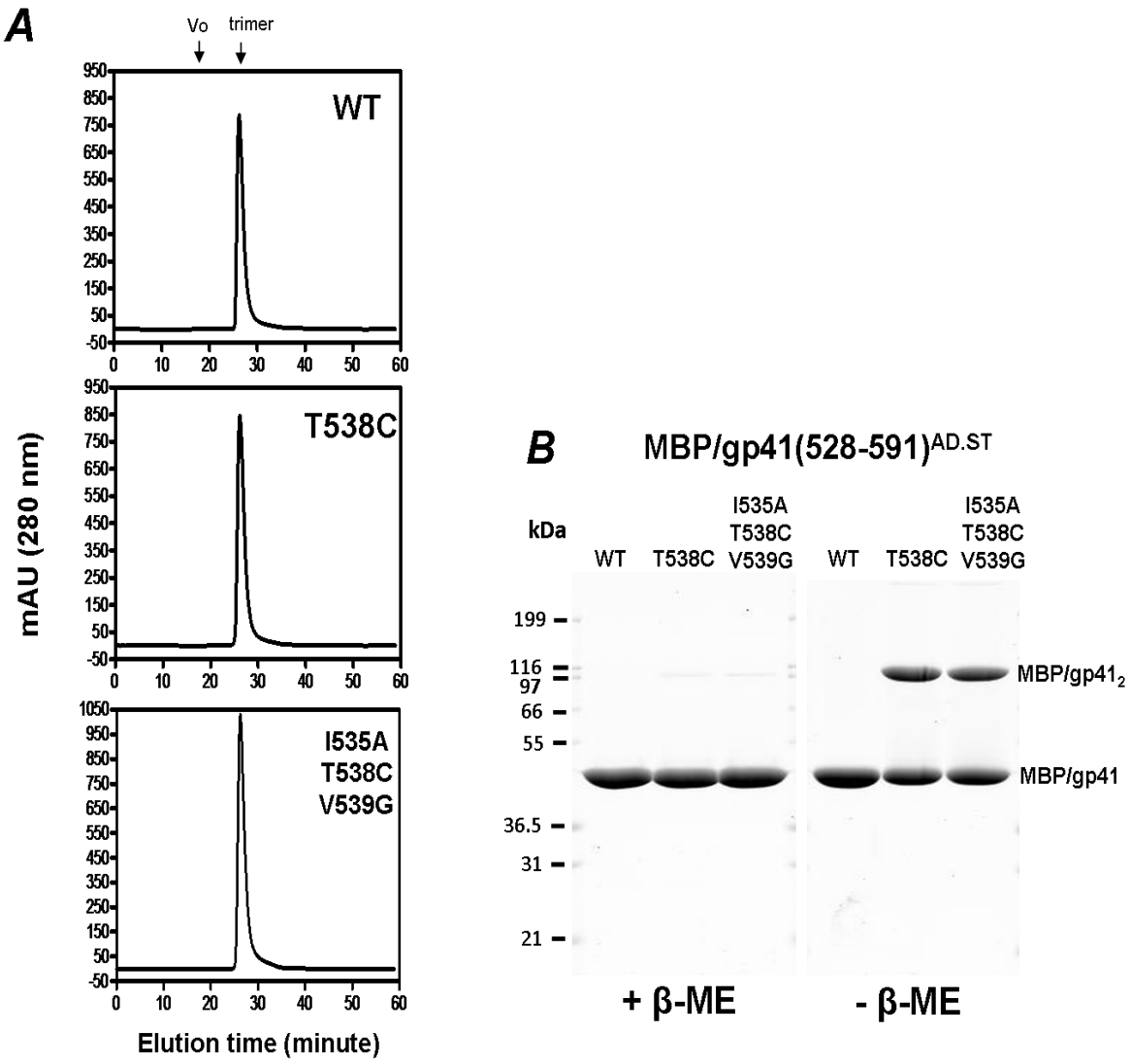


Figure 4.15 The effect of the simultaneous I535A/V539G mutation on Cys-538-mediated disulfide formation in the gp41 polar segment HR1-only chimera, MBP/gp41(528-591)AD.ST

(A) Superdex 200 HR (10/30) gel filtration profiles of MBP/gp41(528-591)$^{AD.ST}$ WT, T538C and I535A/T538C/V539G (0.5 ml/min flow rate). **(B)** The purified, trimeric MBP/gp41(528-L-677) WT, T538C and I535A/T538C/V539G chimeras were then subjected to SDS-PAGE under reducing (+ βME) and non-reducing (- βME) conditions. The chimeras containing T538C formed dimers with similar efficiencies, even the presence of the I535A/V539G.

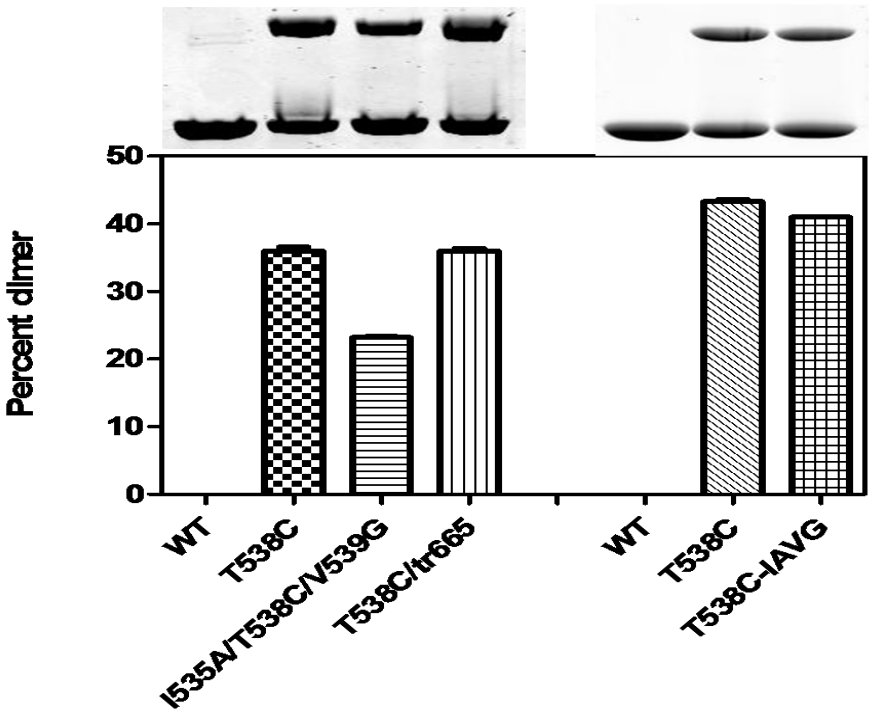


Figure 4.16 Effects of I535A/V539G and MPER deletion (tr665) on Cys-538-dependent disulfide formation

The percentage dimer (versus total MBP/gp41) values for the various chimeras were calculated as described for Figure 4.10. The Left side panel represents the MBP/gp41(528-L-677)$^{AD.ST}$ chimera, while the right side panel represents the MBP/gp41(528-591)$^{AD.ST}$ chimera. The data shown are the means ± standard errors from 3 independent experiments. T538C-IAVG: I535A/T538C/V539G.

I535A/T538C/V539G mutants (Figure 4.15B and Figure 4.16) indicating that the I535A/V539G mutation is significant only in the context of a gp41 trimer of hairpins containing the terminal clasp. The data suggest that the helical extension of the polar segment occurs at the prehairpin stage and is independent of HR2 and the MPER.

4.4 Discussion

Cysteines were introduced into hydrophilic *a* and *d* positions within the polar segment-HR1 sequence of gp41 in order to examine the interprotomer proximity of this region in different gp41 conformations. Disulfide formation was examined in 2 expression systems. First, a mammalian cell-based system employing the pTMenv.2 vector yields gp160 precursor that is cleaved in the *Golgi* into a functional gp120-gp41 complex. In this system, gp41 acquires a prefusion or metastable structure that can be activated by gp120-CD4-CKR interactions to refold into fusion active conformers. The second *E. coli*-based system expresses gp41 segments chimerized via their N-termini to MBP. MBP/gp41(528-L-677) contains the gp41 ectodomain minus the fusion peptide in the low energy trimer of hairpins conformation, whereas MBP/gp41(528-591)$^{AD.ST}$ contains the polar segment-HR1 sequence and represents a prehairpin structure (Kobe *et al.*, 1999; Lay *et al.*, 2004). The data are consistent with a model for gp41 conformational states associated with fusion function where the C-terminal portion of HR1 participates in interprotomer contacts and likely forms a coiled coil in the prefusion protein, and then becomes extended in an N-terminal direction to encompass the polar segment following activation by gp120-receptor interactions. The HR2 and adjacent MPER then pack against the extended coiled coil in an antiparallel orientation, forming a trimer of hairpins with a terminal clasp that catalyses membrane fusion.

Of the 6 pTMenv.2 cysteine mutants examined, S534C, T538C, A541C and T569C retained membrane fusion activity, although at moderately reduced levels with respect to WT. The interpretation of disulfide bonding data was therefore restricted to these 4 functional mutants. The small reductions in fusion are likely a result of subtle structural perturbations occurring in gp41 for accommodation of Cys side chains. For example, hydrophobic contact between Cγ2 atoms of Thr-569 and Val-570 in the gp41 coiled coil

(Weissenhorn *et al.*, 1997a) would be precluded in T569C because Cys lacks Cγ2, potentially leading to decreased core stability (Figure 4.17). By contrast, Ser-534, Thr-538 and Ala-541 do not appear to interact with other residues because the N-terminal region of the trimer of hairpins helices expands away from the 3-fold symmetry axis (Buzon *et al.*, 2010). The decreased fusogenicity of these Cys mutants is therefore likely to occur via a different mechanism in comparison to T569C. For example, decreased fusion could be a result of substitution-induced changes to the structural contexts of these residues in other gp41 conformations for which structural information is not available. The Q552C and Q562C lacked fusion function altogether. These residues are closely packed in the trimer interface (Weissenhorn *et al.*, 1997a; Buzon *et al.*, 2010), therefore Q552C and Q562C mutations would introduce substantial potentially destabilizing cavities in the gp41 core.

The introduction of T569C to the C-terminal half of the HR1 sequence resulted in ~ 66 % of the total prefusion gp41 presenting as dimer under nonreducing conditions, consistent with 2 out of 3 gp41 protomers forming a disulfide in every trimer. Quantitative disulfide formation correlated with the cysteine being introduced to the inward-facing *d* position at 569 but not an outward facing *b* position at 567 of the heptad repeat, consistent with coiled coil formation in the region around Thr-569. The results of previous studies suggest that a component of HIV-1 Env glycoproteins present on the surface of expressing cells are nonfunctional (Poignard *et al.*, 2003; Moore *et al.*, 2006; Crooks *et al.*, 2007; Crooks *et al.*, 2011). Care should therefore be taken in interpreting the T569C data in case disulfide formation segregates to this nonfunctional Env component. It should be noted that this scenario cannot apply to the A534C, T538C, A541C, Q552C and Q562C mutants, irrespective of whether or not fusion function was retained, as disulfide formation was not observed in the prefusion context. Further analysis of the T569C mutant using iodixanol

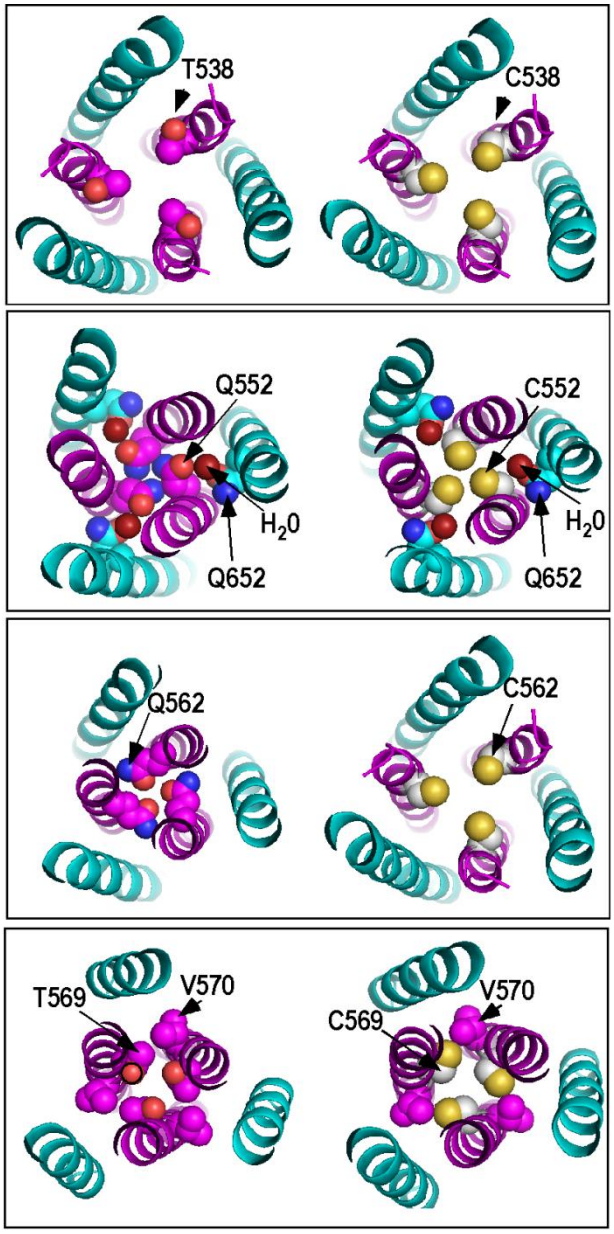


Figure 4.17 Effect of cysteine substitution on gp41 coiled coil residues

Cysteine substitution might lead to cavity formation (Q552C and Q562C) or affect interactions with neighbouring residues (T569C). Polar segment-substitutions do not seem to have any structural effect since the N-terminal region of the trimer of hairpins helices expands away from the 3-fold symmetry axis.

gradient-purified virus (Dettenhofer and Yu, 1999) or glycosidase-protease-treated Env-expressing cells, which removes the "junk" forms of Env (Crooks *et al.*, 2011), is required to rule out this possibility for T569C.

The results of early biochemical and mutational studies on the mechanism of Env oligomerization support the idea that a coiled coil is present in the prefusion form of gp41 and suggest that this oligomerization determinant extends to (at least) Leu-555. First, early biophysical studies with synthetic peptides or recombinant proteins representing the 553-591 sequence indicate high α-helical content and a propensity for self-oligomerization (Wild *et al.*, 1995; Bernstein *et al.*, 1995). Second, the deletion of residues 555-566 of HR1 is sufficient to block gp41 oligomerization (Poumbourios *et al.*, 1995), and residues 548-586 represent the minimum sequence required for hetero-oligomerization between WT and truncated HIV-1envelope glycoproteins (Poumbourios *et al.*, 1997). Further evidence for coiled coil formation in this region is provided by the demonstration that replacement of the Leu-576-Gln-577-Ala-578 triad with Cys-Cys-Gly leads to the production of disulfide-stabilized gp160 (Farzan *et al.*, 1998). We note that the analysis of disulfide formation in gp41 was precluded in this study because the mutation blocked precursor processing. Taken together, these data suggest that the C-terminal portion of the HR1 sequence forms a coiled coil in both gp160 and gp41. In contrast to T569C, the S534C, T538C and A541C mutations introduced to the N-terminal region did not form disulfides in prefusion gp41. Consistent with these data, a glycoprotein truncated at Gln-547, which contains the polar segment and N-terminal portion of HR1, did not form hetero-oligomers with the WT glycoprotein (Poumbourios *et al.*, 1997). These data suggest that the polar segment and N-terminal portion of the HR1 sequence is not apposed in the prefusion structure of gp41 or in gp160.

In contrast to the data obtained with prefusion gp41, S534C, S538C and A541C formed interprotomer disulfides with varying efficiencies in MBP/gp41(528-L-677), which models the fusion activated trimer of hairpins conformation. A low level of disulfide bonding (6 % of total MBP/gp41[528-L-677]) was observed for the *f*-position T536C mutant even though the Cβ-Cβ separation for Thr-536 residues is much greater that the optimal distance of disulfide formation (21.2 Å *versus* 5.28 Å, respectively). These data suggest that a component of the MBP/gp41(528-L-677) trimer preparation contains a disordered N-terminal region, which enables apposition of Cys-536 residues. Because the efficiency of S534C- and A541C-mediated disulfide bonding was also low (16 % and 3 % of total MBP/gp41[528-L-677], respectively), we cannot preclude disorder as the reason for disulfide formation in these cases. By contrast, high-efficiency disulfide formation (36-38 % of total protein) was observed for T538C, indicating position dependence in this case. The Cβ-Cβ separation for Thr-538 is 9.8 Å, indicating that interprotomer disulfide bonding in the context of a coiled coil will only occur if this region is mobile. This idea is consistent with the crystallographic data of Buzon *et al.* (2010) indicating that the Gly-531 - Ala-541 region encompassing the polar segment is associated with high mobility (average *B*-value for Cα atoms = 48.64) **(**Figure 4.2 and Figure 4.18**).** Furthermore, 2 of 3 molecules in the asymmetric unit of the 2X7R PDB entry lack electron density in the polar segment region. It is plausible that the lateral movement of coiled coil helices relative to the three-fold symmetry axis in this N-terminal region allows apposition of Cys-538 residues for disulfide formation. Alternatively transitions between α-helix and less ordered structures in this region may enable appropriate disulfide bonding distances to be achieved.

Interestingly, disulfide bonding was not observed for T569C in MBP/gp41(528-L-677) even though the Cβ-Cβ separation for this residue is 5.4 Å and close to the optimal disulfide bonding distance. This observation could be explained by the well-ordered nature

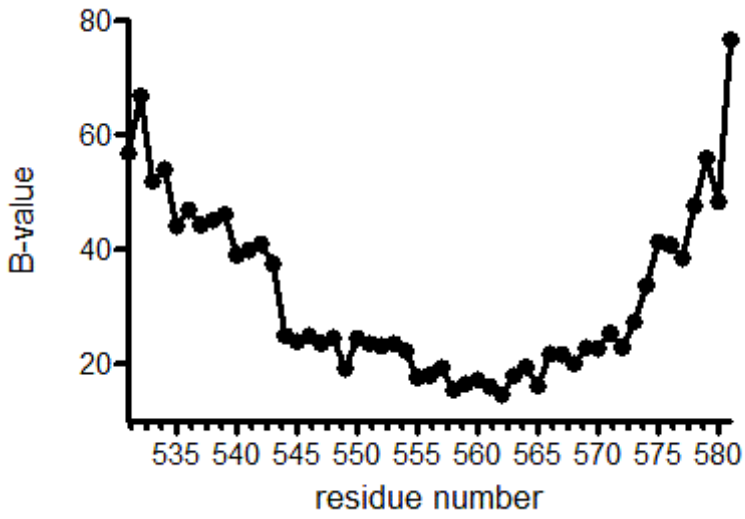


Figure 4.18 C-alpha B-value versus residue number for the polar segment-HR1 helix

B-values for residues 544-572 hover around 20 indicating restricted mobility in this portion of the coiled coil; on the other hand mobility in the polar segment terminal is enhanced as seen from B-values that jump to ~50. Polar segment-HR1 helix was derived from the 2X7R coordinates.

of the trimer of hairpins core domain (average *B*-value for Cα atoms in the 544-572 sequence = 20.88) (Figure 4.2 and 4.18). The region about Thr-569 is stabilized by both knobs-into-holes packing interactions mediated by *a*- and *d*- position hydrophobic residues (Leu-566, Ile-573, Leu-566) of HR1 and extensive interactions between the coiled coil and antiparallel HR2 helices (His-564-Tyr-638, Asn-567-Thr-639, Trp-571-Trp-631-Leu-568-Ile-635, and Lys-574-Asp-632), which may inhibit mobility and closer apposition of the substituted Cys-569 residues. The absence of HR1-HR2 interactions in prefusion gp41 may confer mobility to this region such that Cys-569 side chains are close enough to form an interprotomer disulfide. We note that a 0.06-Å inward displacement per HR1 helix relative to the trimer of hairpins structure would be sufficient for a disulfide to form.

The results of Buzon *et al.* (2010) indicate that the coiled coil extension formed by the polar segment in the trimer of hairpins does not exhibit regular knobs into holes packing of *a*- and *d*- position residues but is stabilized by hydrophobic interactions with adjacent antiparallel MPER helices in a terminal clasp structure (Figure 4.12). We found that the I535A/V539G mutation, which destabilizes the polar segment-MPER clasp in MBP/gp41(528-L-677)$^{AD.ST}$ and blocks membrane fusion function (Lay *et al.*, Manuscript in Preparation) also reduced Cys-538-mediated disulfide formation by ~ 40 %. These data suggest that ablation of hydrophobic interactions between Ile-535 and Ile-675 affects the structural context of Cys-538 such that disulfide formation is less favourable. On the other hand, the efficiency of Cys-538 disulfide formation in MBP/gp41(528-591)$^{AD.ST}$, which corresponds to the gp41 HR1 prehairpin was similar to that observed for the MBP/gp41(528-L-677)$^{AD.ST}$ trimer of hairpins, and was not affected by I535A/V539G. Taken together, these data suggest that the initial extension of the central coiled coil through the polar segment occurs at the "early" prehairpin stage. However, appropriate

interactions between the polar segment and MPER are important to maintain the structure of this region.

The results of these studies allow a model for the conformational states of gp41 that are associated with membrane fusion function to be inferred. *1)* Disulfide bonding via T569C in prefusion gp41 suggests that the C-terminal portion of HR1 participates in interprotomer contacts and likely forms a coiled coil in this context. *2)* Disulfide bonding via T538C in the prehairpin and trimer of hairpins MBP/gp41 models suggests that following activation, the coiled coil is propagated in an N-terminal direction to encompass the polar segment in these "activated" structures. These changes would propel the fusion peptide, which is immediately N-terminal to the polar segment, towards the target cell membrane for insertion. *3)* The extended coiled coil structure of the prehairpin provides a packing surface for both the HR2 sequence and adjacent MPER, forming a trimer of hairpins, which apposes membrane-embedded fusion peptide and TMD sequences, leading to fusion (Figure 4.19).

