Department of Laboratory Medicine,
Division of Clinical Microbiology
Karolinska Institutet,
Stockholm, Sweden

LIGAND-INDUCED STRUCTURAL CHANGES IN HIV-1 ENVELOPE GLYCOPROTEIN

Selina Poon

Karolinska Institutet

Stockholm 2013

Published by Karolinska Institutet. Printed by Larserics Digital Print, AB.

ISBN 978-91-7549-264-3

To my Family,
For their love, support and encouragement

ABSTRACT

The first step in HIV-1 entry is the interaction between viral envelope glycoprotein gp120 and CD4 receptor on permissive host cells. The interaction initiates a series of sequential conformational rearrangements that exposes epitopes involved in interaction with co-receptors. Using x-ray crystallography, the tertiary structural rearrangements are elucidated. However, the envelope glycoprotein on the virus surface is arranged as a trimer consisting of three gp120 and three gp41 subunits (Env) and the tertiary conformational change observed in gp120 has not been characterised in the context of a quaternary structure with sufficient conformity.

In this thesis, the envelope glycoprotein responsible for initiating virus entry, Env, was studied using cryo-electron microscopy. The quaternary structure of Env, in the pre-entry and entry intermediate, provided evidence of conformational shifts upon receptor or ligand-binding. This method was further used to analyse the quaternary structure of another ligand-bound Env, the Tat protein-bound Env. The Tat protein-bound Env was studied due to the postulation that Tat protein contributes, in part, to disease progression. This stems from the observations that combination vaccine trials encompassing both Tat and Env proteins protected macaques from mucosal SHIV challenge, that monoclonal anti-Tat antibodies inhibited HIV-1 replication and that HIV-1 infected individuals who are slow progressors of the disease (have a longer clinical asymptomatic phase) show an inverse relationship in their anti-Tat antibody levels and viral load magnitude. In addition, the combined Tat-Env vaccine induced antibodies that recognised more regions in Env and had antibodies with higher reactivity for epitopes in the V3 loop of Env.

Comparative analysis of the structures obtained showed that after HIV-1 binding to its receptor CD4, the tertiary conformational arrangement of gp120 is translated to the following quaternary changes: gp120 subunit tilt from the z-axis, gp120 subunit rotation along its own axis and rotation of CD4 epitopes with concomitant exposure of co-receptor epitopes. Including the Tat protein-bound Env in the comparison showed that Tat-bound Env has a conformation intermediate between the native and CD4-bound Env, and hence, optimal for virus entry. This structure change could explain the increased disease progression in patients with low anti-Tat antibody to viral load magnitude and slower disease progressions in those patients with high anti-Tat antibody. In addition, the intermediate structure of Tat-gp120 complex could expose epitopes on gp120 that provide better immune protection in slow progressors.

This thesis further evaluated the effect of Tat protein treatment of HIV-1 on infectivity and spread *in vitro*, showing that Tat protein increases infection in target cells and increases spread between infected and non-infected cells through cell-cell contacts. The increased infectivity and spread was proposed to be due to interaction of Tat protein with the V3 loop of gp120, which potentially alters the co-receptor recognition. In this way, HIV-1 of a particular tropism, for example an X4-tropic HIV-1, could infect by using CCR5 as a co-receptor instead. Moreover, the low concentration of circulating Tat protein observed in the serum of HIV-1 patients could be a result of regulation by the virus as Tat protein at high concentration was found to have a negative impact on infectivity rates.

Key words: HIV-1, Env, CD4, Tat, quaternary structure, infectivity

LIST OF PUBLICATIONS

This thesis is based on the following papers:

I. Carlos G. Moscoso, Yide Sun, Selina Poon, Li Xing, Elaine Kan, Loïc Martin, Dominik Green, Frank Lin, Anders G. Vahlne, Susan Barnett, Indresh Srivastava, and R. Holland Cheng**. Quaternary structures of HIV Env immunogen exhibit conformation vicissitudes and interface diminution elicted by ligand binding.** *Proc Natl Acad Sci U S A. 2011. 108(15): 6091-6096.*

II. Selina Poon, Carlos G. Moscoso, Li Xing, Elaine Kan, Yide Sun, Prasanna R. Kolatkar, Anders Vahlne, Indresh Srivastava, Susan W. Barnett, and R. Holland Cheng. **Putative role of Tat-Env interaction in HIV infection.** *AIDS. 2013. In press. Published online ahead of print July 24.*

III. Selina Poon, Carlos G Moscoso, Onur M Yenigun, Prasanna R. Kolatkar, R. Holland Cheng, and Anders Vahlne. **HIV-1 Tat protein induces viral internalization through Env-mediated interactions in dose-dependent manner.** *AIDS. 2013. In press. Published online ahead of print July 24.*

CONTENTS

List of abbreviations

AIDS	**a**cquired **i**mmune **d**eficiency **s**yndrome
APC	**A**llo**p**hyco**c**yanin
ARV	**A**IDS-associated **r**etro**v**irus
CD4m	**CD4 m**imetic miniprotein
CDC	**C**entres for **D**isease, **C**ontrol
CFSE	**c**arboxy**f**luorescein diacetate **s**uccinimidyl **e**ster
CHR	**C**-terminal **h**eptad **r**epeat
CTL	**c**ytotoxic **T l**ymphocyte
Env	**env**elope glycoprotein
Fab	**a**nti**b**ody **F**ragment
FP	**f**usion **p**eptide
gp160	glycoprotein 160 (precursor of gp120 and gp41)
gp120	glycoprotein 120 (surface subunit)
gp41	glycoprotein 41 (membrane spanning subunit)
HAART	**h**ighly **a**ctive **a**nti**r**etroviral **t**herapy
HIV-1	**h**uman **i**mmunodeficiency **v**irus type **1**
HR	**h**eptad **r**epeat
HTLV (I or II)	**h**uman **T**-**l**ymphotropic **v**irus (type **I** or **II**)
IN	**in**tegrase protein
LAV	**l**ymphadenopathy **a**ssociated **v**irus
LTR	**l**ong **t**erminal **r**epeat
MA	**ma**trix protein
MPER	**m**embrane **p**roximal **e**xtracellular **r**egion
NHR	**N**-terminal **h**eptad **r**epeat
NNRI	**n**on-**n**ucleoside **r**everse **t**ranscriptase **i**nhibitor
NRT	**n**ucleoside **r**everse **t**ranscriptase **i**nhibitor
NtRI	**n**ucleo**t**ide **r**everse **t**ranscriptase **i**nhibitor
P-TEFb	**p**ositive **t**ranscription **e**longation **f**actor **b**
$p55^{Gag}$	**55**kDa **p**olyprotein **Gag**
$p160^{Gag/Pol}$	**160**kDa **p**olyprotein **Gag/Pol**
PBMCs	**p**eripheral **b**lood mononuclear **c**ell**s**
PI	**p**rotease **i**nhibitor

PR	viral **pr**otease
rER	**r**ough **e**ndoplasmic **r**eticulum
RNA	**r**ibo**n**ucleic **a**cid
RNAP	**RNA p**olymerase
RT	**r**everse **t**ranscriptase
SIV	**s**imian **i**mmunodeficiency **v**irus
UNAIDS	Joint **U**nited **N**ations Programme on HIV/**AIDS**
VEGF	**v**ascular **e**ndothelial **g**rowth **f**actor
VEGFR-2	**v**ascular **e**ndothelial **g**rowth **f**actor **r**eceptor **2**

HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1)

ORIGIN AND DISCOVERY OF HIV-1

In the summer of 1981, clinicians observed a significant cluster of rare diseases and opportunistic infections among homosexual men in New York and California [1, 2]. They were later found to have severe immune depletion of $CD4^+$ T cells, with common immunological deficit in cell-mediated immunity. Initial observations led to the theory that it was a retrovirus, since human T lymphotropic viruses (HTLV-I and HTLV-II) were the only viruses known then to infect $CD4^+$ T lymphocytes. In 1983, a putative new human retrovirus, termed lymphadenopathy-associated virus (LAV), was identified from the lymph node of a patient suffering from generalised persistent lymphadenopathy. Anti-sera to HTLV-I did not precipitate proteins of LAV, which indicated that LAV was distinct from HTLV-I [3]. The following year, isolation of the new retrovirus with morphology distinct from HTLV-I and HTLV-II but similar to LAV was achieved from two siblings with haemophilia B where one had acquired immune deficiency syndrome (AIDS) while the other was healthy [4]. The term AIDS was first properly defined by the Centers for Disease, Control (CDC) in September 1982 to describe an acquired, compromised immune system that resulted in several clinical manifestations [5]. Despite the early identification and isolation, the new retrovirus was not shown to be the etiological agent of AIDS even though it was clearly associated with the disease. It was only a month later when a series of papers was published that described the new retrovirus, designated HTLV-III, as the etiological agent of AIDS [6-9]. The new retrovirus also had another name, the AIDS-associated retrovirus (ARV), when it was isolated by researchers in San Francisco from patients with AIDS or lymphadenopathy syndrome [10]. Nucleotide sequence analysis of LAV, HTLV-III and ARV showed that these viruses are very similar [11, 12] and the International Committee of Viral Taxonomy renamed them HIV [13]. When another similar yet distinct immunodeficiency virus was discovered in 1986, the terms HIV-1 and HIV-2 were introduced to distinguish the latter from the first HIV discovered [14].

Molecular virology and epidemiology studies identified the oldest known HIV-infected samples to come from the African city of Kinshasa (formerly Leopoldville), in the Democratic Republic of the Congo, in a blood plasma sample and a lymph node biopsy in 1959 and 1960, respectively [15]. Phylogenetic analysis initially estimated HIV-1 to originate in the 1940s or early 1950s [16], but a more recent comparative evolutionary

genetic study proposed the origin to be sometime in the early twentieth century, a period of time that corresponded to rise in urbanisation in the Kinshasa area [17]. Regardless of its origin, the nature of the course of an HIV-1 infection meant that individuals who acquired HIV-1 could go unrecognised for many years before an actual diagnosis is made [18]. In the absence of awareness and preventive measures, it was estimated that by the end of the first decade of HIV discovery, between 5 and 10 million people became infected.

TRANSMISSION AND DYNAMICS OF HIV-1 INFECTION

HIV-1 is a blood-borne virus, i.e. is transmitted mainly through blood exposure and contact of mucosal surfaces to HIV-1-contaminated fluids [19].

Clinically, HIV-1 infection is divided in three phases: An initial phase termed primary infection, a second phase which consists of a long asymptomatic period between the primary infection and the final phase where infection is characterised by the emergence of clinical immunodeficiency (AIDS).

Cells targeted by HIV-1 express both the CD4 receptor and an appropriate co-receptor [20]. The $CD4^+$ T lymphocytes of the immune system appear to be the major target cells [21], but macrophages, monocytes and microglial cells can also be infected. Shortly after infection, the individual develops viremia which is accompanied by infectious mononucleosis-like symptoms in some patients. This usually does not last more than four to six weeks [22, 23]. A sharp decline in $CD4^+$ T lymphocyte is observed and correlates inversely with high copies of viral RNA detected in the blood (Figure 1).

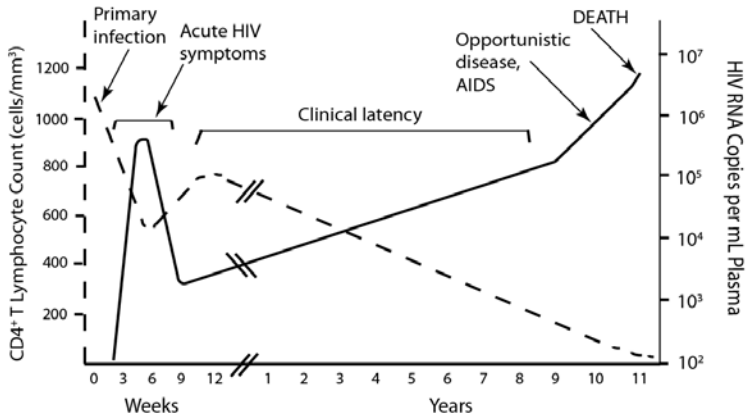


Figure 1. Natural course of HIV-1 infection in the absence of antiretroviral therapy.

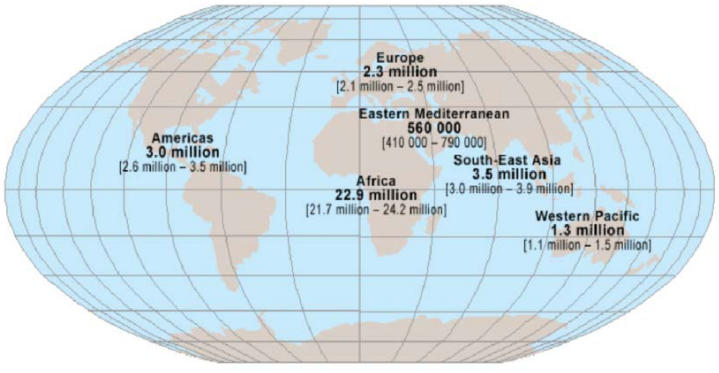
The level of virus in the blood gradually declines within a few weeks. This coincides with an antiviral response to HIV-1 established by the host immune system. As the level of virus in the blood decreases, the $CD4^+$ T lymphocyte level recovers to subnormal levels. The course of infection is now in the second phase which lasts a variable number of years. The individual is usually clinically asymptomatic. However, during this time two pathophysiological processes take place: Continuous virus replication, as evidenced by the detection of free virus in circulation and infected cells in the peripheral lymphoid tissues [24, 25], and a gradual loss of $CD4^+$ T lymphocytes.

In the absence of highly active antiretroviral therapy (HAART), the $CD4^+$ T lymphocyte count continues to decline. A precipitous drop in $CD4^+$ T lymphocytes is observed as the infection reaches the final AIDS phase over a median of 8 to 10 years after infection. Concurrent with the decline in $CD4^+$ T lymphocyte there is a significant increase in the level of HIV-1 particles in the blood [26]. As $CD4^+$ T lymphocyte count drops below 200 cells/µl, AIDS-defining illnesses and other opportunistic infections appear which eventually lead to death (Figure 1).

HIV-1 AND AIDS

The current world situation on HIV-1 and AIDS…

According to numbers provided by the Joint United Nations Programme on HIV/AIDS (UNAIDS) and the World Health Organisation (WHO), an estimated 34 million individuals globally are living with HIV-1, with 2.7 million new HIV-1 infections and 1.8 million AIDS-related deaths in 2010. Currently, the largest number of infected individuals is in Sub-Saharan Africa (Figure 2).



Total: 34.0 million [31.6 million – 35.2 million]

Figure 2. Adults and children estimated to be living with HIV in 2010, by WHO region.
Source: 2011 progress report on global HIV/AIDS, WHO/UNAIDS/UNICEF.

