Transmembrane glycoprotein gp41 of the Human Immunodeficiency Virus Type I: gene synthesis, recombinant expression and immunological characterization

Anjo Steyn

Submitted in partial fulfilment of the requirement for the degree *Magister Scientiae* in the

Faculty of Natural and Agricultural Sciences
Department of Biochemistry
University of Pretoria
Pretoria

July 2011

Acknowledgments

- My supervisors Professors JA Verschoor and CP Kenyon for time and effort.
- The CSIR for awarding me bursaries to enable me to complete my studies.
- My fellow CSIR scientists for helping me through the project especially Lyndon, Robyn, Ilze and Marli.
- Sandra at the University of Pretoria for all her support and guidance.
- Ian for all his love, patience and support.
- My parents for the love and support.

- I declare that this thesis/dissertation, which I herby submit for the degree MSc. Biochemistry at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.

SIGNATURE:..... DATE:.....

Plagiarism Declaration

UNIVERSITY OF PRETORIA

FACULTY OF NATURAL AND AGRICULTURAL SCIENCES

DEPARTMENT OF BIOCHEMISTRY

Full name: _____ Student number: _____

Title of the work: _____

Declaration

1. I understand what plagiarism entails and am aware of the University's policy in this regard.
2. I declare that this _____ (e.g. essay, report, project, assignment, dissertation, thesis etc) is my own, original work. Where someone else's work was used (whether from a printed source, the internet or any other source) due acknowledgement was given and reference was made according to departmental requirements.
3. I did not make use of another student's previous work and submit it as my own.
4. I did not allow and will not allow anyone to copy my work with the intention of presenting it as his or her own work.

Signature _____ Date _____

Contents

List of Figures

List of Tables

List of Abbreviations

aa	Amino acids
Abs	Antibodies
AIDS	Acquired Immune Deficiency Syndrome
ARV	Antiretroviral drug
APC	Antigen presenting cells
AZT	Azidothymidine
BMGY	Buffered Glycerol-complex Medium
BMMY	Buffered Methanol-complex Medium
bp	Base pairs
BSA	Bovine serum albumin
CA	Capsid protein
CaM	Calmodulin
CCR5	C-C chemokine receptor 5
CDC	Centres for Disease Control and Prevention
CPD	Cytoplasmic domain
CRFs	Circulating recombinant forms
CTL	Cytotoxic T Lymphocytes
CXCR4	C-X-C chemokine receptor type 4
DNA	Deoxyribonucleic acid
DSMB	Data and Safety Monitoring Board
E.coli	Escherichia coli
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-linked Immunosorbant Sandwich Assay
Env	Envelope protein
ER	Endoplasmic reticulum
FasL	Fas Ligand
FITC	Fluorescein isothiocyanate
FP	Fusion peptide
FTOC	Fetal Thymic Organ Culture
Gag	Core protein
GP41	Glycoprotein 41
GP120	Glycoprotein 120
HAART	Highly active antiretroviral therapy
HA	Hemagglutinin
HCL	Hydrochloric acid

His	Histidine
HIV	Human Immunodeficiency Virus
HTLV	Leukemia virus
HVTN	HIV Vaccine Trials Network
IFN-γ	Interferon-gamma
IgG	Immunoglobulin G
IL-2	Interleukin-2
IN	Integrase
IPTG	Isopropyl-beta-D-thiogalactopyranoside
kDa	Kilo Dalton
LB	Luria-Bertani
LTP	Lipid Transfer Protein
LTR	Long Terminal Repeat
ms	Millisecond
MA	Matrix protein
MAb	Monoclonal Antibody
MBP	Maltose Binding Protein
MHC	Major histocompatibility complex antigens
MMP	Mitochondrial membrane permeabilization
MRC	Medical Research Council
mRNA	Messenger RNA
Mut^{+}	Methanol positive
Mut^{S}	Methanol sensitive
NC	Nucleocapsid protein
Nef	Negative effector
NIAID	National Institute of Allergy and Infectious Diseases
NIH	National Institutes of Health
NKT	Natural Killer T-cells
NLS	Nuclear localization signal
NSP	HIV strategy Plan
OD	Optical density
OPD	Ortho-Phenylenediamine
PACC	Pre-apoptotic chromatin condensation
PAGE	Poly-acrylamide gel electrophoresis
PBS Buffer	Phosphate Buffered Saline
PCD	Programmed cell death
PCR	Polymerase chain reaction
PEG	Polyethelyne Glycol

Pol	Polymerase
P.pastoris	*Pichia pastoris*
PR	Protease
RRE	Rev Response element
PS	Phosphatidyl serine
PVDF	Polyvinylidene fluoride membrane
Rev	Regulator of Viral gene Expression
RNA	Ribonucleic acid
RT	Reverse transcriptase
SAAVI	South African AIDS Vaccine Initiative
sCD4	Soluble CD4
SDS	Sodium dodecyl-sulphate
SIV	Simian immunodeficiency virus
SU	Surface glycoproteins
TAE	Tris Acetate Electrophoresis
TAR	Trans-activating response element
Tat	Transcriptional activator
TBST	Tris Buffer Saline Tween20 buffer
TCR	T-cell receptor
TM	Transmembrane protein
TNF	Tumor Necrosis Factor
TNFL	Tumor Necrosis Factor Ligand
TRIS	Tris (hydroxymethyl) aminomethane
UNAIDS	Joint United Nations Programme on HIV and AIDS
V	Voltage
V3	Variable loop 3
Vif	Viral infectivity factor
Vpr	Viral protein R
Vpu	Viral protein U
WHO	World Health Organization
YNB	Yeast Nitrogen Base
YPD	Yeast Extract Peptone Dextrose

Abstract

Title: Transmembrane glycoprotein gp41 of the Human Immunodeficiency Virus Type 1: gene synthesis, recombinant expression and immunological characterization

Student: Anjo Steyn

Supervisor: Prof. J.A. Verschoor

Co-supervisor: Prof. C.P. Kenyon

University: University of Pretoria

Department: Biochemistry

HIV, the causative agent of AIDS, is one of the first retroviruses to infect humans. Controlling this epidemic is extremely challenging. Studies that focus on understanding the biology, biochemistry and structural biology of HIV as well as the interactions between viral components and new drug candidates are required to determine an effective strategy to combat HIV/AIDS. Gp41 might be of interest for vaccine development because of its antigenic neutralizing epitope sites previously discovered. Here, the ecto- and endo-domains of HIV-1 gp41 were studied. The endo-domain was truncated into three gene fragments namely gp41T3 (405 bp), gp41T2 (591 bp) and gp41T1(696 bp), all containing the Kennedy domain located on the C-terminal of gp41. The 80 amino acid fragment gp41(80) of the ecto-domain, located on the N-terminal of gp41 was also constructed. The pMAL-gp41T3, pMAL-gp41(80) and gp41T1 constructs were expressed in either *E. coli* or *P. pastoris* and purified. Both the pMAL-gp41T3 and the pMAL-gp41(80) peptides proved to be antigenic towards HIV positive patient sera determined by ELISA, but in the case of gp41T1 no antigenicity was found. Gp41(80) showed increasing antibody-antigen binding activity as the CD4 T cell count decreased, i.e. as HIV infection progressed towards AIDS, while gp41T3 showed no correlation between the CD4 count and the antibody-antigen binding activity. Therefore, irrespective of the decline in the CD4 counts of the patients there is a constant antibody-antigen binding activity. It was also indicated that the hydrophobicity of pMAL-gp41T3 was higher than for pMAL-gp41(80). The high levels of hydrophobicity of this peptide may lead to its presentation by antigen presenting cells on CD1 as a “lipid like”

peptide to CD4/CD8 double negative T-cell or NKT cells. Auto-immunity may be of some concern, because antibodies and other forms of immunity induced by the CD4/CD8 double negative T-cell population are often involved in auto-immunity. Many viruses have, as part of their invasion strategy, the ability to modulate the apoptotic pathways of the host. Preliminary evidence shows that the pMAL-gp41T3 construct has apoptotic properties. Considerable homology was illustrated by the sequence alignment of gp41T3 with apoptotic TNFα, FasL, TRAIL, APRIL, TRANCE and VEGI. Gp41 may be considered as a vaccine target due to its antigenicity shown in this study, but its apoptotic and auto-immune potential argue for care. The ecto-domain seems to elicit antibodies that do not prevent the progression to AIDS. Even if no vaccine can be made from this, the antibody response to the endo-domain of gp41 was found not to be perturbed by the progression towards AIDS. It may therefore, find use as a surrogate marker for HIV infection that may be useful in diagnosis. The immune properties of pMAL-gp41(80) and pMAL-gp41T3 may be exploited to steer immunity away from AIDS, e.g. by selective suppression of pathological auto-immune antibody activity by idiotypic vaccines or suppression of peptide induced apoptosis of CD4 T-cells by pharmaceutical or biological means.