The data suggest that terminal clasp formation involves a 2-step process whereby polar segments initially assemble into a coiled coil that provides a packing surface for the MPER. Assembly of the 6-HB core domain by the antiparallel packing of HR2 helices against the coiled coil of HR1 helices juxtaposes the MPER and the polar segment coiled coil leading to formation of the terminal clasp. The data therefore argue against a cooperative mechanism of terminal clasp assembly where polar segment coiled coil formation occurs concomitantly with polar segment-MPER interactions. A 2-step clasp formation mechanism where a coiled coil of polar segments assembles prior to MPER engagement points to a novel structural determinant that can be exploited for antiviral drug discovery. For example, tr665, which contains a "naked" polar segment, can be used to screen chemical or phage peptide display libraries for agents that bind to this N-terminal

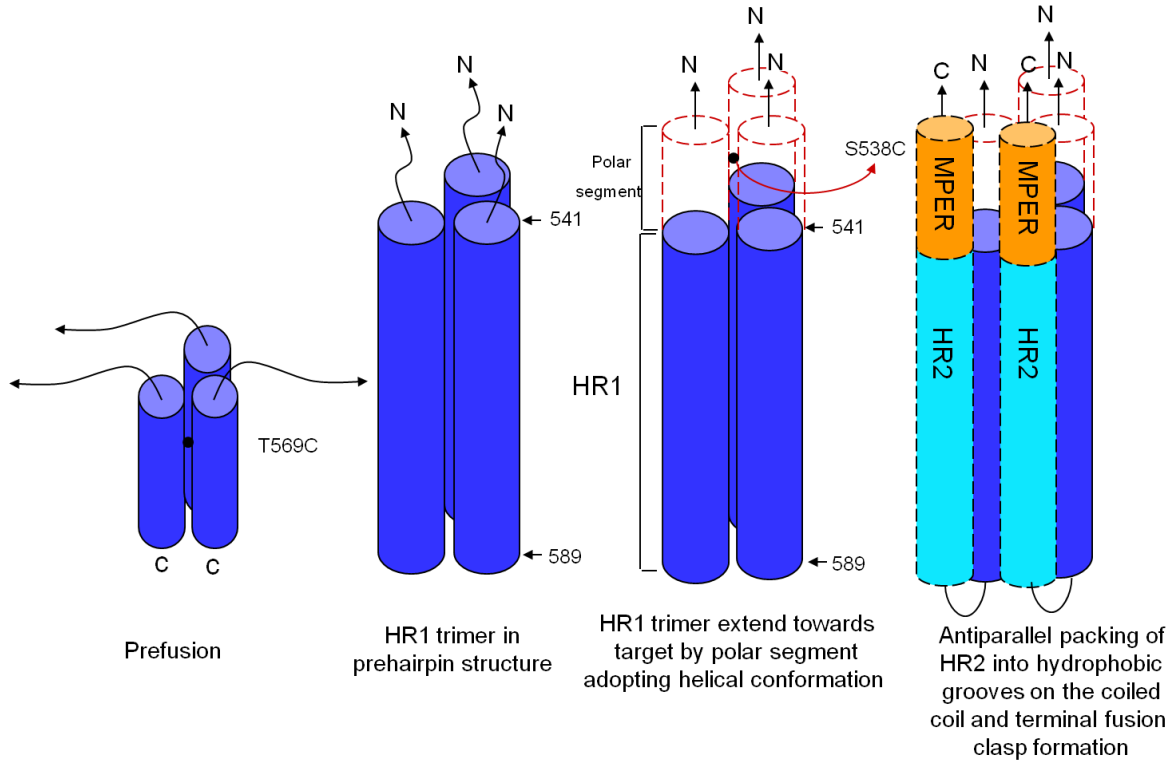


Figure 4.19 A model of the conformational states of gp41 that are associated with membrane fusion function

The antiparallel packing step (last step in the figure) is believed to involve zipper formation along the coiled coil axis that extends into the membranes in a manner similar to the F protein of the parainfluenza virus 5 (PIV5). Membrane fusion in the PIV5 occurs via the lipid centric or the protein pinprick models (Donald *et al.*, 2011). Those two models are likely to provide a better understanding of the fusion process and be applicable to all type I fusion proteins including gp41 of HIV-1.

region and block clasp formation. Binding specificity can be verified with MBP/gp41(528-L-677)$^{AD.ST}$ where the polar segment is occluded due to interactions with the MPER, and the activity of leads can be determined in cell-cell fusion and virus entry assays.

The fusion mechanism employed by class I viral fusion proteins involves their initial folding into a prefusion, metastable conformation that undergoes large conformational rearrangements to a lower energy state in response to a variety of activation triggers. The key feature of this mechanism is the antiparallel association of trimeric N-terminal HR1 and C-terminal HR2 segments in a hairpin-like structure that apposes viral and cell membranes. However, structural variation is observed for the fusion peptides, the inter-helical region separating HR1 and HR2 and the final structure of HR2 when bound to the coiled coil. Structural biological studies of EBOV GP2, influenza virus HA2 and paramyxovirus F in prefusion and fusion activated states have revealed the activation mechanism for three different class I fusion proteins that vary in their structural complexity.

Ebola virus GP2 appears to represent the simplest activation model where the C-terminal portion of HR1 (HR1D) forms a trimeric coiled coil in the prefusion state and upon exposure to the low endosomal pH and cathepsins B and L, the other three segments of HR1 extend the coiled coil structure and projecting the internal fusion loop towards the target cell membrane (Lee *et al.*, 2008; Weissenhorn *et al.*, 1998b). In this case, HR2 packs as a coil-helix-coil structure against the hydrophobic grooves on the outer surface of the coiled coil in an antiparallel orientation. For influenza virus HA2, low pH triggers dramatic refolding events over the entire ectodomain. Prefusion HA2 is composed of a 51-residue central coiled coil (Figure 4.20; orange-pink-red cylinder) that is linked via a random coil (Figure 4.20; yellow) to a shorter 19-residue outer helix (Figure 4.20; salmon cylinder) that packs against it (Wilson *et al.*, 1981). Upon activation, the random coil

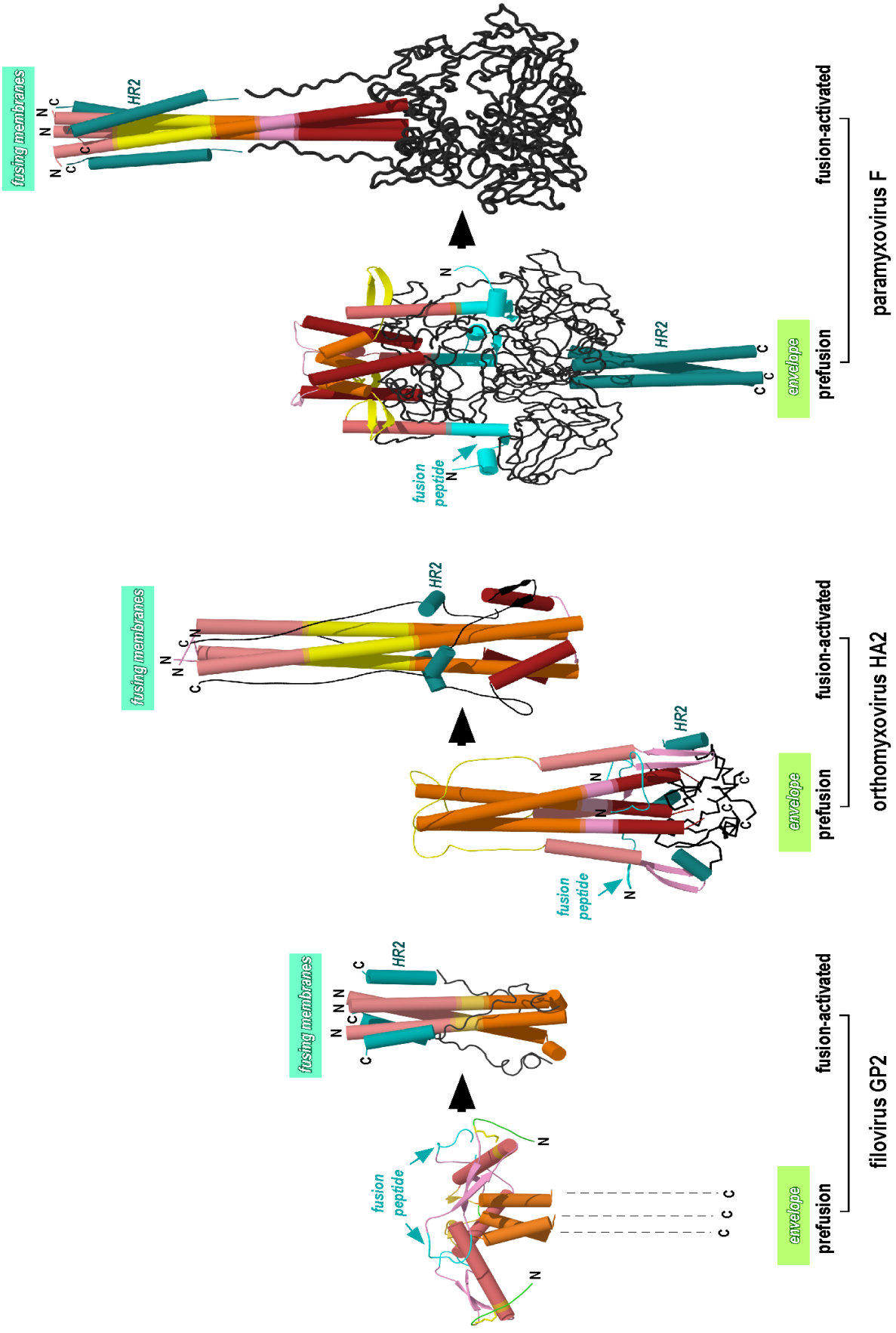

fusing membranes
HR2
fusion peptide
envelope
prefusion
fusion-activated
paramyxovirus F
orthomyxovirus HA2
filovirus GP2

Figure 4.20 Prefusion vs. fusion-activated structures of EBOV GP2, influenza virus HA2 and paramyxovirus F

EBOV GP2 represent the simplest class I fusion protein with a simple extension of HR1D (orange cylinder) towards the target membrane and then antiparallel packing of the HR2 to form the hairpin. Influenza HA2 is more complex involving both an N-terminal extension of the central coiled coil from the orange portion onwards but also a collapse and repacking involving the pink and dark red C-terminal segments to form the chain reversal region. Parainfluenza F is the most complex; this structure would involve major repacking of the HR1 coiled coil core with expulsion of the small orange coiled coil from the core to enable coiled coil formation by the pink and dark red C-terminal segments of HR1 and then followed by the N-terminal extension. HR2 also forms a coiled coil at the envelope-proximal end of the prefusion structure, which is separated from HR1 sequences by a large domain. The HR2 coiled coil appears to dissociate and then HR2 sequences repack as the antiparallel outer layer in the activated structure.

acquires helical structure and, together with the outer helix, extends the coiled coil in an N-terminal direction, projecting the fusion peptide toward the target cell. Together with these changes, the C-terminal segment of the coiled coil (Figure 4.20; pink and red segments of the cylinder) melts to form a region of chain reversal, enabling the antiparallel packing of the C-terminal region onto the newly formed coiled coil as a coil-helix-coil structure (Bullough *et al.*, 1994; Chen *et al.*, 1999).

The F protein of paramyxoviruses represents the most complex class I fusion protein so far revealed by x-ray crystallography. A monomer of the prefusion structure of parainfluenza virus 5 (PIV5) F protein is shown in figure 4.21A. It consists of a globular head region connected to a helical stalk. The globular head comprises 3 domains: DI and DII are rich in β-strands while DIII forms a core structure around which HR1 (or "HRA") is wrapped (Yin *et al.*, 2006). The helical stalk comprises HR2 segments, which form a coiled coil in the F trimer. The structure of human parainfluenza virus 3 (PIV3) F in a fusion-activated conformation indicates that major refolding events occur in HR1 (or "HRA") and HR2 (or "HRB") upon receptor binding by the viral HN protein (Dorig *et al.*, 1993; Yanagi *et al.*, 2002; Negrete *et al.*, 2005; Bonaparte *et al.*, 2005; Feldman *et al.*, 1999), while the structure of DI-III remains constant throughout the refolding process. The structural changes occurring in HR1 during transition of the prefusion F trimer to a fusion-activated form are compared in figure 4.20. A mini-domain comprising a short coiled coil (HR1 residues 168-176; orange cylinder) that is linked via a random coil (HR1 residues 178-183; pink) to an antiparallel outer helical layer (HR1 residues 184-203, red cylinder) resides at the top of the DI-III globular domain (black). In the activated structure, the linker acquires helical structure, the orientation of the short coiled coil forming helices is reversed and the outer layer helices now form the base of the central coiled coil. The HR1 (or "HRA") helices, β-strands and linkers reorganize to form an extended 72-residue coiled

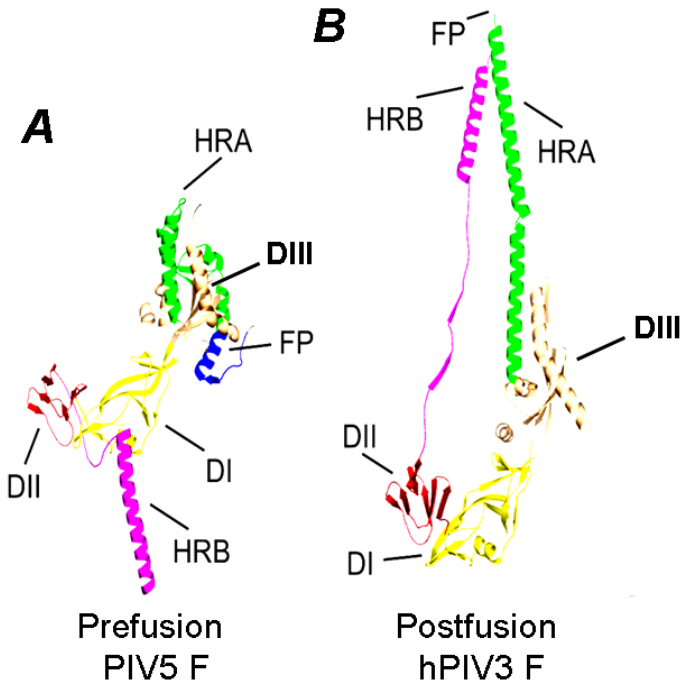


Figure 4.21 Structure of paramyxovirus F monomer in the pre and postfusion conformation

Helices, β-strands and linkers of HR1 ("HRA" green) form a mini-domain atop the DI-DIII globular domain (brown, yellow, red), while HR2 ("HRB", magenta) forms a helical stem in the prefusion state. In the fusion-activated or postfusion state, HR1 elements are reorganised into an extended α-helix that forms a trimeric coiled coil and projects the fusion peptide (blue) towards the target cell. Domains I-III retain their structural organization during the prefusion to fusion-activated/postfusion transition. The prefusion trimer is additionally stabilized by DI and II through interactions with HR2, while a short DIII helix (gold) interacts with the HR1 coiled coil in both prefusion and fusion-activated conformations. Adapted from Lamb and Jardetzky (2007).

coil, which projects the fusion peptide towards the target cell and provides a packing surface for the HR2 stalk helices (Yin *et al.*, 2005) (see also Figure 4.21B). The results of this study suggest that retroviruses have adopted a simple activation mechanism that is analogous to EBOV, where a small part of the coiled coil is formed in the prefusion structure that upon receptor binding is extended in an N-terminal direction to include the polar segment, moving the fusion peptide closer to the target cell. This step is followed by the antiparallel packing of HR2 into hydrophobic grooves on the outer surface of the coiled coil to form the trimer of hairpins structure.

Chapter 5: Forced Evolution Reveals a Functional Linkage Between the DSR and MPER of gp41

5.1 Aim

To identify domains of the gp120-gp41 complex that are involved in transmitting the activation signal from receptor-bound gp120 to gp41.

5.2 Introduction

The mechanism whereby conformational signals evoked in retroviral SU glycoproteins by receptors(s) are transmitted to their partner TM glycoproteins, to trigger their refolding into fusion-active forms is being elucidated. Evidence is accumulating to suggest that the association site formed by the DSR of HIV-1 gp41 and the terminal C1 and C5 regions of gp120 (Figure 5.1A) could act a synapse for SU-to-TM conformational signalling (Poumbourios *et al.*, 2003; Maerz *et al.*, 2001; Binley *et al.*, 2003). The highly conserved 30-residue segment representing the DSR of gp41 links the HR1 and HR2 (Figure 5.1B, C). It is concealed in the prefusion conformation of Env, but becomes recognisable by antibody following gp120 binding to the CD4 receptor (Sattentau *et al.*, 1995; Finnegan *et al.*, 2002). In the trimer of hairpins structure of gp41, the DSR forms a mobile linker connecting the HR1 coiled coil and HR2 outer helical layer (Lay *et al.*, 2004). The findings that W596L and W610F mutations in the DSR of HIV-$1_{QH1549.13}$ block the formation of the gp41 prehairpin intermediate, that is induced by gp120-CD4 interactions, and the initial lipid mixing (hemifusion) phase of the membrane fusion cascade led to the proposal that the DSR acts as a sensor of receptor-induced conformational changes in gp120 that lead to the fusion activation of gp41 (Bellamy-McIntyre *et al.*, 2010).

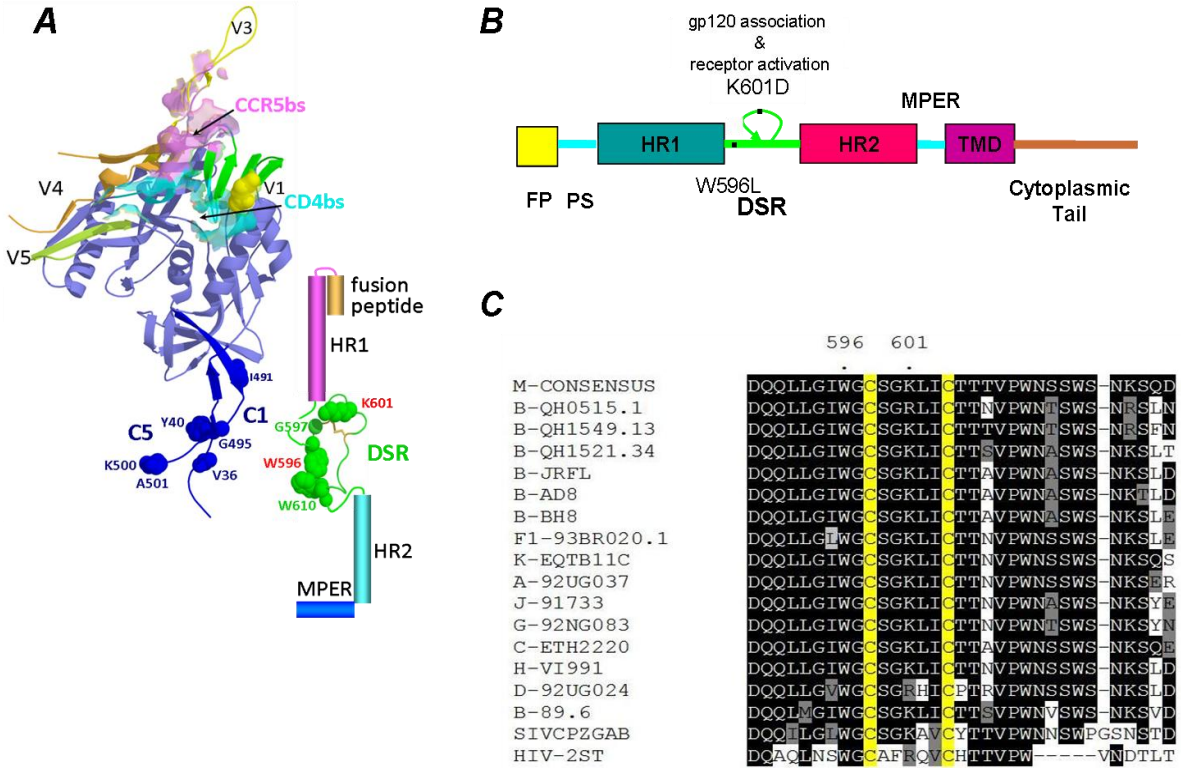


Figure 5.1 Location of the DSR mutations W596L and K601D within gp41

(A) gp120-gp41 structural schematic showing the W596L and K601D mutation site. (**B)** Schematic of gp41 denoting the location of W596L and K601D DSR mutations. FP, fusion peptide; PS, polar segment. (**C)** Alignment of DSR amino acid sequences showing the conservation of Trp-596 and Lys-601 among different HIV-1 clades.

The aim of this study was to identify determinants in the gp120-gp41 complex that are functionally linked to the activation mechanism by the method of forced sequence evolution. The experimental approach involves mutating the sequence or structural element under investigation, which is the DSR in this case, and then characterising the mutant phenotype. The mutant virus is then subjected to sequential passage in permissive target cells to generate revertant viruses that have overcome or "suppress" the introduced defect by the emergence of additional mutations (suppressors) in the viral genome. The error-prone nature of the reverse transcription step is responsible for the generation of variable species called quasispecies (Preston, 1997; Mansky and Temin, 1995). The variant with the best fitness will outgrow the other variants.

In this study, the macrophage-adapted R5 strain, HIV-1_{AD8} (Theodore *et al.*, 1996), bearing a W596L/K601D double mutation in the DSR was subjected to sequential culture in U87.CD4.CCR5 cells until revertant viruses were obtained. The W596L/K601D mutation in this viral strain leads to a gp120 shedding defect and a block in cell-cell fusion and virus replication. The defective gp120-gp41 association and viral replication phenotype of the W596L.K601D mutant was suppressed by two pathways: *1)* second site mutations led to deletions at Thr-389 and Trp-390 dyad in V4, operating in conjunction with a D601H (KH) pseudoreversion in the DSR and D674N in the MPER. *2)* The D601H (KH) pseudoreversion again appeared but in this case was linked to D674E in the MPER. The D601H pseudoreversion restored gp120-gp41 association and partially restored viral replication competence, whereas D674E further optimised replication. These data indicate, for the first time, that the MPER is functionally linked to the association/activation synapse of gp120-gp41.

5.3 Results

5.3.1 Phenotype of the W596L.K601D HIV-1_{AD8} DSR mutant

The DSR of gp41 is associated with the C1 and C5 regions of gp120, and forms a potential pathway for transmitting the fusion activation signal which results from gp120 binding to CD4 and chemokine receptors (Binley *et al.*, 2000; Abrahamyan *et al.*, 2003; Binley *et al.*, 2003; Helseth *et al.*, 1991; Maerz *et al.*, 2001; Poumbourios *et al.*, 2003). For this study, DSR mutants with disrupted gp120 association and/or gp120-gp41 conformational signalling were identified for serial passaging in U87.CD4.CCR5 cells to select second site suppressor mutation(s). The location of the suppressor mutation(s) would suggest a functional link with the DSR and enable a study of how the different regions cooperate in the fusion process.

To this end, K601D and W596L/K601D mutations were introduced to the HIV-1_{AD8} Env expression vector, pCDNA3.1-AD8*env*, and their effects on gp120-gp41 association and cell-cell fusion determined. gp120-gp41 association phenotype was determined by immunoprecipitation of biosynthetically labelled glycoproteins expressed in 293T cells using IgG14. The K601D mutation was associated with a slight increase in gp120 being shed into the culture supernatant when compared to the WT, whereas ~ 90 % of total gp120 was shed into the supernatant of W596L/K601D (Figure 5.2A). These data contrast the lack of a shedding phenotype for the single mutants, W596L and K601E, which were included as controls. Consistent with the gp120-shedding phenotypes observed for K601D and W596L/K601D, cell-cell fusion function was markedly diminished for both mutants across an Env concentration range (Figure 5.2B). The gp120-shedding phenotype was recapitulated in a viral context for W596L/K601D (Figure 5.3A) and the mutant virus was unable to replicate in U87.CD4.CCR5 cells (Figure 5.3B). Thus W596L/K601D has a gp120-shedding phenotype, which translates to a fusion and replication defect.