Since its identification in 1983, elaborate studies performed in three decades have helped us to understand the virus infection mechanisms and its mode of transmission [27]. Today, more than twenty drugs are approved for clinical use [28, 29]. These drugs target different stages of the virus replication cycle and fall under the following categories: Nucleoside reverse transcriptase inhibitor (NRTI), nucleotide reverse transcriptase inhibitor (NtRTI), non-nucleoside reverse transcriptase inhibitor (NNRTI), protease inhibitor (PI), integration inhibitor and fusion inhibitor [29]. Highly active anti-retroviral therapy (HAART) is a regimen that uses combinations of these drugs to treat patients. Such combination of different anti-retrovirals poses multiple obstacles to virus replication, reducing the possibility for the development of drug resistant virus mutants.

An important factor behind successful drug control of the infection and reduced risk of resistance development is compliance. However, despite impressive developments in anti-retroviral therapy and control of the disease, a universal anti-HIV vaccine still eludes researchers [30-32]. Persistent quiescent HIV-1 infection within long-lived $CD4^+$ T cells is a major obstacle to the goal of eradicating HIV-1 infection (reviewed in [33]). Even in highly compliant HAART patients where viremia is suppressed for long periods, latently infected cells were still identified [34-36]. Increasing duration of HAART also did not lead to decline in the frequency of these cells [35]. Thus, it seems that the virus cannot be cleared and an infected person carries the virus throughout his or her lifetime.

VIRUS STRUCTURE

HIV-1 is a lentivirus belonging to the retrovirus family. With an average diameter of 110 nm, the virion is roughly spherical and its conical capsid is surrounded by a host-derived lipid bilayer called the envelope. Embedded in the viral envelope are the envelope glycoproteins (Env) which is a complex of two proteins, the external, glycosylated surface subunit gp120 and the membrane spanning subunit gp41. They are synthesized as a single polyprotein, gp160, which is proteolytically cleaved during its transport to the surface of the infected cell. The surface subunits gp120 and gp41 are glycosylated during their transport to the surface of the infected cell. The processed proteins remain associated, with gp120 interacting non-covalently with the ectodomain

of gp41. On the surface, the gp120/gp41 complex is arranged in trimers, giving the Env a spike appearance.

The matrix protein (MA) lines the inner layer of the lipid membrane. Together with the lipid bilayer, the MA envelopes the capsid, the latter being a conical shaped structure built from the capsid protein (CA). The HIV-1 genome is found within the capsid and consists of two copies of single-stranded ribonucleic acid (RNA) tightly bound to the nucleocapsid protein (NC). Virus reverse transcriptase (RT) and integrase (IN) are also found within the capsid.

Newly released HIV-1 particles are immature and differ from the mature HIV-1 virion in its internal morphology characterised by the distinct absence of a conical-shaped capsid (Figure 3).

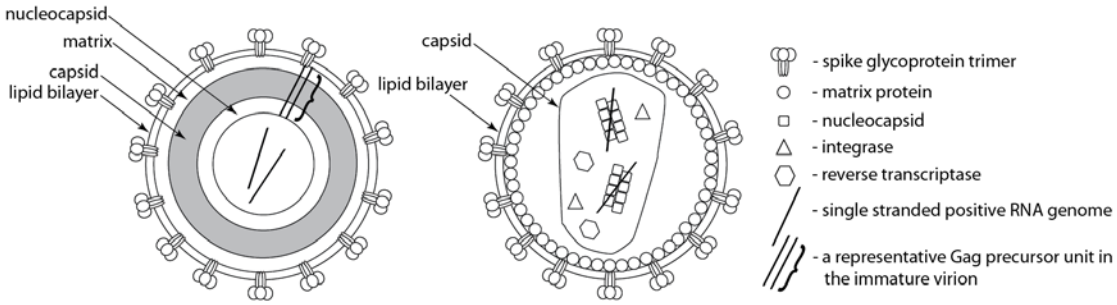


Figure 3. Schematic drawing of an immature HIV-1 virion (left) and a mature HIV-1 virion (right).

In transmission electron microscopy, these particles have a doughnut/crescent-shaped morphology [37] which arises from a radial arrangement of unprocessed 55kDa Gag polyprotein ($p55^{Gag}$) and 160kDa Gag/Pol polyprotein ($p160^{Gag/Pol}$), with the N-terminal at the viral membrane and the C-terminal located towards the centre of the particle [38]. The $p55^{Gag}$ comprises the MA, CA, NC, spacer peptides 1 and 2, and a p6 protein. The $p160^{Gag/Pol}$ comprises all components of the $p55^{Gag}$ in addition to the viral protease (PR), RT and IN proteins. During maturation, PR is activated which leads to the cleavage of $p55^{Gag}$ and $p160^{Gag/Pol}$ [39] with subsequent morphological rearrangement of the particle's interior (Figure 3).

VIRUS INFECTION AND REPLICATION CYCLE

Virus entry in a permissive host cell represents the first step in initiating a virus infection and is briefly described in Figure 4.

For HIV-1, entry is a complex series of events that begins with interaction of the gp120 with the primary receptor CD4 [40-42]. This interaction induces conformation changes in gp120 [43] that exposes the co-receptor binding site [44] which binds either the CCR5 [45-49] or CXCR4 chemokine receptors [50], depending on sequence and charge of the variable loop three (V3) of gp120 [51]. CCR5 and CXCR4 are the main chemokine receptors that HIV-1 uses as the co-receptors for entry, although other chemokine receptors have also been reported [20, 46, 48]. Receptor and co-receptor interactions bring the virus close to the target cell, and the large conformational changes eventually lead to dissociation of the gp120 from the gp120/gp41 complex and subsequent release of the metastable gp41 in an extended pre-hairpin intermediate state [52]. The ectodomain of gp41 contains several functional domains responsible for driving virus-host membrane fusion, including the fusion peptide, the N-terminal heptad repeat (NHR), the C-terminal heptad repeat (CHR) and the membrane proximal extracellular region (MPER) [53, 54]. Fusion begins when the fusion peptide is inserted into the host membrane brought into close proximity by the earlier receptor and co-receptor interactions. gp41 now folds into a six-helix bundle where the NHR and CHR are brought together in a hairpin conformation. Transition of gp41 into the six-helix bundle is believed to provide the energy necessary to facilitate apposition of virus and host membranes and to drive membrane fusion and virus entry [55-58].

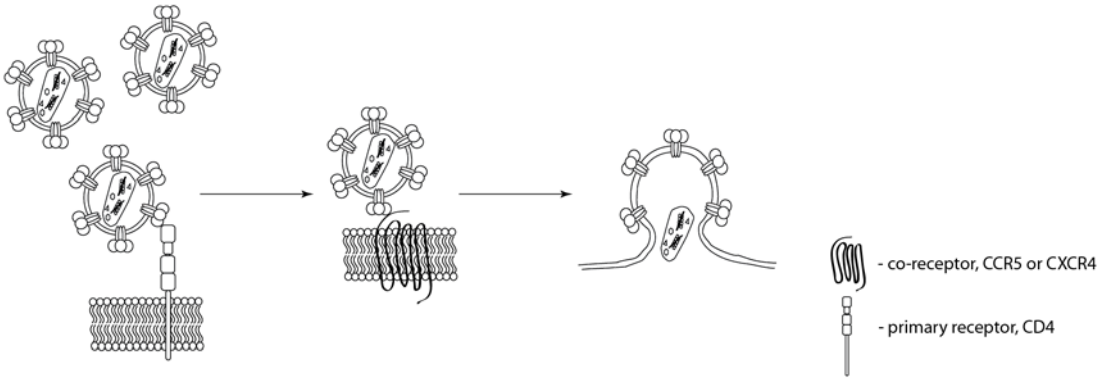


Figure 4. Schematic of virus entry by HIV-1. Virus entry utilises two different receptors: the primary receptor, CD4, and a co-receptor, CCR5 or CXCR4.

Once within the host cell, the viral RNA genome is reversely transcribed by virus RT, firstly to single-stranded DNA, and then to double-stranded DNA. The virus RT lacks proof-reading capability and is responsible for introducing mutations into the virus genome at this step which has consequences for virus evolution. The capsid shell subsequently disintegrates. The newly synthesized, double-stranded DNA genome is translocated across the nuclear pore as part of a high-molecular-weight complex known

as the pre-integration complex. The double-stranded virus DNA is integrated into the host genome by virus IN. The integrated virus DNA is referred to as the provirus.

The provirus is transcribed by host transcription machinery to mRNA. Splicing of viral mRNA [59] gives multiply-spliced transcripts and the expression of regulatory proteins which are crucial for production of infectious progeny virions. Later in replication, full-length transcripts are produced which are either unspliced or singly-spliced. Translation from unspliced mRNA leads to the synthesis of a $p55^{Gag}$. A frameshift event during translation of the unspliced mRNA results in the synthesis of a $p160^{Gag/Pol}$, which consists of the PR, IN and RT proteins in addition to the $p55^{Gag}$ component proteins.

Both Gag and Gag-Pol are transported to the host cell membrane where they assemble and lead to budding of the virions. Env is translated from the singly-spliced mRNA as a 160kDa glycoprotein (gp160) precursor. The transport of Env to the cell membrane occurs via the endoplasmic reticulum (ER)-Golgi network. During its transport, Env is cleaved into the gp120 and gp41 subunits and remain associated via non-covalent interactions. At the same time the gp120 subunit becomes highly glycosylated.

The assembled virion buds from the host cell membrane as an immature virus, with radial arrangement of $p55^{Gag}$ and $p160^{Gag/Pol}$. After virus has bud off from the host cell, the viral protease becomes active and catalyses the cleavage and processing of the precursor $p160^{Gag/Pol}$ and $p55^{Gag}$ at specific sites to release the component proteins required for viral maturation and assembly [60, 61]. The immature virus undergoes huge rearrangement within its envelope that results in the mature capsid core structure [62].

HIV-1 ENV

BIOSYNTHESIS

The Env protein is a type I membrane protein encoded by the *env* gene. The protein is synthesized as a 160kDa glycoprotein (gp160) precursor from a singly-spliced mRNA on the rough endoplasmic reticulum (rER). Glycosylation of gp160 by *N*-linked and *O*-linked oligosaccharides occurs concomitantly during translation in the rER [63, 64]. In addition, the gp160 monomers oligomerise into trimers in the ER before it is trafficked to the Golgi network [65-67]. The oligosaccharide side chains acquire complex modifications as gp160 is trafficked through the Golgi network. The gp160 is cleaved by cellular furin or furin-like proteases [68] in the Golgi to yield gp120 and gp41 but remains non-covalently associated.

STRUCTURE

By crystallisation and X-ray analysis, in 1998 Kwong et al. reported on the atomic structure of gp120. However, it was a truncated form of gp120, with deletions in the N- and C-termini, Gly-Ala-Gly substitutions of the V1/V2 and V3 loop residues and over 90% deglycosylated to facilitate its crystallization. In addition, gp120 was in a ternary complex with the CD4 ectodomain and the antibody fragment (Fab) of 17b. This structure of gp120 in the ternary complex should represent a CD4-binding induced structure due to the interaction with the primary receptor and 17b, which interacts with an epitope made accessible by CD4-induced structural change of gp120 [69]. Extreme difficulties in reporting the structure, as described by the authors in preparation of stable complexes for structural studies, alluded to the structural flexibility of gp120 [44].

Subsequent from Kwong's structure in 1998, gp120 in complex with other ligands were achieved. These studies reported the atomic structure of a V3 loop containing gp120 core in complex with the ectodomain of CD4 and the Fab of the neutralizing X5 antibody [70], and gp120 core in complex with the Fab of b12, a CD4-binding site antibody [71]. These structural analyses of CD4-induced gp120 core showed the V3 loop to form an extended finger-like projection, while b12-bound gp120 probably

indicated a gp120 that was in structural transition. The atomic structure analysis gave us information on the location of crucial epitopes involved in host receptor recognition.

The gp120 directs recognition of target cells, while gp41 promotes fusion between viral and cellular membranes. Limited proteolysis of a recombinant gp41 ectodomain fragment resulted in a stable, soluble complex composed of two peptide fragments from the gp41 ectodomain. One fragment corresponded to the region adjacent to the fusion peptide (denoted the N-helix), while the other fragment corresponded to the region prior to the transmembrane segment (denoted the C-helix) [72]. X-ray analysis of crystals of the two-peptide complex showed a six-stranded helical bundle. The centre of the helical bundle was composed of a parallel arrangement of three N-helices, while three C-helices wrap the central, trimeric N-helices in an anti-parallel arrangement [55]. At the same time, X-ray analysis of crystals obtained from gp41 solubilised with GCN4 in place of the amino-terminal fusion peptide reported that the core of gp41 was an extended, triple-stranded α-helix [56]. Biochemical evidence suggests that these atomic structures were the fusogenic state of gp41. These structures were obtained in the absence of gp120 and are extremely stable against thermal denaturation. The binding of Env to receptors results in dissociation of gp120 from Env and increased exposure of gp41. Thus, gp41 produced in the absence of gp120 would be expected to adopt the fusion-active state. Moreover, peptides derived from either the N- or C-helices were able to inhibit infection, and point mutations in the gp41 residues that forms the core bundle abolished infectivity in HIV-1.

The X-ray crystallography analysis of the fusogenic state of gp41 and of gp120 in a ternary complex with the CD4 ectodomain and the Fab of 17b provided snapshots of individual subunits of Env during virus entry. It was not until the structure of a simian immunodeficiency virus (SIV) gp120 subunit in the native state was reported when comparisons made it possible to show the extensive structural reorganisation of the protein [73].

Since the trimeric Env structure could be important for vaccine development strategies, researchers pushed towards the study of the Env trimer. The use of cryo-electron tomography to analyse the Env spike on vitrified virus particles provided further details that were missing from the atomic structure of gp120 and gp41 made by X-ray analysis [74-76]. In their analysis, the Env spike reported by Zanetti et al. and Zhu et al. were reconstructed from cryo-electron tomograms of SIV particles. The reason for this was

that SIV had a greater number of Env spikes per virion compared to HIV-1. Furthermore, SIV is more stable than HIV-1 [77]. The authors proposed that SIV provides a good working model for HIV-1 due to the high sequence identity and similarity. Although both groups reported three-fold symmetry in the native Env spike, the reconstructed density map revealed vastly different structural features. Zanetti et al. reported a native Env spike that consisted of a stem capped by three globular densities that folded over the stem. As several orientations were possible, the authors fitted the atomic model of the native SIV gp120 into the density map based on reported biophysical and biochemical properties, and proposed two models for the Env spike in the unbound, native state. The two fits had common features in the exposed CD4 binding surface, outward projection of glycans and V1/V2 loops and positioning of V4 and V5 variable regions on the top surface of the Env complex. However, they varied in the position of the V3 loop. One model showed the V3 pointed towards the trimer interface, while in the other, V3 was exposed to solvent. While the models proposed by Zanetti et al. suggested some degree of flexibility, the reconstructed density map reported by Zhu et al. showed a distinctively compact head standing on tripod legs. Although fitting of gp120 into this map was possible, it resulted in many steric clashes at the three-fold as missing residues in the N- and C-termini of the gp120 atomic structure were not accounted for in the fitting.