Chapter 1

Literature Review

1.1 HIV Epidemic

The origin of the Human Immunodeficiency Virus (HIV) and Acquired Immune Deficiency Syndrome (AIDS) has puzzled scientists ever since the disease emerged in the early 1980s. For more than twenty years it has been the subject of fierce debate and the cause of countless arguments, with everything from a promiscuous flight attendant to a corrupt vaccine programme being blamed (UNAIDS 2006). International awareness of HIV infection began in 1981, but the presence of the virus was later detected in archived blood and tissue samples, revealing that it had been silently spreading in gay male population in the US and Sweden and heterosexuals in Africa and Haiti in the late 1970s. Three of the earliest known instances of HIV infection are 1) a plasma sample taken in 1959 from an adult male living in the Democratic Republic of Congo (Zhu *et al*., 1998), 2) HIV found in tissue samples from an American teenager who died in St. Louis in 1969 and 3) HIV found in tissue samples from a Norwegian sailor who died around 1976 (Zhu *et al*., 1998).

A 1998 analysis of the plasma sample from 1959 has suggested that HIV-1 was introduced into humans around the 1940s or the early 1950s, much earlier than previously thought (Zhu *et al*., 1998). In January 2000, the results of a study presented at the 17th Conference on Retroviruses and Opportunistic Infections, suggested that the first case of HIV-1 infection occurred around 1930 in West Africa. The study was carried out by Dr. Bette Korber of the Los Alamos National Laboratory. The estimate of 1930 was based on a complex computer model of HIV's evolution. If accurate, it means that HIV was in existence before many scenarios such as the oral polio vaccine theory and conspiracy theories suggest (Korber *et al.,* 2000).

It is likely that we will never know who the first person was to be infected with HIV, or exactly how it spread from that initial person. The spread of AIDS could quite conceivably have been induced by a combination of many different events;

whether through injections, travel, wars, colonial practices or genetic engineering. Nevertheless, perhaps a more pressing concern for scientists today should not be how the AIDS epidemic originated, but how those it affects can be treated, how the further the spread of HIV can be prevented and how the world can change to ensure that a similar pandemic never occurs again.

The fact is that South Africa is currently experiencing one of the most severe HIV epidemics in the world. By the end of 2007, there were approximately 5.7 million people living with HIV in South Africa and almost a thousand AIDS deaths occurring every day (UNAIDS/WHO, 2009). The prevalence among South Africans is stated in Figure 1.1.

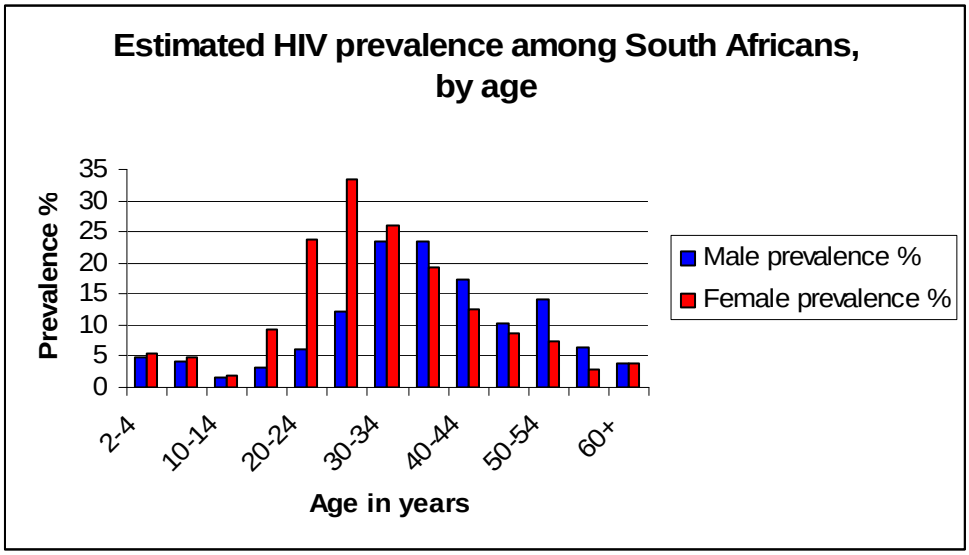

Figure 1.1: **Estimated HIV prevalence among South Africans in the different age groups.** Blue represent male prevalence and red represent female prevalence for the year 2008 (UNAIDS/WHO, 2009).

It is not only South Africa that is currently experiencing such a severe HIV epidemic, the rest of Africa is also suffering from this epidemic as indicated in Figure 1.2.

Figure 1.2: HIV prevalence in adults and children in sub-Saharan Africa, 2008 (UNAIDS Report 2009).

1.2 Etiology of AIDS

AIDS is characterized by a gradual deterioration of immune function - most notably, the depletion of crucial immune cells like CD4 T-cells that play a central function during the typical course of infection. CD4 T-cells initiate and coordinate the adaptive immune response to particular antigens by secretion of specific cytokines, but cannot bind directly to free antigen. Antigenic peptides are presented by major histocompatibility complex antigens (MHC) on the surface of antigen presenting cells (APC) (Delves *et al*., 2000). CD4 T-cells recognize these peptides combined with MHC class II antigens through their T-cell receptor (TCR). Upon stimulation, CD4 T-cells signal other cells in the immune system to perform their specific function.

The delay in the discovery that HIV causes AIDS was due to an unfavourable mind-set during the late 1970s. Many people thought that epidemic diseases caused by microbes, including viruses, no longer posed a threat in industrialized countries. Other prevailing beliefs were that viruses did not cause any human cancers and that there was no such thing as a retrovirus that infected humans.

In 1982 AIDS was recognised as a unique long-lasting disease, with an extremely long lag time between exposure to the agent and the profound state of immune suppression characterized by the occurrence of opportunistic infections or cancer. Many factors were proposed as possible causes of AIDS, including fungi, chemicals and even auto-immunity to leukocytes. Several clues lead to the eventual discovery of the real causative agent. First, the various manifestations of AIDS were unified by one biologic marker: a decrease in the levels of a specific subgroup of T-cell that contained the CD4 surface antigen. CD4 and other CD surface proteins had been identified only a few years earlier with the use of specific monoclonal antibodies by Milstein and Köhler (Köhler *et al*., 1975). The findings regarding the T-cell subgroup suggested an agent that specifically targeted $CD4^+$ T-cells, and the Human T-cell leukemia virus (HTLV) was one such agent. Furthermore, HTLV transmitted through blood and sexual activity, as well as from mother to infant was consistent with the spreading of AIDS. Finally, the Centres for Disease Control and Prevention (CDC) reported cases of AIDS in patients with hemophilia who had received only filtered clotting factors, which seemed to eliminate the possibility that the agent was a microorganism larger than a virus.

In early 1983, a clear-cut isolate was obtained in Paris, with the help of interleukin-2 and anti-interferon serum, from cultured T lymphocytes derived from a lymph node biopsy specimen from a patient with lymphadenopathy, a syndrome that was considered to be a precursor of AIDS. The only retroviruses then known were the human T-cell leukemia viruses, HTLV-1 and HTLV-2, identified by Gallo. Montagnier and co-workers tested if the proteins in the supernatant obtained from the biopsy specimen could be recognized by Gallo's antibodies against HTLV. Surprisingly, the labeled viral supernatant could not be immune precipitated with the HTLV antibodies, but could be precipitated with the patient's own serum (Montagnier *et al*., 2002).This virus proved to be different from HTLV in terms of antigenicity and morphology (Levy *et al*., 1984; Coffin *et al*., 1986).

Identifying the cause of AIDS presented a unique challenge, because unlike other viruses responsible for past epidemics, AIDS was characterize by clinical signs that developed years after the infection had occurred and by then patients usually had numerous other infections. Thus, an exceptional linkage of agent to disease had to be established. This linkage was made through the repeated isolation of HIV from patients with AIDS and through the development of a readily reproducible blood test. The blood test also helped in the cloning and molecular characterization of the genetic material of the virus at the end of 1984, which clearly proved that the new virus belonged to the superfamily of lentiviruses. This finding in turn, opened the way for the design of specific drugs and vaccines.

Like all viruses, HIV can replicate only inside cells, commandeering the cell's machinery to reproduce. However, only HIV and other retroviruses, once inside a cell, use reverse transcriptase to convert their RNA into DNA, which can be incorporated into the host cell's genes. The course of infection with these viruses is characterized by a long interval between initial infection and the onset of serious symptoms. Primary infection with HIV may be accompanied by transient illness similar to glandular fever, with malaise, muscle pains, swollen lymph nodes, sore throat and rash. There is transient depletion of peripheral $CD4^+$ T-cells, expansion of $CD8^+$ T-cells and high plasma levels of HIV. By 2-6 weeks, antibodies to core and surface proteins can be detected by enzyme-linked immunoassays. A chronic infection ensues without illness, but about 33% of patients have swollen lymph nodes. Fifty percent of those infected develop AIDS within 7 to 10 years. Later in infection, non-specific constitutional symptoms such as fever, night sweats and weight loss occur, together with 'minor' conditions that largely affect the mucous membranes and skin. These conditions often signal the development of serious opportunistic infections and tumours, which constitute AIDS when the $CD4^+$ T-cell count is, usually, below 200 (Roitt *et al*., 2002).