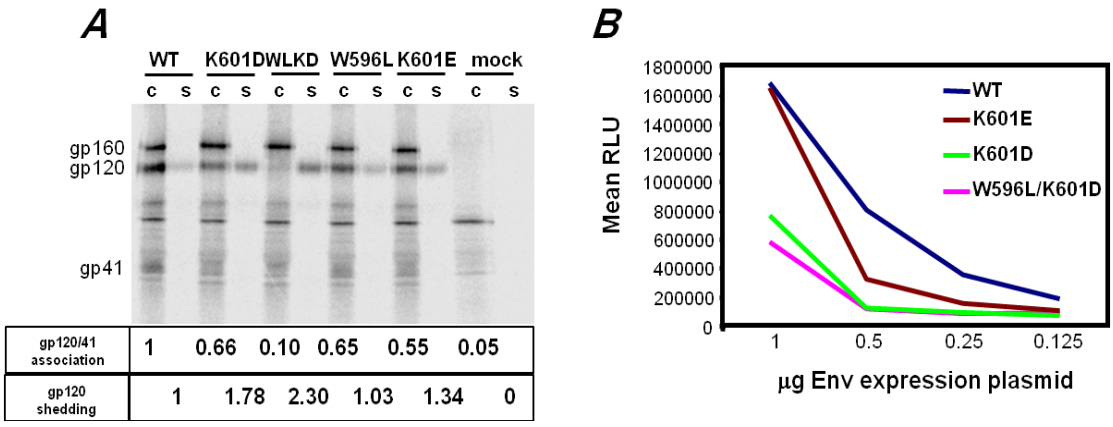


Figure 5.2 Association and fusion phenotypes of HIV-1AD8-WLKD

(A) gp120-anchoring ability of gp41 mutants. 293T cells were cotransfected with pcDNA3.1-AD8*env* and pCAG-T7 vectors. At 48 h post-transfection, cells were labeled with Tran-^{35}S-label for 45 min and chased in complete medium for 6 h before lysis. Cell lysates (*C*) of metabolically labelled WT, K601D, W596L.K601D (WLKD), K601E or empty vector (mock) and corresponding clarified culture supernatants (*S*) were immunoprecipitated with IgG14 and protein G sepharose. Proteins were analysed under reducing conditions in 8-12% SDS-PAGE gels and scanned in a Fuji phosphorimager. **(B)** Cell-cell fusion activity. 293T effector cells were cotransfected with 1 ug pCAG-T7 plus 1, 0.5, 0.25 or 0.125 ug of pcDNA3.1-AD8*env* and then cocultured (18 h, 37 ˚C) with BHK21 target cells that had been cotransfected with pc.CCR5 and pT4*luc*. The mean relative light units (RLU) of a representative experiment is shown. Data in this figure were obtained from Dr. P. Poumbourios.

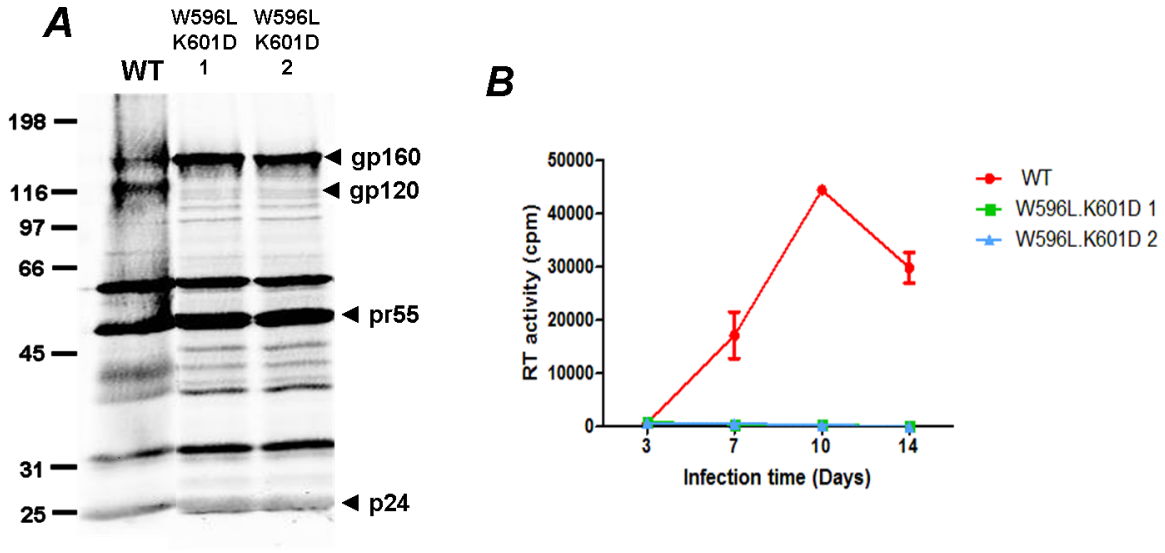


Figure 5.3 Virion protein gel and replication kinetics

(A) Virion characterization. Wild type and mutated HIV-1_{AD8} virions produced by transfected 293T cells were pelleted through a 25 % sucrose cushion and analysed by SDS-PAGE under reducing conditions in 7.5-15% polyacrylamide gradient gel and western blotting with IgG14. (**B)** 14-day replication kinetics in U87.CD4.CCR5 cells. Virus produced in 293T cells was normalised for RT activity and used to infect U87.CD4.CCR5 cells. RT activity was measured in the culture fluid at days 3, 7, 10 and 14. The mean RT activity ± standard deviation of three samples is shown.

5.3.2 Long term culture of the W596L.K601D HIV-1_{AD8} DSR mutant

Two independent W596L/K601D infectious HIV-1_{AD8} clones were subjected to long-term culture in U87.CD4.CCR5 cells with serial passage of cell-free virus onto fresh U87.CD4.CCR5 cells every 10 days. Evidence of replication was not observed for either clone even after 50 days of culture (Figure 5.4A). We reasoned that the lack of gp120 in mutant viruses may block entry and therefore reverse transcription, which is the critical step required for the generation of second site mutations. In order to initiate entry of W596L/K601D, the mutant virus was pseudotyped with vesicular stomatitis virus glycoprotein G (VSV G), which directs a single cycle of infection via the endosomal pathway. The VSV G-pseudotyped viruses were allowed to infect U87.CD4 CCR5 for 24 h prior to extensive washing and trypsinisation to remove any residual pseudotyped virus. The cell-free virus was then sequentially passaged on fresh cells every 10 days. This approach resulted in the restoration of infectivity after 47 and 30 days in culture for W596L.K601D cultures 1 and 2, respectively (Figure 5.4B).

5.3.3 Dominant genotypes of W596L.K601D HIV-1_{AD8} revertants

In order to identify potential suppressors of W596L.K601D, the *env* region was PCR-amplified from genomic DNA isolated from infected cells at days 10, 20, 30, 40 and 50. The PCR products were cloned into the pΔKAD8*env* expression vector and the entire *env* region subjected to DNA sequencing. ***W596L.K601D clone 1 (WLKD 1).*** A D601H pseudoreversion emerged at day 10 (2/6 clones) and preceded the dominant W596L.K601H.D674E genotype, which emerged at day 20 (3/6 clones) and was maintained to day 50; 2 of 6 day-50 clones contained W596L.K601H.D674G. A minority of clones containing L85M in addition to W596L.K601H.D674E were observed at days 20, 30 and 40 but not at day 50. The emergence of replication-competent virus in the

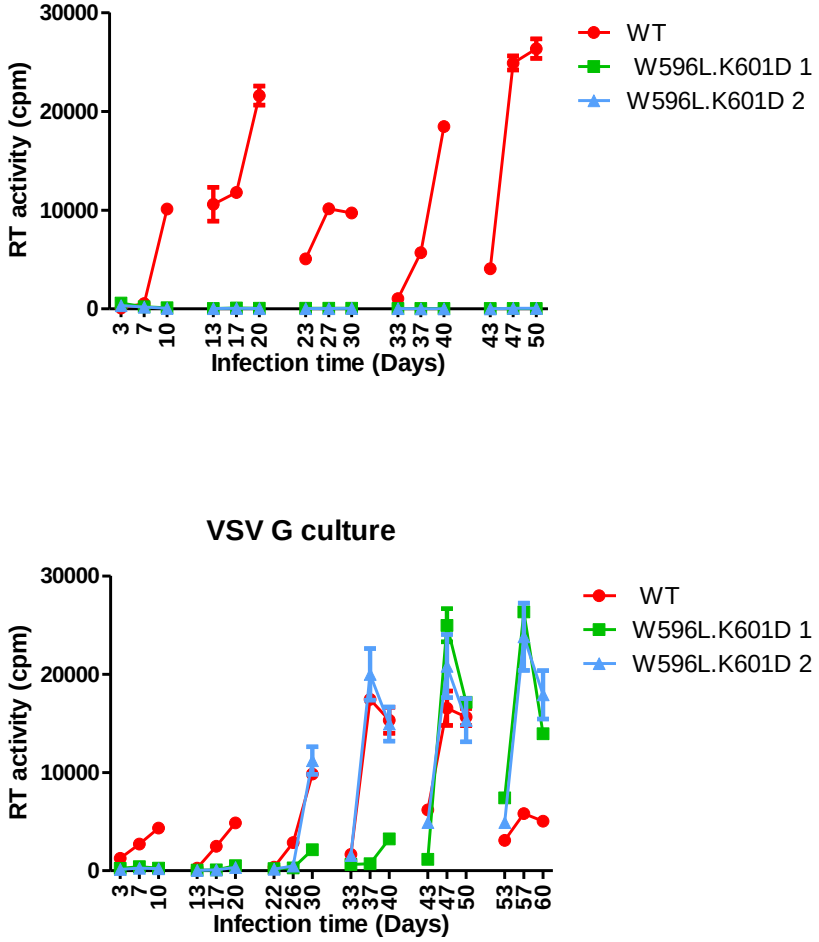


Figure 5.4 Long Term Culture of HIV-1AD8-WT and HIV-1AD8-WL.KD in U87.CD4.CCR5

(A) Wild type and W596L.K601D-mutated HIV-1_{AD8} virus stocks produced by transfected 293T cells were normalized according to RT activity and used to infect U87.CD4.CCR5 cells. The cell-free virus obtained at day 10 was filtered through a 0.45 µm nitrocellulose filter, normalised according to RT activity and used to infect fresh U87.CD4.CCR5 cells. This procedure was used to subject viruses to 5 sequential passages in total. Supernatants from the infected cell cultures were removed at days 3, 7, and 10 for RT assay as a measure of viral replication. **(B)** As for (A) except that VSV G-pseudotyped virus was used to initiate infection in the first passage. The results shown represent the mean RT activity ± standard deviation of triplicate samples.

WL.KD 1 culture occurred during the 5th passage and correlated with the presence of the following genotypes recovered at days 40 and 50: WL.KH (7/17), WL.KH.DE (7/17), WL.KH.DG (2/17) and LM.WL.KH.DE (1/17) (Tables 5.1-5.5 and Figure 5.5A). ***W596L.K601D clone 2.*** At day 10, 3/6 clones contained the D601H pseudoreversion, while 3 others contained the Thr-389-Trp-390 deletion in V4 (ΔTW 389/90), the D601H pseudoreversion and D674N in the MPER; the 6 clones also contained other mutations that were randomly distributed throughout *env*. The ΔTW389/90.W596L.K601H.D674N genotype appeared to be under purifying selection between days 20 and 30 but by days 40 and 50, the dominant genotype was W596L.K601H.D674N (Table 5.1-5.5 and Figure 5.5B).

5.3.4 Replication kinetics of representative revertant virus genotypes

To determine how the W596L.K601D replication defect became suppressed in the long-term WLKD1 and WLKD2 cultures, the dominant genotypes were reconstructed in the context of the pAD8 proviral clone and the pΔKAD8*env* expression vector. In the case of WLKD1, viral replication competence in U87.CD4.CCR5 was partially restored by D601H in the DSR and was optimised further by inclusion of D674E in the MPER (Figure 5.6). The addition of L85M to W596L.K601H.D674E did not improve replication relative to W596L.K601H.D674E, suggesting that this mutation did not confer any functional advantage. Interestingly, the addition of the D674G MPER mutation to W596L.K601H resulted in a replication block. These data suggest that D601H and D674E act synergistically to suppress W596L.K601D. The analysis of WLKD2 genotypes indicated step-wise improvements in replication competence for W596L.K601H.D674N, ΔTW389/90.W596L.K601H.D674N and W596L.K601H, respectively. These data suggest that, in contrast to D674E, D674N is inhibitory in the context of W596L.K601H and requires the Thr-389-Trp-390 deletion in V4 to partially relieve this inhibition. The G145E

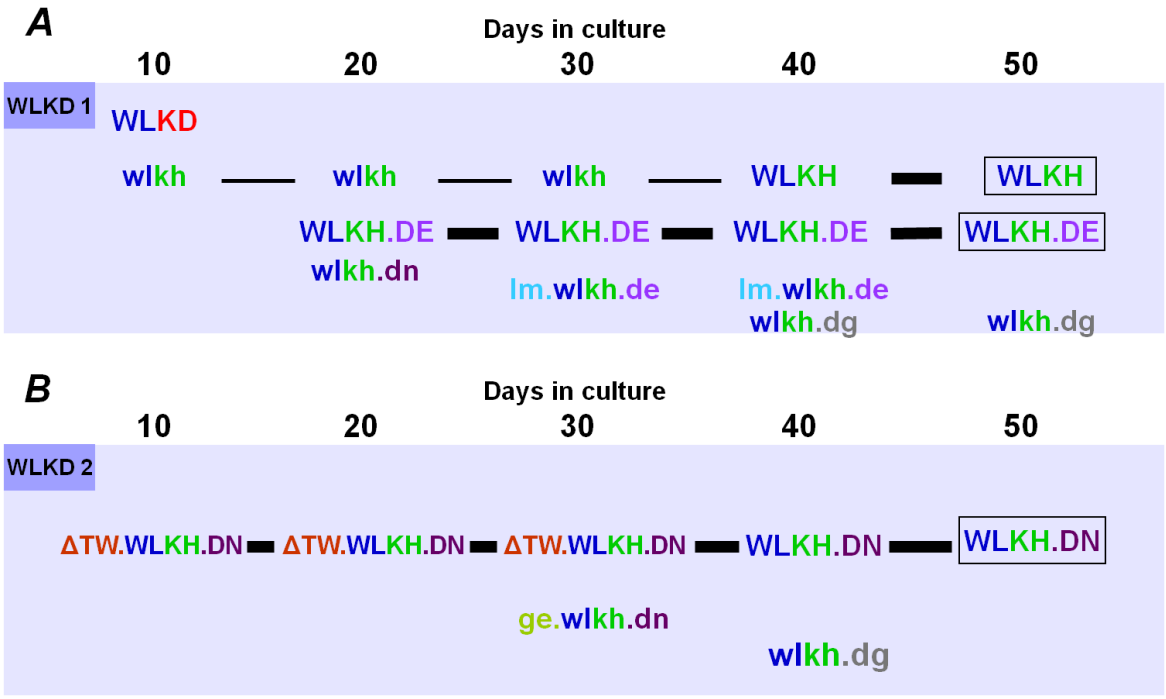


Figure 5.5 Dominant *env* genotypes in WLKD 1 and 2 cultures

The *env* region of proviral DNA isolated at days 10, 20, 30, 40 and 50 from WLKD 1 and 2 was PCR amplified and cloned into pΔKAD8*env*. The entire *env* region present in individual clones was sequenced using Bigdye terminator 3.1. The reversion pathways in WLKD1 and WLKD2 are shown in **(A)** and **(B)** respectively. Upper case and ▬ denotes a major evolutionary pathway, while lower case and —— represent a minor pathway. Lower case only genotypes represent non-dominant genotypes arising at the specified days. WL, W596L; KD, K601D; KH, K601H; DE, D674E; DN, D674N; LM, L85M; DG, D674G; ΔTW, deletion of T389 and W390; GE, G145E.

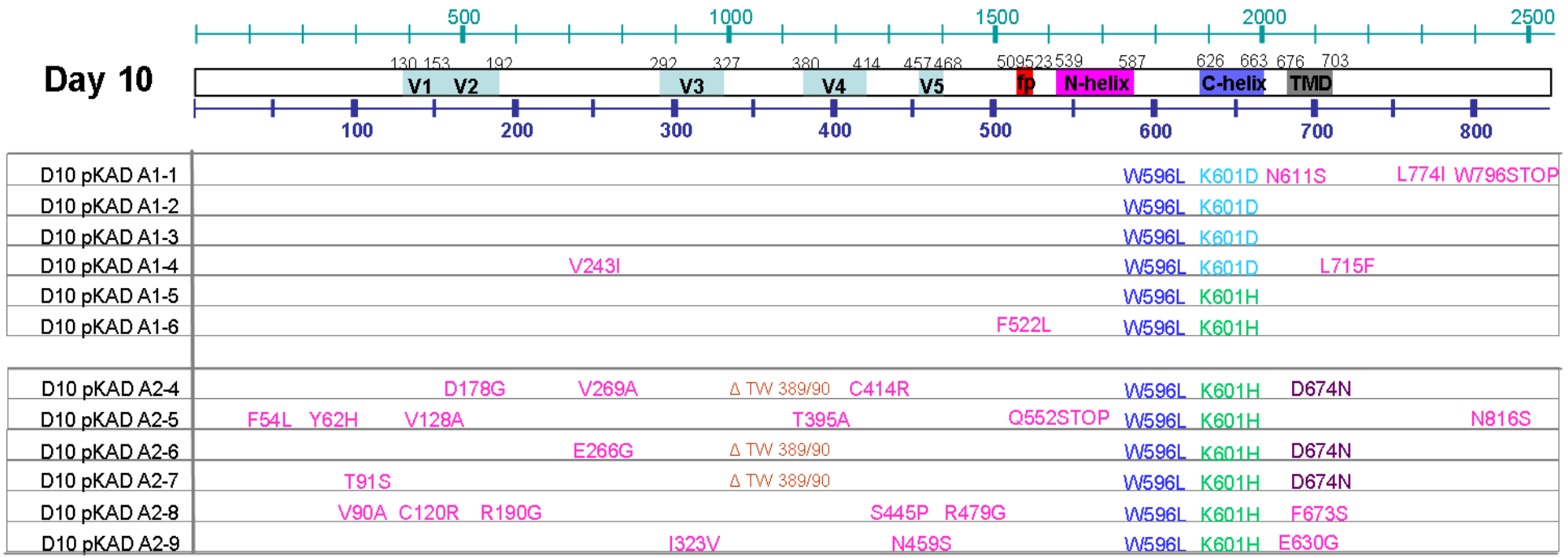


Table 5.1 Day 10 genotypes of W596L.K601D HIV-1AD8 revertants

Sequence analysis of revertant clones obtained at day 10 in figure 5.4B. A1 clones derived from WL.KD1; A2 clones derived from WL.KD2. The *env* region of proviral DNA isolated at day 10 was PCR amplified and cloned into pΔKAD8*env*. The entire *env* region present in individual clones was sequenced using Bigdye terminator 3.1. Note the presence of mutations in several clones, for example ΔTW389-390, K601H and D674N. Sequence analysis was conducted using Vector NTI 10.3.0. Numbering is based on HXB2.

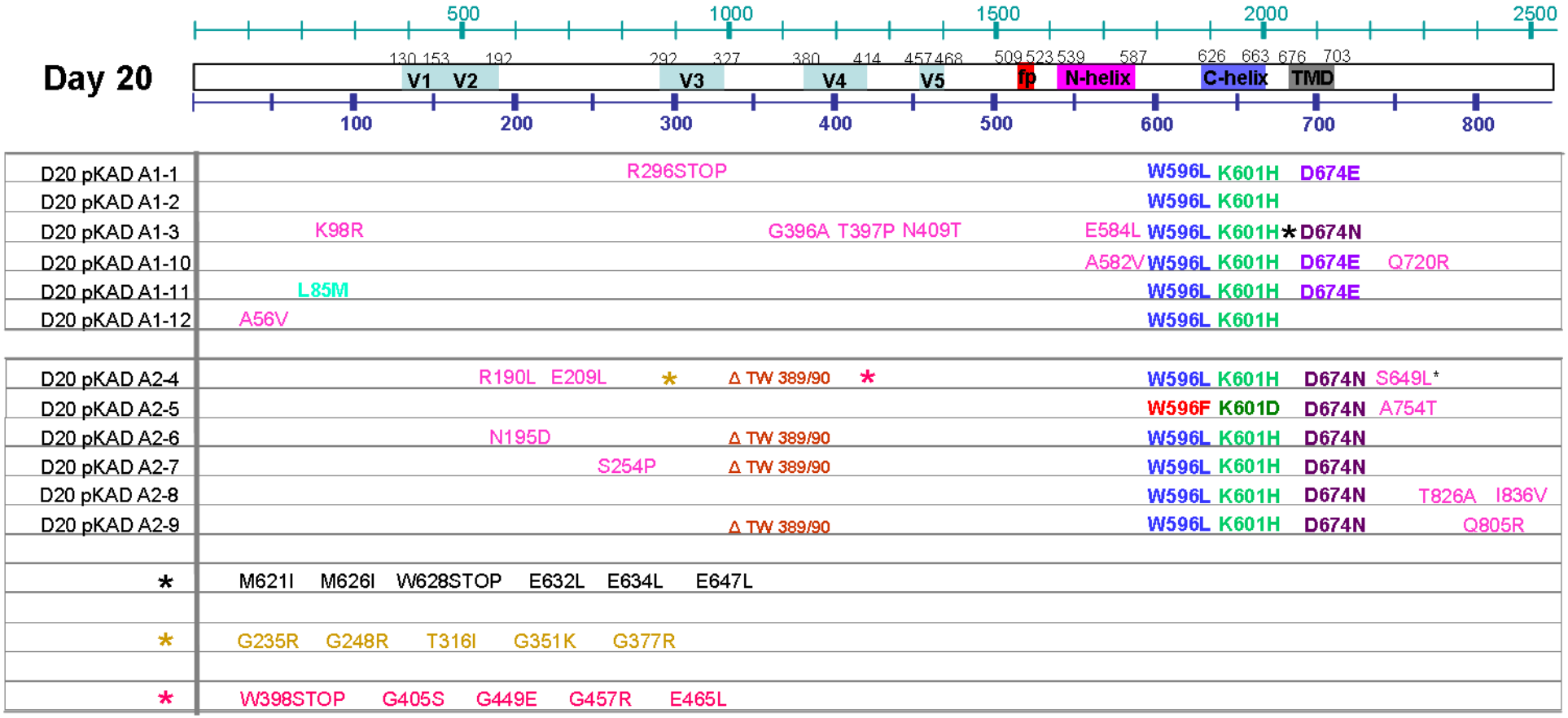


Table 5.2 Day 20 genotypes of W596L.K601D HIV-1AD8 revertants

Sequence analysis of revertant clones obtained at day 20 in figure 5.4B. Note the presence of mutations in several clones, for example L85M, ΔTW389-390, W596F, K601H, D674E and D674N. Small asterisk (*) denotes a frameshift for alignment purposes. Numbering is based on HXB2.