Liu et al., on the other hand, reported three separate density maps of Env reconstructed from cryo-electron tomograms of HIV-1 virus particles in the native state, complexed with the Fab of b12 antibody, and complexed with the ectodomain of CD4 and Fab of 17b antibody [76]. The native Env b12-complexed Env were reported as a mushroom-shaped spike connected to the virus membrane by a narrow stalk. Glycosylated residues were found to be solvent-accessible and CD4 binding site was exposed. In comparison, the Env trimer bound to CD4 and 17b showed dramatic conformational changes in the gp120 and gp41 regions. According to Liu et al., each gp120 monomer rotates 45° clockwise along an axis parallel to the three-fold axis, and 15° out of plane. As a result, CD4-binding is seen to lead to an opening of the trimer that likely promotes gp41 exposure. The gp120 monomer rotation shifts the V1/V2 loops from its position at the trimer apex seen in the native and b12-bound Env with concomitant release of V3 loop from the lateral edge of the trimer apex to point directly at the target cell. However, whether the CD4-bound Env structure is affected by the presence of the Fab of 17b is

unclear as it was reported that 17b stabilises and locks the CD4-bound conformation of gp120 [78].

Based on the observations made by different groups on the structure of the gp120 and Env, several conformations of Env appear possible. Inherent flexibility of gp120 and gp41, coupled with the use of truncated forms of the proteins which are more stable in X-ray crystallography analysis, has hampered a complete understanding. Moreover, the results obtained from cryo-electron tomography analysis of Env on virus particles, despite carrying the full sequence of the proteins, are conflicting. This could arise from the presence of fundamental differences in the conformation of monomeric HIV-1 and SIV gp120, despite their sequence identity and similarity. Indeed, a first report to describe the atomic structure of the native, unliganded HIV-1 gp120 core structure shows this [79]. The unliganded gp120 core of the HIV-1 was shown to be more similar to the CD4-bound form than was previously reported for SIV gp120 core, with the additional finding that the bridging sheet (which contributes to part of the co-receptor binding site) is formed. It should, however, be reiterated that gp120 on the virus exists in complex with two other gp120 monomers and three gp41 units. The structure of proteins obtained in crystals is highly static and usually in a ‘preferred’ conformation, and this could account for the lack of neutralisation of unliganded HIV-1 despite the presence of the co-receptor binding site. A recent work on cryo-electron microscopy and single particle reconstruction analysis of detergent-solubilised Env trimers extracted from virus particles showed a cage-like morphology [80]. Whether the presence of detergent in the preparation affected the native quaternary structure remains unknown, as the minimal gp120-gp41 contacts reported in the structure conflicts with biochemical and mutagenesis data reporting close gp120-gp41 contacts.

HIV-1 TAT

ROLE IN VIRUS REPLICATION

Shortly after integration the proviral DNA genome is transcribed by host RNA polymerase (RNAP) II. The Tat protein, a regulatory protein encoded by the virus *tat* gene, is expressed early after integration and is constantly produced throughout all stages of viral replication. It is encoded by two exons and comprises 86-101 amino acids (exon 1: residues 1-72; exon 2: residues 73-86/101). The first exon consists of five domains which is sufficient for its function in trans-activating the long terminal repeat (LTR) promoter of HIV-1. In HIV-1 infected cells , Tat interacts with host positive transcription elongation factor b (P-TEFb).

P-TEFb is a heterodimer complex assembled with cyclin-dependent kinase 9 (CDK9) and cyclin-T1. In normal cellular activity, P-TEFb is required for generation of mRNA and it functions by regulating the elongation phase of RNAP II activity during transcription [81, 82]. During HIV-1 infection, Tat protein interacts with P-TEFb by inserting in a groove at the heterodimer interface. The interaction augments the interaction surface between CDK9 and cyclin-T1 [83], resulting in a more stable and more active P-TEFb [84]. One of the domains in the N-terminal of Tat, coded by the first exon, is an RNA-binding domain which interacts with the transactivation response (TAR) element found in the LTR. Hence, by interacting with the TAR, Tat recruits P-TEFb to the LTR where host RNAP II is paused. The CDK9 of P-TEFb hyper-phosphorylates the carboxy-terminal domain of RNAP II [85] and increases its processivity [81]. In this way, Tat enabled production of full-length viral transcripts [86] that are poly-adenylated. Full-length viral transcripts are needed for expression of structural proteins. The expression of structural proteins precedes progeny virion assembly. Virus particles with a non-functional *tat* gene are replication defective, producing short viral transcripts and are not able to assemble progeny virions [87, 88].

TAT AND PATHOGENESIS

Tat contains a basic region rich in lysine and arginine residues. This confers efficient membrane transduction property to the protein and allows Tat to readily cross cell membranes and easily be released into the cellular microenvironment into the

extracellular space. In fact, the transduction domain has been used to design the delivery of potential therapeutic agents such as proteins, peptides and nanoparticles into cells [89].

Circulating Tat can interact with receptors on cells and bring about a plethora of effects that are not related to its regulatory role in virus replication. Tat is able to bind and activate the vascular endothelial growth factor receptor 2 (VEGFR-2). The native ligand of VEGFR-2 is the vascular endothelial growth factor (VEGF). Under normal physiological and biochemical homeostasis in the body, VEGF ligand binding to VEGFR-2 is involved in vasculogenesis and angiogenesis, playing an important role in the early development of the vasculature. This was demonstrated in animals where targeted gene disruption of VEGF and VEGFR-2 in mice resulted in death before birth due to defects in the development of the cardiovascular system (reviewed in [90]). In HIV-1 infected patients, the interaction between Tat and VEGFR-2 activates the receptor [91]. This interaction probably potentiates the angiogenic stimuli provided by Kaposi's sarcoma herpes virus-encoded proteins [92], resulting in the development of Kaposi's sarcoma in HIV-1 positive patients.

Cytokine expression in immune cells in HIV-1 infected patients is often dysregulated (reviewed in [93]). The dysregulation is a result of a cascade of events that begin with interaction of HIV-1 Tat with receptors on the cell surface. For example, Tat protein upregulates expression of tumour necrosis factor-α (TNF-α), interleukin-6 (IL-6), IL-10, interferon-α (IFN-α), but downregulates IFN-γ. HIV-1 Tat dysregulation of TNF-α cytokine expression could lead to neurotoxicity, cell death and viral replication; however, IFN-α expression in primary human mononuclear cells induced by HIV-1 Tat could suppress the immune response by antigen-stimulated monocytic cells, leading to immunosuppressive effects in T cells.

Cytokine dysregulation by HIV-1 Tat could contribute to pathogenesis in HIV-1 infected patients by creating an environment that supports opportunistic infections. In addition, HIV-1 Tat could also interact with various co-infecting pathogens during disease progression in HIV-1 infected patients, leading to the development of HIV-associated diseases and severe microbial infections.

BACKGROUND OF THE THESIS

The structural characterisation of Env by established methods has been separated into gp120 and gp41. In particular, structures of gp120 observed by x-ray crystallography have involved monomeric recombinant proteins with truncations in the protein sequence. Neither the full gp41 protein structure has been available. Nor the tertiary conformational change in gp120 observed in x-ray crystallography has been definitively characterised in the context of a quaternary structure of the complex found on the virus particle. Recent developments in cryo-electron tomography allow biological material to be observed directly without heavy metal staining. Frozen hydrated samples can be visualised, thus avoiding artefacts notoriously found in chemical fixation and dehydration of samples for electron microscopy. In addition, it also circumvents problems in interpretation caused by the accumulation of staining material. However, despite the advantages, tomograms of the Env trimer so far have given contradictory results [74-76].

The use of cryo-electron microscopy imaging and structure determination of Env as a trimeric complex, as opposed to the imaging of Env purified from virions, can provide evidence of quaternary conformational shifts upon receptor or ligand binding. The higher resolution of structures achieved could hopefully serve to address the discrepancies between previous observations. To mimic Env on the virus surface and investigate its structure by cryo-electron microscopy imaging, the soluble trimeric form of Env was used [94, 95]. These Env trimers are also referred to as gp140 trimers, due to a partial V2 loop truncation, fusion of gp120 subunit to the ectodomain of gp41, and the absence of the transmembrane domain of gp41. These trimers display increased immunogenic properties and are suitable for use as models to study the quaternary conformational change in Env. Furthermore, immunochemical characterisation has shown that the topology of gp140 trimers is similar to native Env. (For ease of reference in this thesis, I would also refer to these gp140 trimers as Env.)

In order to elicit the CD4-induced conformation of Env, the CD4-mimetic miniprotein (CD4m) was used. The CD4m consists of the gp120-binding surface of CD4 transplanted into a scorpion toxin scaffold [96]. In addition, CD4m binds gp120 with nanomolar affinity [97] and induced a gp120 conformation that was nearly identical to the conformation induced by soluble CD4 (sCD4) [98]. Since CD4m has a small size (with only 27 amino acid residues), it is easier to focus on the quaternary changes in

Env during structural analysis and circumvent possible steric aggregation induced by sCD4-Env binding.

In vitro studies have shown that anti-Tat monoclonal antibodies inhibit HIV-1 replication in experimentally infected cell culture experiments [99, 100]. The inverse relationship between anti-Tat antibody level in HIV-1 infected patients and viral load magnitude [101] suggests that Tat protein and/or anti-Tat antibody contribute, in part, to influence disease progression *in vivo*. Hence, controlling extracellular Tat protein in HIV-1 infected individuals could slow down disease development.

To study how Tat protein possibly could influence virus infectivity, the Tat and Env protein complex (Tat-Env) was imaged by cryo-electron microscopy and compared to native and CD4m-bound Env structures. The effects of Tat protein on virus infectivity and spread, and a possible mechanism of action by Tat protein in influencing virus infectivity were also investigated *in vitro*.

AIMS OF THE THESIS

In general, the aim of this thesis was to investigate the various quaternary structures of the HIV-1 Env trimer, and to bring the knowledge learned from the structural investigations to subsequent functional, biological investigations. Specifically, this thesis studied the following:

- The quaternary structure of the native Env glycoprotein trimer presented on the virus surface
- Compared the quaternary structure of the CD4-induced state of the Env glycoprotein trimer with the native state
- The interactions between the Env glycoprotein and regulatory Tat protein and the resulting quaternary structure
- Effect of the regulatory Tat protein on HIV-1 infectivity rates and the underlying mechanism(s)

COMMENTS ON METHODOLOGY

Some of the methods used in this thesis will be discussed in this section.

CRYO-ELECTRON MICROSCOPY AND IMAGE PROCESSING (PAPERS I AND II)

Established methods of protein structure determination, such as x-ray crystallography and nuclear magnetic resonance spectroscopy, have been used extensively as a means to identify the mechanism of biological processes. However, limitations inherent to these methodologies place restrictions when attempting to study the macromolecular structures of proteins, which are often large complexes. For example, the most precise method of structure determination is by x-ray crystallography; however, as it requires crystals of the protein to be studied, it is a highly disadvantageous method when it comes to multi-component complexes, as often large complexes resist attempts at crystallisation. In addition, crystal packing favours particular modes of interaction and may interfere with the study of macromolecular complexes that exist in nature. Finally, because the examination of proteins in the crystal structure is a solid phase approach, the behaviour of the protein in solution cannot be investigated, and the study of proteins in motion is not possible. In nuclear magnetic resonance spectroscopy, protein structure in solution can be examined, but only information regarding the degree of probability of the protein segment in a given conformation can be achieved. Moreover, while a lot of data can be extracted from a system for accurate determination of a structure, determining higher molecular mass structures is extremely difficult.

The use of cryo-electron microscopy imaging of macromolecular protein structures was aimed at circumventing the problems faced in cryo-electron tomography, as well as, the limitations in nuclear magnetic resonance spectroscopy and x-ray crystallography, presenting an approach that is not limited by size and does not require crystals. In particular where analysis of Env is concerned (which has a molecular weight of approximately 324 kDa), the complex is free to assume functional states without any steric constraints. The Env was incubated overnight with CD4m at 4°C to obtain the CD4m-bound Env complex. To obtain Tat-bound Env complex, Env was incubated for 5 to 10 minutes with Tat protein at room temperature. To image proteins by cryo-electron microscopy, protein samples, native Env, CD4m-bound Env or Tat-bound Env, were applied to holey carbon grids prepared using 300 mesh grids base for 10 s,

blotted, then immediately plunge-frozen into liquid ethane cooled in a liquid nitrogen bath. Throughout image collection, the frozen protein samples were maintained at liquid nitrogen temperature of approximately -196°C.

Each micrograph collected on the CCD camera contains many images of the protein particles. The selection of these particles (also referred to as images) from the micrographs was done manually as well as semi-automatically. Manual selection involved visual inspection of the micrographs, and particles that represented strong trimeric top and oblique views were selected. Semi-automated selection was done using reference-based autoboxing, and projections generated from a starting reference density map were used. At the end of the particle image selection process, several thousand images were usually obtained. To process these images, the principal components of variation were identified using multivariate statistical analysis and correspondence analysis in the Spider software [102]. Hierarchical clustering was then done to group images that appeared similar together, and an average of each cluster was calculated. This average provided a characteristic view of the cluster by boosting the common signal by a factor of $n^{½}$ while suppressing random noise [103, 104]. This facilitated the exclusion of particles with low quality factor scores and clusters that represented trimeric top and oblique views were retained for further processing.

ASSAYS TO STUDY EFFECT OF TAT PROTEIN ON VIRUS INFECTIVITY (PAPER III)

There are several ways to identify HIV-1 infection *in vitro*. Peripheral blood mononuclear cells (PBMCs) are used in most infectivity studies as they are broadly susceptible to infection and its use in infectivity assays has been a gold standard for many years. However, the use of PBMCs is both time-consuming and labour-intensive, expensive, lacks precision and is difficult to validate due to variations in PBMCs isolated from different blood donors. Following infection by HIV-1, RT activity, p24 antigen production and virus RNA copies can be quantified.

To study the effect of Tat on HIV-1 infectivity, the TZM-bl cell line was used. The TZM-bl cell line has the reporter luciferase gene integrated and under the control of the HIV-1 LTR promoter. TZM-bl was chosen as it stably expresses CD4, CCR5 and CXCR4, receptor and co-receptors of HIV-1. Moreover, it is highly sensitive to

infection with HIV-1. TZM-bl can be maintained in culture and the effect of infection can be quickly and directly monitored by luciferase production, i.e. by luminescence, two days post-infection. Single-cycle infectivity by HIV-1 in TZM-bl cell line can also be easily achieved by the addition of the protease inhibitor indinavir [105].