1.2.1 Classification of HIV

The causative agent of AIDS is HIV. About 85% of the HIV isolates from humans can be grouped into two types: the more pathogenic and widespread HIV-1 and the relatively uncommon HIV-2. Both types are transmitted via sexual contact, blood and from mother to child. HIV-2 is less easily transmitted and the period between initial infection and illness is double that for HIV-1. HIV-2 originates from and is still focused to regions in Western and Central Africa (Chan *et al*., 1997). HIV-1 can be classified into three groups: the “major” group M, the “outlier” group O and the “new” group N as indicated in Figure 1.3. Group M, O and N viruses are members of the primate *Lentivirus* lineage that includes also Simian immunodeficiency virus (SIV) strains (Requejo *et al*., 2006).

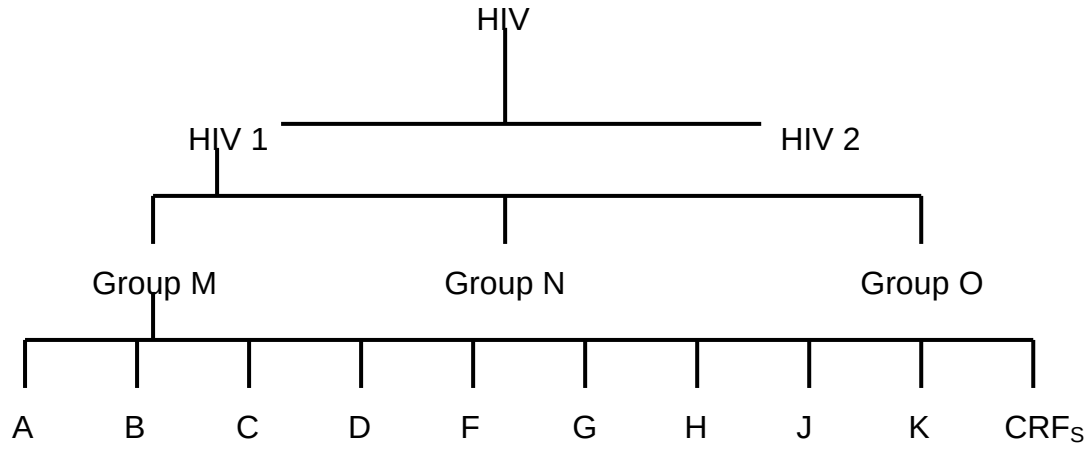

Figure 1.3: **Different levels of HIV classification subtypes.** HIV-1: A, B, C, D, F, G, H, J and K and more than 15 circulating recombinant forms (CRFs) (Requejo *et al*., 2006).

Group O appears to be most prevalent to West Central Africa, mainly in Cameroon, Gabon, Nigeria and Equatorial Guinea. Group O is phylogenetically removed from the other strains and accounts for less than 10% of HIV-1 infections worldwide. The extremely rare group N was discovered in 1998 in Cameroon. More than 90% of HIV-1 infections belong to HIV-1 group M. Within group M there are known to be at least nine genetically distinct subtypes of HIV-1. These are Subtypes A, B, C, D, F, G, H, J and K and more than 15 circulating recombinant forms (CRFs).

If an individual is infected by two different HIV-1 subtypes, they circulate as different distinct social networks. However, these barriers eventually can collapse, establishing conditions for inter-subtype recombination (Artenstein *et al.*, 1995). When the genetic combination is satisfactorily established in the individual, then the new hybrid is designated a CRF.

The many subtypes and CRFs of HIV contribute to serious obstacles in developing a successful vaccine. In addition, HIV infections induce immune dysfunction in patients that lead to abnormalities in every arm of the immune system.

1.2.2 HIV replication cycle

HIV replication, shown in Figure 1.4, can be divided into several viral replication steps starting with attachment and entry, followed by reverse transcription and DNA synthesis, transport to nucleus, integration, viral transcription, viral protein synthesis, assembly of virus, release of virus and maturation. The replication can also be divided into two phases: the early and the late phase. The early phase begins with the recognition of the target cell ($CD4^+$ T-cell mainly, but also CD4 macrophage) by the mature virion and involves all processes leading to and including integration of the genomic DNA into the chromosome of the host cell. The late phase begins with the regulated expression of the integrated proviral genome and involves all processes up to and including virus budding and maturation (Turner and Summers*.,* 1999).

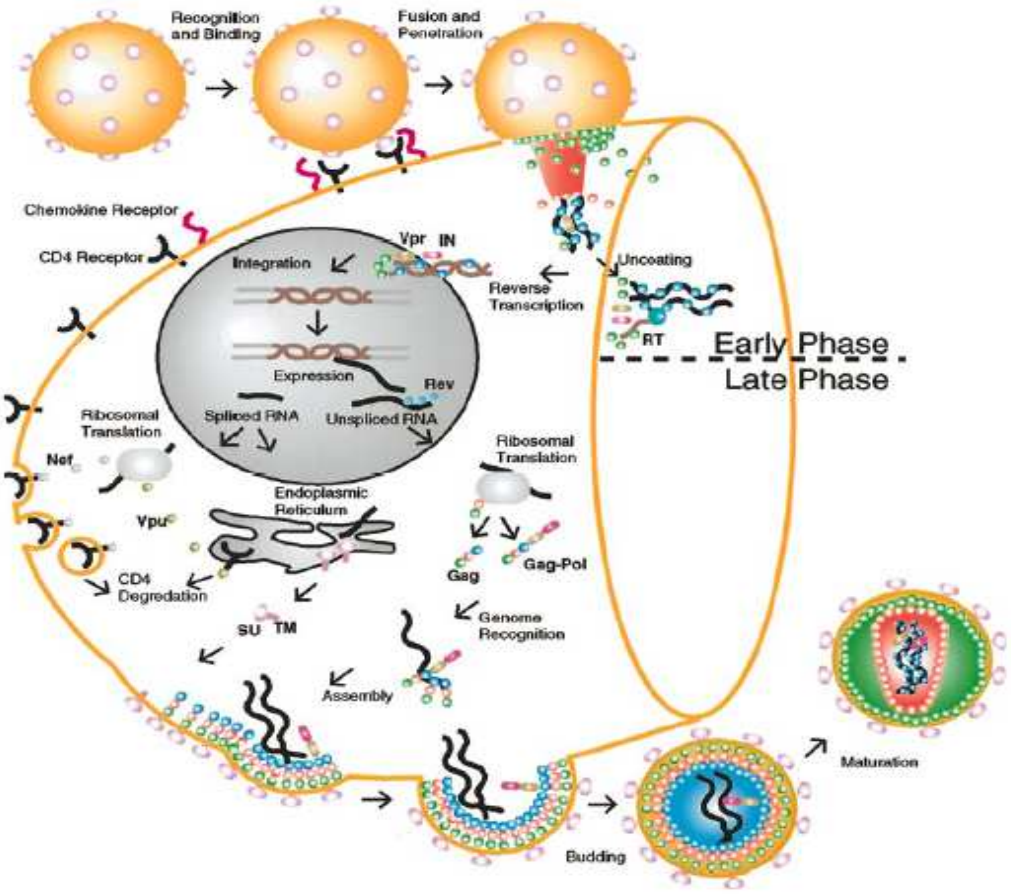

Figure 1.4: **General features of the HIV-1 replication cycle.** Indicated are the two phases of the replication cycle as well as the events that take place during each phase (Turner *et al.,* 1999).

1.2.3 Anatomy of HIV

The mature HIV virion shown in Figure 1.5 is enveloped by a lipid bilayer that is derived from the membrane of the host cell. The transmembrane protein (TM, gp41) anchores the exposed surface glycoproteins (SU, gp120) to the virus. Several cellular membrane proteins (major histocompatibility antigens, actin and ubiquitin) are included in the lipid bilayer and are derived from the host cell. The inner surface of the viral membrane is lined with approximately 2000 copies of the matrix protein (MA, p17) and a conical capsid core particle comprising 2000 copies of the capsid protein (CA, p24) is located in the centre of the virus. The capsid particle encapsulates the two copies of the unspliced viral genome. The viral genome is stabilized as a ribonucleoprotein complex with 2000 copies of the nucleocapsid protein (NC, p7). It also contains three essential virally encoded enzymes: protease (PR), reverse transcriptase (RT) and integrase (IN). The accessory proteins, Nef, Vif and Vpr (not shown) are also packaged by the virus

particles. Three additional accessory proteins that function in the host cell, Rev, Tat and Vpu do not appear to be packaged (Turner *et al*., 1999).