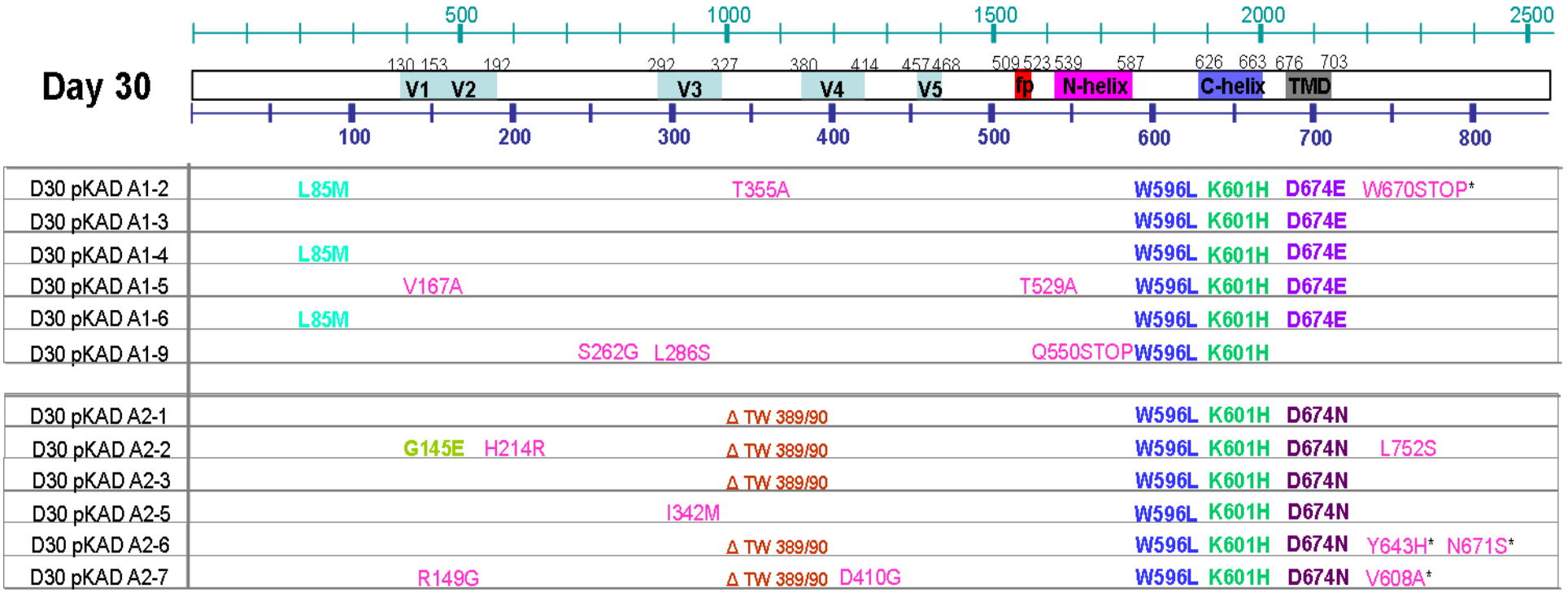


Table 5.3 Day 30 genotypes of W596L.K601D HIV-1AD8 revertants

Sequence analysis of revertant clones obtained at day 30 in figure 5.4B. Note the presence of mutations in several clones, for example L85M, G145E, ΔTW389-390, K601H, D674E and D674N. Small asterisk (*) denotes a frame shift. Numbering is based on HXB2.

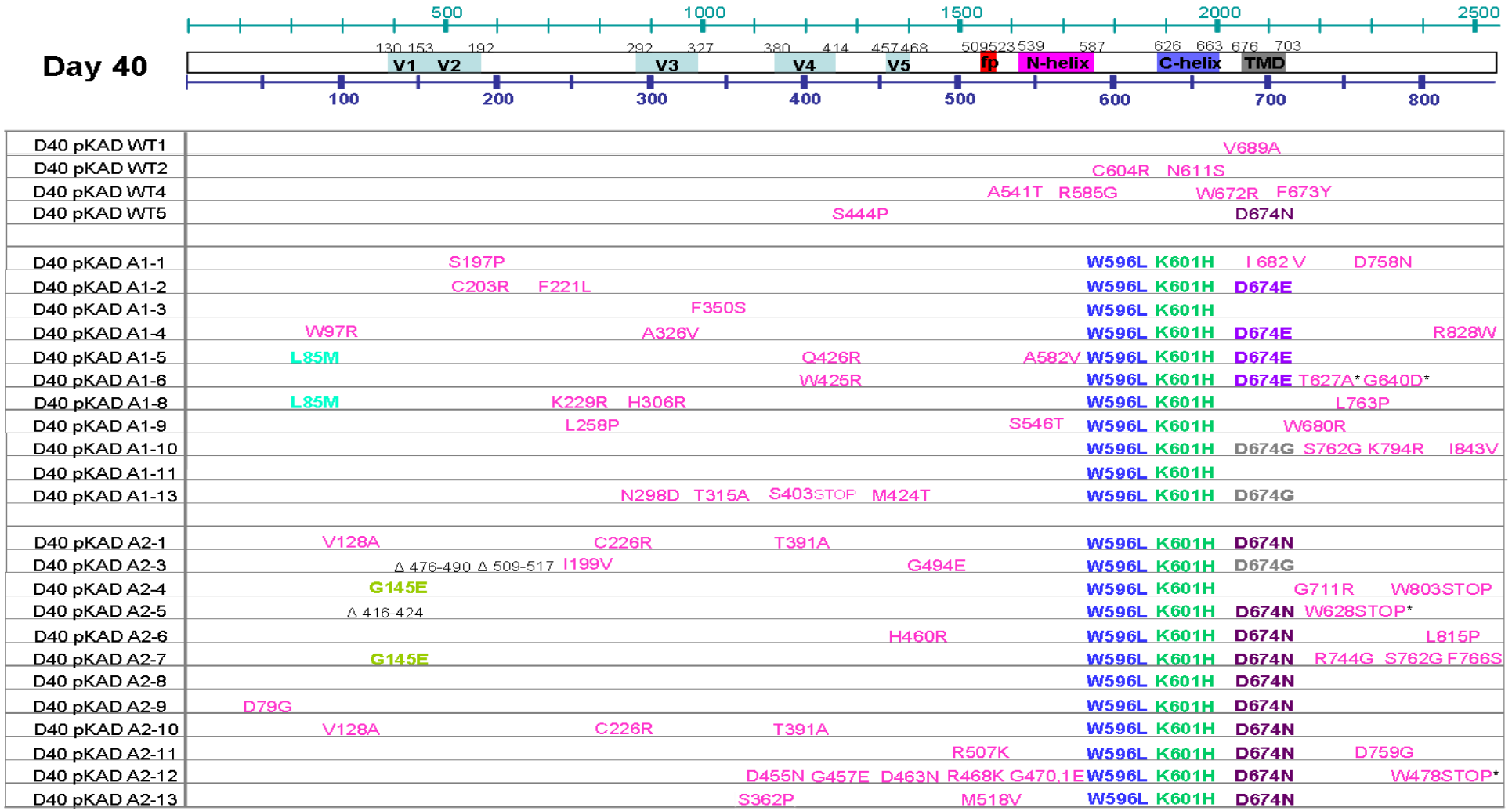


Table 5.4 Day 40 genotypes of W596L.K601D HIV-1AD8 revertants

Sequence analysis of revertant clones obtained at day 40 in figure 5.4B. Note the presence of mutations in several clones, for example L85M, G145E, K601H, D674E, D674G and D674N.

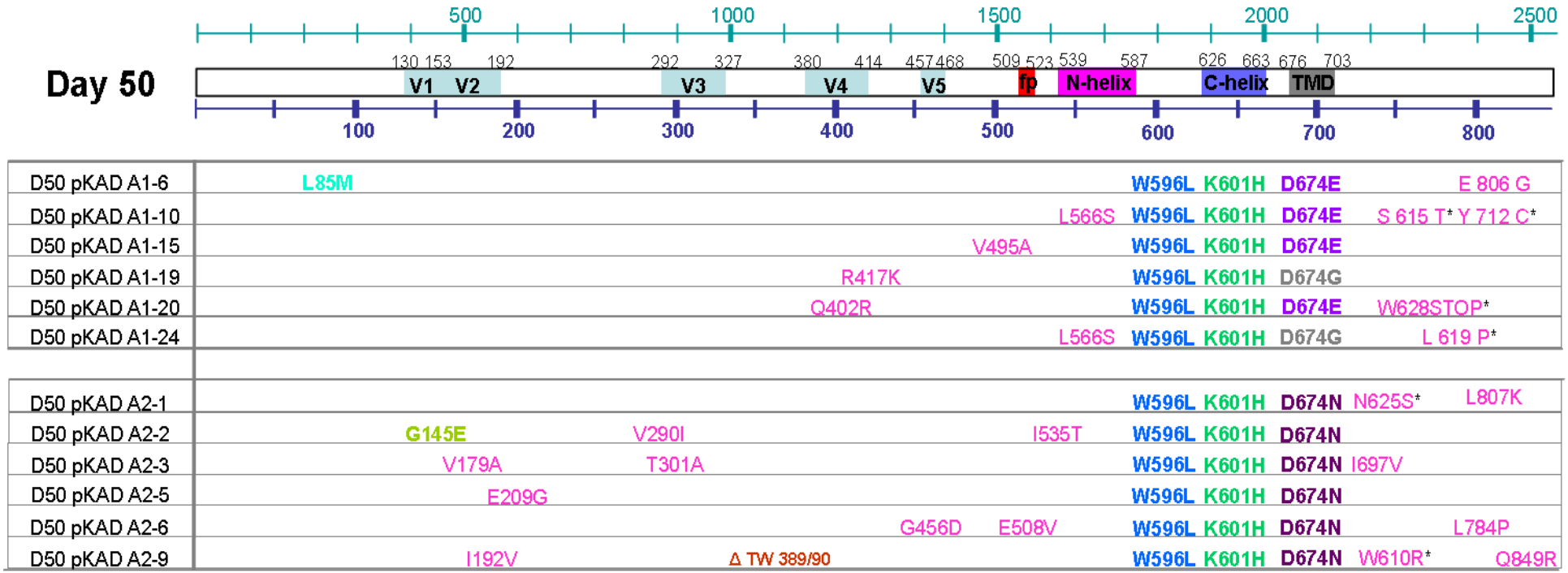


Table 5.5 Day 50 Genotypes of W596L.K601D HIV-1AD8 revertants

Sequence analysis of revertant clones obtained at day 50 in figure 5.4B. Note the presence of mutations in several clones, for example L85M, G145E, ΔTW389-390, K601H, D674E, D674G and D674N. Small asterisk (*) denotes a frameshift. Numbering is based on HXB2.

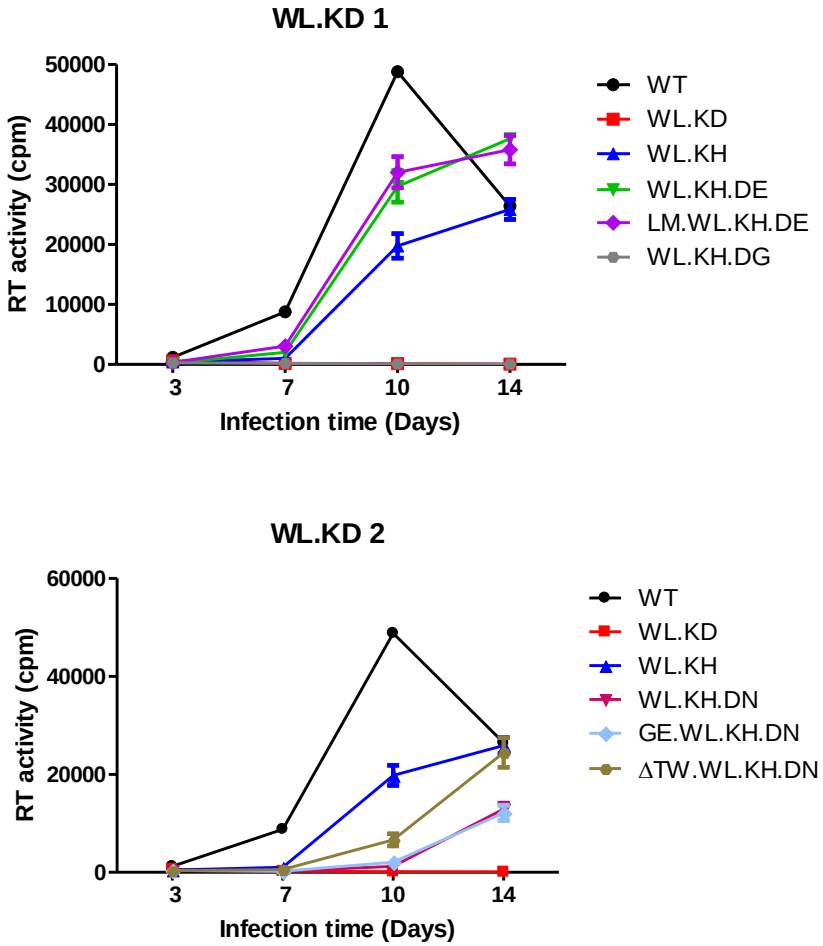


Figure 5.6 14-day replication kinetics of representative WLKD1 and WLKD2 genotypes in U87.CD4.CCR5 cells

Virus produced in 293T cells was normalised according to RT activity and used to infect U87.CD4.CCR5 cells. RT activity was measured in the culture fluid at days 3, 7, 10 and 14. (A) and (B) represent replication kinetics of WL.KD 1 and WL.KD 2 revertants respectively. The mean RT activity ± standard deviation of three samples is shown.

V1 mutation, which occurred once or twice at days 20 – 30, did not confer a replication advantage to W596L.K601H.D674N. Interestingly, the replication competence of reconstructed genotypes derived from the WL.KD2 culture were inferior to the W596L.K601H.D674E genotype obtained from WLKD1, even though replication-competent virus emerged in the WLKD2 culture first. This observation suggests that additional mechanisms of suppression of the WL.KD low-fusion phenotype may exist in the WLKD2 culture, which have not been identified at this stage.

The expression and processing of Env derived from the genotypes analysed above was examined by western blotting of 293T cells transfected with the corresponding pΔKAD8*env* expression vectors (Figure 5.7). The gp160 precursor was expressed at WT levels for all revertant clones and was successfully processed to gp120 and gp41 except for W596L.K601D, which is known to have a gp120/gp41 association defect. Notably, the D601H pseudoreversion alone was sufficient to restore the levels of cell-associated gp120 to that of the WT. While similar levels of processed gp41 were observed for wild type and the revertant clones, the relative abundance of the 4 known gp41 glycoforms differed. An equal abundance of the 4 glycoforms were observed for the wild type whereas di- and tetraglycosylated forms appeared to predominate for W596L.K601H and the other revertant genotypes. These data suggest that reversion is associated with a subtle change in the initial fold or conformation of the gp41 ectodomain.

5.3.5 Dissection of the potential functional linkages between the DSR, Thr-389-Trp-390 in V4, and Asp-674 in MPER

The functional linkages between Trp-596 and Lys-601 within the DSR, the Thr-389-Trp-390 dyad in V4 and Asp-674 in the MPER were further dissected in a single cycle entry assay. The results presented in figure 5.8 indicate that W596L.K601D led to an ~

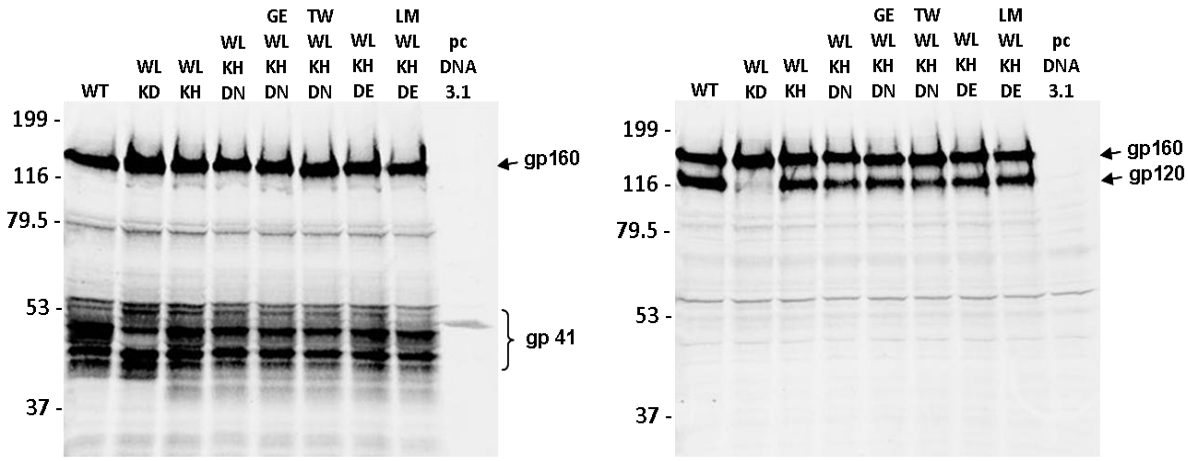


Figure 5.7 Western blotting of selected revertant clones

293T cells were transfected with 1 µg of pΔKAD8*env* plus 1ug pCAGT7. At 48 h posttransfection the cells were lysed and subjected to reducing SDS-PAGE in 10% polyacrylamide gels. The HIV-1 glycoproteins were visualized by western blotting with the mAb C8 (left) and DV012 (right).

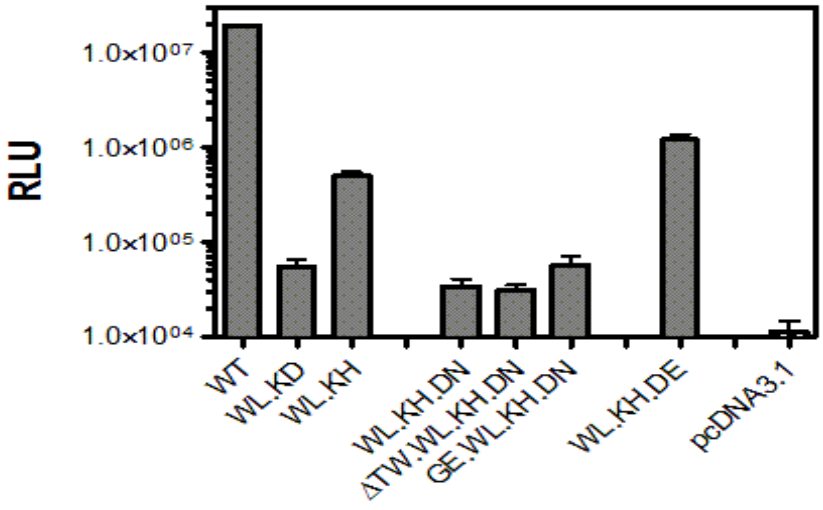


Figure 5.8 Single-cycle entry activities of representative revertant genotypes

Luciferase reporter viruses were prepared by co-transfecting 293T cells with 1 µg Env expression vector plus 1 µg pNL4.3LucR-E-. Viral entry was assessed by measuring luciferase activity in U87.CD4.CCR5 target cells at 48 h postinfection. Luciferase activity was normalised against reverse transcriptase activity present in each virus inoculum. The data presented are the means ± standard errors of 3 independent assays.

2.5log_{10}–fold reduction in single-cycle virus infectivity. The D601H pseudoreversion increased infectivity by ~ 10-fold relative to W596L.K601D while the addition of D674E to W596L.K601H increased infectivity by a further 2-fold. Even though W596L.K601H.D674E represented the revertant *env* genotype that conferred the highest level of single cycle infectivity, this was still 20-fold less than that of WT. The alternate D674N mutation was inhibitory on a W596L.K601H background, however the addition of the ΔTW389/90 V4 deletion to W596L.K601H.D674N did not improve entry competence. While the single cycle infectivity data largely reflected the observations with infectious clones in 14-day replication assays (Fig. 5.6), the functional advantages conferred by the D601H pseudoreversion and the additional suppressor mutations such as D674E, were more obvious in the latter assay. These differences could be attributed to the fact that the single-cycle virus infectivity assay relies on a single cell-free virus entry event, while virus transmission in 14-day replication assays is a sum of multiple rounds of cell-free virus infection and cell-to-cell viral spread across virological synapses, the latter of which is a more efficient process (Martin *et al.*, 2010) and has not been accounted for in the single-cycle assay (Figure 5.8).

5.3.6 Examination of the functional linkage between position 601 of the DSR and 674 of the MPER

The data indicates that the side chain at position 674 in the MPER can control virus infectivity when the DSR genotype is W596L.K601H with D674E enhancing infectivity in both single cycle and 14-day replication assays, whereas D674N and D674G are inhibitory. A survey of HIV-1 strains in the Los Alamos National Laboratory HIV sequence database indicates that Asp-674 is the dominant polymorphism in clade B strains (62 %), Asn-674 is the second most frequent polymorphism (9 %) while Glu-674 occurs rarely (1 %). Glycine is not observed at this position in clade B strains but occurs at low frequency in clades A and C. To further examine how Asp-674 mutations modulate Env

function, D674N, D674E and D674G were introduced to the wild type Env background in vector pΔKAD8*env* and their ability to mediate a single cycle of infection determined in U87.CD4.CCR5 cells. Aspartic acid-674 was also mutated to Gln, which in a sense represents a structural intermediate of Glu and Asn, in order to determine whether the functional effects arising from Asp-674 mutations could be titrated.

Western blotting was first used to confirm that the mutants were expressed and processed to gp120 and gp41 in a similar manner to the WT (Figure 5.9). Luciferase reporter viruses that had been pseudotyped with the Asp-674 mutants were then titrated onto U87.CD4.CCR5 cells in order to compare their infectivity. The results show clearly that the Asp-674 mutations in the wild type Env background did not significantly affect viral entry activity (Figure 5.10). By contrast, a gradation of effects on entry was observed with the different Asp mutations in the W596L.K601H background. As observed in figure 5.11, D674G completely blocked virus infectivity in a W596L.K601H context, D674E led to an increase in virus infectivity whereas D674N was inhibitory (Figure 5.11). The entry capacity of the W596L.K601H.D674Q combination was not significantly different to W596L.K601H ($p > 0.05$) indicating that the effects of this change were intermediate between those obtained with D674E and D674N. The expression and processing of the mutants was again found to be similar to that of the WT following western blot analysis (Figure 5.12).

5.3.7 Cell-to-cell fusion activity of Envs with revertant genotypes

The membrane fusion activities of selected revertant Env sequences were examined in a cell-to-cell fusion assay at limiting Env concentrations. This assay allows the intrinsic fusogenicity of the glycoproteins to be measured, as Env in this context is not subjected to conformational constraints that may be imposed by virus structure, for example, matrix-gp41 cytoplasmic tail interactions.

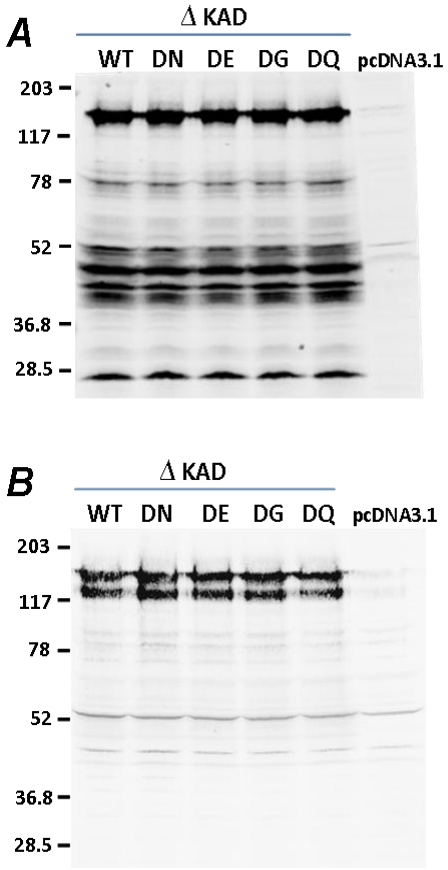


Figure 5.9 Western blotting of Asp-674 different mutants in the wild type Env background

293T cells were transfected with 1 µg of pΔKAD*env* plus 1ug pCAGT7. At 48 h posttransfection the cells were lysed and subjected to reducing SDS-PAGE in 10% polyacrylamide gels. The HIV-1 glycoproteins were visualized by western blotting with the mAb C8 **(A)** and DV012 **(B)**.