The experiments were designed with the aim of studying the effect of Tat protein on HIV-1 infectivity. The incubation of Tat protein with HIV-1 is referred to as treatment and the incubation of HIV-1 (whether treated or not) with cells is referred to as infection. HIV-1 was treated with Tat protein prior to infecting target cells, TZM-bl cells, and any effect Tat protein would have on virus infectivity would be compared to target cells infected with untreated HIV-1. An important parameter was queried: Time. Did the length of infection by Tat protein treated HIV-1 affect virus infectivity, and is virus infectivity affected by the duration of HIV-1 treatment by Tat protein? In addition, to avoid possible effects of Tat protein on TZM-bl which may influence the results, control experiments were performed. In these control experiments, only Tat protein was used to treat the cells. Luciferase production by TZM-bl cells treated with Tat protein was analysed. The expression of co-receptors CCR5 and CXCR4 in TZM-bl cells treated with Tat protein was analysed by flow cytometry.

The effect of Tat protein on HIV-1 spread was studied using a cell-cell fusion model. The ACH-2 cell line, which carries a single integrated copy of HIV-1, was used as a model of chronic infection. The ACH-2 cell line constitutively produces low amount of virus [106], which mimics the physiological condition in an HIV-1 infected individual during the latent phase of infection. The ACH-2 cell line was stained with carboxyfluorescein diacetate succinimidyl ester (CFSE). CFSE passively diffuses into cells; reactions between the CFSE and cellular proteins cause CFSE to stain the ACH-2 cells fluorescent green. The target of virus spread was the SupT1 cell line, a non-infected T cell line which is sensitive to HIV-1 infection and syncytium-induction [107]. Fusion between ACH-2 and SupT1 cell lines occur when they are co-cultured and fusion leads to multinucleated cells called syncytia. As ACH-2 cells are stained fluorescent green by CFSE, syncytia (the fused cells) can be identified by counter-staining the SupT1 cells. This was done by labelling the CD4 receptors on SupT1 with an anti-hCD4 antibody conjugated with allophycocyanin (APC). The fused cells can be monitored and quantified by a flow cytometer, which would identify syncytia as having both fluorescent CFSE and APC dyes (Figure 5). As the SupT1 cell line is sensitive to infection and syncytium-induction by HIV-1 particles released from the ACH-2 cells,

they were grown with indinavir before and during co-culture with SupT1 cells, thus preventing release of infectious HIV-1.

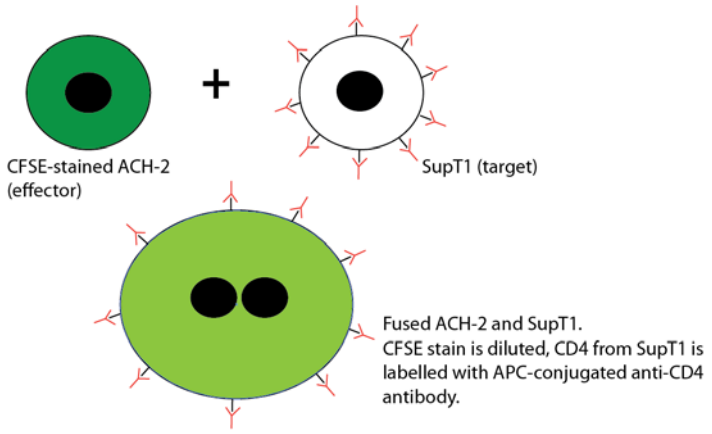


Figure 5. Schematic of cell-cell fusion assay.

Similar to infectivity assays conducted in the TZM-bl cell line, control experiments were performed in SupT1 cells to ensure that the Tat protein did not have effects on SupT1 cells which would influence the results. Co-receptor expression CXCR4 and CCR5 on SupT1 treated with Tat protein alone were analysed by flow cytometry.

RESULTS

QUATERNARY STRUCTURE OF NATIVE AND CD4M-BOUND ENV (PAPER I)

Atomic structures obtained from crystals of unliganded simian immunodeficiency virus (SIV) gp120 [73] and CD4-bound HIV-1 gp120 [44] have revealed a large conformational rearrangement in gp120. The significance of the conformational rearrangement, however, has been unclear in the quaternary context as the structure of Env could not be resolved by x-ray crystallography. Several groups have reported the quaternary structure of Env using cryo-electron tomography, but observations made were contradictory and, hence, inconclusive. In this study, cryo-electron microscopy and structure determination of a soluble trimeric complex of gp140 (referred to as Env) provided evidence of quaternary conformational shifts upon CD4-binding. To elicit the CD4-bound conformation, the CD4-miniprotein (CD4m) was used.

Following image processing of particles selected from the micrographs, reconstruction of retained images was performed in EMAN [108]. Comparative analysis of the density maps of native Env and CD4m-bound Env (CD4m-Env) suggested three events to occur upon receptor interaction: Outward density shift away from the threefold axis, subunit tilt away from *z*-axis that flattens the trimer, and gp120 subunit rotation about subunit axes normal to the threefold axis (Figure 6).

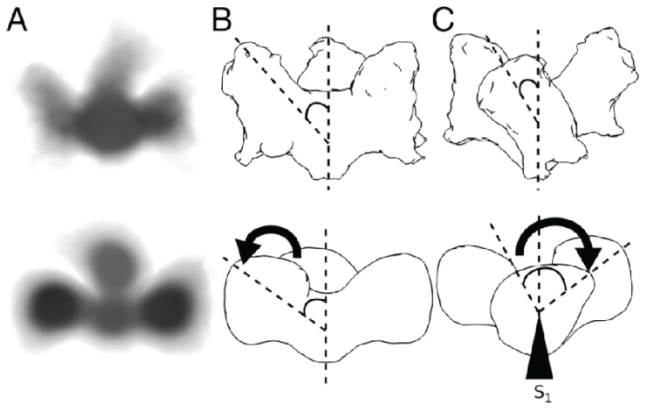


Figure 6. Summary of quaternary changes induced by CD4 binding. Top panel: unliganded gp140; bottom panel: CD4m-bound gp140. (A) Density shift away from threefold axis. (B) Tilt away from *z*-axis flattens the trimer. (C) Subunit gp120 rotation along designated subunit axis S_1.

a) *Outward density shift*

Isosurface density analysis of the native Env and CD4m-Env at 3σ above the mean contour value shows that the gp120 subunit connections in native Env remained contiguous with the threefold axis. This was in contrast with the CD4m-Env where density of the subunit arms appeared more bulbous at the

distal tip. Segmentation of both density maps revealed a weakened density at the threefold axis in CD4m-Env. The distance between dense regions in the subunit arms were measured to determine the extent of the density shift. The distance between threefold axis to the centre of the dense region in native and CD4m-Env was 25 Å and 40 Å, respectively. In addition, the distance between the dense regions of adjacent subunit arms in native and CD4m-Env was 58 Å and 72 Å, respectively (Figure 7).

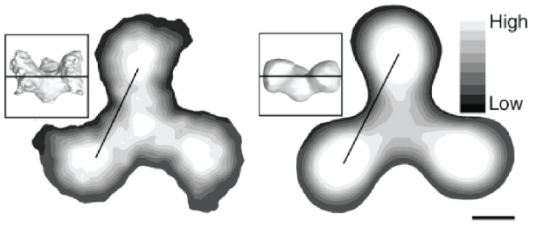


Figure 7. Segmentation of density maps shows weakend density at the threefold axis of CD4m-Env (right panel) compared to native Env (left panel). Distance between gravity centres (dense region in subunit arms) is 58 Å and 72 Å in native Env and CD4m-bound Env, respectively. Inset shows the plane at which density maps were segmented. Scale bar: 25 Å.

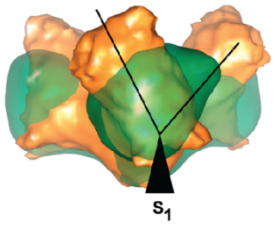


Figure 8. Superimposed native Env (orange) and CD4m-Env (green). The S_1 axis refers to the subunit axes and lies perpendicular to the threefold axis. Viewing through the S_1 axis, the gp120 subunit is rotated clockwise by approximately 60°.

b) *Subunit gp120 tilts away from z-axis due to rotation about its axes normal to threefold axis*

Rotation of the gp120 subunit in the CD4m-bound state leads to a subunit tilt away from the threefold *z*-axis (Figure 6B). The tilt is approximately 25° from the *z*-axis, and possibly exposes the fusion peptide on gp41. Moreover, its juxtaposition with the host membrane facilitates fusion peptide insertion into the host.

c) *Shift of epitopes due to gp120 rotation about its axis normal to threefold axis*

Docking the atomic coordinates of SIV gp120 and CD4-bound gp120 into the respective density maps led to another observation that was a result of the gp120 subunit rotations. Superimposing the native and CD4m-bound density map and viewing through the trimer axes shows that the gp120 subunit is rotated clockwise by approximately 60° (Figure 8). A closer study of the crucial epitopes in the density map docked with atomic coordinates of gp120 revealed several observations (Figure 9). The location of the CD4 binding site on the outermost trimer apex allows gp120 to approach CD4 at an angle that is parallel to the *z*-axis. The V3 loop, a determinant of co-receptor interaction, is located on the outer edge of each trimer arm in the native Env trimer.

Interaction with CD4m causes subunit rotation in gp120 (mentioned in *point b*). As a result, the CD4 binding site on gp120 now faces the adjacent counter-clockwise gp120 subunit. In addition, the V3 loop is also rotated to a position that is maximally exposed on the trimer surface. The co-receptor binding site, comprising the bridging sheet and the V3 loop, are located proximal to the three-fold axis and more accessible once gp120 in the CD4m-bound state rotates. The V2 loop in the native Env trimer is located on the leading edge of each trimer arm, while in the CD4m-bound Env it is pointing towards the adjacent counter-clockwise gp120 subunit. Its location in the docked density maps suggests a vertical translocation and rotation following CD4 interaction.

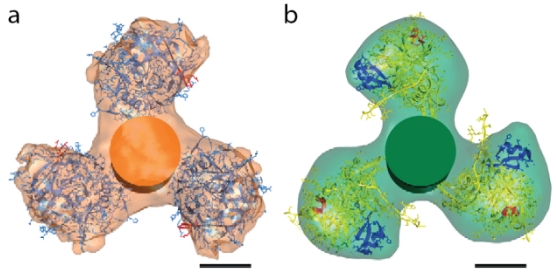


Figure 9. Docking of gp120 coordinates into density maps. (a) Native Env docked with unliganded SIV gp120 coordinates (PDB: 2BF1, blue). An orange cylinder is used to represent gp41. (b) CD4m-Env docked with CD4-bound HIV-1 gp120 coordinates (PDB: 1GC1, yellow) and CD4m coordinates (PDB: 2I5Y, blue). A green cylinder is used to represent gp41. V3 loop position is coloured red. Scale bar = 25 Å.

The major glycosylation sites of gp120 were situated on the trimer surface. This renders them fully exposed to the solvent and serves an important function in masking Env from the host immune response. Comparing the trimer arms of the maps also revealed that the native Env trimer contains a centre of high density which corresponds to the location of the inner domain of gp120 when docked with the atomic coordinates of unliganded SIV gp120. In the CD4m-Env, there is no perceptible difference in the density: When docked with CD4-bound gp120 atomic coordinates, it suggests closer association between the inner and outer domains of gp120.

QUATERNARY STRUCTURE OF TAT PROTEIN-BOUND ENV (PAPER II)

Evidence that Tat protein interacts with Env has previously been shown through cell biology and biochemical studies [109]. The presence of anti-Tat antibody *in vitro* suppresses HIV-1 replication, while *in vivo* HIV-1 infected individuals show slow disease progression [100, 101, 110]. These observations led to the hypothesis of how Tat protein may influence virus infectivity. With the previous experience of studying the quaternary structure of native Env and CD4m-Env by cryo-electron microscopy, this method was extended to study the structure of Tat protein-bound Env (Tat-Env).

Interesting observations were made when the density map of the Tat-Env complex was compared with the native and CD4m-Env density maps from Paper I. Firstly, the Tat-Env density map resembled the native Env density map with the addition of a bulge present in the trimer arms. This bulge was present in the location of the V3 loop that was identified in the native Env trimer. Isosurface density comparison of the density maps at 3σ above the mean contour value indicated that the connections of the subunit arms remained contiguous with the threefold axis and the subunit arms are tilted away from the threefold axis. The distance between the threefold axis and the dense domain is 32 Å, while the distance between the dense domains of adjacent subunits arms is 62 Å. These distances measured were intermediate when compared with the corresponding distances in native Env and CD4m-Env (Figure 10).

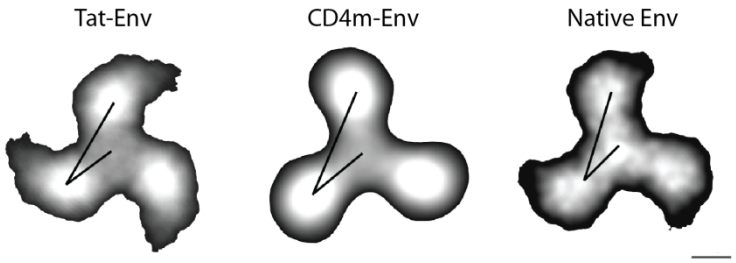


Figure 10. Comparison of the distance measured in Tat-Env (left panel), CD4m-Env (middle panel), and native Env (right panel). Scale bar: 25 Å.

Further observations were made when the Tat-Env density map was docked with a structural model of gp120 in complex with Tat (Figure 11). The location of CD4 binding site and V3 loop in the docked Tat-Env density map were compared with those in the docked native and CD4m-bound Env density maps. Although the CD4 binding site residues in both the native Env and Tat-Env density maps are oriented towards the trimer surface, distance between these residues on adjacent gp120 arms were 60 Å and 75 Å, respectively. This was in contrast to the orientation of the CD4 binding site in the

CD4m-bound Env, where the trimer arm is rotated from its native position to face the adjacent counter-clockwise gp120 subunit. While the CD4 binding site in Tat protein-bound Env resembled that in the native Env trimer, the V3 loop, on the other hand, was oriented on the trimer surface, a position that is similar to that found in the CD4m-bound Env trimer. In addition, the Tat-gp120 structural model showed close interaction between Tat protein and the V3 loop, an interaction that was also reported recently [111].

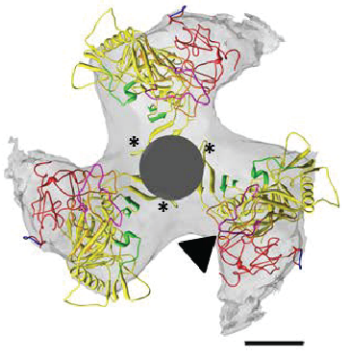


Figure 11. Tat-Env density map docked with Tat-gp120 structural model. Position of the V1/V2 stem is marked by an asterisk (*), and position of V3 loop is marked by an arrowhead. gp120: yellow. Tat: red. Integrin recognition sequence: blue. Scale bar = 25 Å.