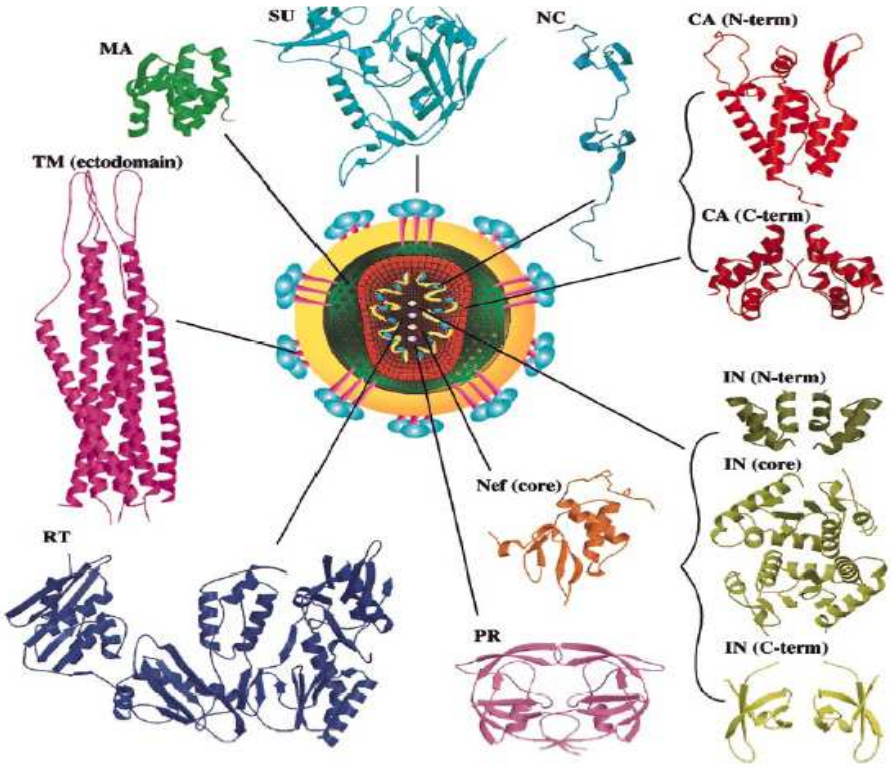

Figure 1.5: **The mature HIV virion surrounded by ribbon representations of the structurally characterized viral proteins and protein fragments.** Lipid bilayer in yellow, Exposed surface glycoproteins (SU, gp120; cyan), transmembrane protein (TM, gp41; violet), matrix protein (MA, p17; green), capsid protein (CA, p24; red), nucleocapsid protein (NC, p7; blue), protease (PR; pink), reverse transcriptase (RT; purple) and integrase (IN; olive) and the accessory protein Nef (orange) (Turner *et al*., 1999).

HIV and Simian immunodeficiency virus (SIV) are genetically related members of the *Lentivirus* genus of the Retroviridae family. Lentiviruses include complex exogenous viruses responsible for a variety of neurological and immunological diseases, but are not directly implicated in any malignancies. Genomes of these viruses are characterized by three structural genes: *gag* (core protein), *pol* (polymerase/reverse transcriptase) and *env* (envelope protein) genes. The retrovirus genome is unique among viruses in several aspects, including physical organization, mode of synthesis and replication. The diploid virus genome is composed of two identical copies of single-strand ribonucleic acid (RNA). The RNA is synthesized and processed by the host cell messenger RNA (mRNA) handling machinery. The HIV genome can be characterized by the three

structural genes *gag* (core protein), *pol* (polymerase/reverse transcriptase) and *env* (envelope protein). Some additional genes and their functions are indicated in Figure 1.6 (Greene *et al.,* 1993).

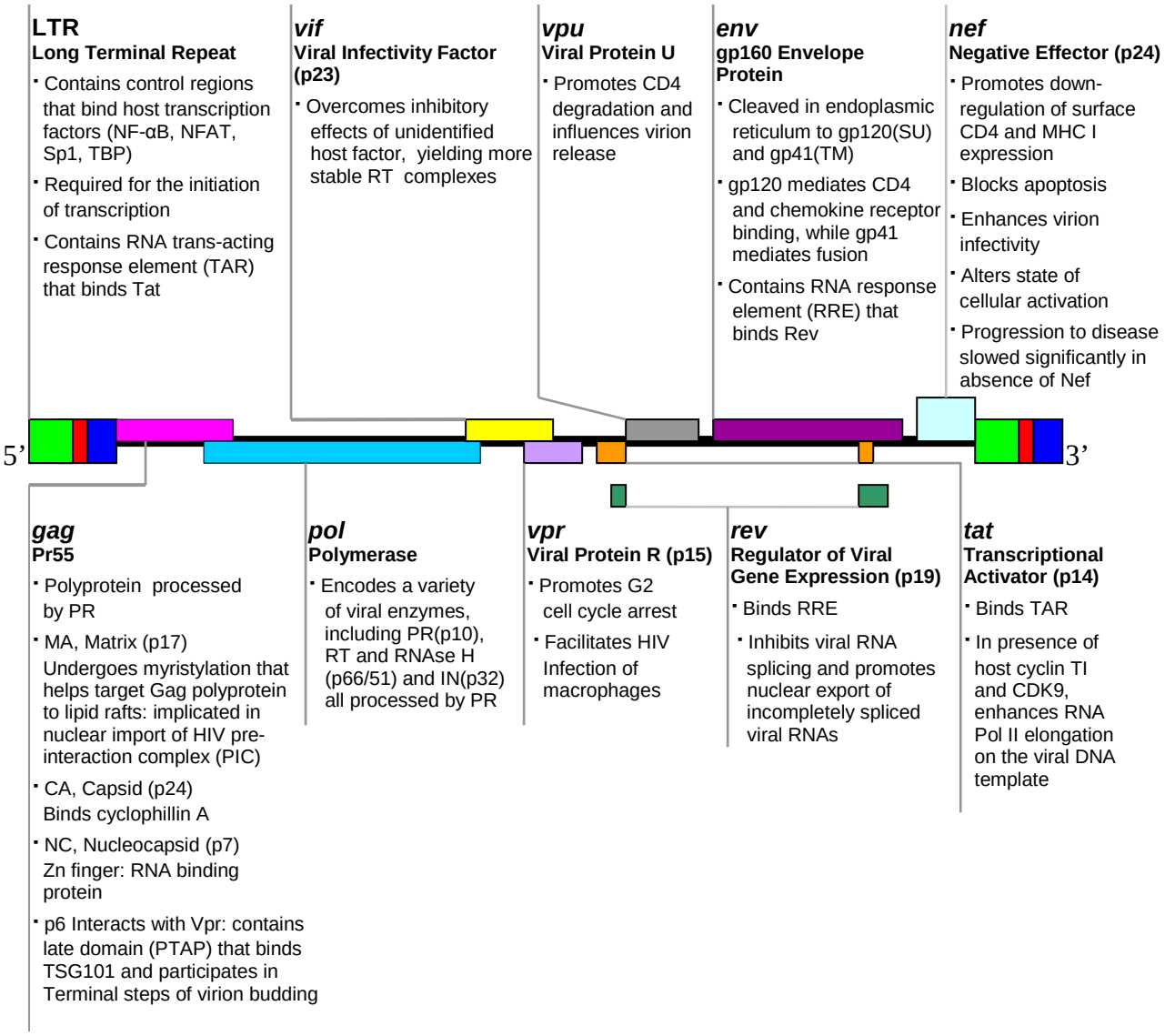

Figure 1.6: **The organization of the HIV genome.** Schematic illustration of the gene components of HIV and their specific functions (Greene *et al.,* 1993).

1.3 Immunology of HIV infection and AIDS

HIV infection induces a profound immune dysfunction, with a whole range of abnormalities. These abnormalities take place in both the humoral and cellular immune response.

In humoral immune response, HIV causes B-cell dysfunction and this is mediated by viral proteins and cytokine deregulation (Patke *et al.,* 2000). In HIV infected patients the following symptoms occur namely: B-cell hyperplasia, circulating immune complexes and elevated auto-antibodies. Ultimately, HIV infection leads to B-cell depletion. A subpopulation of B-cells with low CD21 expression has been described in high viremia patients. These cells are enhanced immunoglobulin secretors and poor antibody responders and might be partly responsible for the humoral defects. HIV activates complement through alternative and classical pathways (Reisinger *et al.,* 1990). Although complement C3 is deposited on the viral surface, there is poor function of the complement C5-C9 membrane attack complex. HIV might infect cells by means of the complement receptors. Soluble CD16 has been shown to inhibit C3 receptor mediated HIV-1 infection of monocytes (Bouhlal *et al.,* 2001).