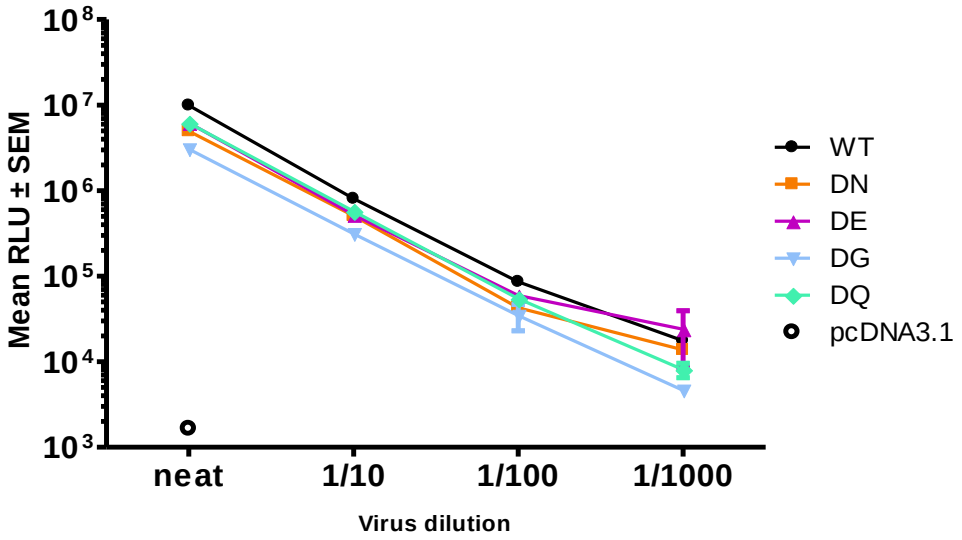


Figure 5.10 Single cycle entry mediated by Asp-674 mutants in the wild type Env background

Env-pseudotyped luciferase reporter viruses were prepared by co-transfecting 293T cells with 1 µg Env expression vector plus 1 µg pNL4.3LucR-E-. Viral entry was assessed after the addition of 4 different concentrations (neat, 1/10, 1/100, 1/1000) of the virus supernatant by measuring luciferase activity in U87.CD4.CCR5 target cells at 48 h post infection. Luciferase activity was normalised against reverse transcriptase activity present in each virus inoculum. The data presented are the means ± standard errors from 3 independent assays.

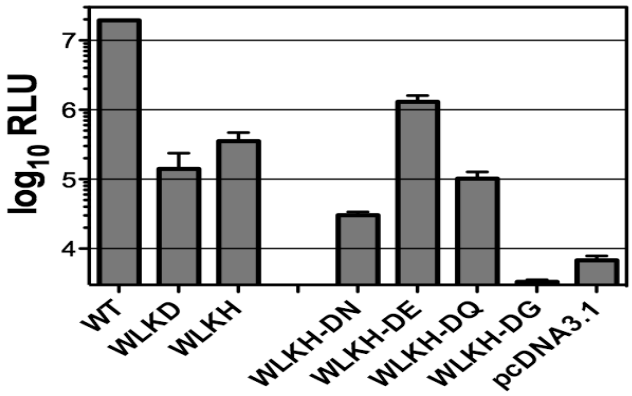


Figure 5.11 Single-cycle entry mediated by Asp-674 mutants in the W596L.K601H Env background

Luciferase reporter viruses were prepared by co-transfecting 293T cells with 1 µg Env expression vector plus 1 µg pNL4.3LucR-E-. Viral entry was assessed by measuring luciferase activity in U87.CD4.CCR5 target cells at 48 h postinfection. Luciferase activity was normalised against reverse transcriptase activity present in each virus inoculum. The data presented are the means ± standard errors of 3 independent assays.

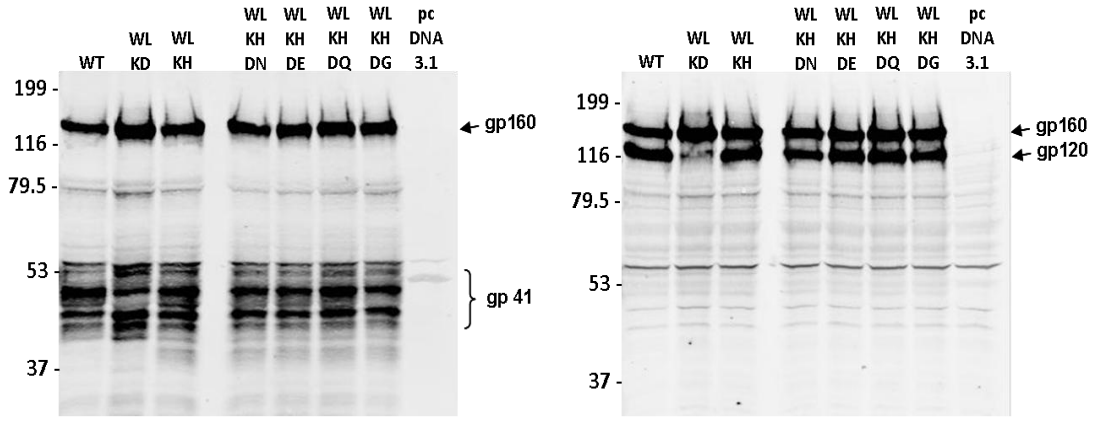


Figure 5.12 Western blotting of Asp-674 different mutants in the W596L.K601H Env background

293T cells were transfected with 1 µg of pΔKAD8*env* plus 1ug pCAGT7. At 48 h posttransfection the cells were lysed and subjected to reducing SDS-PAGE in 10% polyacrylamide gels. The HIV-1 glycoproteins were visualized by western blotting with the mAb C8 (left) and DV012 (right).

The assay was conducted at limiting Env concentrations (0.25 μg pΔKAD8*env*; Figure 5.2B) in order to enable detection of subtle changes in fusion function. The data indicate that W596L.K601D blocks cell-cell fusion activity and that the D601H pseudoreversion partially restores this function (Fig. 5.13). As was observed in single cycle infectivity and 14-day replication experiments, the D674N and D674G mutations were inhibitory in a W596L.K601H context. However, in contrast to the infectivity data, the addition of D674E did not enhance the fusogenicity of W596L.K601H. Furthermore, D674Q was inhibitory to membrane fusion in a W596L.K601H background, whereas this mutation did not significantly affect the ability of W596L.K601H to mediate a single cycle of infection. These data suggest that the functional interaction between His-601 and Glu-674 operates in the context of a virus particle.

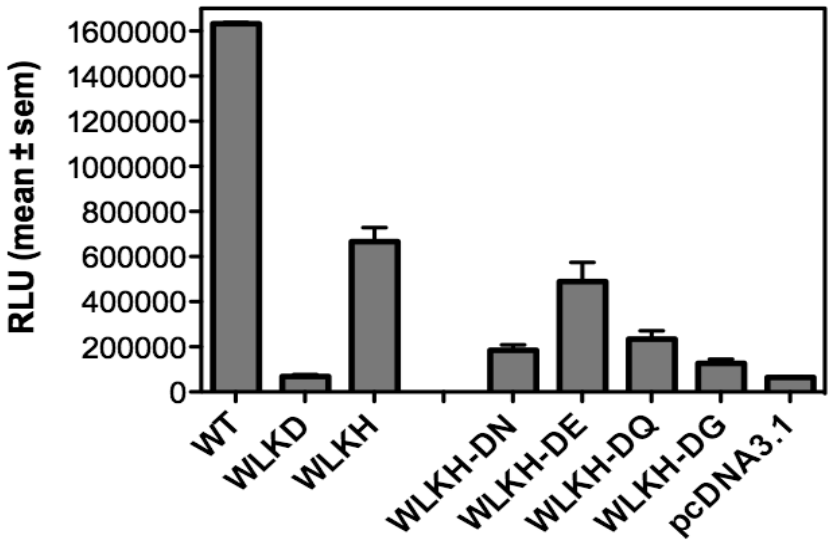


Figure 5.13 Cell-cell fusion activity of the different Asp-674 mutants in the W596L.K601H Env background using a luciferase reporter assay

293T effector cells were cotransfected with the T7 polymerase expression vector, pCAG-T7, together with pΔKAD8*env* vectors. BHK21 target cells were cotransfected with pT4luc and pc.CCR5. At 18 h posttransfection,the effectors and targets were mixed and co-cultured for 24 h prior to lysis and assay for luciferase activity. Mean relative light units (RLU) ± standard error is shown from 3 independent experiments.

5.4 Discussion

The serial passaging of the W596L.K601D mutant virus, which was defective in gp120-gp41 association and viral replication, enabled the selection of revertants with second and third site suppressor mutation(s) that overcame the attenuation. For clone WLKD 1, 47 days of culture were required to restore replication competence. The dominant suppressors included a D601H pseudoreversion in the DSR, plus D674E in the MPER. The D601H change restored gp120-gp41 association, while the combination of D601H with D674E resulted in a 20-fold enhancement of single cycle virus infectivity over the original W596L.K601D mutant and near-wild type replication kinetics. For the WL.KD 2 clone, only 30 days in culture were required to restore replication. The dominant suppressor combination in this case was D601H plus D674N (30 of 36 clones obtained from days 10-50) with a Thr-389-Trp-390 deletion (ΔTW389/90) in V4 also being present in 12 clones obtained between days 10 and 30. Despite W596L.K601H exhibiting better replication fitness than ΔTW389/90.W596L.K601H.D674N and W596L.K601H.D674N, DNA sequencing indicated that the least fit genotype (W596L.K601H.D674N) dominated days 40 and 50 of the culture, which follow the emergence of replication competent virus. Conservative substitutions at Asp-674 were found to modulate virus entry in the context of W596L.K601H but not wild type Env, indicating that the MPER is functionally linked to the association/activation synapse of gp120-gp41.

The W596L mutant has a very subtle gp120-shedding phenotype but retains wild type levels of cell-cell fusion and virus entry function (Poumbourios *et al.*, 2003) whereas a more pronounced shedding and fusion phenotype was observed for K601D. Combining the individual mutations in W596L.K601D resulted in a severe gp120-gp41 association site defect with a virtual absence of cell- and virion-associated gp120 observed. These data suggest that Trp-596 and Lys-601 contribute to gp120 association in a synergistic manner.

The W596L.K601D-mutated virus remained non-infectious even after 5 passages in U87.CD4.CCR5 cells, indicating severely attenuated replication capacity. The use of VSV G pseudotyping *in trans* to initiate the first round of infection by W596L.K601D virus successfully overcame this obstacle, leading to the generation of infectious variants that could undergo purifying selection with sequential passaging. Thus the replication defect observed for W596L.K601D is at the level of viral entry.

The phenotypic analysis of Env clones indicated that the D601H pseudoreversion is a key evolutionary step that restores gp120-gp41 association. The D601H pseudoreversion would reintroduce a bulky sidechain of a basic nature to the 601 site, perhaps restoring a key contact with gp120. Interestingly, Leu-596 was maintained over the 50 days in both cultures, suggesting that a smaller hydrophobic sidechain relative to Trp is tolerated at this position in the presence of His-601. Despite the restoration of gp120-gp41 association by D601H, membrane fusion, viral entry and virus replication functions were only partially restored, suggesting that suppression of the association defect came with a fitness cost in fusogenicity. This fitness cost was overcome to a degree by D674E in the MPER, which boosted single cycle entry and virus replication capacity. Importantly, D674E did not alter entry efficiency in a wild type Env context, suggesting that the effects seen with this mutation are specifically linked to the changes in the DSR. While His-601 enables efficient gp120-gp41 association when Leu-596 is present, the conformation of Env may nevertheless differ from the WT such that its ability to mediate membrane fusion is compromised and requires the other change for further optimisation. Notably, the D674E change in the MPER appeared to optimise gp41 conformation in a virion-specific manner, since improvements in cell-cell fusion function were not observed. In contrast to D674E, Gln, Asn and Gly substitutions at this site were inhibitory in a W596L.K601H context (but not a wild type Env context), with the severity of the

inhibition having an inverse correlation with sidechain length. These data provide further support for a correlation between the MPER and DSR in the fusion mechanism. It is possible that both the MPER and the DSR together play an important role in stabilizing the gp120-gp41 complex and in the fusion mechanism. It could be also possible that the MPER and DSR exert independent effects on stabilizing the gp120-gp41 complex, but that the MPER effect can only be observed when the DSR effect is sufficiently disrupted.

Nuclear magnetic resonance has revealed the structure of a MPER peptide in association with a dodecylphosphocholine micelle as a metastable L-shaped structure comprised of an N-terminal α-helix (residues 664-672) connected to a C-terminal helix (residues 675-683) via a flexible hinge (Figure 5.14A) (Sun *et al.*, 2008). The C-terminal helix is likely to interact with the TMD and cholesterol in the lipid phase via the Leu-679-Trp-Tye-Ile-Lys cholesterol recruitment motif (Chen *et al.*, 2009b; Epand *et al.*, 2006; Vincent *et al.*, 2002)**,** while the N-terminal helix represents a more flexible segment that might be in a metastable state and is noteworthy for having 3 Trp residues which are critical for membrane fusion (Salzwedel *et al.*, 1999). The flexible hinge comprises Phe-673, which is buried in the lipid phase, and Asn-674 (the HXB2R sequence was used for the NMR structure), which hydrogen bonds with Asn-677 via Oδ1 above the polar headgroup layer. The modelling of Asp-674 into this structure suggests that it is out of hydrogen-bonding range of Asn-677, while the introduction of Glu-674 predicts that Oε forms a hydrogen bond with a backbone amide within the polar headgroup layer (Figure 5.14B). These models suggest that D674E affects hinge stability and that transition of a flexible MPER structure to a more rigid one is linked to improved W596L.K601H Env functionality. It is notable that the presence of Glu and Asp at position 674 conferred better replication and/or single cycle entry properties to W596L.K601H in comparison to their neutral amino acid counterparts, Gln-674 and Asn-674, respectively. Negative charge

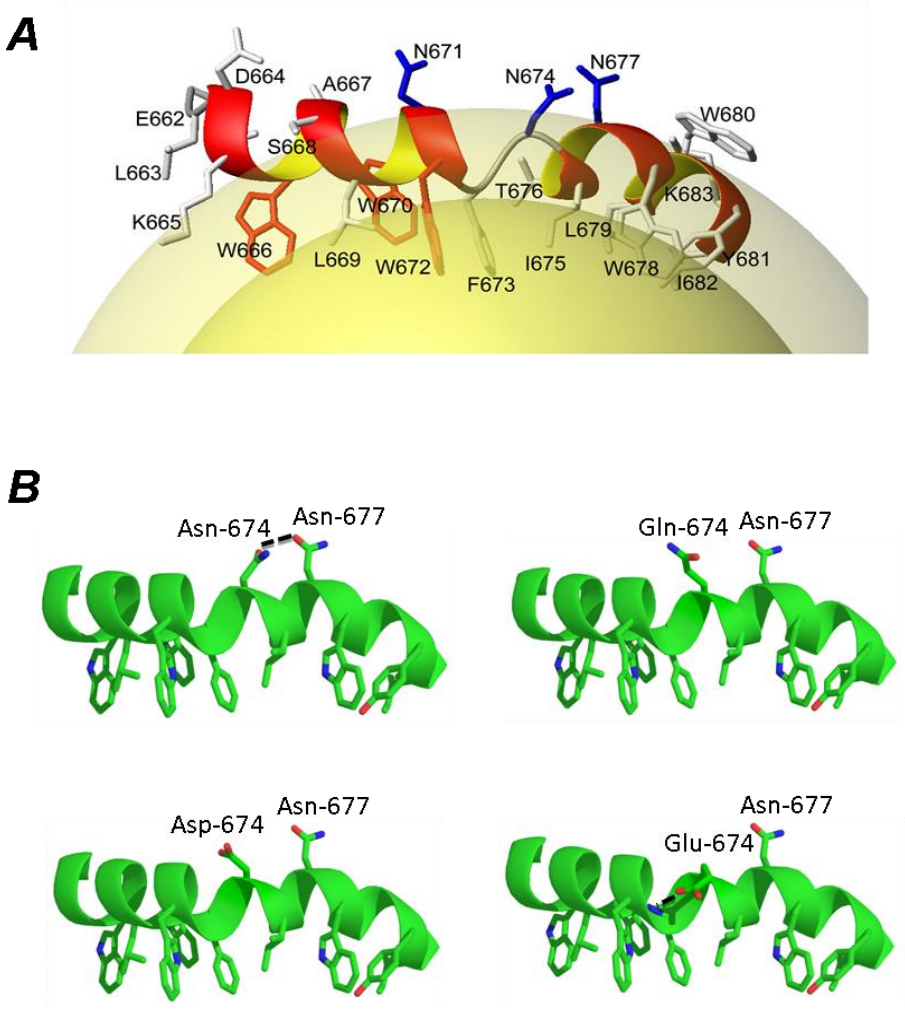


Figure 5.14. NMR structure of dodecylphosphocholine-associated MPER peptide

(A) The polar headgroup region is coloured light yellow; the lipid acyl-chain region is deep yellow. The three red Trp residues (666, 670 and 672) play an important role in virus-mediated fusion, while the solvent-exposed Asn residues (671, 674 and 677) are colored blue. Note that this peptide was derived from the HXB2R sequence, which contains Asn at position 674. Figure obtained from Sun *et al.* (2008). **(B)** Amino acid changes at position 674 modelled on the dodecylphosphocholine-associated MPER peptide (PDB entry 2PV6) using Pymol.

at position 674 therefore plays an important role in improving the function of W596L.K601H, thereby implying that Glu-674 and Asp-674 mediate a salt bridge that is crucial for the functional boost.

Interestingly, the D674E-mediated enhancement in W596L.K601H virus entry and replication was not reflected in cell-cell fusion function, suggesting that this component of the reversion mechanism operates in a virion context and not simply at the level of the intrinsic fusogenicity of Env. Evidence is accumulating to indicate that interactions between the gp41 cytoplasmic tail and the underlying viral MA layer modulate the structure and function of the Env ectodomain. For example, a gp120-gp41 association defect caused by a mutation in MA can be rescued by the decoupling of the 2 proteins via truncation of the cytoplasmic tail (Davis *et al.*, 2006). Furthermore, the gp41 cytoplasmic tail appears to be the key mediator of the 5- to 10-fold enhancement in fusogenicity of the viral gp120-gp41 complex induced by Gag maturation (Wyma *et al.*, 2004; Jiang and Aiken, 2007). Thus D674E may modulate the function of W596L.K601H Env in the virion specific conformation, which is dictated in part by MA-gp41 cytoplasmic tail interactions. It would be interesting to determine whether the decoupling of gp41 from MA by cytoplasmic tail truncation reverses the enhancing effects of D674E.

It is perplexing that the W596L.K601H.D674N combination, which conferred inferior single cycle entry and/or replication fitness with respect to W596L.K601H and ΔTW389/90.W596L.K601H.D674N, dominated time points that followed the emergence of replication competent virus in the WLKD 2 culture. Additional or alternative suppressor mutations may be necessary to boost replication. The emergence of the ΔTW389/90 deletion in V4 coincided with the initial spike of virus replication at day 30 and enhanced the replication competence of W596L.K601H.D674N, however this deletion was repaired by later time points. The G145E and D674G mutations appeared at low frequencies at

days 40 and 50 but they were not associated with replication advantages. It is possible that a low-frequency genotype associated with very high replication competence that has not been detected as yet is responsible for the WT levels of virus production observed from day 30 onward. However, this scenario appears unlikely because it is not clear why such a virus has not outgrown W596L.K601H.D674N by days 40 and 50. Deep sequencing methods such as 454, which enable the detection of low frequency genotypes, would help to address this possibility. Alternatively, the W596L.K601H.D674N combination may operate in conjunction with additional suppressors in non-*env* regions of the viral genome. Matrix is a likely candidate region given its structural and functional association with gp41 via the cytoplasmic tail. The sequencing of *gag* amplified from genomic DNA derived from the various time points may shed light on this possibility.

A recent study by Ruprecht *et al.* has shown that prolonged treatment of cell-free HIV-1 particles with MPER-specific neutralizing mAbs (2F5 and 4E10) can induce gp120 shedding (Ruprecht *et al.*, 2011). These workers put forward 2 explanations for this observation: 1) MPER-mAb binding affects the labile association between gp41 and gp120, leading to shedding; 2) conformations that expose the MPER and allow for the binding of 2F5 and 4E10 induce trimer instability. The results presented in this chapter indicate that the DSR, which forms part of the gp120-gp41 association site, is linked functionally to the MPER support the first hypothesis where structural changes induced in the MPER by 2F5 and 4E10 in turn induce a conformational change in the DSR, which leads to gp120 dissociation and the neutralization of infectivity. Structural biological studies of MPER-mAb interactions indicate that 2F5 and 4E10 binding does indeed alter the structure of the MPER by extracting paratope-interactive residues such as Trp-666 and Phe-673, respectively, from the hydrophobic phase (Song *et al.*, 2009; Sun *et al.*, 2008; Ofek *et al.*, 2004; Cardoso *et al.*, 2005).

Early mutational studies indicated that the DSR is associated with the C1 and C5 regions of gp120 (Binley *et al.*, 2003; Helseth *et al.*, 1991; Maerz *et al.*, 2001; Poumbourios *et al.*, 2003). The C1 and C5 regions are connected to the gp120 core via a 7-stranded β-sandwich, forming projections at the base of gp120 (Pancera *et al.*, 2010). The modelling of these projections in a trimeric context suggests that they will form a layer that encases the DSR, occluding it from antibody recognition (Bellamy-McIntyre *et al.*, 2010). The MPER is believed to occupy a spatially distinct location at the base of the gp41 trimer, partially embedded in the envelope and proximal to the TMD but available for antibody binding (Sattentau *et al.*, 1995; Finnegan *et al.*, 2002). Thus the DSR and MPER are likely to occupy spatially separate locations in the prefusion viral gp120-gp41 complex. The MPER-DSR functional linkage is therefore likely to operate via an allosteric mechanism that is mediated via other structural elements of the ectodomain. Recent work from the Overbaugh laboratory implicates the HR1 segment of gp41 as one such element. This idea is suggested by the finding that I675V (MPER) and T569A (HR1) polymorphism synergise in conferring a neutralization sensitive Env conformation (Blish *et al.*, 2008; Wang *et al.*, 2011). The results of the cysteine mutagenesis study described in Chapter 4 indicated that Thr-569 forms part of the trimer interface of prefusion gp41, since T569C enables disulfide bonding between protomers. Thus the transmission of structural signals between the MPER and DSR may involve the central coiled coil core domain of gp41. The polar segment represents another structural element that might be involved, since simultaneous mutations in the polar segment and the MPER was shown to induce gp120 shedding. This shedding defect ablates the responsiveness of the residual gp120-gp41 complex to CD4 ligation and the initiation of membrane fusion (Bellamy-McIntyre *et al.*, 2007). These data indicate that the CD4-induced conformational activation of gp41 involves cooperative interactions between multiple regions of the ectodomain.

Bibliography:

ABACIOGLU, Y. H., FOUTS, T. R., LAMAN, J. D., CLAASSEN, E., PINCUS, S. H., MOORE, J. P., ROBY, C. A., KAMIN-LEWIS, R. & LEWIS, G. K. 1994. Epitope mapping and topology of baculovirus-expressed HIV-1 gp160 determined with a panel of murine monoclonal antibodies. *AIDS Res Hum Retroviruses,* 10**,** 371-81.

ABRAHAMYAN, L. G., MARKOSYAN, R. M., MOORE, J. P., COHEN, F. S. & MELIKYAN, G. B. 2003. Human immunodeficiency virus type 1 Env with an intersubunit disulfide bond engages coreceptors but requires bond reduction after engagement to induce fusion. *J Virol,* 77**,** 5829-36.