Surface plasmon resonance analysis was performed to determine the optimal pH for Tat binding to an anti-Tat protein antibody, 2A4.1. It showed that Tat protein binding to the antibody was optimal at an acidic pH 5.7 compared to at physiological pH. As Tat protein is able to bind several receptors at the plasma membrane of T cells which are endocytosed into endosomes, this observation suggested that Tat protein is more stable and active in a low pH environment. Tat protein internalised into the low pH endosomes could assist in virus entry by introducing the conformational change to the intermediate structure observed. Additionally, that Tat protein is localised to the V3 loop of gp120 could also have a potential effect on the co-receptor binding site gp120.

TAT PROTEIN INCREASES HIV-1 INFECTIVITY AND SPREAD (PAPER III)

By bringing the structural observations made in Paper II to a functional, biological study, the effect of Tat protein on HIV-1 infectivity and spread was studied using cell infectivity and cell-cell fusion assays, respectively. Treating HIV-1 with low to moderate levels of exogenous Tat protein (between 1 nM and 100 nM) prior to infecting TZM-bl cells led to increased infectivity. The effect could be seen whether the duration of infection by treated virus was 2 h or 24 h (Figure 12a). Moreover, the increased HIV-1 infectivity was only seen when HIV-1, and not target cells, were treated with Tat protein prior to infection (Figure 12b).

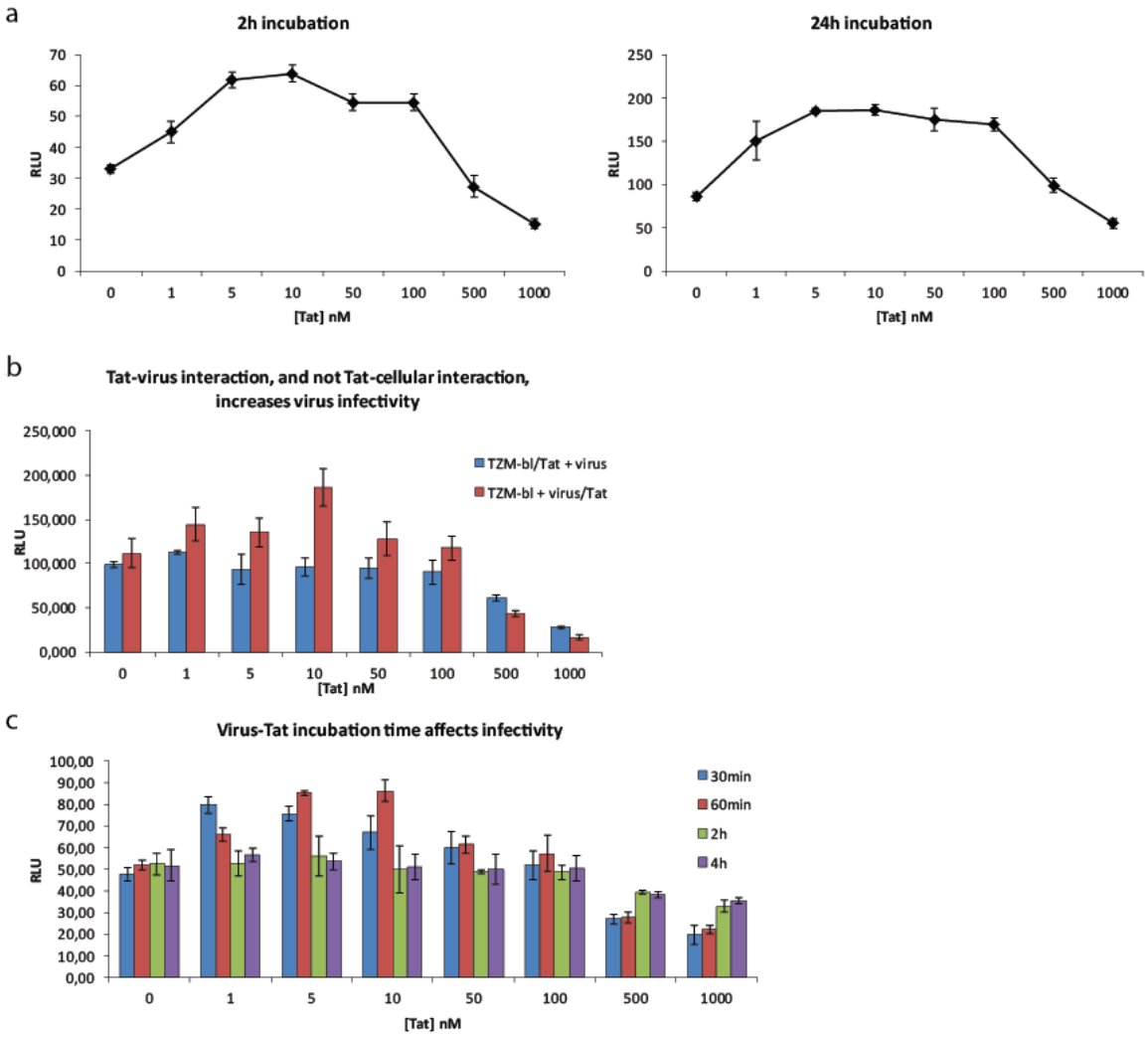


Figure 12. Single-cycle HIV-1 infectivity assays in TZM-bl cells. (a) Treating HIV-1 with Tat protein increased HIV-1 infectivity and the effect could be seen when infection was allowed to proceed for either 2 h (left panel) or 24 h (right panel). (b) The increased HIV-1 infectivity is observed when virus was treated. (c) A shorter duration virus with Tat protein increased HIV-1 infectivity as opposed to longer duration of treatment with did not show any effect. RLU = relative luminiscence units.

The duration of HIV-1 treatment by Tat protein was also a factor that affected virus infectivity: Treatment duration longer than 2 h had no effect on virus infectivity (Figure

12c). An interesting observation was the reduced infectivity when HIV-1 was treated with excessive amounts of Tat protein (500 nM or 1000 nM).

The spread of HIV-1, observed from cell-cell fusion assays, was also affected by Tat protein treatment of ACH-2 cells prior to co-culture with SupT1 cells. Cell-cell fusion was significantly increased when ACH-2 cells were treated with Tat protein above 50 nM (Figure 13a). The receptor of HIV-1 entry on SupT1 cells, CD4, can be blocked by an anti-CD4 antibody prior to co-culture with Tat protein treated ACH-2, and resulted in that the cell-cell fusion was inhibited (Figure 13b).

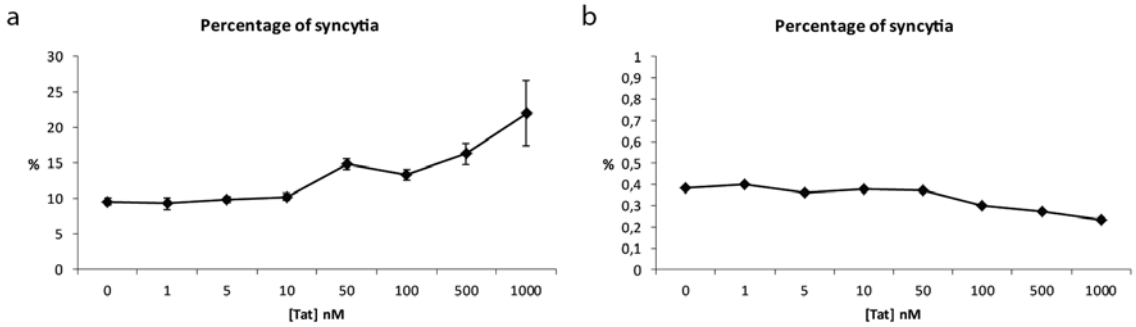


Figure 13. (a) Cell-cell fusion between ACH-2 and SupT1 in co-culture in the presence of Tat protein at respective concentrations. (b) Cell-cell fusion was inhibited when an antibody against the HIV-1 receptor, CD4, on SupT1 cells was added.

HIV-1 infection was thus dependent on CD4 receptor regardless of Tat protein treatment. Then, I looked at whether blocking either of the co-receptors on TZM-bl cells affected infection by HIV-1 previously treated with the Tat protein. When the CXCR4 co-receptor on TZM-bl cells was blocked with anti-CXCR4 antibody, infectivity of untreated HIV-1 was decreased (Figure 14); however, treating the virus with Tat protein up to 100 nM increased its infectivity by 20 – 25% (Figure 15). This change in infectivity could be observed both when the CXCR4 co-receptor inhibited cells were infected with treated virus for either 2 h or 24 h.

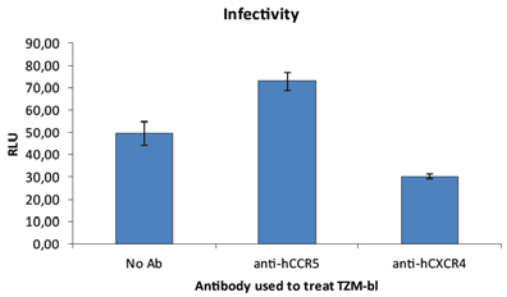


Figure 14. HIV-1 infectivity in permissive host (TZM-bl cells) when respective co-receptors were inhibited by antibodies.

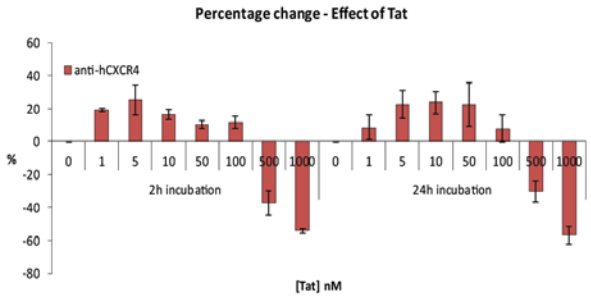


Figure 15. Infectivity assay of Tat protein-treated HIV-1 in CXCR4-inhibited TZM-bl cells. The infectivity was normalised to infection in CXCR4-inhibited TZM-bl cells by untreated HIV-1. Values represent percentage change in infectivity.

On the other hand, when the CCR5 co-receptor on TZM-bl cells was blocked with anti-CCR5 antibody, infectivity of untreated HIV-1 was increased (Figure 14). When comparing infection of these CCR5 co-receptor inhibited cells with control cells where co-receptor was not inhibited, the level of infectivity by Tat-treated virus reached similar levels in both cells (Figure 16).

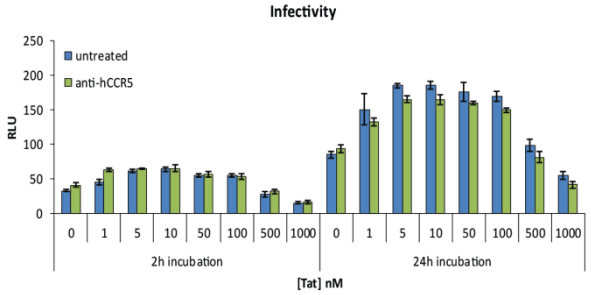


Figure 16. Infectivity assay of Tat protein-treated HIV-1 in CCR5-inhibited TZM-bl cells. Values represent luciferase activity measured in relative luminiscence units (RLU).

Control experiments were performed to ascertain that Tat protein did not induce other effects which might influence the results obtained. Firstly, the concentration of exogenously added Tat protein that would influence the LTR promoter controlled luciferase production in TZM-bl cells was shown to be above 5 µM (Figure 17).

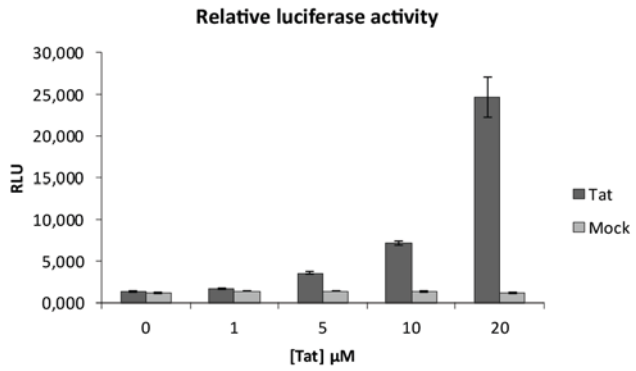


Figure 17. Relative luciferase production in TZM-bl cells treated with Tat protein at 0 µM, 1 µM, 5 µM, 10 µM and 20 µM. Signal is measured as relative luminiscence units, RLU.

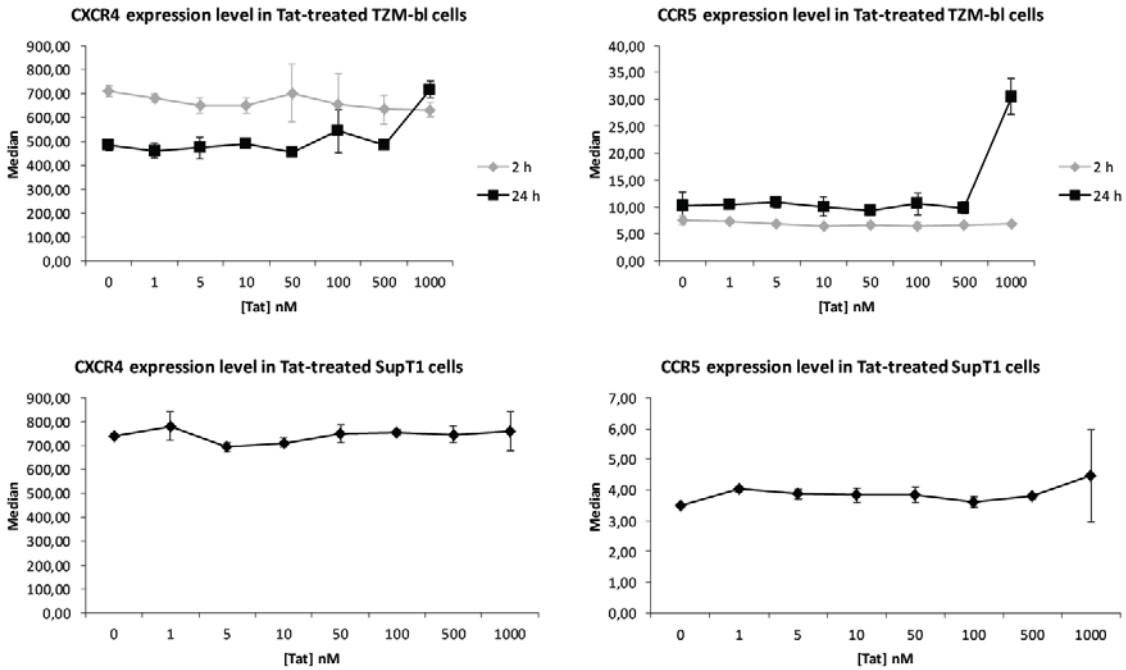


Figure 18. Expression level of co-receptors, CXCR4 (left panel) and CCR5 (right panel), in TZM-bl cells (top panel) and SupT1 cells (bottom panel) cultured in the presence of Tat protein for 2 h or 24 h in TZM-bl cells, and cultured in the presence of Tat protein for 24 h in SupT1 cells.