The cellular immune response is strongly induced by HIV. However, the infection is only partially controlled. This apparent contradiction can be explained by several viral mechanisms of immune evasion: HIV provirus latency, sequestration reservoirs, down-regulation of MHC molecules, up-regulation of Fas ligand and viral protein epitope mutations to mention a few (Mc Micheal *et al.,* 1998). The study of long-term non-progressors (HIV-infected patients who are asymptomatic and have normal $CD4^+$ T-cell counts in the absence of treatment) has revealed that several immune mechanisms are significant in controlling HIV infection. Such patients might have low but detectable viremia, which seems to be important in maintaining the host-specific immune response. These mechanisms include: (1) increased production of TH1-type cytokines, such as IL-2 and IFN-γ, (2) HIV-specific $CD4^+$ T-cell proliferative responses and cytotoxic $CD8^+$ T-cell activity and (3) increased synthesis of $CD8^+$ T-cell suppressive factors and β-chemokines. When T-cell responses to HIV were first examined in infected persons, there was a striking disparity between the

apparently vigorous $CD8^+$ T-cell response and an almost absent $CD4^+$ T-cell response (Plata *et al.,* 1987). The $CD8^+$ T-cell response was surprising in its magnitude and seemed much larger than for other viruses (Walker *et al.,* 1989).

The CD4 molecule is a receptor for HIV attachment and entry into the cell (Connors *et al.,* 1997). The population of $CD4^+$ T-cells is significantly and preferentially targeted. Although the direct infection of $CD4^+$ T-cells with HIV is well documented, the actual mechanism of cell destruction is not clear. One possibility is a change in autoimmunity. Non-polymorphic determinants of the MHC II share some degree of structural homology with gp120 of HIV-1 (Golding *et al.,* 1988) and antibodies to these HIV proteins could therefore, cross-react with MHC II molecules. In fact, antibodies that react with class II molecules have been found in the serum of patients with HIV infection. These antibodies could prevent interaction between CD4 and class II molecules expressed on the antigen-presenting cells, thus impairing the cellular interaction required for efficient antigen presentation, thereby inhibiting antigen-specific functions mediated by $CD4^+$ T-cells (Golding *et al.*, 1988).

The typical course of HIV infection is indicated in Figure 1.7. It is clear that with the increase of HIV RNA copies there is a steep decrease in the number of $CD4^+$ T-cells.

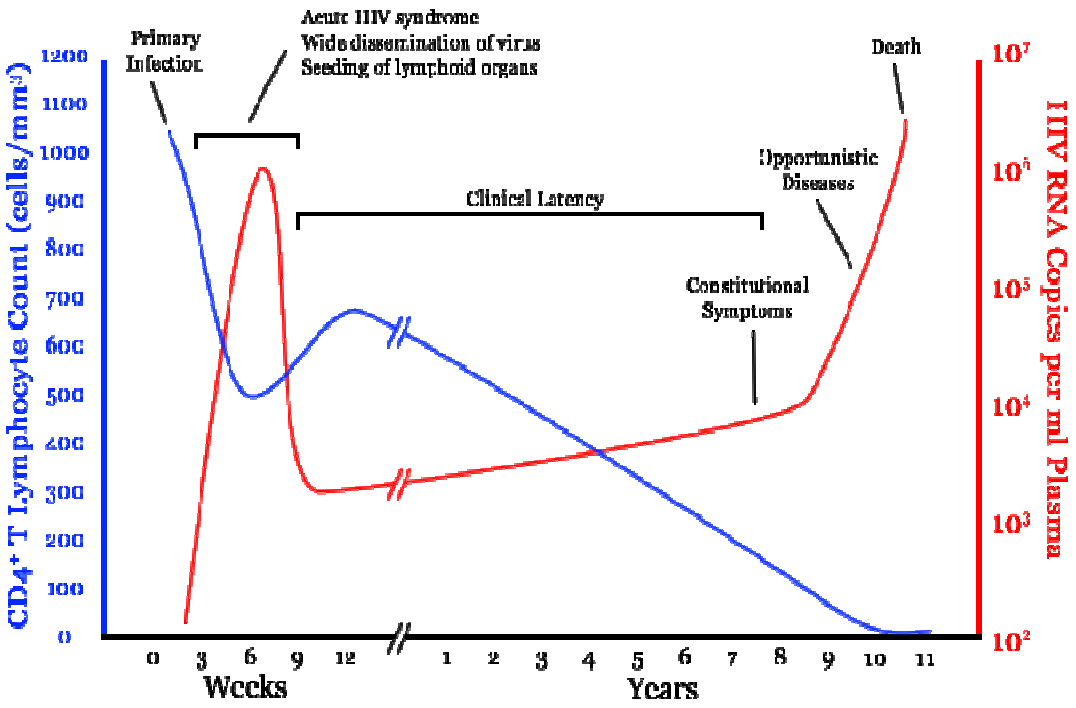

Figure 1.7: **Typical course of HIV infection.** During the early period after primary infection there is widespread dissemination of virus and a sharp decrease in the number of CD4 T-cells in peripheral blood. An immune response to HIV ensues, with a decrease in detectable viremia followed by a prolonged period of clinical latency. The CD4 T-cell count continues to decrease during the following years, until it reaches a critical level, below which there is a substantial risk of opportunistic diseases (Greene *et al.,* 1993).

The CD4 T-cell count continues to decrease until it reaches levels were the risk for opportunistic diseases occurs.

1.4 Treatment of HIV

Efforts to control the HIV/AIDS epidemic have focused heavily on studies including biology, biochemistry and structural biology of HIV and on interactions between viral components and new drug candidates. For the reason that HIV is the first retrovirus that infects humans, it is unlikely that conventional approaches to combat viral diseases will necessarily work. First, anti-retroviral drugs could control the retrovirus in humans only temporarily, due to the incorporation of the RNA of the virus into the human host cell. The infection can also not be treated with conventional anti-viral drugs that relieve symptoms rather than ridding the host of the virus.

A number of anti-retrovirals do exist. However, they have their limitations, like drug resistance, side effects and cost. These drugs delay the progress of the disease but there is no cure yet. Current anti-retroviral drugs and their prime functions are as follows: Azidothymidine (AZT) is a nucleoside analog which inhibits viral reverse transcriptase (RT). Saquinavir contains protease inhibitors that target the HIV protease (PR). HIV proteases are required to cleave the viral proteins into smaller peptides. Highly active antiretroviral therapy (HAART), which is a combination drug containing both PR and RT inhibitors, can reduce viral loads to undetectable levels. These low levels can be maintained for a period of two years or more. Enfuviritide/T-20 is a class of drug that prevents viral fusion and entry into cells; therefore, prevents the HIV virus from entering the CD4 cells by binding to the protein on the viral membrane that facilitates membrane fusion with CD4 cells (UNAIDS/WHO, 2009).

In South Africa, the National HIV strategy Plan (NSP) for 2007-2011 stated that the current state of the AIDS pandemic calls for drastic measures. However, antiviral treatment numbers have continually missed operational plan targets. The gap has been reduced significantly in recent years. Whereas in 2004, actual antiretroviral treatment represented seven percent of the operational plan target, the actual antiretroviral treatment rollout at February 2010 represented 36 percent. Among the primary aims of the NSP are to reduce the rate of new HIV infection by 50% by 2011 and to minimize the impact of HIV/AIDS on individuals, families, communities and society. It seeks to accomplish these goals via expanding access to appropriate treatment, care and support to 80% of all HIV positive individuals by 2011. The four key areas of interventions needed to reach NSP's goals are: prevention; treatment, care and support; research, monitoring and surveillance; human rights and access to justice. The plan seeks to reduce mother-to-child transmission of HIV to less than 5% of cases, increase HIV testing, and scale up access to food and medicine (www.doh.gov.za/docs/misc/stratplan/2007-2011/part1.pdf). Despite all this, to date no chemical treatment or therapy can claim a complete cure for AIDS.

1.5 Prevention of AIDS

HIV has several inherent strategies by which to escape the host immune response and continue replicating. The most studied of these strategies are antigenic variation (Borrow *et al.,* 1997, Goulder *et al.,* 1997) down-regulation of the surface expression of MHC molecules, and reduction of specific CD8+ T-cells (Pantaleo *et al.,* 1997). Despite this, several attempts were made to develop a vaccine against AIDS.