ADAM, B., LINS, L., STROOBANT, V., THOMAS, A. & BRASSEUR, R. 2004. Distribution of hydrophobic residues is crucial for the fusogenic properties of the Ebola virus GP2 fusion peptide. *J Virol,* 78**,** 2131-6.

ALKHATIB, G. 2009. The biology of CCR5 and CXCR4. *Current Opinion in HIV and AIDS,* 4**,** 96-103.

ALKHATIB, G., COMBADIERE, C., BRODER, C. C., FENG, Y., KENNEDY, P. E., MURPHY, P. M. & BERGER, E. A. 1996. CC CKR5: a RANTES, MIP-1alpha, MIP-1beta receptor as a fusion cofactor for macrophage-tropic HIV-1. *Science,* 272**,** 1955-8.

BANNERT, N. & KURTH, R. 2004. Retroelements and the human genome: new perspectives on an old relation. *Proc Natl Acad Sci U S A,* 101 Suppl 2**,** 14572-9.

BARRE-SINOUSSI, F., CHERMANN, J. C., REY, F., NUGEYRE, M. T., CHAMARET, S., GRUEST, J., DAUGUET, C., AXLER-BLIN, C., VEZINET-BRUN, F., ROUZIOUX, C., ROZENBAUM, W. & MONTAGNIER, L. 1983. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science,* 220**,** 868-71.

BELLAMY-MCINTYRE, A. K., BAR, S., LUDLOW, L., DRUMMER, H. E. & POUMBOURIOS, P. 2010. Role for the disulfide-bonded region of human immunodeficiency virus type 1 gp41 in receptor-triggered activation of membrane fusion function. *Biochem Biophys Res Commun,* 394**,** 904-8.

BELLAMY-MCINTYRE, A. K., LAY, C. S., BAAR, S., MAERZ, A. L., TALBO, G. H., DRUMMER, H. E. & POUMBOURIOS, P. 2007. Functional links between the fusion peptide-proximal polar segment and membrane-proximal region of human immunodeficiency virus gp41 in distinct phases of membrane fusion. *J Biol Chem,* 282**,** 23104-16.

BERNSTEIN, H. B., TUCKER, S. P., KAR, S. R., MCPHERSON, S. A., MCPHERSON, D. T., DUBAY, J. W., LEBOWITZ, J., COMPANS, R. W. & HUNTER, E. 1995. Oligomerization of the hydrophobic heptad repeat of gp41. *J Virol,* 69**,** 2745-50.

BINLEY, J. M., CAYANAN, C. S., WILEY, C., SCHULKE, N., OLSON, W. C. & BURTON, D. R. 2003. Redox-triggered infection by disulfide-shackled human immunodeficiency virus type 1 pseudovirions. *J Virol,* 77**,** 5678-84.

BINLEY, J. M., SANDERS, R. W., CLAS, B., SCHUELKE, N., MASTER, A., GUO, Y., KAJUMO, F., ANSELMA, D. J., MADDON, P. J., OLSON, W. C. & MOORE, J. P. 2000. A recombinant human immunodeficiency virus type 1 envelope glycoprotein complex stabilized by an intermolecular disulfide bond between the gp120 and gp41 subunits is an antigenic mimic of the trimeric virion-associated structure. *J Virol,* 74**,** 627-43.

BITTNER, J. J. 1936. Some Possible Effects of Nursing on the Mammary Gland Tumor Incidence in Mice. *Science,* 84**,** 162.

BJORNDAL, A., DENG, H., JANSSON, M., FIORE, J. R., COLOGNESI, C., KARLSSON, A., ALBERT, J., SCARLATTI, G., LITTMAN, D. R. & FENYO, E. M. 1997. Coreceptor usage of primary human immunodeficiency virus type 1 isolates varies according to biological phenotype. *J Virol,* 71**,** 7478-87.

BLISH, C. A., NGUYEN, M. A. & OVERBAUGH, J. 2008. Enhancing exposure of HIV-1 neutralization epitopes through mutations in gp41. *PLoS Med,* 5**,** e9.

BLUMENTHAL, R., SARKAR, D. P., DURELL, S., HOWARD, D. E. & MORRIS, S. J. 1996. Dilation of the influenza hemagglutinin fusion pore revealed by the kinetics of individual cell-cell fusion events. *J Cell Biol,* 135**,** 63-71.

BONAPARTE, M. I., DIMITROV, A. S., BOSSART, K. N., CRAMERI, G., MUNGALL, B. A., BISHOP, K. A., CHOUDHRY, V., DIMITROV, D. S., WANG, L. F., EATON, B. T. & BRODER, C. C. 2005. Ephrin-B2 ligand is a functional receptor for Hendra virus and Nipah virus. *Proc Natl Acad Sci U S A,* 102**,** 10652-7.

BORGGREN, M., REPITS, J., KUYLENSTIERNA, C., STERJOVSKI, J., CHURCHILL, M. J., PURCELL, D. F., KARLSSON, A., ALBERT, J., GORRY, P. R. & JANSSON, M. 2008. Evolution of DC-SIGN use revealed by fitness studies of R5 HIV-1 variants emerging during AIDS progression. *Retrovirology,* 5**,** 28.

BORREGO-DIAZ, E., PEEPLES, M. E., MARKOSYAN, R. M., MELIKYAN, G. B. & COHEN, F. S. 2003. Completion of trimeric hairpin formation of influenza virus hemagglutinin promotes fusion pore opening and enlargement. *Virology,* 316**,** 234-44.

BULLOUGH, P. A., HUGHSON, F. M., SKEHEL, J. J. & WILEY, D. C. 1994. Structure of influenza haemagglutinin at the pH of membrane fusion. *Nature,* 371**,** 37-43.

BUSHMAN, F. D. 2003. Targeting survival: integration site selection by retroviruses and LTR-retrotransposons. *Cell,* 115**,** 135-8.

BUZON, V., NATRAJAN, G., SCHIBLI, D., CAMPELO, F., KOZLOV, M. M. & WEISSENHORN, W. 2010. Crystal structure of HIV-1 gp41 including both fusion peptide and membrane proximal external regions. *PLoS Pathog,* 6**,** e1000880.

CAFFREY, M., CAI, M., KAUFMAN, J., STAHL, S. J., WINGFIELD, P. T., COVELL, D. G., GRONENBORN, A. M. & CLORE, G. M. 1998. Three-dimensional solution structure of the 44 kDa ectodomain of SIV gp41. *Embo J,* 17**,** 4572-84.

CARDOSO, R. M., ZWICK, M. B., STANFIELD, R. L., KUNERT, R., BINLEY, J. M., KATINGER, H., BURTON, D. R. & WILSON, I. A. 2005. Broadly neutralizing anti-HIV antibody 4E10 recognizes a helical conformation of a highly conserved fusion-associated motif in gp41. *Immunity,* 22**,** 163-73.

CENTER, R. J., KEMP, B. E. & POUMBOURIOS, P. 1997. Human immunodeficiency virus type 1 and 2 envelope glycoproteins oligomerize through conserved sequences. *J Virol,* 71**,** 5706-11.

CENTER, R. J., KOBE, B., WILSON, K. A., TEH, T., HOWLETT, G. J., KEMP, B. E. & POUMBOURIOS, P. 1998. Crystallization of a trimeric human T cell leukemia virus type 1 gp21 ectodomain fragment as a chimera with maltose-binding protein. *Protein Sci,* 7**,** 1612-9.

CENTER, R. J., LEAPMAN, R. D., LEBOWITZ, J., ARTHUR, L. O., EARL, P. L. & MOSS, B. 2002. Oligomeric structure of the human immunodeficiency virus type 1 envelope protein on the virion surface. *J Virol,* 76**,** 7863-7.

CHAN, D. C., CHUTKOWSKI, C. T. & KIM, P. S. 1998. Evidence that a prominent cavity in the coiled coil of HIV type 1 gp41 is an attractive drug target. *Proc Natl Acad Sci U S A,* 95**,** 15613-7.

CHAN, D. C., FASS, D., BERGER, J. M. & KIM, P. S. 1997. Core structure of gp41 from the HIV envelope glycoprotein. *Cell,* 89**,** 263-73.

CHAN, D. C. & KIM, P. S. 1998. HIV entry and its inhibition. *Cell,* 93**,** 681-4.

CHANDRAN, K., SULLIVAN, N. J., FELBOR, U., WHELAN, S. P. & CUNNINGHAM, J. M. 2005. Endosomal proteolysis of the Ebola virus glycoprotein is necessary for infection. *Science,* 308**,** 1643-5.

CHANEL-VOS, C. & KIELIAN, M. 2004. A Conserved Histidine in the ij Loop of the Semliki Forest Virus E1 Protein Plays an Important Role in Membrane Fusion. *J. Virol.,* 78**,** 13543-13552.

CHANG, L. J., URLACHER, V., IWAKUMA, T., CUI, Y. & ZUCALI, J. 1999. Efficacy and safety analyses of a recombinant human immunodeficiency virus type 1 derived vector system. *Gene Ther,* 6**,** 715-28.

CHEN, B., VOGAN, E. M., GONG, H., SKEHEL, J. J., WILEY, D. C. & HARRISON, S. C. 2005. Structure of an unliganded simian immunodeficiency virus gp120 core. *Nature,* 433**,** 834-41.

CHEN, J., SKEHEL, J. J. & WILEY, D. C. 1999. N- and C-terminal residues combine in the fusion-pH influenza hemagglutinin HA(2) subunit to form an N cap that terminates the triple-stranded coiled coil. *Proc Natl Acad Sci U S A,* 96**,** 8967-72.

CHEN, L., KWON, Y. D., ZHOU, T., WU, X., O'DELL, S., CAVACINI, L., HESSELL, A. J., PANCERA, M., TANG, M., XU, L., YANG, Z. Y., ZHANG, M. Y., ARTHOS, J., BURTON, D. R., DIMITROV, D. S., NABEL, G. J., POSNER, M. R., SODROSKI, J., WYATT, R., MASCOLA, J. R. & KWONG, P. D. 2009a.

Structural basis of immune evasion at the site of CD4 attachment on HIV-1 gp120. *Science,* 326**,** 1123-7.

CHEN, S. S., YANG, P., KE, P. Y., LI, H. F., CHAN, W. E., CHANG, D. K., CHUANG, C. K., TSAI, Y. & HUANG, S. C. 2009b. Identification of the LWYIK motif located in the human immunodeficiency virus type 1 transmembrane gp41 protein as a distinct determinant for viral infection. *J Virol,* 83**,** 870-83.

CHENG, S. F., WU, C. W., KANTCHEV, E. A. & CHANG, D. K. 2004. Structure and membrane interaction of the internal fusion peptide of avian sarcoma leukosis virus. *Eur J Biochem,* 271**,** 4725-36.

CHO, M. W., LEE, M. K., CARNEY, M. C., BERSON, J. F., DOMS, R. W. & MARTIN, M. A. 1998. Identification of determinants on a dualtropic human immunodeficiency virus type 1 envelope glycoprotein that confer usage of CXCR4. *J Virol,* 72**,** 2509-15.

CHOE, H., FARZAN, M., KONKEL, M., MARTIN, K., SUN, Y., MARCON, L., CAYABYAB, M., BERMAN, M., DORF, M. E., GERARD, N., GERARD, C. & SODROSKI, J. 1998. The orphan seven-transmembrane receptor apj supports the entry of primary T-cell-line-tropic and dualtropic human immunodeficiency virus type 1. *J Virol,* 72**,** 6113-8.

CHOE, H., FARZAN, M., SUN, Y., SULLIVAN, N., ROLLINS, B., PONATH, P. D., WU, L., MACKAY, C. R., LAROSA, G., NEWMAN, W., GERARD, N., GERARD, C. & SODROSKI, J. 1996. The beta-chemokine receptors CCR3 and CCR5 facilitate infection by primary HIV-1 isolates. *Cell,* 85**,** 1135-48.

CHU, V. C., MCELROY, L. J., CHU, V., BAUMAN, B. E. & WHITTAKER, G. R. 2006. The avian coronavirus infectious bronchitis virus undergoes direct low-pH-dependent fusion activation during entry into host cells. *J Virol,* 80**,** 3180-8.

COETZER, M., NEDELLEC, R., SALKOWITZ, J., MCLAUGHLIN, S., LIU, Y., HEATH, L., MULLINS, J. I. & MOSIER, D. E. 2008. Evolution of CCR5 use before and during coreceptor switching. *J Virol,* 82**,** 11758-66.

COFFIN, J., HAASE, A., LEVY, J. A., MONTAGNIER, L., OROSZLAN, S., TEICH, N., TEMIN, H., TOYOSHIMA, K., VARMUS, H., VOGT, P. & ET AL. 1986. What to call the AIDS virus? *Nature,* 321**,** 10.

COFFIN, J. M. 1990. Molecular mechanisms of nucleic acid integration. *J Med Virol,* 31**,** 43-9.

COFFIN, J. M. 1992. *Structure and classification of retroviruses,* New York, Plenum Press.

COLMAN, P. M. & LAWRENCE, M. C. 2003. The structural biology of type I viral membrane fusion. *Nat Rev Mol Cell Biol,* 4**,** 309-19.

COMBADIERE, C., SALZWEDEL, K., SMITH, E. D., TIFFANY, H. L., BERGER, E. A. & MURPHY, P. M. 1998. Identification of CX3CR1. A chemotactic receptor for

the human CX3C chemokine fractalkine and a fusion coreceptor for HIV-1. *J Biol Chem,* 273**,** 23799-804.

CONNOR, R. I., SHERIDAN, K. E., CERADINI, D., CHOE, S. & LANDAU, N. R. 1997. Change in coreceptor use correlates with disease progression in HIV-1--infected individuals. *J Exp Med,* 185**,** 621-8.

CROOKS, E. T., MOORE, P. L., FRANTI, M., CAYANAN, C. S., ZHU, P., JIANG, P., DE VRIES, R. P., WILEY, C., ZHARKIKH, I., SCHULKE, N., ROUX, K. H., MONTEFIORI, D. C., BURTON, D. R. & BINLEY, J. M. 2007. A comparative immunogenicity study of HIV-1 virus-like particles bearing various forms of envelope proteins, particles bearing no envelope and soluble monomeric gp120. *Virology,* 366**,** 245-62.

CROOKS, E. T., TONG, T., OSAWA, K. & BINLEY, J. M. 2011. Enzyme digests eliminate nonfunctional Env from HIV-1 particle surfaces, leaving native Env trimers intact and viral infectivity unaffected. *J Virol,* 85**,** 5825-39.

DAMICO, R. L., CRANE, J. & BATES, P. 1998. Receptor-triggered membrane association of a model retroviral glycoprotein. *Proc Natl Acad Sci U S A,* 95**,** 2580-5.

DANIELI, T., PELLETIER, S. L., HENIS, Y. I. & WHITE, J. M. 1996. Membrane fusion mediated by the influenza virus hemagglutinin requires the concerted action of at least three hemagglutinin trimers. *J Cell Biol,* 133**,** 559-69.

DAVIS, M. R., JIANG, J., ZHOU, J., FREED, E. O. & AIKEN, C. 2006. A mutation in the human immunodeficiency virus type 1 Gag protein destabilizes the interaction of the envelope protein subunits gp120 and gp41. *J Virol,* 80**,** 2405-17.

DELBOY, M. G., PATTERSON, J. L., HOLLANDER, A. M. & NICOLA, A. V. 2006. Nectin-2-mediated entry of a syncytial strain of herpes simplex virus via pH-independent fusion with the plasma membrane of Chinese hamster ovary cells. *Virol J,* 3**,** 105.

DELOS, S. E., LA, B., GILMARTIN, A. & WHITE, J. M. 2010. Studies of the "chain reversal regions" of the avian sarcoma/leukosis virus (ASLV) and ebolavirus fusion proteins: analogous residues are important, and a His residue unique to EnvA affects the pH dependence of ASLV entry. *J Virol,* 84**,** 5687-94.

DENG, H., LIU, R., ELLMEIER, W., CHOE, S., UNUTMAZ, D., BURKHART, M., DI MARZIO, P., MARMON, S., SUTTON, R. E., HILL, C. M., DAVIS, C. B., PEIPER, S. C., SCHALL, T. J., LITTMAN, D. R. & LANDAU, N. R. 1996. Identification of a major co-receptor for primary isolates of HIV-1. *Nature,* 381**,** 661-6.

DETTENHOFER, M. & YU, X. F. 1999. Highly purified human immunodeficiency virus type 1 reveals a virtual absence of Vif in virions. *J Virol,* 73**,** 1460-7.

DEWANNIEUX, M., DUPRESSOIR, A., HARPER, F., PIERRON, G. & HEIDMANN, T. 2004. Identification of autonomous IAP LTR retrotransposons mobile in mammalian cells. *Nat Genet,* 36**,** 534-9.

DONALD, J. E., ZHANG, Y., FIORIN, G., CARNEVALE, V., SLOCHOWER, D. R., GAI, F., KLEIN, M. L. & DEGRADO, W. F. 2011. Transmembrane orientation and possible role of the fusogenic peptide from parainfluenza virus 5 (PIV5) in promoting fusion. *Proc Natl Acad Sci U S A,* 108**,** 3958-63.

DONEGAN, E., LEE, H., OPERSKALSKI, E. A., SHAW, G. M., KLEINMAN, S. H., BUSCH, M. P., STEVENS, C. E., SCHIFF, E. R., NOWICKI, M. J., HOLLINGSWORTH, C. G. & ET AL. 1994. Transfusion transmission of retroviruses: human T-lymphotropic virus types I and II compared with human immunodeficiency virus type 1. *Transfusion,* 34**,** 478-83.

DORANZ, B. J., RUCKER, J., YI, Y., SMYTH, R. J., SAMSON, M., PEIPER, S. C., PARMENTIER, M., COLLMAN, R. G. & DOMS, R. W. 1996. A dual-tropic primary HIV-1 isolate that uses fusin and the beta-chemokine receptors CKR-5, CKR-3, and CKR-2b as fusion cofactors. *Cell,* 85**,** 1149-58.

DORFMAN, T., BUKOVSKY, A., OHAGEN, A., HOGLUND, S. & GOTTLINGER, H. G. 1994. Functional domains of the capsid protein of human immunodeficiency virus type 1. *J Virol,* 68**,** 8180-7.

DORIG, R. E., MARCIL, A., CHOPRA, A. & RICHARDSON, C. D. 1993. The human CD46 molecule is a receptor for measles virus (Edmonston strain). *Cell,* 75**,** 295-305.

DRAGIC, T., LITWIN, V., ALLAWAY, G. P., MARTIN, S. R., HUANG, Y., NAGASHIMA, K. A., CAYANAN, C., MADDON, P. J., KOUP, R. A., MOORE, J. P. & PAXTON, W. A. 1996. HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5. *Nature,* 381**,** 667-73.

EARL, P. L., MOSS, B. & DOMS, R. W. 1991. Folding, interaction with GRP78-BiP, assembly, and transport of the human immunodeficiency virus type 1 envelope protein. *J Virol,* 65**,** 2047-55.

ECKERT, D. M. & KIM, P. S. 2001. Mechanisms of viral membrane fusion and its inhibition. *Annu Rev Biochem,* 70**,** 777-810.

EDINGER, A. L., HOFFMAN, T. L., SHARRON, M., LEE, B., YI, Y., CHOE, W., KOLSON, D. L., MITROVIC, B., ZHOU, Y., FAULDS, D., COLLMAN, R. G., HESSELGESSER, J., HORUK, R. & DOMS, R. W. 1998. An orphan seven-transmembrane domain receptor expressed widely in the brain functions as a coreceptor for human immunodeficiency virus type 1 and simian immunodeficiency virus. *J Virol,* 72**,** 7934-40.

ELLERMAN, V. & BANG, O. 1908. Experimentelle Leukamie bei Huhnern. *Centralbl. Bakteriol.,* 46**,** 595-609.

EMERMAN, M. & MALIM, M. H. 1998. HIV-1 regulatory/accessory genes: keys to unraveling viral and host cell biology. *Science,* 280**,** 1880-4.

EPAND, R. F., THOMAS, A., BRASSEUR, R., VISHWANATHAN, S. A., HUNTER, E. & EPAND, R. M. 2006. Juxtamembrane protein segments that contribute to recruitment of cholesterol into domains. *Biochemistry,* 45**,** 6105-14.

FAN, N., GAVALCHIN, J., PAUL, B., WELLS, K. H., LANE, M. J. & POIESZ, B. J. 1992. Infection of peripheral blood mononuclear cells and cell lines by cell-free human T-cell lymphoma/leukemia virus type I. *J Clin Microbiol,* 30**,** 905-10.

FARZAN, M., CHOE, H., DESJARDINS, E., SUN, Y., KUHN, J., CAO, J., ARCHAMBAULT, D., KOLCHINSKY, P., KOCH, M., WYATT, R. & SODROSKI, J. 1998. Stabilization of human immunodeficiency virus type 1 envelope glycoprotein trimers by disulfide bonds introduced into the gp41 glycoprotein ectodomain. *J Virol,* 72**,** 7620-5.

FASS, D., HARRISON, S. C. & KIM, P. S. 1996. Retrovirus envelope domain at 1.7 angstrom resolution. *Nat Struct Biol,* 3**,** 465-9.

FELBER, B. K., DERSE, D., ATHANASSOPOULOS, A., CAMPBELL, M. & PAVLAKIS, G. N. 1989. Cross-activation of the Rex proteins of HTLV-I and BLV and of the Rev protein of HIV-1 and nonreciprocal interactions with their RNA responsive elements. *New Biol,* 1**,** 318-28.

FELDMAN, S. A., HENDRY, R. M. & BEELER, J. A. 1999. Identification of a linear heparin binding domain for human respiratory syncytial virus attachment glycoprotein G. *J Virol,* 73**,** 6610-7.

FENG, Y., BRODER, C. C., KENNEDY, P. E. & BERGER, E. A. 1996. HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science,* 272**,** 872-7.

FINNEGAN, C. M., BERG, W., LEWIS, G. K. & DEVICO, A. L. 2002. Antigenic properties of the human immunodeficiency virus transmembrane glycoprotein during cell-cell fusion. *J Virol,* 76**,** 12123-34.

FINZI, A., XIANG, S. H., PACHECO, B., WANG, L., HAIGHT, J., KASSA, A., DANEK, B., PANCERA, M., KWONG, P. D. & SODROSKI, J. 2010. Topological layers in the HIV-1 gp120 inner domain regulate gp41 interaction and CD4-triggered conformational transitions. *Mol Cell,* 37**,** 656-67.

FOUCHIER, R., BROUWER, M., BROERSEN, S. & SCHUITEMAKER, H. 1995. Simple determination of human immunodeficiency virus type 1 syncytium-inducing V3 genotype by PCR. *J. Clin. Microbiol.,* 33**,** 906-911.

FREED, E. O. 2001. HIV-1 replication. *Somat Cell Mol Genet,* 26**,** 13-33.

FREED, E. O. & MARTIN, M. A. 1995. The role of human immunodeficiency virus type 1 envelope glycoproteins in virus infection. *J Biol Chem,* 270**,** 23883-6.

FREED, E. O., MYERS, D. J. & RISSER, R. 1990. Characterization of the fusion domain of the human immunodeficiency virus type 1 envelope glycoprotein gp41. *Proc Natl Acad Sci U S A,* 87**,** 4650-4.