Secondly, a series of co-receptor expression analysis on TZM-bl and SupT1 cells (the target cells of infection or virus spread, respectively) were performed after they were cultured in the presence of Tat protein. The reason for this was because it was previously reported that Tat protein could up-regulate the cell surface expression of co-receptors, leading to increased virus infection [112]. As these target cells were exposed to exogenous Tat protein for the duration of virus infection (in the case of TZM-bl

cells) or co-culture (in the case of SupT1 cells), the cells were treated with Tat protein for the same duration. Except for the increased CXCR4 and CCR5 cell surface expression in TZM-bl cells exposed to 1000 nM Tat protein for 24 h, co-receptor expression level was constant under all other conditions (Figure 18). Despite the increased expression level of co-receptors in TZM-bl after 24 h exposure to Tat protein, it was not correlated with increased virus infectivity.

Thirdly, because Tat protein interacts with cellular transcription factors and the LTR promoter to enhance the rate of virus genome transcription and the production of full length virus transcripts, virus production in the presence of exogenous Tat protein was monitored and compared to virus production in the absence of exogenous Tat protein. HIV-1 production by ACH-2 cells, measured by p24 amount in the culture media, was relatively constant (Figure 19).

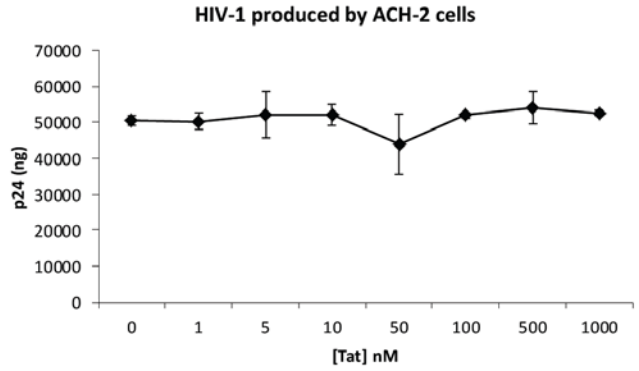


Figure 19. Virus production from ACH-2 cells cultured in the presence of Tat protein, measured by p24 content in nanogrammes (ng).

TAT PROTEIN AT HIGH CONCENTRATION CAUSES VIRUS AGGREGATION (PAPER III)

In single cycle infectivity assays of HIV-1 infectivity, a recurring observation was the reduced infectivity when high concentrations of Tat protein (500 nM and 1000 nM) were used to treat virus before infection (Figure 12). Hence, the effect of excessive amounts of Tat protein on Env and HIV-1 was investigated and the observations documented by transmission electron microscopy imaging.

At low concentration of Tat protein (100 nM), negatively stained Tat-Env complexes exhibited relatively monodisperse distribution when imaged by transmission electron microscopy (Figure 20a). HIV-1 virus-like particles demonstrate similar distribution at 100 nM of Tat protein (Figure 20d). At 500 nM Tat protein, several aggregates of Tat-Env (Figure 20b) and virus-like particles (Figure 20e) could be observed, and the extent of aggregation increased further at 1000 nM Tat protein for both Tat-Env (Figure 20c) and virus-like particles (Figure 20f).

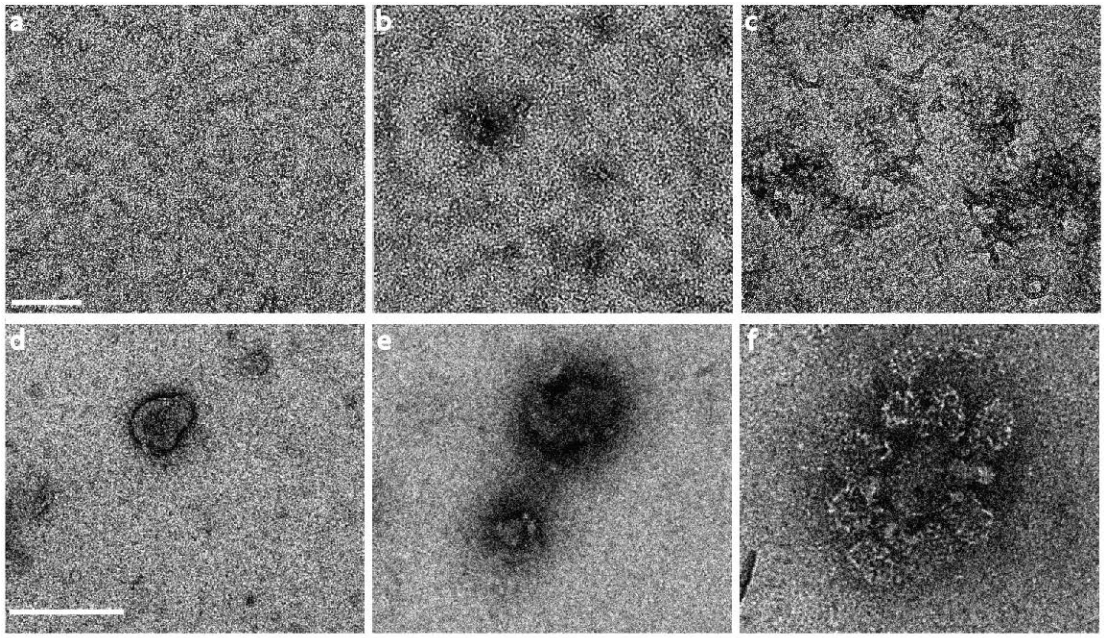


Figure 20. Transmission electron microscopy images of negatively stained Tat-Env complexes at (a) 100 nM, (b) 500 nM and (c) 1000 nM Tat protein, and of negatively stained HIV-1 virus-like particles at (d) 100 nM, (e) 500 nM and (f) 1000 nM Tat protein. Scale bar in (a), (b) and (c): 100 nm; scale bar in (d), (e) and (f): 250 nm.

DISCUSSION

The conformational change in the tertiary structure of gp120 following interaction with the CD4 receptor has been a central theme in understanding HIV-1 entry, as well as, the design of entry inhibitors. A comparison of the atomic structure of native SIV gp120 and CD4-bound HIV-1 gp120 obtained from x-ray crystallography has shown substantial rearrangement in gp120. Investigating the pre- and post-CD4 interacting states of Env trimer, and incorporating previous knowledge from gp120 atomic structures, provide the opportunity to assess the impact of tertiary quaternary changes in a single gp120 subunit on the overall quaternary arrangement since each Env trimer has three gp120 subunits. Because native and CD4-bound Env tertiary and quaternary structures are conformational states that the Env achieves during virus entry, they serve as structural models or "standards" of classical HIV-1 entry. Subsequent comparison of ligand-bound Env trimer with the "standards" would aid in determining possible deviations in the mechanism of virus entry in the presence of these ligands. Using this rationale, the quaternary structure of the Tat protein-bound Env trimer was compared with native and CD4m-bound Env trimers to determine how Tat protein could contribute to increased infectivity and spread during HIV-1 infection.

An initial comparison between native Env and CD4m-Env density maps have helped in the understanding of how tertiary structural changes of gp120 in HIV-1 Env translates to the quaternary structural change during virus entry. The sequence of events that occur upon CD4 interaction by HIV-1 gp120 could be summarised as such:

i. Upon CD4 interaction, structural rearrangement in individual gp120 subunits lead to rearrangement of inner and outer domains that was observed from x-ray crystallography structures of native and CD4-bound gp120. On the density map this is observed as a change in density distribution in the trimer arms where a closer association between the inner and outer domains is observed.

ii. These changes could lead to both the 60° rotation of the gp120 subunit along its axis and at the same time diminish the gp120/gp41 contact at the threefold axis, as evidenced by the weakened density at the threefold axis in the density map.

iii. Rotation, along with gp120 inner/outer domain rearrangements, lead to CD4 binding site epitopes facing the counter-clockwise subunit arm, as well as,

subsequent exposure of V3 loop and bridging sheet, which forms after CD4 interaction, to the perpendicular surface for interaction with co-receptors on the host.

iv. Density shift weakens the threefold axis and this leads to exposure of fusion peptide on gp41 at the trimer core. Weakening of threefold axis could also signify shedding of gp120 from the virus that has been reported during virus entry after CD4 interaction [43, 113, 114].

The docked gp120 coordinates in the native Env trimer is in agreement with the model proposed by Chen et al. with respect to its orientation and the angle at which CD4 approaches [73]. This corroborates with the previous work which established that all three CD4 binding sites are exposed [115]. Moreover, the orientation of gp120 docked in the CD4m-bound Env also agrees with the prediction by Kwong et al. on trimeric arrangement of gp120 [44, 116]. The rotation of gp120 subunit would ensure that CD4 does not sterically occlude the co-receptor binding site and this maximises the accessibility to the bridging sheet and V3 loop. The native and CD4m-bound Env density maps corroborates with the native and liganded Env reported by Liu et al. with respect to the subunit rotations observed. However, they differ in the reporting of a cap structure by Liu et al., which is absent in the structures here. The cap structure at the three-fold axis was also seen in the structures reported by Zanetti et al., Zhu et al., and Wu et al.. Could the difference in the structures reported in this thesis and with previous work be due to the source of material used in structure determination? Previous Env structures were determined through cryo-electron tomography of vitrified virus particles, or cryo-electron microscopy of Env extracted from HIV-1 by detergent solubilisation. The Env used for structure determination in this thesis were obtained from mammalian cell culture transfected to express soluble Env lacking the gp41 transmembrane domain, and subsequently purified. Truncations to the gp120 sequence were made to stabilise the Env trimer produced and could introduce differences in the structures obtained from earlier studies.

The observations of the native Env and CD4m-bound Env provide a basis for comparison with the Tat-Env quaternary structure, as well as, the Tat-gp120 atomic model. The conformation of Tat-Env density map was found to be distinct from that of the native Env. Comparison between the measurements taken in the Tat-Env density map shows that the distance between adjacent gp120 subunits and the distance between

gp120 and the three-fold axis is intermediate between the native and CD4m-bound Env trimers. This suggests that interaction with Tat protein elicits a conformation change that is intermediate between the native and CD4m-bound Env structures, while retaining exposure of CD4 binding site. Fusion between HIV-1 and the target host cell is a complex process that requires three to six Env trimers [117, 118]. The intermediate quaternary structure induced by Tat protein could effectively reduce steric clash between adjacent gp120-CD4 interactions. In addition, thermodynamic studies have shown that the interaction between gp120 and CD4 releases large bonding energy that is exploited to drive substantial rearrangement of the gp120 structure [119]. Considering the distances measured, as well as, the accessibility of the CD4 binding site, the Tat protein likely elicits a gp120/Env conformation that is optimal for CD4 interaction. The large bonding energy released following CD4 interaction could then be redirected to downstream conformational changes in gp120 or gp41 to aid HIV-1 entry. While the intermediate distances measured suggest that Tat protein interaction induces a conformational change in Env that is in structural transition, it could also explain for the induction of plasma IgG that recognised more epitopes in Env with higher reactivity for epitopes within the V3 loop when Tat and Env were used in a combination vaccine [120]. The increased distance between subunits would essentially appear to open up the trimer and expose more epitopes to the host immune response.

The observations made in the Tat protein induced Env structure also raised questions which were investigated in a subsequent biological study where HIV-1 was treated with exogenous Tat protein prior to infection. I showed that Tat protein increased HIV-1 infectivity in *in vitro* studies and it also likely promoted spread through the formation of cell-cell contacts. Interaction with the CD4 receptor was found to be necessary, indirectly supporting the observation that Tat protein elicits an Env conformation optimal for CD4 interaction.

Another question was also raised: Besides eliciting an intermediate quaternary Env conformation that is optimal for CD4 interaction, what other effects result from interaction with Tat proteins? The Tat protein has been shown to interact with the V3 loop of gp120 [111]. The V3 loop is important for co-receptor interaction. Sequence and charge on the V3 loop influences co-receptor tropism [121]. Because Tat protein is highly charged and natively unfolded [122], the way HIV-1 co-receptor tropism was affected in the presence of Tat protein during virus infection became a highly interesting topic to investigate. Blocking the CXCR4 co-receptor reduced infection by

an X4-tropic HIV-1 strain which was in agreement with published findings [20]. The X4 virus used here is HIV-1_{LAI}. However, treating the virus with Tat protein recovered infectivity by 20 – 25%. While trying to understand the results I looked at the infectivity of HIV-1_{LAI} in TZM-bl cells when either of the co-receptors was blocked separately and made an interesting observation: When the CCR5 co-receptor was blocked, infection by X4-tropic HIV-1_{LAI} was almost 50% higher than when neither co-receptor was blocked. This observation suggested that during natural infection, HIV-1, regardless of tropism, could have interacted with either of the co-receptors but in order for the eventual fusion event to occur, HIV-1 co-receptor interaction site and the co-receptor that it is interacting with somehow must match, else entry is arrested. Hence, virus that could have been arrested through interaction with CCR5 became "re-routed" to interact with only CXCR4 co-receptors, resulting in the increased infectivity. However, Tat protein interacting at the V3 loop likely introduces changes to the V3 loop, allowing an X4-tropic virus to use CCR5 as a co-receptor for entry.

In order to ensure that increased infectivity was not due to Tat protein activating the LTR promoter in the target reporter TZM-bl cells, the concentrations of Tat protein used for infectivity studies were assayed. In the absence of HIV-1 infection, Tat protein efficiently increased luciferase expression in TZM-bl cells but only at concentrations above 5 μM (Figure 17). In addition, that HIV-1 production in chronically infected ACH-2 cell line in the presence of exogenous Tat protein was not affected also showed that the concentration used was not capable of stimulating the LTR promoter.

The results of control experiments performed demonstrated that exposure of target cells to Tat protein during infection was not the reason for increasing HIV-1 infectivity or spread. Rather, it was the effect of Tat protein on the virion, and specifically on Env, that caused the increase. Yet, HIV-1 treated with high concentration of Tat protein, 500 nM or more, had reduced infectivity. Since Tat protein neither reduces nor increases the replication of an enveloped DNA virus, the HSV-1, the reduced HIV-1 infectivity is unlikely due to Tat protein induced virus aggregation through insertion of the protein's transduction domain into virus envelope membrane previously described [89], and that the increased infectivity observed is specific to HIV-1.