The South African AIDS Vaccine Initiative (SAAVI) was formed in 1999 as a lead programme of the Medical Research Council (MRC) of South Africa. SAAVI was established to co-ordinate the research, development and testing of AIDS vaccines in South Africa. Several gp120-based vaccines were trialled in South Africa, most of which were developed elsewhere, mainly the USA. On September 2007, the United States National Institute of Allergy and Infectious Diseases (NIAID), part of the National Institutes of Health (NIH), the pharmaceutical company Merck and Co. Inc., and the NIAID-funded HIV Vaccine Trials Network (HVTN) announced that immunisations in the HIV vaccine clinical trial known as the STEP study, also referred to as the HVTN 502 or Merck V520-023 study, would be discontinued. The decision was based on recommendations made by an independent Data and Safety Monitoring Board (DSMB), which concluded that the vaccine did not prevent HIV infection, nor reduced the amount of virus in those who became infected with HIV. The STEP study was conducted in the USA, Australia, South America and the Caribbean; countries where the predominant circulating HIV subtype is B. As a result of those findings, immunisations and enrolment in a separate clinical trial in South Africa, known as Phambili, which was evaluating the same Merck test HIV vaccine, was paused to allow for further analysis of these findings. Based on the additional review of the STEP data, an independent DSMB has concluded in 2008 that there is no basis for anticipating more favourable results in the South African clinical trial known as VTN 503 or the Phambili study. Therefore, the HVTN 503 oversight committee has permanently suspended immunisations and enrolment in the study in South Africa (Watkins *et al.,* 2008). It is clear that a standard approach to vaccination against AIDS is not feasible.

1.6 Back to the drawing board: The biology of HIV/AIDS

The challenge to develop an efficient vaccine or find a cure against AIDS demands the recognition of the unique and complex nature of HIV-host interaction. AIDS vaccine development requires more than just the elucidation of the appropriate viral envelope structure and the identification of the epitopes that induce protective antibodies. Similarly, a chemotherapeutic cure is not simply affected by the identification of a virus specific lethal target enzyme and the development of an appropriate selective chemical inhibitor, such as the antiretroviral compounds. Real and permanent solutions will only come by a better understanding of the general biology of HIV. What follows is a summary of the current understanding:

The interaction between HIV and the human host cell can be divided into two phases namely the early and late phase. The early phase starts off with the binding of the envelope glycoprotein to the host cell via a specific interaction between gp120 and the amino-terminal immunoglobulin domain of CD4 T-cell. Additional cell-surface proteins are necessary to promote fusion of the viral and cellular membrane, including the chemokine receptors CXCR4 and CCR5 (Moore *et al*., 1997). Membrane fusion is followed by a poorly understood uncoating event that releases an intracellular reverse transcription complex. Reverse transcription is catalyzed in the cytosol by reverse transcriptase (RT). The accessory protein Vif appears to be important during one or more of these early events, perhaps by facilitating the initial stages of reverse transcription. RT-dependent DNA synthesis relies on the viral (nucleocapsid (NC) proteins and is initiated by the binding of a cellular tRNALys primer. Although the process of reverse transcription is complex, the mechanism of RT dependent DNA synthesis has emerged from extensive *in vitro* and *in vivo* studies. Once synthesized, the viral DNA is transported to the nucleus as part of a pre-integration complex that appears to include the integrase (IN), matrix (MA), RT, and Vpr proteins, as well as the cellular host protein HMG-I(Y) (Miller *et al*., 1997). The HIV capsid (CA) proteins do not appear to be part of the pre-integration complex, although they contribute to the structure of other retroviral pre-integration complexes (Bowerman *et al*., 1989). Nuclear localization of the pre-integration complex is directed by the accessory protein Vpr (Fouchier *et al*.,

1996; Freed *et al*., 1995; Nie *et al*., 1998), which does not contain a nuclear localization signal, but appears to function by connecting the pre-integration complex to the cellular nuclear import machinery, including importin-α and the nucleoporins (Fouchier *et al*., 1997; Popov *et al*., 1998; Vodicka *et al*., 1998). Vpr also interferes with normal cell cycle control by arresting the growth of infected cells in the G2 phase (Jowett *et al*., 1995; Re *et al*., 1995; Rogel *et al*., 1995). Nuclear localization may be facilitated by the MA proteins (Bukrinsky *et al*., 1993; von Schwedler *et al.*, 1994), although this proposal has been questioned (Nie *et al*., 1998; Reil *et al*., 1998). After active transport to the nucleus, the viral DNA is covalently integrated into the host genome by the catalytic activity of IN.

The late phase of the virus life cycle begins with the synthesis of unspliced and spliced mRNA transcripts, which are transported out of the nucleus for translation (Figure 1.4). Initially, short spliced RNA species that encode the regulatory proteins Tat, Rev and Nef are synthesized. Tat is an essential transcriptional activator that binds to a stemloop element of the nascent RNA transcript (TAR, for trans-activating response element) and recruits the cellular proteins cyclin T and cyclin-dependent protein kinase-9 (Cdk9; previously called TAK or PITALRE). Studies indicate that cyclin T binds directly to Tat, enhancing its affinity and altering its specificity for the TAR RNA (Wei *et al*., 1998). Cdk9 then phosphorylates the RNA polymerase II transcription complex, stimulating transcription elongation (Reines *et al*., 1996; Wei *et al*., 1998). Ordinarily, unspliced cellular mRNAs are retained in the nucleus where they can be further processed or degraded. However, full length and singly spliced HIV mRNA transcripts that contain functional introns are needed in the cytoplasm for Gag and Gag-Pol synthesis and packaging. Their export is mediated by the essential HIV accessory protein Rev. Rev binds as an oligomer to the Rev response element (RRE) of nascent unspliced mRNAs and recruits the cellular nuclear shuttling protein exportin-1 (XPO) (Ohno *et al*., 1998) and the nuclear export factor Ran guanosine triphosphatase in its GTP-bound form (Pollard and Malim, 1998). This complex is then transported through the nuclear pore to the cytosol where GTP is hydrolyzed to GDP, the complex dissociates, and the amino-terminal nuclear localization signal (NLS) of Rev directs its import back into the nucleus (Emerman and Malim, 1998). In this manner, Rev functions as a switch

between the early synthesis of highly spliced mRNAs (encoding Tat, Rev and Nef) and the later synthesis of unspliced (encoding the Gag and Gag-Pol proteins) and singly spliced (encoding Env, Vpu, Vif and Vpr) mRNAs. The Env precursor polyprotein (gp160) is synthesized in the endoplasmic reticulum (ER) using the spliced Env mRNA gene as the message. The protein appears to oligomerize to a trimeric structure in the ER, and is heavily glycosylated (Chan *et al*., 1997; Earl *et al*., 1991). Env is posttranslationally modified in the ER and Golgi apparatus and is cleaved to produce the non-covalently associated (TM-SU) 3 trimeric glycoprotein complex. The heterogeneously glycosylated TM-SU trimer is then transported to the cell membrane for virus assembly. Env and CD4 molecules are both synthesized in the ER, and the premature binding of CD4 to Env in the ER can inhibit translocation of Env to the cell membrane or the formation of a fully functional TM-SU complex (Hoxie *et al*., 1986). Thus, CD4 is targeted for removal from the ER by the viral accessory protein Vpu, which binds CD4 molecules and signals their degradation via the ubiquitin-proteasome pathway (Crise *et al*., 1990; Margottin *et al*., 1998; Schubert *et al*., 1998). Similarly, cell-surface CD4 molecules are targeted for endosomal degradation by the binding of the accessory protein Nef, which also binds to the AP-2 adapter complex and stimulates the formation of clathrin-coated pits (Foti *et al*., 1997). The down-regulation of CD4 molecules on the surface of infected cells may also serve as a means for avoiding an immune response. The Gag polyprotein is synthesized in the ribosomes from the unspliced mRNA. A translational frameshift results in the generation of smaller amounts of Gag-Pol precursor proteins, which associate with the Gag polyprotein at the cellular membrane. The N-terminally myristoylated MA domain of the polyproteins directs binding to the cellular membrane. Approximately 1200 to 2000 copies of Gag bud to form an immature particle, which encapsidates two copies of the unspliced viral genome. Subsequent to budding, the polyproteins are cleaved by PR to produce the independent enzymes, as well as the MA, CA and NC structural proteins. The structural proteins rearrange via a process called maturation to form the infectious virus particle. Cleavage of Gag appears to occur via an ordered, sequential cleavage process that is controlled by different intrinsic proteolysis rates at the different cleavage sites (Wiegers *et al*., 1998). Other factors may also be important for infectivity. For example, HIV-1 requires the packaging of the

cellular protein cyclophilin A, whereas HIV-2 and most other primate lentiviruses do not (Franke *et al*., 1994b). In addition, Vif is required for the production of infectious virions from some, but not all, cell lines (Courcoul *et al*., 1995; Sova *et al*., 1995; von Schwedler *et al*., 1993). Although virions package 7-100 copies of Vif (Camaur and Trono, 1996; Fouchier *et al*., 1996; Liu *et al*., 1995), it is not known if this packaging is essential (Camaur and Trono, 1996). Vif may function by interacting with cellular factors rather than with viral components (Simon *et al*., 1998).