FURUTA, R. A., WILD, C. T., WENG, Y. & WEISS, C. D. 1998. Capture of an early fusion-active conformation of HIV-1 gp41. *Nat Struct Biol,* 5**,** 276-9.

GALLO, R. C., SALAHUDDIN, S. Z., POPOVIC, M., SHEARER, G. M., KAPLAN, M., HAYNES, B. F., PALKER, T. J., REDFIELD, R., OLESKE, J., SAFAI, B. & ET AL. 1984. Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *Science,* 224**,** 500-3.

GARCIA, J. A., WU, F. K., MITSUYASU, R. & GAYNOR, R. B. 1987. Interactions of cellular proteins involved in the transcriptional regulation of the human immunodeficiency virus. *Embo J,* 6**,** 3761-70.

GARTNER, S., MARKOVITS, P., MARKOVITZ, D. M., BETTS, R. F. & POPOVIC, M. 1986. Virus isolation from and identification of HTLV-III/LAV-producing cells in brain tissue from a patient with AIDS. *Jama,* 256**,** 2365-71.

GHEZ, D., LEPELLETIER, Y., LAMBERT, S., FOURNEAU, J. M., BLOT, V., JANVIER, S., ARNULF, B., VAN ENDERT, P. M., HEVEKER, N., PIQUE, C. & HERMINE, O. 2006. Neuropilin-1 is involved in human T-cell lymphotropic virus type 1 entry. *J Virol,* 80**,** 6844-54.

GORRY, P. R., BRISTOL, G., ZACK, J. A., RITOLA, K., SWANSTROM, R., BIRCH, C. J., BELL, J. E., BANNERT, N., CRAWFORD, K., WANG, H., SCHOLS, D., DE CLERCQ, E., KUNSTMAN, K., WOLINSKY, S. M. & GABUZDA, D. 2001. Macrophage tropism of human immunodeficiency virus type 1 isolates from brain and lymphoid tissues predicts neurotropism independent of coreceptor specificity. *J Virol,* 75**,** 10073-89.

GRIVEL, J. C. & MARGOLIS, L. B. 1999. CCR5- and CXCR4-tropic HIV-1 are equally cytopathic for their T-cell targets in human lymphoid tissue. *Nat Med,* 5**,** 344-6.

GROSS, L. 1951. "Spontaneous" leukemia developing in C3H mice following inoculation in infancy, with AK-leukemic extracts, or AK-embrvos. *Proc Soc Exp Biol Med,* 76**,** 27-32.

HATCH, W. C., TANAKA, K. E., CALVELLI, T., RASHBAUM, W. K., KRESS, Y. & LYMAN, W. D. 1992. Persistent productive HIV-1 infection of a CD4- human fetal thymocyte line. *J Immunol,* 148**,** 3055-61.

HELSETH, E., OLSHEVSKY, U., FURMAN, C. & SODROSKI, J. 1991. Human immunodeficiency virus type 1 gp120 envelope glycoprotein regions important for association with the gp41 transmembrane glycoprotein. *J Virol,* 65**,** 2119-23.

HERNANDEZ, L. D. & WHITE, J. M. 1998. Mutational analysis of the candidate internal fusion peptide of the avian leukosis and sarcoma virus subgroup A envelope glycoprotein. *J Virol,* 72**,** 3259-67.

HERRERA, C., KLASSE, P. J., MICHAEL, E., KAKE, S., BARNES, K., KIBLER, C. W., CAMPBELL-GARDENER, L., SI, Z., SODROSKI, J., MOORE, J. P. & BEDDOWS, S. 2005. The impact of envelope glycoprotein cleavage on the antigenicity, infectivity, and neutralization sensitivity of Env-pseudotyped human immunodeficiency virus type 1 particles. *Virology,* 338**,** 154-72.

HOFFMAN, N. G., SEILLIER-MOISEIWITSCH, F., AHN, J., WALKER, J. M. & SWANSTROM, R. 2002. Variability in the human immunodeficiency virus type 1

gp120 Env protein linked to phenotype-associated changes in the V3 loop. *J Virol,* 76**,** 3852-64.

HOOD, C. L., ABRAHAM, J., BOYINGTON, J. C., LEUNG, K., KWONG, P. D. & NABEL, G. J. 2009. Biochemical and structural characterization of cathepsin L-processed Ebola virus glycoprotein: implications for viral entry and immunogenicity. *J Virol,* 84**,** 2972-82.

HOUT, D. R., MULCAHY, E. R., PACYNIAK, E., GOMEZ, L. M., GOMEZ, M. L. & STEPHENS, E. B. 2004. Vpu: a multifunctional protein that enhances the pathogenesis of human immunodeficiency virus type 1. *Curr HIV Res,* 2**,** 255-70.

HU, D. J., DONDERO, T. J., RAYFIELD, M. A., GEORGE, J. R., SCHOCHETMAN, G., JAFFE, H. W., LUO, C. C., KALISH, M. L., WENIGER, B. G., PAU, C. P., SCHABLE, C. A. & CURRAN, J. W. 1996. The emerging genetic diversity of HIV. The importance of global surveillance for diagnostics, research, and prevention. *Jama,* 275**,** 210-6.

HUANG, C. C., LAM, S. N., ACHARYA, P., TANG, M., XIANG, S. H., HUSSAN, S. S., STANFIELD, R. L., ROBINSON, J., SODROSKI, J., WILSON, I. A., WYATT, R., BEWLEY, C. A. & KWONG, P. D. 2007. Structures of the CCR5 N terminus and of a tyrosine-sulfated antibody with HIV-1 gp120 and CD4. *Science,* 317**,** 1930-4.

HUANG, C. C., TANG, M., ZHANG, M. Y., MAJEED, S., MONTABANA, E., STANFIELD, R. L., DIMITROV, D. S., KORBER, B., SODROSKI, J., WILSON, I. A., WYATT, R. & KWONG, P. D. 2005. Structure of a V3-containing HIV-1 gp120 core. *Science,* 310**,** 1025-8.

HUGHSON, F. M. 1997. Enveloped viruses: a common mode of membrane fusion? *Curr Biol,* 7**,** R565-9.

IGAKURA, T., STINCHCOMBE, J. C., GOON, P. K., TAYLOR, G. P., WEBER, J. N., GRIFFITHS, G. M., TANAKA, Y., OSAME, M. & BANGHAM, C. R. 2003. Spread of HTLV-I between lymphocytes by virus-induced polarization of the cytoskeleton. *Science,* 299**,** 1713-6.

JANSSON, M., BACKSTROM, E., BJORNDAL, A., HOLMBERG, V., ROSSI, P., FENYO, E. M., POPOVIC, M., ALBERT, J. & WIGZELL, H. 1999. Coreceptor usage and RANTES sensitivity of non-syncytium-inducing HIV-1 isolates obtained from patients with AIDS. *J Hum Virol,* 2**,** 325-38.

JIANG, J. & AIKEN, C. 2007. Maturation-dependent human immunodeficiency virus type 1 particle fusion requires a carboxyl-terminal region of the gp41 cytoplasmic tail. *J Virol,* 81**,** 9999-10008.

JIANG, S., ZHAO, Q. & DEBNATH, A. K. 2002. Peptide and non-peptide HIV fusion inhibitors. *Curr Pharm Des,* 8**,** 563-80.

JOLLY, C. & SATTENTAU, Q. J. 2004. Retroviral spread by induction of virological synapses. *Traffic,* 5**,** 643-50.

KAMPMANN, T., MUELLER, D. S., MARK, A. E., YOUNG, P. R. & KOBE, B. 2006. The Role of Histidine Residues in Low-pH-Mediated Viral Membrane Fusion. *Structure,* 14**,** 1481-1487.

KARLSSON HEDESTAM, G. B., FOUCHIER, R. A. M., PHOGAT, S., BURTON, D. R., SODROSKI, J. & WYATT, R. T. 2008. The challenges of eliciting neutralizing antibodies to HIV-1 and to influenza virus. *Nat Rev Micro,* 6**,** 143-155.

KILGORE, N. R., SALZWEDEL, K., REDDICK, M., ALLAWAY, G. P. & WILD, C. T. 2003. Direct evidence that C-peptide inhibitors of human immunodeficiency virus type 1 entry bind to the gp41 N-helical domain in receptor-activated viral envelope. *J Virol,* 77**,** 7669-72.

KOBE, B., CENTER, R. J., KEMP, B. E. & POUMBOURIOS, P. 1999. Crystal structure of human T cell leukemia virus type 1 gp21 ectodomain crystallized as a maltose-binding protein chimera reveals structural evolution of retroviral transmembrane proteins. *Proc Natl Acad Sci U S A,* 96**,** 4319-24.

KONDRATOWICZ, A. S., LENNEMANN, N. J., SINN, P. L., DAVEY, R. A., HUNT, C. L., MOLLER-TANK, S., MEYERHOLZ, D. K., RENNERT, P., MULLINS, R. F., BRINDLEY, M., SANDERSFELD, L. M., QUINN, K., WELLER, M., MCCRAY, P. B., JR., CHIORINI, J. & MAURY, W. 2011. From the Cover: T-cell immunoglobulin and mucin domain 1 (TIM-1) is a receptor for Zaire Ebolavirus and Lake Victoria Marburgvirus. *Proc Natl Acad Sci U S A,* 108**,** 8426-31.

KONG, L., HUANG, C. C., COALES, S. J., MOLNAR, K. S., SKINNER, J., HAMURO, Y. & KWONG, P. D. 2010. Local conformational stability of HIV-1 gp120 in unliganded and CD4-bound states as defined by amide hydrogen/deuterium exchange. *J Virol,* 84**,** 10311-21.

KOSHIBA, T. & CHAN, D. C. 2003. The prefusogenic intermediate of HIV-1 gp41 contains exposed C-peptide regions. *J Biol Chem,* 278**,** 7573-9.

KOZLOV, M. M. & CHERNOMORDIK, L. V. 2002. The protein coat in membrane fusion: lessons from fission. *Traffic,* 3**,** 256-67.

KUHN, J. H., RADOSHITZKY, S. R., GUTH, A. C., WARFIELD, K. L., LI, W., VINCENT, M. J., TOWNER, J. S., NICHOL, S. T., BAVARI, S., CHOE, H., AMAN, M. J. & FARZAN, M. 2006. Conserved receptor-binding domains of Lake Victoria marburgvirus and Zaire ebolavirus bind a common receptor. *J Biol Chem,* 281**,** 15951-8.

KUIKEN, C. L., FOLEY, B., HAHN, B., KORBER, B., MARX, P., MCCUTCHAN, F., MELLORS, J. W., MULLINS, J. I., SODROSKI, J. & WOLINSKY, S. 2000. Human retrovirus and AIDS 2000: A compilation and analysis of nucleic acid and amino acid sequences. *2000, Los Alamos National Laboratory*.

KWONG, P. D., DOYLE, M. L., CASPER, D. J., CICALA, C., LEAVITT, S. A., MAJEED, S., STEENBEKE, T. D., VENTURI, M., CHAIKEN, I., FUNG, M., KATINGER, H., PARREN, P. W., ROBINSON, J., VAN RYK, D., WANG, L., BURTON, D. R., FREIRE, E., WYATT, R., SODROSKI, J., HENDRICKSON, W.

A. & ARTHOS, J. 2002. HIV-1 evades antibody-mediated neutralization through conformational masking of receptor-binding sites. *Nature,* 420**,** 678-82.

KWONG, P. D., WYATT, R., MAJEED, S., ROBINSON, J., SWEET, R. W., SODROSKI, J. & HENDRICKSON, W. A. 2000. Structures of HIV-1 gp120 envelope glycoproteins from laboratory-adapted and primary isolates. *Structure,* 8**,** 1329-39.

KWONG, P. D., WYATT, R., ROBINSON, J., SWEET, R. W., SODROSKI, J. & HENDRICKSON, W. A. 1998. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature,* 393**,** 648-59.

LAMB, R. A. & JARDETZKY, T. S. 2007. Structural basis of viral invasion: lessons from paramyxovirus F. *Current Opinion in Structural Biology,* 17**,** 427-436.

LANDER, E. S., LINTON, L. M., BIRREN, B., NUSBAUM, C., ZODY, M. C., BALDWIN, J., DEVON, K., DEWAR, K., DOYLE, M., FITZHUGH, W., FUNKE, R., GAGE, D., HARRIS, K., HEAFORD, A., HOWLAND, J., KANN, L., LEHOCZKY, J., LEVINE, R., MCEWAN, P., MCKERNAN, K., MELDRIM, J., MESIROV, J. P., MIRANDA, C., MORRIS, W., NAYLOR, J., RAYMOND, C., ROSETTI, M., SANTOS, R., SHERIDAN, A., SOUGNEZ, C., STANGE-THOMANN, N., STOJANOVIC, N., SUBRAMANIAN, A., WYMAN, D., ROGERS, J., SULSTON, J., AINSCOUGH, R., BECK, S., BENTLEY, D., BURTON, J., CLEE, C., CARTER, N., COULSON, A., DEADMAN, R., DELOUKAS, P., DUNHAM, A., DUNHAM, I., DURBIN, R., FRENCH, L., GRAFHAM, D., GREGORY, S., HUBBARD, T., HUMPHRAY, S., HUNT, A., JONES, M., LLOYD, C., MCMURRAY, A., MATTHEWS, L., MERCER, S., MILNE, S., MULLIKIN, J. C., MUNGALL, A., PLUMB, R., ROSS, M., SHOWNKEEN, R., SIMS, S., WATERSTON, R. H., WILSON, R. K., HILLIER, L. W., MCPHERSON, J. D., MARRA, M. A., MARDIS, E. R., FULTON, L. A., CHINWALLA, A. T., PEPIN, K. H., GISH, W. R., CHISSOE, S. L., WENDL, M. C., DELEHAUNTY, K. D., MINER, T. L., DELEHAUNTY, A., KRAMER, J. B., COOK, L. L., FULTON, R. S., JOHNSON, D. L., MINX, P. J., CLIFTON, S. W., HAWKINS, T., BRANSCOMB, E., PREDKI, P., RICHARDSON, P., WENNING, S., SLEZAK, T., DOGGETT, N., CHENG, J. F., OLSEN, A., LUCAS, S., ELKIN, C., UBERBACHER, E., FRAZIER, M., et al. 2001. Initial sequencing and analysis of the human genome. *Nature,* 409**,** 860-921.

LAY, C. S., WILSON, K. A., KOBE, B., KEMP, B. E., DRUMMER, H. E. & POUMBOURIOS, P. 2004. Expression and biochemical analysis of the entire HIV-2 gp41 ectodomain: determinants of stability map to N- and C-terminal sequences outside the 6-helix bundle core. *FEBS Lett,* 567**,** 183-8.

LE ROUZIC, E. & BENICHOU, S. 2005. The Vpr protein from HIV-1: distinct roles along the viral life cycle. *Retrovirology,* 2**,** 11.

LEE, J. E., FUSCO, M. L., HESSELL, A. J., OSWALD, W. B., BURTON, D. R. & SAPHIRE, E. O. 2008. Structure of the Ebola virus glycoprotein bound to an antibody from a human survivor. *Nature,* 454**,** 177-82.

LEE, T. H., COLIGAN, J. E., SODROSKI, J. G., HASELTINE, W. A., SALAHUDDIN, S. Z., WONG-STAAL, F., GALLO, R. C. & ESSEX, M. 1984. Antigens encoded by the 3'-terminal region of human T-cell leukemia virus: evidence for a functional gene. *Science,* 226**,** 57-61.

LEVESQUE, K., FINZI, A., BINETTE, J. & COHEN, E. A. 2004. Role of CD4 receptor down-regulation during HIV-1 infection. *Curr HIV Res,* 2**,** 51-9.

LEVY, J. A., HOFFMAN, A. D., KRAMER, S. M., LANDIS, J. A., SHIMABUKURO, J. M. & OSHIRO, L. S. 1984. Isolation of lymphocytopathic retroviruses from San Francisco patients with AIDS. *Science,* 225**,** 840-2.

LI, K., ZHANG, S., KRONQVIST, M., WALLIN, M., EKSTROM, M., DERSE, D. & GAROFF, H. 2008. Intersubunit disulfide isomerization controls membrane fusion of human T-cell leukemia virus Env. *J Virol,* 82**,** 7135-43.

LI, S., JUAREZ, J., ALALI, M., DWYER, D., COLLMAN, R., CUNNINGHAM, A. & NAIF, H. M. 1999. Persistent CCR5 utilization and enhanced macrophage tropism by primary blood human immunodeficiency virus type 1 isolates from advanced stages of disease and comparison to tissue-derived isolates. *J Virol,* 73**,** 9741-55.

LI, W., ZHANG, P., FELLERS, J. P., FRIEBE, B. & GILL, B. S. 2004. Sequence composition, organization, and evolution of the core Triticeae genome. *Plant J,* 40**,** 500-11.

LIU, J., BARTESAGHI, A., BORGNIA, M. J., SAPIRO, G. & SUBRAMANIAM, S. 2008. Molecular architecture of native HIV-1 gp120 trimers. *Nature,* 455**,** 109-13.

LIU, S., LU, H., NIU, J., XU, Y., WU, S. & JIANG, S. 2005. Different from the HIV fusion inhibitor C34, the anti-HIV drug Fuzeon (T-20) inhibits HIV-1 entry by targeting multiple sites in gp41 and gp120. *J Biol Chem,* 280**,** 11259-73.

LLORENS, C., FUTAMI, R., COVELLI, L., DOMINGUEZ-ESCRIBA, L., VIU, J. M., TAMARIT, D., AGUILAR-RODRIGUEZ, J., VICENTE-RIPOLLES, M., FUSTER, G., BERNET, G. P., MAUMUS, F., MUNOZ-POMER, A., SEMPERE, J. M., LATORRE, A. & MOYA, A. 2010. The Gypsy Database (GyDB) of mobile genetic elements: release 2.0. *Nucleic Acids Res,* 39**,** D70-4.

LOW, A. J., SWENSON, L. C. & HARRIGAN, P. R. 2008. HIV coreceptor phenotyping in the clinical setting. *AIDS Rev,* 10**,** 143-51.

MAERZ, A. L., CENTER, R. J., KEMP, B. E., KOBE, B. & POUMBOURIOS, P. 2000. Functional implications of the human T-lymphotropic virus type 1 transmembrane glycoprotein helical hairpin structure. *J Virol,* 74**,** 6614-21.

MAERZ, A. L., DRUMMER, H. E., WILSON, K. A. & POUMBOURIOS, P. 2001. Functional analysis of the disulfide-bonded loop/chain reversal region of human immunodeficiency virus type 1 gp41 reveals a critical role in gp120-gp41 association. *J Virol,* 75**,** 6635-44.

MALASHKEVICH, V. N., CHAN, D. C., CHUTKOWSKI, C. T. & KIM, P. S. 1998. Crystal structure of the simian immunodeficiency virus (SIV) gp41 core: conserved

helical interactions underlie the broad inhibitory activity of gp41 peptides. *Proc Natl Acad Sci U S A,* 95**,** 9134-9.

MALASHKEVICH, V. N., SCHNEIDER, B. J., MCNALLY, M. L., MILHOLLEN, M. A., PANG, J. X. & KIM, P. S. 1999. Core structure of the envelope glycoprotein GP2 from Ebola virus at 1.9-A resolution. *Proc Natl Acad Sci U S A,* 96**,** 2662-7.

MANEL, N., KIM, F. J., KINET, S., TAYLOR, N., SITBON, M. & BATTINI, J. L. 2003. The ubiquitous glucose transporter GLUT-1 is a receptor for HTLV. *Cell,* 115**,** 449-59.

MANSKY, L. M. & TEMIN, H. M. 1995. Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. *J Virol,* 69**,** 5087-94.

MARTIN, N., WELSCH, S., JOLLY, C., BRIGGS, J. A., VAUX, D. & SATTENTAU, Q. J. 2010. Virological synapse-mediated spread of human immunodeficiency virus type 1 between T cells is sensitive to entry inhibition. *J Virol,* 84**,** 3516-27.

MEDSTRAND, P., VAN DE LAGEMAAT, L. N. & MAGER, D. L. 2002. Retroelement distributions in the human genome: variations associated with age and proximity to genes. *Genome Res,* 12**,** 1483-95.

MELIKYAN, G. B., MARKOSYAN, R. M., HEMMATI, H., DELMEDICO, M. K., LAMBERT, D. M. & COHEN, F. S. 2000. Evidence that the transition of HIV-1 gp41 into a six-helix bundle, not the bundle configuration, induces membrane fusion. *J Cell Biol,* 151**,** 413-23.

MILICH, L., MARGOLIN, B. H. & SWANSTROM, R. 1997. Patterns of amino acid variability in NSI-like and SI-like V3 sequences and a linked change in the CD4-binding domain of the HIV-1 Env protein. *Virology,* 239**,** 108-18.

MOORE, P. L., CROOKS, E. T., PORTER, L., ZHU, P., CAYANAN, C. S., GRISE, H., CORCORAN, P., ZWICK, M. B., FRANTI, M., MORRIS, L., ROUX, K. H., BURTON, D. R. & BINLEY, J. M. 2006. Nature of nonfunctional envelope proteins on the surface of human immunodeficiency virus type 1. *J Virol,* 80**,** 2515-28.

MOSS, B., ELROY-STEIN, O., MIZUKAMI, T., ALEXANDER, W. A. & FUERST, T. R. 1990. Product review. New mammalian expression vectors. *Nature,* 348**,** 91-2.

NARAYAN, S., BARNARD, R. J. & YOUNG, J. A. 2003. Two retroviral entry pathways distinguished by lipid raft association of the viral receptor and differences in viral infectivity. *J Virol,* 77**,** 1977-83.

NEGRETE, O. A., LEVRONEY, E. L., AGUILAR, H. C., BERTOLOTTI-CIARLET, A., NAZARIAN, R., TAJYAR, S. & LEE, B. 2005. EphrinB2 is the entry receptor for Nipah virus, an emergent deadly paramyxovirus. *Nature,* 436**,** 401-5.

OFEK, G., TANG, M., SAMBOR, A., KATINGER, H., MASCOLA, J. R., WYATT, R. & KWONG, P. D. 2004. Structure and mechanistic analysis of the anti-human

immunodeficiency virus type 1 antibody 2F5 in complex with its gp41 epitope. *J Virol,* 78**,** 10724-37.

OSAME, M., USUKU, K., IZUMO, S., IJICHI, N., AMITANI, H., IGATA, A., MATSUMOTO, M. & TARA, M. 1986. HTLV-I associated myelopathy, a new clinical entity. *Lancet,* 1**,** 1031-2.

PANCERA, M., MAJEED, S., BAN, Y. E., CHEN, L., HUANG, C. C., KONG, L., KWON, Y. D., STUCKEY, J., ZHOU, T., ROBINSON, J. E., SCHIEF, W. R., SODROSKI, J., WYATT, R. & KWONG, P. D. 2010. Structure of HIV-1 gp120 with gp41-interactive region reveals layered envelope architecture and basis of conformational mobility. *Proc Natl Acad Sci U S A,* 107**,** 1166-71.

PANTOPHLET, R. & BURTON, D. R. 2006. GP120: target for neutralizing HIV-1 antibodies. *Annu Rev Immunol,* 24**,** 739-69.

PARK, H. E., GRUENKE, J. A. & WHITE, J. M. 2003. Leash in the groove mechanism of membrane fusion. *Nat Struct Biol,* 10**,** 1048-53.

PENN, M. L., GRIVEL, J. C., SCHRAMM, B., GOLDSMITH, M. A. & MARGOLIS, L. 1999. CXCR4 utilization is sufficient to trigger CD4+ T cell depletion in HIV-1-infected human lymphoid tissue. *Proc Natl Acad Sci U S A,* 96**,** 663-8.