However, the observations from transmission electron microscopy images of Tat-Env complexes and HIV-1 virus-like particles showed that indeed aggregation occurred at high concentration of Tat. Hence, it is hypothesized that the Tat protein might cause

HIV-1 aggregation through indiscriminate interaction with Env at high concentrations. Tat protein carries a cysteine-rich domain that gives rise to Tat dimers [123, 124]. In addition, it is natively unfolded [122]. A quick calculation using HADDOCK [125] shows that multiple Tat-gp120 interactions were possible. The lowest energy, optimal binding interaction is shown in the docking of Tat-Env density map. Therefore, at low Tat protein concentrations, Tat protein possibly interacts with gp120 in its lowest energy, i.e. stable conformation to form Tat-Env complexes that promotes infectivity. When these low energy, stable conformations have been occupied, excess Tat protein could begin to interact indiscriminately. Virus aggregation could occur through dimerization of the Tat proteins that bind Env indiscriminately, with the effect that infectivity is drastically reduced.

CONCLUSION REMARKS AND FUTURE PERSPECTIVES

The structural resolution of the HIV-1 Env trimer via cryo-electron microscopy enabled us to understand the tertiary structural changes in gp120 in a quaternary context, particularly since the Env trimer is a complex macromolecule. We observe that during HIV-1 entry, interaction between gp120 subunit on Env and CD4 “opens” up to reveal the fusion peptide on gp41. From previous biochemical studies, we learn that gp120 shedding events occur after CD4 interaction [43, 113, 114], and this could be attributed to the weakening interaction between gp120 and gp41 observed from our results.

The studies with Tat protein interaction with Env also showed the possibility of ligands, other than CD4, influencing Env during virus entry. That Tat protein elicits an intermediate Env structure which contributes positively to virus entry warrants further attempts to understand the Tat-gp120 interaction to be made, e.g. through x-ray crystallography to add to the current repertoire of gp120 atomic structures. Improving this repertoire might aid the rationale design of vaccines or therapies against HIV-1 that target virus entry. On this note, it is possible that HIV-1 infected patients presenting decreased progression towards late stage of infection could have antibodies against unique epitopes in Env, gp120, or Tat-gp120 elicited by Tat protein, in addition to the high levels of anti-Tat antibodies found, but which could not be detected by standard screening assays.

Infectivity studies and cell-cell fusion assays have confirmed that Tat protein increases HIV-1 infection and spread. Because Tat protein interacts with the V3 loop on gp120, the hypothesis is that Tat protein could affect co-receptor recognition and thus influence infectivity was investigated. Our studies showed that Tat protein could increase an X4-tropic HIV-1_{LAI} infectivity in CXCR4-blocked TZM-bl cells. It would be prudent to investigate if the reverse occurs for a CCR5-tropic HIV-1 strain: Whether a CCR5-tropic HIV-1 infects better in CXCR4-blocked TZM-bl cells, and whether infectivity of this virus in CCR5-blocked TZM-bl cells improves in after treatment by exogenous Tat protein. The reduced HIV-1 infectivity in the presence of high concentrations of Tat protein suggests a negative influence for virus infection and could be a reason for low concentration of circulating Tat protein (between 0.01 and 0.1 nM) detected in the sera of HIV-1 infected individuals [126]. In this way, the virus could

influence its propagation and spread in an infected person, and possibly provides another avenue of regulative function to control viral replication.

ACKNOWLEDGEMENTS

I am immensely grateful to everyone who has supported me and my work during these years, especially to:

Professor R Holland Cheng, my supervisor, for giving me the opportunity to work in his lab and on the Env project. Thank you for introducing me to the study of protein structure in quaternary contexts, cryo-electron microscopy imaging and reconstruction.

Associate Professor Prasanna Kolatkar, my supervisor, for accepting me in your lab in this A*STAR-KI research partnership. I know it had not been easy for you but I appreciate the advice and support you have given me.

Professor Anders Vahlne, my supervisor, for your scientific guidance, expertise and knowledge. I am deeply grateful to you for the scientific freedom, for your critical comments, and most importantly for believing in me. It gave me the confidence and made me strive to be a better scientist.

Alenka Jejcic, for all the help you have so selflessly provided from my first moment in the lab. Thank you for showing me how to work WITH HIV-1 and your patience whenever I had questions in the P3 lab. I will miss the times when we were stuck working long hours in the P3-plasmid lab.

Maria Perdomo, for your encouragements and optimism. You have infected me with your positive attitude and to never give up even when things are not looking up. Thank you for your suggestions and help.

The individuals whom I share the office with: Alenka Jejcic, Maria Perdomo and Samir Abdurahman. Thank you for the light-hearted moments that happen (quite frequently, I would say), the meaningful discussions, and for the crazy scramble for cookies, candies and anything resembling sweet desserts when they are around.

My co-authors and collaborators for their active and fruitful contributions and collaborations.

Li Xing, Dominik Green and Carlos Moscoso at the University of California, Davis, for teaching me the principles in EM imaging and single particle reconstruction. Thank you for your patience.

Fredrik, Thore and Krister, the IT guys at LABMED, for your help whenever I had problems with the Linux “toki” computer and for any Linux shortcuts.

Hernán Concha and Anette Hofmann at the Centre for Infectious Medicine (CIM), for kindly allowing me to use the equipment in their department.

Special thanks and gratitude to Gudrun Rafi, for your help and patience whenever I came to you with administrative matters.

To all the past and present students and co-workers at the Division of Clinical Microbiology, for creating a warm and lively working and learning environment.

To the personnel at Clinical Virology laboratory (F68) at Karolinska University Hospital, especially to Charles and those in the P3 lab, Eileen and Eva for the help you have given.

Charis, Sarene, Michelle, Jamie, Sam, Kai Er, Xiaohui, Jimmie, Yinghui, Samantha, for the friendship in our journey away from home.

And most importantly, to Mum, Dad, Alvin (my brother), Auntie Norica, Uncle Edwin and Auntie Irene, for the love, support and confidence you have in me. Words cannot fully express the gratitude and love I feel. Thank you for always being a phone call or text message away.

REFERENCES

1. Centers for Disease, C., *Kaposi's sarcoma and Pneumocystis pneumonia among homosexual men--New York City and California.* MMWR Morb Mortal Wkly Rep, 1981. **30**(25): p. 305-8.
2. Centers for Disease, C., *Pneumocystis pneumonia--Los Angeles.* MMWR Morb Mortal Wkly Rep, 1981. **30**(21): p. 250-2.
3. Barre-Sinoussi, F., et al., *Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS).* Science, 1983. **220**(4599): p. 868-71.
4. Vilmer, E., et al., *Isolation of new lymphotropic retrovirus from two siblings with haemophilia B, one with AIDS.* Lancet, 1984. **1**(8380): p. 753-7.
5. Centers for Disease, C., *Update on acquired immune deficiency syndrome (AIDS) among patients with hemophilia A.* MMWR Morb Mortal Wkly Rep, 1982. **31**(48): p. 644-6, 652.
6. Gallo, R.C., et al., *Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS.* Science, 1984. **224**(4648): p. 500-3.
7. Popovic, M., et al., *Detection, isolation, and continuous production of cytopathic retroviruses (HTLV-III) from patients with AIDS and pre-AIDS.* Science, 1984. **224**(4648): p. 497-500.
8. Sarngadharan, M.G., et al., *Antibodies reactive with human T-lymphotropic retroviruses (HTLV-III) in the serum of patients with AIDS.* Science, 1984. **224**(4648): p. 506-8.
9. Schupbach, J., et al., *Serological analysis of a subgroup of human T-lymphotropic retroviruses (HTLV-III) associated with AIDS.* Science, 1984. **224**(4648): p. 503-5.
10. Levy, J.A., et al., *Isolation of lymphocytopathic retroviruses from San Francisco patients with AIDS.* Science, 1984. **225**(4664): p. 840-2.
11. Ratner, L., R.C. Gallo, and F. Wong-Staal, *HTLV-III, LAV, ARV are variants of same AIDS virus.* Nature, 1985. **313**(6004): p. 636-7.
12. Wain-Hobson, S., et al., *Nucleotide sequence of the AIDS virus, LAV.* Cell, 1985. **40**(1): p. 9-17.
13. Coffin, J., et al., *Human immunodeficiency viruses.* Science, 1986. **232**(4751): p. 697.
14. Clavel, F., et al., *Isolation of a new human retrovirus from West African patients with AIDS.* Science, 1986. **233**(4761): p. 343-6.
15. Sharp, P.M. and B.H. Hahn, *AIDS: prehistory of HIV-1.* Nature, 2008. **455**(7213): p. 605-6.
16. Zhu, T., et al., *An African HIV-1 sequence from 1959 and implications for the origin of the epidemic.* Nature, 1998. **391**(6667): p. 594-7.
17. Worobey, M., et al., *Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960.* Nature, 2008. **455**(7213): p. 661-4.
18. Cohen, M.S., et al., *The detection of acute HIV infection.* J Infect Dis, 2010. **202 Suppl 2**: p. S270-7.
19. Pope, M. and A.T. Haase, *Transmission, acute HIV-1 infection and the quest for strategies to prevent infection.* Nat Med, 2003. **9**(7): p. 847-52.
20. Berger, E.A., P.M. Murphy, and J.M. Farber, *Chemokine receptors as HIV-1 coreceptors: roles in viral entry, tropism, and disease.* Annu Rev Immunol, 1999. **17**: p. 657-700.
21. Haase, A.T., *Population biology of HIV-1 infection: viral and CD4+ T cell demographics and dynamics in lymphatic tissues.* Annu Rev Immunol, 1999. **17**: p. 625-56.
22. Cates, W., Jr., M.A. Chesney, and M.S. Cohen, *Primary HIV infection--a public health opportunity.* Am J Public Health, 1997. **87**(12): p. 1928-30.
23. Quinn, T.C., *Acute primary HIV infection.* JAMA, 1997. **278**(1): p. 58-62.

24. Piatak, M., Jr., et al., *High levels of HIV-1 in plasma during all stages of infection determined by competitive PCR.* Science, 1993. **259**(5102): p. 1749-54.
25. Pantaleo, G., et al., *HIV infection is active and progressive in lymphoid tissue during the clinically latent stage of disease.* Nature, 1993. **362**(6418): p. 355-8.
26. Connor, R.I., et al., *Increased viral burden and cytopathicity correlate temporally with CD4+ T-lymphocyte decline and clinical progression in human immunodeficiency virus type 1-infected individuals.* J Virol, 1993. **67**(4): p. 1772-7.
27. Freed, E.O., *HIV-1 replication.* Somat Cell Mol Genet, 2001. **26**(1-6): p. 13-33.
28. De Clercq, E., *Emerging anti-HIV drugs.* Expert Opin Emerg Drugs, 2005. **10**(2): p. 241-73.
29. De Clercq, E., *Anti-HIV drugs: 25 compounds approved within 25 years after the discovery of HIV.* Int J Antimicrob Agents, 2009. **33**(4): p. 307-20.
30. Sekaly, R.P., *The failed HIV Merck vaccine study: a step back or a launching point for future vaccine development?* J Exp Med, 2008. **205**(1): p. 7-12.
31. Flynn, N.M., et al., *Placebo-controlled phase 3 trial of a recombinant glycoprotein 120 vaccine to prevent HIV-1 infection.* J Infect Dis, 2005. **191**(5): p. 654-65.
32. Gilbert, P.B., et al., *Correlation between immunologic responses to a recombinant glycoprotein 120 vaccine and incidence of HIV-1 infection in a phase 3 HIV-1 preventive vaccine trial.* J Infect Dis, 2005. **191**(5): p. 666-77.
33. Blankson, J.N., D. Persaud, and R.F. Siliciano, *The challenge of viral reservoirs in HIV-1 infection.* Annu Rev Med, 2002. **53**: p. 557-93.
34. Chun, T.W., et al., *Presence of an inducible HIV-1 latent reservoir during highly active antiretroviral therapy.* Proc Natl Acad Sci U S A, 1997. **94**(24): p. 13193-7.
35. Finzi, D., et al., *Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy.* Science, 1997. **278**(5341): p. 1295-300.
36. Wong, J.K., et al., *Recovery of replication-competent HIV despite prolonged suppression of plasma viremia.* Science, 1997. **278**(5341): p. 1291-5.
37. Ohagen, A., et al., *The morphology of the immature HIV-1 virion.* Virology, 1997. **228**(1): p. 112-4.
38. Wilk, T., et al., *Organization of immature human immunodeficiency virus type 1.* J Virol, 2001. **75**(2): p. 759-71.
39. Peng, C., et al., *Role of human immunodeficiency virus type 1-specific protease in core protein maturation and viral infectivity.* J Virol, 1989. **63**(6): p. 2550-6.
40. Dalgleish, A.G., et al., *The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus.* Nature, 1984. **312**(5996): p. 763-7.
41. Klatzmann, D., et al., *T-lymphocyte T4 molecule behaves as the receptor for human retrovirus LAV.* Nature, 1984. **312**(5996): p. 767-8.
42. Lasky, L.A., et al., *Delineation of a region of the human immunodeficiency virus type 1 gp120 glycoprotein critical for interaction with the CD4 receptor.* Cell, 1987. **50**(6): p. 975-85.
43. Sattentau, Q.J., et al., *Conformational changes induced in the envelope glycoproteins of the human and simian immunodeficiency viruses by soluble receptor binding.* J Virol, 1993. **67**(12): p. 7383-93.
44. Kwong, P.D., et al., *Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody.* Nature, 1998. **393**(6686): p. 648-59.
45. Alkhatib, G., et al., *CC CKR5: a RANTES, MIP-1alpha, MIP-1beta receptor as a fusion cofactor for macrophage-tropic HIV-1.* Science, 1996. **272**(5270): p. 1955-8.
46. Choe, H., et al., *The beta-chemokine receptors CCR3 and CCR5 facilitate infection by primary HIV-1 isolates.* Cell, 1996. **85**(7): p. 1135-48.
47. Deng, H., et al., *Identification of a major co-receptor for primary isolates of HIV-1.* Nature, 1996. **381**(6584): p. 661-6.
48. Doranz, B.J., et al., *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, 1996. **85**(7): p. 1149-58.