1.6.1 The Envelope Proteins

The envelope glycoproteins of HIV-1 are synthesized as precursors, known as gp160. This precursor is proteolytically cleaved to generate two associated subunits, gp120 and gp41 (Allen *et al*., 1985). These two subunits are non-covalently associated. The surface glycoproteins play a critical role in the events that lead up to viral infection of the cell. These include virion attachment to the cell and fusion of the viral and cellular membranes. The envelope glycoproteins are also the major targets for the antiviral immuno-response in infected hosts.

The surface subunit (gp120) and the transmembrane subunit (gp41) remain non-covalently associated and most likely oligomerize as trimers (trimeric spikes) on the surface of the virion (Chan *et al*., 1998). During virion attachment, the gp120 subunit binds to the CD4 receptor and a co-receptor (one of the seven-transmembrane proteins of the chemokine family), which are present on susceptible cells such as T-lymphocytes and macrophages. Binding to the CD4 receptor leads to conformational changes in both the gp120 and gp41 subunits. The gp120 sequences of different immunodeficiency viruses have five variable regions (V1-V5) (Starcich *et al*., 1986). The first four variable regions form surface-exposed loops that contain disulphide bonds at their bases. Whereas the V3 loop contributes to co-receptor binding (Rizzuto and Sodroski, 2000), the V1-V2 loops are dispensable for viral infection of target cells (Kolchinsky *et al.,* 2001). The conserved gp120 regions (C1-C5) contain discontinuous structures that interact with the gp41 exodomain and with viral receptors on the target cells. The five constant regions are functionally indispensable as they constitute the

core molecule and contribute most of tne interactions with CD4, co-receptors and gp41 (Kwong *et al.,* 1998). Both the variable and conserved regions of the gp120 are extensively glycosylated. The variable loops present a myriad of different epitopes to the immune system and therefore, many of the protein's determinants contained in the variable loops serve as "decoy" targets, directing the antibody response away from the functionally conserved regions of gp120. The variable loops also shield neutralization-sensitive surfaces, such as the conserved receptor binding regions, from antibodies by steric blockade (Wyatt and Sodroski, 1998). Therefore, deletion of the variable loops V1, V2 or V3 has been suggested as a means to generate envelope glycoprotein-based immunogens that preferentially elicit antibodies against conserved regions of gp120 such as the CD4 binding site (Barnett *et al.,* 2001).

1.6.2 Conformational changes in gp120 and gp41

The conformational changes in both the HIV-1 envelope glycoproteins, which are induced by the gp120/CD4 binding, are likely to promote the subsequent steps in virus entry (Sattentau *et al*., 1993). These conformational changes were found to create a high-affinity binding site on gp120 for the CCR5 chemokine receptor. Some of the gp120 variable loops are believed to change conformation or become more exposed upon the binding of soluble CD4 (sCD4) (Sattentau *et al*., 1995). Using sCD4, which is thought to mimic membrane anchored CD4, and two cross-competing monoclonal antibodies, (17b and CG10), which recognise CD4-inducible gp120 epitopes, the nature and functional significance of gp120 changes, which are initiated by the CD4 binding, were investigated (Sullivan *et al*., 1995). The CD4 induced exposure of the 17b epitope on the oligomeric envelope glycoprotein complex was found to involve the movement of the gp120 V1/V2 variable loops.

The gp41 subunit undergoes conformational changes that mediate fusion of the viral membrane and the target cell membrane (Allen *et al*., 1985). The structure of the gp41 has been proposed and reviewed (Chan and Kim, 1998). The schematic structure of gp41 is illustrated in Figure 1.8. The HIV envelope glycoprotein complex has several features in common with other viral membrane-fusion proteins, which include the hemagglutinin (HA) protein of

influenza virus. In the case of HIV gp160, it is proteolytically processed to generate a receptor binding subunit (gp120 like HA_1) and a membrane spanning subunit (gp41 like HA_2) which contains the "fusion peptide" (FP). In both gp41 and HA_2, the fusion peptide region begins immediately at the new amino terminus that arises as a result of the proteolytic processing of the influenza virus and HIV precursors, HA and gp160 respectively. The fusion peptide is hydrophobic and glycine rich. Secondly, there are two regions with 4-3 hydrophobic heptad repeats. This sequence motif is characteristic of coiled coils. Between the two heptad repeats is a loop region containing two cysteins.

The immunodeficiency that defines AIDS is due primarily to the progressive decline in the number and function of $CD4^+$ T-cells. A number of mechanisms for this decline have been investigated and demonstrated. Direct viral destruction of the $CD4^+$ cells is not regarded as contributing significantly to this decline in the number of $CD4^+$ cells (Terai *et al.,* 1991). There is increasing data that, within HIV-infected individuals, the depletion of $CD4^+$ T-cells is secondary to enhanced lymphocyte apoptosis mediated by HIV, which is probably caused by the interference in the physiological mechanisms that are maintained in lymphocytes by the Fas and Tumor Necrosis Factor (TNF) signalling pathways, initiated by the engagement of the respective ligands, Fas ligand (FasL) and TNF ligand (TNFL) (Krammer *et al*., 1994; Sytwu *et al.,* 1996; Van Parijs *et al.,* 1996). In HIV, a large proportion of the peripheral blood $CD4^+$ T-cells are found to be highly susceptible to FasL-mediated apoptosis (Katsikis *et al.,* 1995; Badley *et al.,* 2000). It has also been shown that spontaneous and anti-Fas-induced apoptosis of $CD4^+$ T-cells from HIV positive individuals is higher in symptomatic than in asymptomatic individuals, with symptomatic individuals being defined as those which present with Kaposi's sarcoma or opportunistic infections indicating an association with disease severity (Badley *et al.*, 2000). No difference could be shown for $CD8^+$ T-cell counts. Similarly, it has been demonstrated that there is an inverse correlation with the peripheral blood $CD4^+$ T-cell counts and the spontaneous and anti-Fas-induced apoptosis of the $CD4^+$ T-cells in both the asymptomatic subset of the population and the combined asymptomatic and symptomatic individuals, indicating that apoptosis may be involved in disease progression. The induction of apoptosis by the HIV envelope protein gp120 in

$CD4^+$ T-cells has also been demonstrated (Katsikis *et al.,* 1995).

1.7 The gp41 fragment of HIV-1

The structure of the gp41 was elucidated by Chan *et al*., (1997). HIV-1 gp41 is composed of three domains, an extracellular domain (ecto-domain), a transmembrane domain, and an intracellular domain (endo-domain). The ecto-domain contains three major functional regions, namely the fusion peptide located on the N-terminus of gp41, followed by two 4-3 hydrophobic heptad repeats, designated N36 and C34. Protein dissection studies demonstrated that the two 4-3 hydrophobic heptad repeat regions within gp41 form a helical trimer of antiparallel dimers (Lu *et al*., 1995). The crystallographic analysis confirmed that the gp41 core comprises a six helix bundle in which the N- and C- helices (N36 and C34) are arranged into three hairpins (Chan *et al*., 1997). Three central helices comprising N-peptides are arranged in a trimeric coiled-coil, and three outer helices comprising the C-peptides, are packed in anti-parallel, hydrophobic grooves on the surface of the coiled coil illustrated in Figure 1.8.

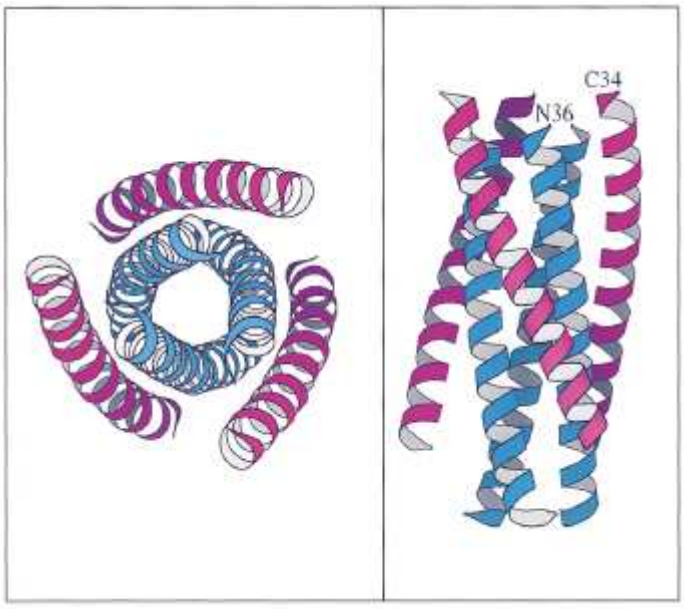

Figure 1.8: **Overall views of the N36/C34 complex of gp41**. The left panel shows an end-on view of the N36/C34 complex looking down the three-fold axis of the trimer. The right panel shows a side view with one N36 and one C34 helix labelled (Chan *et al.,* 1997).