PICCHIO, G. R., GULIZIA, R. J., WEHRLY, K., CHESEBRO, B. & MOSIER, D. E. 1998. The cell tropism of human immunodeficiency virus type 1 determines the kinetics of plasma viremia in SCID mice reconstituted with human peripheral blood leukocytes. *J Virol,* 72**,** 2002-9.

PINTER, A., KOPELMAN, R., LI, Z., KAYMAN, S. C. & SANDERS, D. A. 1997. Localization of the labile disulfide bond between SU and TM of the murine leukemia virus envelope protein complex to a highly conserved CWLC motif in SU that resembles the active-site sequence of thiol-disulfide exchange enzymes. *J Virol,* 71**,** 8073-7.

PIQUE, C., PHAM, D., TURSZ, T. & DOKHELAR, M. C. 1992. Human T-cell leukemia virus type I envelope protein maturation process: requirements for syncytium formation. *J Virol,* 66**,** 906-13.

POIESZ, B. J., RUSCETTI, F. W., GAZDAR, A. F., BUNN, P. A., MINNA, J. D. & GALLO, R. C. 1980. Detection and isolation of type C retrovirus particles from fresh and cultured lymphocytes of a patient with cutaneous T-cell lymphoma. *Proc Natl Acad Sci U S A,* 77**,** 7415-9.

POIGNARD, P., MOULARD, M., GOLEZ, E., VIVONA, V., FRANTI, M., VENTURINI, S., WANG, M., PARREN, P. W. & BURTON, D. R. 2003. Heterogeneity of envelope molecules expressed on primary human immunodeficiency virus type 1 particles as probed by the binding of neutralizing and nonneutralizing antibodies. *J Virol,* 77**,** 353-65.

POPOVIC, M., SARNGADHARAN, M. G., READ, E. & GALLO, R. C. 1984. Detection, isolation, and continuous production of cytopathic retroviruses (HTLV-III) from patients with AIDS and pre-AIDS. *Science,* 224**,** 497-500.

POUMBOURIOS, P., CENTER, R. J., WILSON, K. A., KEMP, B. E. & KOBE, B. 1999. Evolutionary conservation of the membrane fusion machine. *IUBMB Life,* 48**,** 151-6.

POUMBOURIOS, P., EL AHMAR, W., MCPHEE, D. A. & KEMP, B. E. 1995. Determinants of human immunodeficiency virus type 1 envelope glycoprotein oligomeric structure. *J Virol,* 69**,** 1209-18.

POUMBOURIOS, P., MAERZ, A. L. & DRUMMER, H. E. 2003. Functional evolution of the HIV-1 envelope glycoprotein 120 association site of glycoprotein 41. *J Biol Chem,* 278**,** 42149-60.

POUMBOURIOS, P., WILSON, K. A., CENTER, R. J., EL AHMAR, W. & KEMP, B. E. 1997. Human immunodeficiency virus type 1 envelope glycoprotein oligomerization requires the gp41 amphipathic alpha-helical/leucine zipper-like sequence. *J Virol,* 71**,** 2041-9.

PRESTON, B. D. 1997. Reverse transcriptase fidelity and HIV-1 variation. *Science,* 275**,** 228-9; author reply 230-1.

RATNER, L., HASELTINE, W., PATARCA, R., LIVAK, K. J., STARCICH, B., JOSEPHS, S. F., DORAN, E. R., RAFALSKI, J. A., WHITEHORN, E. A., BAUMEISTER, K. & ET AL. 1985. Complete nucleotide sequence of the AIDS virus, HTLV-III. *Nature,* 313**,** 277-84.

RAY, N. & DOMS, R. W. 2006. HIV-1 coreceptors and their inhibitors. *Curr Top Microbiol Immunol,* 303**,** 97-120.

RE, F. & LUBAN, J. 1997. HIV-1 Vpr: G2 cell cycle arrest, macrophages and nuclear transport. *Prog Cell Cycle Res,* 3**,** 21-7.

REEVES, J. D. & DOMS, R. W. 2002. Human immunodeficiency virus type 2. *J Gen Virol,* 83**,** 1253-65.

REITER, Y., BRINKMANN, U., JUNG, S. H., PASTAN, I. & LEE, B. 1995. Disulfide stabilization of antibody Fv: computer predictions and experimental evaluation. *Protein Eng,* 8**,** 1323-31.

REPITS, J., STERJOVSKI, J., BADIA-MARTINEZ, D., MILD, M., GRAY, L., CHURCHILL, M. J., PURCELL, D. F., KARLSSON, A., ALBERT, J., FENYO, E. M., ACHOUR, A., GORRY, P. R. & JANSSON, M. 2008. Primary HIV-1 R5 isolates from end-stage disease display enhanced viral fitness in parallel with increased gp120 net charge. *Virology,* 379**,** 125-34.

RIBET, D., DEWANNIEUX, M. & HEIDMANN, T. 2004. An active murine transposon family pair: retrotransposition of "master" MusD copies and ETn trans-mobilization. *Genome Res,* 14**,** 2261-7.

RIBET, D., HARPER, F., DUPRESSOIR, A., DEWANNIEUX, M., PIERRON, G. & HEIDMANN, T. 2008. An infectious progenitor for the murine IAP retrotransposon: emergence of an intracellular genetic parasite from an ancient retrovirus. *Genome Res,* 18**,** 597-609.

ROUS, P. 1911. A Sarcoma of the Fowl Transmissible by an Agent Separable from the Tumor Cells. *J Exp Med,* 13**,** 397-411.

RUPRECHT, C. R., KRARUP, A., REYNELL, L., MANN, A. M., BRANDENBERG, O. F., BERLINGER, L., ABELA, I. A., REGOES, R. R., GUNTHARD, H. F., RUSERT, P. & TRKOLA, A. 2011. MPER-specific antibodies induce gp120 shedding and irreversibly neutralize HIV-1. *J Exp Med,* 208**,** 439-54.

SAEED, M. F., KOLOKOLTSOV, A. A., ALBRECHT, T. & DAVEY, R. A. 2010. Cellular entry of ebola virus involves uptake by a macropinocytosis-like mechanism and subsequent trafficking through early and late endosomes. *PLoS Pathog,* 6.

SALZWEDEL, K., WEST, J. T. & HUNTER, E. 1999. A conserved tryptophan-rich motif in the membrane-proximal region of the human immunodeficiency virus type 1 gp41 ectodomain is important for Env-mediated fusion and virus infectivity. *J Virol,* 73**,** 2469-80.

SANCHEZ, A. 2007. Analysis of filovirus entry into vero e6 cells, using inhibitors of endocytosis, endosomal acidification, structural integrity, and cathepsin (B and L) activity. *J Infect Dis,* 196 Suppl 2**,** S251-8.

SANCHEZ, A., TRAPPIER, S. G., MAHY, B. W., PETERS, C. J. & NICHOL, S. T. 1996. The virion glycoproteins of Ebola viruses are encoded in two reading frames and are expressed through transcriptional editing. *Proc Natl Acad Sci U S A,* 93**,** 3602-7.

SANMIGUEL, P. & BENNETZEN, J. L. 1998. Evidence that a Recent Increase in Maize Genome Size was Caused by the Massive Amplification of Intergene Retrotransposons. *Annals of Botany,* 82**,** 37-44.

SATTENTAU, Q. J., ZOLLA-PAZNER, S. & POIGNARD, P. 1995. Epitope exposure on functional, oligomeric HIV-1 gp41 molecules. *Virology,* 206**,** 713-7.

SCARLATTI, G., TRESOLDI, E., BJORNDAL, A., FREDRIKSSON, R., COLOGNESI, C., DENG, H. K., MALNATI, M. S., PLEBANI, A., SICCARDI, A. G., LITTMAN, D. R., FENYO, E. M. & LUSSO, P. 1997. In vivo evolution of HIV-1 co-receptor usage and sensitivity to chemokine-mediated suppression. *Nat Med,* 3**,** 1259-65.

SCHNEIDER, J., KAADEN, O., COPELAND, T. D., OROSZLAN, S. & HUNSMANN, G. 1986. Shedding and interspecies type sero-reactivity of the envelope glycopolypeptide gp120 of the human immunodeficiency virus. *J Gen Virol,* 67 (Pt 11)**,** 2533-8.

SCHORNBERG, K., MATSUYAMA, S., KABSCH, K., DELOS, S., BOUTON, A. & WHITE, J. 2006. Role of endosomal cathepsins in entry mediated by the Ebola virus glycoprotein. *J Virol,* 80**,** 4174-8.

SEN, J., JACOBS, A. & CAFFREY, M. 2008. Role of the HIV gp120 Conserved Domain 5 in Processing and Viral Entry†. *Biochemistry,* 47**,** 7788-7795.

SI, Z., MADANI, N., COX, J. M., CHRUMA, J. J., KLEIN, J. C., SCHON, A., PHAN, N., WANG, L., BIORN, A. C., COCKLIN, S., CHAIKEN, I., FREIRE, E., SMITH, A. B., 3RD & SODROSKI, J. G. 2004. Small-molecule inhibitors of HIV-1 entry block receptor-induced conformational changes in the viral envelope glycoproteins. *Proc Natl Acad Sci U S A,* 101**,** 5036-41.

SKEHEL, J. J., BAYLEY, P. M., BROWN, E. B., MARTIN, S. R., WATERFIELD, M. D., WHITE, J. M., WILSON, I. A. & WILEY, D. C. 1982. Changes in the conformation of influenza virus hemagglutinin at the pH optimum of virus-mediated membrane fusion. *Proc Natl Acad Sci U S A,* 79**,** 968-72.

SKEHEL, J. J. & WILEY, D. C. 2000. Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. *Annu Rev Biochem,* 69**,** 531-69.

SONG, L., SUN, Z. Y., COLEMAN, K. E., ZWICK, M. B., GACH, J. S., WANG, J. H., REINHERZ, E. L., WAGNER, G. & KIM, M. 2009. Broadly neutralizing anti-HIV-1 antibodies disrupt a hinge-related function of gp41 at the membrane interface. *Proc Natl Acad Sci U S A,* 106**,** 9057-62.

SOWDHAMINI, R., SRINIVASAN, N., SHOICHET, B., SANTI, D. V., RAMAKRISHNAN, C. & BALARAM, P. 1989. Stereochemical modeling of disulfide bridges. Criteria for introduction into proteins by site-directed mutagenesis. *Protein Eng,* 3**,** 95-103.

STARCICH, B., RATNER, L., JOSEPHS, S. F., OKAMOTO, T., GALLO, R. C. & WONG-STAAL, F. 1985. Characterization of long terminal repeat sequences of HTLV-III. *Science,* 227**,** 538-40.

STEHELIN, D., FUJITA, D. J., PADGETT, T., VARMUS, H. E. & BISHOP, J. M. 1977. Detection and enumeration of transformation-defective strains of avian sarcoma virus with molecular hybridization. *Virology,* 76**,** 675-84.

STEHELIN, D., VARMUS, H. E., BISHOP, J. M. & VOGT, P. K. 1976. DNA related to the transforming gene(s) of avian sarcoma viruses is present in normal avian DNA. *Nature,* 260**,** 170-3.

STEIN, B. S. & ENGLEMAN, E. G. 1990. Intracellular processing of the gp160 HIV-1 envelope precursor. Endoproteolytic cleavage occurs in a cis or medial compartment of the Golgi complex. *J Biol Chem,* 265**,** 2640-9.

STERJOVSKI, J., CHURCHILL, M. J., ELLETT, A., GRAY, L. R., ROCHE, M. J., DUNFEE, R. L., PURCELL, D. F., SAKSENA, N., WANG, B., SONZA, S., WESSELINGH, S. L., KARLSSON, I., FENYO, E. M., GABUZDA, D., CUNNINGHAM, A. L. & GORRY, P. R. 2007. Asn 362 in gp120 contributes to enhanced fusogenicity by CCR5-restricted HIV-1 envelope glycoprotein variants from patients with AIDS. *Retrovirology,* 4**,** 89.

STOYE, J. P. & COFFIN, J. M. 1987. The four classes of endogenous murine leukemia virus: structural relationships and potential for recombination. *J Virol,* 61**,** 2659-69.

SUN, Z. Y., OH, K. J., KIM, M., YU, J., BRUSIC, V., SONG, L., QIAO, Z., WANG, J. H., WAGNER, G. & REINHERZ, E. L. 2008. HIV-1 broadly neutralizing antibody

extracts its epitope from a kinked gp41 ectodomain region on the viral membrane. *Immunity,* 28**,** 52-63.

TAKADA, A., ROBISON, C., GOTO, H., SANCHEZ, A., MURTI, K. G., WHITT, M. A. & KAWAOKA, Y. 1997. A system for functional analysis of Ebola virus glycoprotein. *Proc Natl Acad Sci U S A,* 94**,** 14764-9.

TAKATSUKI, K., UCHIYAMA, T. & SAGAWA, K. 1978. [Subsets of normal lymphocytes and malignant lymphoid cells, with special reference to adult T-cell leukemia (author's transl)]. *Rinsho Byori,* 26**,** 188-93.

TAKIKAWA, S., ISHII, K., AIZAKI, H., SUZUKI, T., ASAKURA, H., MATSUURA, Y. & MIYAMURA, T. 2000. Cell fusion activity of hepatitis C virus envelope proteins. *J Virol,* 74**,** 5066-74.

TAN, K., LIU, J., WANG, J., SHEN, S. & LU, M. 1997. Atomic structure of a thermostable subdomain of HIV-1 gp41. *Proc Natl Acad Sci U S A,* 94**,** 12303-8.

TAYLOR, B. M., FOULKE, J. S., FLINKO, R., HEREDIA, A., DEVICO, A. & REITZ, M. 2008. An alteration of human immunodeficiency virus gp41 leads to reduced CCR5 dependence and CD4 independence. *J Virol,* 82**,** 5460-71.

TEMIN, H. M. 1980. Origin of retroviruses from cellular moveable genetic elements. *Cell,* 21**,** 599-600.

THEODORE, T. S., ENGLUND, G., BUCKLER-WHITE, A., BUCKLER, C. E., MARTIN, M. A. & PEDEN, K. W. 1996. Construction and characterization of a stable full-length macrophage-tropic HIV type 1 molecular clone that directs the production of high titers of progeny virions. *AIDS Res Hum Retroviruses,* 12**,** 191-4.

THOENNES, S., LI, Z.-N., LEE, B.-J., LANGLEY, W. A., SKEHEL, J. J., RUSSELL, R. J. & STEINHAUER, D. A. 2008. Analysis of residues near the fusion peptide in the influenza hemagglutinin structure for roles in triggering membrane fusion. *Virology,* 370**,** 403-414.

TRKOLA, A., DRAGIC, T., ARTHOS, J., BINLEY, J. M., OLSON, W. C., ALLAWAY, G. P., CHENG-MAYER, C., ROBINSON, J., MADDON, P. J. & MOORE, J. P. 1996. CD4-dependent, antibody-sensitive interactions between HIV-1 and its co-receptor CCR-5. *Nature,* 384**,** 184-7.

TUTTLE, D. L., ANDERS, C. B., AQUINO-DE JESUS, M. J., POOLE, P. P., LAMERS, S. L., BRIGGS, D. R., POMEROY, S. M., ALEXANDER, L., PEDEN, K. W., ANDIMAN, W. A., SLEASMAN, J. W. & GOODENOW, M. M. 2002. Increased replication of non-syncytium-inducing HIV type 1 isolates in monocyte-derived macrophages is linked to advanced disease in infected children. *AIDS Res Hum Retroviruses,* 18**,** 353-62.

UCHIYAMA, T., YODOI, J., SAGAWA, K., TAKATSUKI, K. & UCHINO, H. 1977. Adult T-cell leukemia: clinical and hematologic features of 16 cases. *Blood,* 50**,** 481-92.

UNAIDS/WHO 2010. Report on the global AIDS epidemic: 2010.

VINCENT, N., GENIN, C. & MALVOISIN, E. 2002. Identification of a conserved domain of the HIV-1 transmembrane protein gp41 which interacts with cholesteryl groups. *Biochim Biophys Acta,* 1567**,** 157-64.

WALLIN, M., EKSTROM, M. & GAROFF, H. 2004. Isomerization of the intersubunit disulphide-bond in Env controls retrovirus fusion. *EMBO J,* 23**,** 54-65.

WANG, J., SEN, J., RONG, L. & CAFFREY, M. 2008. Role of the HIV gp120 conserved domain 1 in processing and viral entry. *J Biol Chem,* 283**,** 32644-9.

WANG, J., TONG, P., LU, L., ZHOU, L., XU, L., JIANG, S. & CHEN, Y. H. 2011. HIV-1 gp41 core with exposed membrane-proximal external region inducing broad HIV-1 neutralizing antibodies. *PLoS One,* 6**,** e18233.

WEIS, W., BROWN, J. H., CUSACK, S., PAULSON, J. C., SKEHEL, J. J. & WILEY, D. C. 1988. Structure of the influenza virus haemagglutinin complexed with its receptor, sialic acid. *Nature,* 333**,** 426-31.

WEISSENHORN, W., CALDER, L. J., DESSEN, A., LAUE, T., SKEHEL, J. J. & WILEY, D. C. 1997a. Assembly of a rod-shaped chimera of a trimeric GCN4 zipper and the HIV-1 gp41 ectodomain expressed in Escherichia coli. *Proc Natl Acad Sci U S A,* 94**,** 6065-9.

WEISSENHORN, W., CALDER, L. J., WHARTON, S. A., SKEHEL, J. J. & WILEY, D. C. 1998a. The central structural feature of the membrane fusion protein subunit from the Ebola virus glycoprotein is a long triple-stranded coiled coil. *Proc Natl Acad Sci U S A,* 95**,** 6032-6.

WEISSENHORN, W., CARFI, A., LEE, K. H., SKEHEL, J. J. & WILEY, D. C. 1998b. Crystal structure of the Ebola virus membrane fusion subunit, GP2, from the envelope glycoprotein ectodomain. *Mol Cell,* 2**,** 605-16.

WEISSENHORN, W., DESSEN, A., HARRISON, S. C., SKEHEL, J. J. & WILEY, D. C. 1997b. Atomic structure of the ectodomain from HIV-1 gp41. *Nature,* 387**,** 426-30.

WEISSENHORN, W., WHARTON, S. A., CALDER, L. J., EARL, P. L., MOSS, B., ALIPRANDIS, E., SKEHEL, J. J. & WILEY, D. C. 1996. The ectodomain of HIV-1 env subunit gp41 forms a soluble, alpha-helical, rod-like oligomer in the absence of gp120 and the N-terminal fusion peptide. *Embo J,* 15**,** 1507-14.

WILD, C., DUBAY, J. W., GREENWELL, T., BAIRD, T., JR., OAS, T. G., MCDANAL, C., HUNTER, E. & MATTHEWS, T. 1994. Propensity for a leucine zipper-like domain of human immunodeficiency virus type 1 gp41 to form oligomers correlates with a role in virus-induced fusion rather than assembly of the glycoprotein complex. *Proc Natl Acad Sci U S A,* 91**,** 12676-80.

WILD, C., GREENWELL, T., SHUGARS, D., RIMSKY-CLARKE, L. & MATTHEWS, T. 1995. The inhibitory activity of an HIV type 1 peptide correlates with its ability to interact with a leucine zipper structure. *AIDS Res Hum Retroviruses,* 11**,** 323-5.

WILEY, D. C., WILSON, I. A. & SKEHEL, J. J. 1981. Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation. *Nature,* 289**,** 373-8.

WILLEY, R. L., BONIFACINO, J. S., POTTS, B. J., MARTIN, M. A. & KLAUSNER, R. D. 1988. Biosynthesis, cleavage, and degradation of the human immunodeficiency virus 1 envelope glycoprotein gp160. *Proc Natl Acad Sci U S A,* 85**,** 9580-4.

WILSON, I. A., SKEHEL, J. J. & WILEY, D. C. 1981. Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 A resolution. *Nature,* 289**,** 366-73.

WYATT, R., KWONG, P. D., DESJARDINS, E., SWEET, R. W., ROBINSON, J., HENDRICKSON, W. A. & SODROSKI, J. G. 1998. The antigenic structure of the HIV gp120 envelope glycoprotein. *Nature,* 393**,** 705-11.

WYMA, D. J., JIANG, J., SHI, J., ZHOU, J., LINEBERGER, J. E., MILLER, M. D. & AIKEN, C. 2004. Coupling of human immunodeficiency virus type 1 fusion to virion maturation: a novel role of the gp41 cytoplasmic tail. *J Virol,* 78**,** 3429-35.

YANAGI, Y., ONO, N., TATSUO, H., HASHIMOTO, K. & MINAGAWA, H. 2002. Measles virus receptor SLAM (CD150). *Virology,* 299**,** 155-61.

YANG, X., KURTEVA, S., REN, X., LEE, S. & SODROSKI, J. 2005. Stoichiometry of envelope glycoprotein trimers in the entry of human immunodeficiency virus type 1. *J Virol,* 79**,** 12132-47.

YANG, X., MAHONY, E., HOLM, G. H., KASSA, A. & SODROSKI, J. 2003. Role of the gp120 inner domain [beta]-sandwich in the interaction between the human immunodeficiency virus envelope glycoprotein subunits. *Virology,* 313**,** 117-125.

YANG, Z. N., MUESER, T. C., KAUFMAN, J., STAHL, S. J., WINGFIELD, P. T. & HYDE, C. C. 1999. The crystal structure of the SIV gp41 ectodomain at 1.47 A resolution. *J Struct Biol,* 126**,** 131-44.

YIN, H. S., PATERSON, R. G., WEN, X., LAMB, R. A. & JARDETZKY, T. S. 2005. Structure of the uncleaved ectodomain of the paramyxovirus (hPIV3) fusion protein. *Proc Natl Acad Sci U S A,* 102**,** 9288-93.

YIN, H. S., WEN, X., PATERSON, R. G., LAMB, R. A. & JARDETZKY, T. S. 2006. Structure of the parainfluenza virus 5 F protein in its metastable, prefusion conformation. *Nature,* 439**,** 38-44.

ZANETTI, G., BRIGGS, J. A., GRUNEWALD, K., SATTENTAU, Q. J. & FULLER, S. D. 2006. Cryo-electron tomographic structure of an immunodeficiency virus envelope complex in situ. *PLoS Pathog,* 2**,** e83.

ZHOU, T., XU, L., DEY, B., HESSELL, A. J., VAN RYK, D., XIANG, S. H., YANG, X., ZHANG, M. Y., ZWICK, M. B., ARTHOS, J., BURTON, D. R., DIMITROV, D. S., SODROSKI, J., WYATT, R., NABEL, G. J. & KWONG, P. D. 2007. Structural definition of a conserved neutralization epitope on HIV-1 gp120. *Nature,* 445**,** 732-7.

ZHOU, Y., SHEN, L., YANG, H. C. & SILICIANO, R. F. 2008. Preferential cytolysis of peripheral memory CD4+ T cells by in vitro X4-tropic human immunodeficiency virus type 1 infection before the completion of reverse transcription. *J Virol,* 82**,** 9154-63.

ZHU, P., LIU, J., BESS, J., JR., CHERTOVA, E., LIFSON, J. D., GRISE, H., OFEK, G. A., TAYLOR, K. A. & ROUX, K. H. 2006. Distribution and three-dimensional structure of AIDS virus envelope spikes. *Nature,* 441**,** 847-52.