49. Dragic, T., et al., *HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5.* Nature, 1996. **381**(6584): p. 667-73.
50. Feng, Y., et al., *HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor.* Science, 1996. **272**(5263): p. 872-7.
51. Cardozo, T., et al., *Structural basis for coreceptor selectivity by the HIV type 1 V3 loop.* AIDS Res Hum Retroviruses, 2007. **23**(3): p. 415-26.
52. Pancera, M., et al., *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, 2010. **107**(3): p. 1166-71.
53. Weiss, C.D., *HIV-1 gp41: mediator of fusion and target for inhibition.* AIDS Rev, 2003. **5**(4): p. 214-21.
54. Ashkenazi, A. and Y. Shai, *Insights into the mechanism of HIV-1 envelope induced membrane fusion as revealed by its inhibitory peptides.* Eur Biophys J, 2011. **40**(4): p. 349-57.
55. Chan, D.C., et al., *Core structure of gp41 from the HIV envelope glycoprotein.* Cell, 1997. **89**(2): p. 263-73.
56. Weissenhorn, W., et al., *Atomic structure of the ectodomain from HIV-1 gp41.* Nature, 1997. **387**(6631): p. 426-30.
57. Eckert, D.M. and P.S. Kim, *Mechanisms of viral membrane fusion and its inhibition.* Annu Rev Biochem, 2001. **70**: p. 777-810.
58. Root, M.J. and H.K. Steger, *HIV-1 gp41 as a target for viral entry inhibition.* Curr Pharm Des, 2004. **10**(15): p. 1805-25.
59. Smith, J., A. Azad, and N. Deacon, *Identification of two novel human immunodeficiency virus type 1 splice acceptor sites in infected T cell lines.* J Gen Virol, 1992. **73 (Pt 7)**: p. 1825-8.
60. Louis, J.M., G.M. Clore, and A.M. Gronenborn, *Autoprocessing of HIV-1 protease is tightly coupled to protein folding.* Nat Struct Biol, 1999. **6**(9): p. 868-75.
61. Pettit, S.C., et al., *Initial cleavage of the human immunodeficiency virus type 1 GagPol precursor by its activated protease occurs by an intramolecular mechanism.* J Virol, 2004. **78**(16): p. 8477-85.
62. Gelderblom, H.R., *Assembly and morphology of HIV: potential effect of structure on viral function.* AIDS, 1991. **5**(6): p. 617-37.
63. Leonard, C.K., et al., *Assignment of intrachain disulfide bonds and characterization of potential glycosylation sites of the type 1 recombinant human immunodeficiency virus envelope glycoprotein (gp120) expressed in Chinese hamster ovary cells.* J Biol Chem, 1990. **265**(18): p. 10373-82.
64. Bernstein, H.B., et al., *Human immunodeficiency virus type 1 envelope glycoprotein is modified by O-linked oligosaccharides.* J Virol, 1994. **68**(1): p. 463-8.
65. Pinter, A., et al., *Oligomeric structure of gp41, the transmembrane protein of human immunodeficiency virus type 1.* J Virol, 1989. **63**(6): p. 2674-9.
66. Schawaller, M., et al., *Studies with crosslinking reagents on the oligomeric structure of the env glycoprotein of HIV.* Virology, 1989. **172**(1): p. 367-9.
67. Earl, P.L., R.W. Doms, and B. Moss, *Oligomeric structure of the human immunodeficiency virus type 1 envelope glycoprotein.* Proc Natl Acad Sci U S A, 1990. **87**(2): p. 648-52.
68. Hallenberger, S., et al., *Inhibition of furin-mediated cleavage activation of HIV-1 glycoprotein gp160.* Nature, 1992. **360**(6402): p. 358-61.
69. Thali, M., et al., *Characterization of conserved human immunodeficiency virus type 1 gp120 neutralization epitopes exposed upon gp120-CD4 binding.* J Virol, 1993. **67**(7): p. 3978-88.
70. Huang, C.C., et al., *Structure of a V3-containing HIV-1 gp120 core.* Science, 2005. **310**(5750): p. 1025-8.
71. Zhou, T., et al., *Structural definition of a conserved neutralization epitope on HIV-1 gp120.* Nature, 2007. **445**(7129): p. 732-7.
72. Lu, M., S.C. Blacklow, and P.S. Kim, *A trimeric structural domain of the HIV-1 transmembrane glycoprotein.* Nat Struct Biol, 1995. **2**(12): p. 1075-82.

73. Chen, B., et al., *Structure of an unliganded simian immunodeficiency virus gp120 core.* Nature, 2005. **433**(7028): p. 834-41.
74. Zanetti, G., et al., *Cryo-electron tomographic structure of an immunodeficiency virus envelope complex in situ.* PLoS Pathog, 2006. **2**(8): p. e83.
75. Zhu, P., et al., *Distribution and three-dimensional structure of AIDS virus envelope spikes.* Nature, 2006. **441**(7095): p. 847-52.
76. Liu, J., et al., *Molecular architecture of native HIV-1 gp120 trimers.* Nature, 2008. **455**(7209): p. 109-13.
77. Chertova, E., et al., *Envelope glycoprotein incorporation, not shedding of surface envelope glycoprotein (gp120/SU), Is the primary determinant of SU content of purified human immunodeficiency virus type 1 and simian immunodeficiency virus.* J Virol, 2002. **76**(11): p. 5315-25.
78. Kwong, P.D., et al., *HIV-1 evades antibody-mediated neutralization through conformational masking of receptor-binding sites.* Nature, 2002. **420**(6916): p. 678-82.
79. Kwon, Y.D., et al., *Unliganded HIV-1 gp120 core structures assume the CD4-bound conformation with regulation by quaternary interactions and variable loops.* Proc Natl Acad Sci U S A, 2012. **109**(15): p. 5663-8.
80. Wu, S.R., et al., *Single-particle cryoelectron microscopy analysis reveals the HIV-1 spike as a tripod structure.* Proc Natl Acad Sci U S A, 2010. **107**(44): p. 18844-9.
81. Peterlin, B.M. and D.H. Price, *Controlling the elongation phase of transcription with P-TEFb.* Mol Cell, 2006. **23**(3): p. 297-305.
82. Bres, V., S.M. Yoh, and K.A. Jones, *The multi-tasking P-TEFb complex.* Curr Opin Cell Biol, 2008. **20**(3): p. 334-40.
83. Tahirov, T.H., et al., *Crystal structure of HIV-1 Tat complexed with human P-TEFb.* Nature, 2010. **465**(7299): p. 747-51.
84. Zhou, M., et al., *Tat modifies the activity of CDK9 to phosphorylate serine 5 of the RNA polymerase II carboxyl-terminal domain during human immunodeficiency virus type 1 transcription.* Mol Cell Biol, 2000. **20**(14): p. 5077-86.
85. Strebel, K., *Virus-host interactions: role of HIV proteins Vif, Tat, and Rev.* AIDS, 2003. **17 Suppl 4**: p. S25-34.
86. Kao, S.Y., et al., *Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product.* Nature, 1987. **330**(6147): p. 489-93.
87. Dayton, A.I., et al., *The trans-activator gene of the human T cell lymphotropic virus type III is required for replication.* Cell, 1986. **44**(6): p. 941-7.
88. Fisher, A.G., et al., *The trans-activator gene of HTLV-III is essential for virus replication.* Nature, 1986. **320**(6060): p. 367-71.
89. Schwarze, S.R., et al., *In vivo protein transduction: delivery of a biologically active protein into the mouse.* Science, 1999. **285**(5433): p. 1569-72.
90. Neufeld, G., et al., *Vascular endothelial growth factor (VEGF) and its receptors.* FASEB J, 1999. **13**(1): p. 9-22.
91. Albini, A., et al., *The angiogenesis induced by HIV-1 tat protein is mediated by the Flk-1/KDR receptor on vascular endothelial cells.* Nat Med, 1996. **2**(12): p. 1371-5.
92. Ganem, D., *KSHV and Kaposi's sarcoma: the end of the beginning?* Cell, 1997. **91**(2): p. 157-60.
93. Li, J.C., H.C. Yim, and A.S. Lau, *Role of HIV-1 Tat in AIDS pathogenesis: its effects on cytokine dysregulation and contributions to the pathogenesis of opportunistic infection.* AIDS, 2010. **24**(11): p. 1609-23.
94. Srivastava, I.K., et al., *Purification, characterization, and immunogenicity of a soluble trimeric envelope protein containing a partial deletion of the V2 loop derived from SF162, an R5-tropic human immunodeficiency virus type 1 isolate.* J Virol, 2003. **77**(20): p. 11244-59.
95. Lian, Y., et al., *Evaluation of envelope vaccines derived from the South African subtype C human immunodeficiency virus type 1 TV1 strain.* J Virol, 2005. **79**(21): p. 13338-49.

96. Vita, C., et al., *Rational engineering of a miniprotein that reproduces the core of the CD4 site interacting with HIV-1 envelope glycoprotein.* Proc Natl Acad Sci U S A, 1999. **96**(23): p. 13091-6.
97. Van Herrewege, Y., et al., *CD4 mimetic miniproteins: potent anti-HIV compounds with promising activity as microbicides.* J Antimicrob Chemother, 2008. **61**(4): p. 818-26.
98. Stricher, F., et al., *Combinatorial optimization of a CD4-mimetic miniprotein and cocrystal structures with HIV-1 gp120 envelope glycoprotein.* J Mol Biol, 2008. **382**(2): p. 510-24.
99. Steinaa, L., et al., *Antibody to HIV-1 Tat protein inhibits the replication of virus in culture.* Arch Virol, 1994. **139**(3-4): p. 263-71.
100. Re, M.C., et al., *Effect of antibody to HIV-1 Tat protein on viral replication in vitro and progression of HIV-1 disease in vivo.* J Acquir Immune Defic Syndr Hum Retrovirol, 1995. **10**(4): p. 408-16.
101. Re, M.C., et al., *Antibody against human immunodeficiency virus type 1 (HIV-1) Tat protein may have influenced the progression of AIDS in HIV-1-infected hemophiliac patients.* Clin Diagn Lab Immunol, 1996. **3**(2): p. 230-2.
102. Shaikh, T.R., et al., *SPIDER image processing for single-particle reconstruction of biological macromolecules from electron micrographs.* Nat Protoc, 2008. **3**(12): p. 1941-74.
103. Chen, S., A.M. Roseman, and H.R. Saibil, *Electron microscopy of chaperonins.* Methods Enzymol, 1998. **290**: p. 242-53.
104. Ruprecht, J. and J. Nield, *Determining the structure of biological macromolecules by transmission electron microscopy, single particle analysis and 3D reconstruction.* Prog Biophys Mol Biol, 2001. **75**(3): p. 121-64.
105. Korin, Y.D. and J.A. Zack, *Nonproductive human immunodeficiency virus type 1 infection in nucleoside-treated G0 lymphocytes.* J Virol, 1999. **73**(8): p. 6526-32.
106. Clouse, K.A., et al., *Monokine regulation of human immunodeficiency virus-1 expression in a chronically infected human T cell clone.* J Immunol, 1989. **142**(2): p. 431-8.
107. Sei, Y., et al., *Characterization of human B cell (DK) and promonocyte (U937) clones after HIV-1 exposure: accumulation of viral reverse transcriptase activity in cells and early syncytia induction against SupT1 cells.* Cell Immunol, 1990. **125**(1): p. 1-13.
108. Ludtke, S.J., P.R. Baldwin, and W. Chiu, *EMAN: semiautomated software for high-resolution single-particle reconstructions.* J Struct Biol, 1999. **128**(1): p. 82-97.
109. Marchio, S., et al., *Cell surface-associated Tat modulates HIV-1 infection and spreading through a specific interaction with gp120 viral envelope protein.* Blood, 2005. **105**(7): p. 2802-11.
110. Re, M.C., et al., *Antibodies against full-length Tat protein and some low-molecular-weight Tat-peptides correlate with low or undetectable viral load in HIV-1 seropositive patients.* J Clin Virol, 2001. **21**(1): p. 81-9.
111. Monini, P., et al., *HIV-1 Tat Promotes Integrin-Mediated HIV Transmission to Dendritic Cells by Binding Env Spikes and Competes Neutralization by Anti-HIV Antibodies.* PLoS One, 2012. **7**(11): p. e48781.
112. Huang, L., et al., *Tat protein induces human immunodeficiency virus type 1 (HIV-1) coreceptors and promotes infection with both macrophage-tropic and T-lymphotropic HIV-1 strains.* J Virol, 1998. **72**(11): p. 8952-60.
113. Moore, J.P., et al., *Dissociation of gp120 from HIV-1 virions induced by soluble CD4.* Science, 1990. **250**(4984): p. 1139-42.
114. Willey, R.L., M.A. Martin, and K.W. Peden, *Increase in soluble CD4 binding to and CD4-induced dissociation of gp120 from virions correlates with infectivity of human immunodeficiency virus type 1.* J Virol, 1994. **68**(2): p. 1029-39.
115. Sharma, V.A., et al., *Structural characteristics correlate with immune responses induced by HIV envelope glycoprotein vaccines.* Virology, 2006. **352**(1): p. 131-44.

116. Kwong, P.D., et al., *Oligomeric modeling and electrostatic analysis of the gp120 envelope glycoprotein of human immunodeficiency virus.* J Virol, 2000. **74**(4): p. 1961-72.
117. Doms, R.W., *Beyond receptor expression: the influence of receptor conformation, density, and affinity in HIV-1 infection.* Virology, 2000. **276**(2): p. 229-37.
118. Vermeire, K. and D. Schols, *Anti-HIV agents targeting the interaction of gp120 with the cellular CD4 receptor.* Expert Opin Investig Drugs, 2005. **14**(10): p. 1199-212.
119. Myszka, D.G., et al., *Energetics of the HIV gp120-CD4 binding reaction.* Proc Natl Acad Sci U S A, 2000. **97**(16): p. 9026-31.
120. Ferrantelli, F., et al., *A combination HIV vaccine based on Tat and Env proteins was immunogenic and protected macaques from mucosal SHIV challenge in a pilot study.* Vaccine, 2011. **29**(16): p. 2918-32.
121. Speck, R.F., et al., *Selective employment of chemokine receptors as human immunodeficiency virus type 1 coreceptors determined by individual amino acids within the envelope V3 loop.* J Virol, 1997. **71**(9): p. 7136-9.
122. Shojania, S. and J.D. O'Neil, *HIV-1 Tat is a natively unfolded protein: the solution conformation and dynamics of reduced HIV-1 Tat-(1-72) by NMR spectroscopy.* J Biol Chem, 2006. **281**(13): p. 8347-56.
123. Frankel, A.D., D.S. Bredt, and C.O. Pabo, *Tat protein from human immunodeficiency virus forms a metal-linked dimer.* Science, 1988. **240**(4848): p. 70-3.
124. Frankel, A.D., et al., *Dimerization of the tat protein from human immunodeficiency virus: a cysteine-rich peptide mimics the normal metal-linked dimer interface.* Proc Natl Acad Sci U S A, 1988. **85**(17): p. 6297-300.
125. Dominguez, C., R. Boelens, and A.M. Bonvin, *HADDOCK: a protein-protein docking approach based on biochemical or biophysical information.* J Am Chem Soc, 2003. **125**(7): p. 1731-7.
126. Westendorp, M.O., et al., *Sensitization of T cells to CD95-mediated apoptosis by HIV-1 Tat and gp120.* Nature, 1995. **375**(6531): p. 497-500.