The C peptides are highly conserved. It is believed that this structure represents the "fusion-active" conformation of gp41. The transmembrane domain consists of twelve hydrophobic amino acids extended C-terminally with 3 hydrophobic amino acids. The endo-domain contains three major functional regions, namely the Kennedy domain located on the the N-terminus of gp41, followed by two amphipathic calmodulin binding domains, designated LPP2 and LPP1.

1.7.1 Importance of gp41

There are 2 major questions relating to the mechanism of action of viral gp41: How they are involved with other components to facilitate the fusion of HIV and the host T lymphocytes, and how they are involved in antigenicity.

1.7.1.1 Gp41 facilitated fusion

The high-resolution determination of the gp41 core structure of HIV-1 provides well defined landmarks in the cadre of how the viral envelope glycoproteins navigate fusion after CD4 and co-receptor-induced conformational changes. A number of steps in the fusion cascade are revealed by inhibitory conditions that block fusion. The first step is the initial binding and engagement with CD4 receptors. The gp120/41 complex then undergoes conformational changes, allowing for the N-terminal heptad repeat region of gp41 to be exposed. This conformational change in gp41 is now defined as the pre-hairpin formation. Thirdly, co-receptors are recruited to the binding site, presumably into rafts rich in cholesterol and glycosphingolipids. Self-aggregation and/or interaction of several envelope/CD4 complexes may also occur at this step. The fourth step is the engagement of the co-receptor. Concomitant with co-receptor engagement, gp120 and gp41 dissociate and the six-helix bundle configuration is formed. This is also presumed to be the point at which the fusion peptide is inserted into the target and/or host membrane, leading to the establishment of a fusion pore. The final step is fusion pore formation (Salzwedel *et al*., 1999).

1.7.1.2 Apoptotic function

Programmed cell death (PCD), as its name suggests, is an orchestrated biochemical process that leads ultimately to the demise of the cell. Included under this term are three broad categories by which a cell may die. Firstly and most typically is apoptosis, often characterized by chromatin condensation, phophatidylserine exposure, cytoplasmic shrinkage, membrane blebbing and caspase activation. Secondly, in a process quite analogous, 'apoptosis-like PCD' was identified, which presents with some of the apoptotic features but lacks the condensed packed chromatin. The third category is 'necrosis-like PCD', which is characterized by the general absence of chromatin condensation but is distinguished from necrosis (death resulting from injury) through its use of a signalling pathway. Apoptosis is the best characterized form of PCD. It is a crucial component of normal multicellular life, playing a key role in development and immunity (Leist and Jaattela 2001; Shi *et al.,* 2001; Tilly *et al.,* 2001).

The envelope glycoprotein complex (Env), encoded by HIV-1, can induce apoptosis via at least four independent mechanisms. A brief overview of the four mechanisms is given below:

The env glycoprotein precursor protein (gp160) undergoes proteolytic maturation to generate gp41 and gp120. Soluble gp120 can stimulate a pro-apoptotic signal via an action on chemokine receptors namely CXCR4 and CCR5 (Twu *et al.,* 2002). It can induce apoptosis in lymphocytes, but also in neurons, cardiomyocytes, kidney epithelial cells and hepatocytes. In lymphocytes, the effect of gp120 mainly involves interactions with CD4 and CXCR4 for lymphotropic env variants and CCR5 for monocytotropic env variants (Figure 1.9). Although soluble gp120 has been detected in body fluids, it is doubtful whether it could reach concentrations high enough to induce cell death *in vivo* (Klasse *et al.*, 2004).

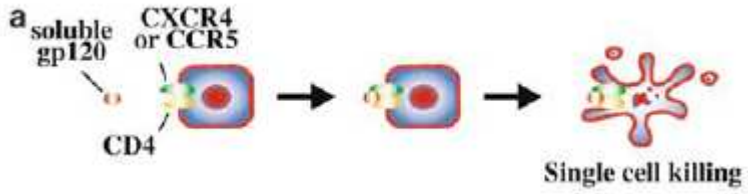

Figure 1.9: **Cell killing by soluble gp120**. Soluble gp120 interacting with CD4 and chemokine receptors (XCR4 or CCR5) can cause a single cell killing (Perfettini *et al.*, 2004).

The membrane-anchored gp120/gp41 complex expressed on the surface of HIV-1 infected cells can induce apoptosis through an interaction with uninfected cells expressing the CD4 receptor and the chemokine co-receptors CXCR4 or CCR5. This type of bystander killing is obtained by at least three distinct mechanisms. First, single cell killing by membrane-anchored Env called 'the kiss of death'. The two interacting cells in Figure 1.10 may not fuse entirely, but simply exchange plasma membrane lipids, after a sort of hemifusion process, followed by rapid death (Blanco *et al.*, 2003). It is associated with gp41 mediated transfer of plasma membrane lipids from the membrane of Env expressing cells to the target cell. It occurs without detectable cytoplasmic mixing and thus involves a hemifusion-like event. It is still not clear how this event takes place. It is not clear if surface proteins are transferred during the cell-cell contact, or whether a simple membrane perturbation or more specific signals lead to cell death. The exact cellular signals involved in target cell killing remain largely unexplored apart from the fact that they do not involve CD95 which is an important part of Fas and FasL mediated apoptosis. (Roggero *et al.,* 2001)

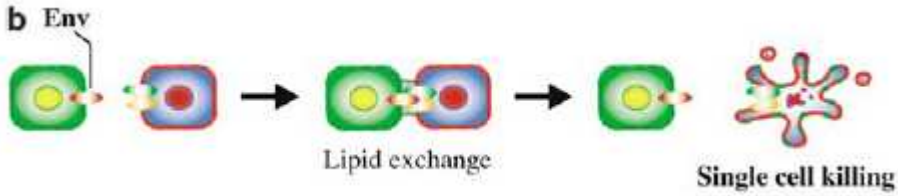

Figure 1.10.**Single cell killing by membrane-anchored Env (The kiss of death).** Interaction between Env and CD4/CXR4 (or CCR5) over-expressing cells can trigger single cell killing through transient interactions with hemifusion-mediated exchange of lipids on the membranes of the interacting cells (Perfettini *et al.*, 2004).

Second, the syncytial apoptosis induced by env arises by the interaction between the two cells which can induce cellular fusion (Sylwester *et al.,* 1997) followed by nuclear fusion within the syncytium (Ferri *et al.,* 2000b). Syncytia are condemned to die from apoptosis after latency phase, as depicted in Figure 1.11, presumably when conflicts in cell cycle between the daughter nuclei are detected or when the polyploidy checkpoint is activated (Zhivotovsky *et al.,* 2004). This mode of cell death induction may participate in the AIDS-associated depletion of $CD4^+$ T-cells (Lifson *et al*., 1986a). Specific signals mediated by plasma membrane receptors and/or a perturbation of cellular metabolism induced by the fusion event can trigger a series of pro-apoptotic kinases. Phosphorylation of the inhibitor of NF-KappaB on Ser 32 and on Ser 36 induces its degradation and initiates NF-B-dependent transcription. Cyclin B-dependent kinase-1 (Cdk1) is induced and causes an abortive entry into the mitotic prophase, resulting in the fusion of several nuclei within the same cytoplasm. mTOR and p38MAPK, two kinases activated by cell fusion, phosphorylate p53 on Ser 15 and 46, thereby facilitating p53-dependent transcription. Puma and Bax, two p53 target genes, are induced and trigger mitochondrial membrane permeabilization (MMP) and consequent caspase activation (Perfettin *et al.,* 2004).

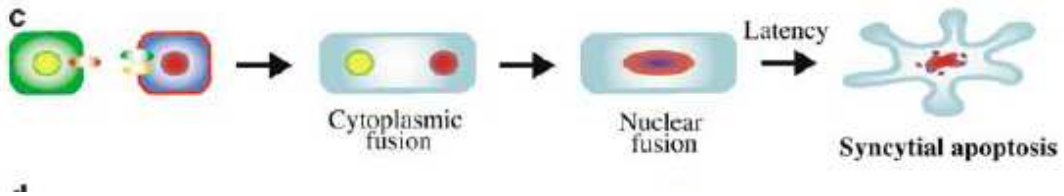

Figure 1.11: **Syncytial apoptosis induced by Env**. Env-expressing cells can fuse with Env-negative cells expressing suitable combinations of receptors and co-receptors resulting in syncytium formation and death after a period of latency (Perfettini *et al.,* 2004).

The third mechanism, contagious apoptosis facilitated by the HIV-1 envelope is indicated in Figure 1.12. It is possible that virus-infected cells, on the point of undergoing apoptosis, fuse with CD4-expressing cells, in which case apoptosis is rapidly transmitted from one cell to the other and thus occurs in contagious fashion (Andreau *et al.,* 2004).

Figure 1.12: **Contagious apoptosis facilitated by the HIV-1 envelope**. Pre-apoptotic Env-positive cells can kill $CD4^+$ T-cells upon fusion, in a rapid contagious process (Perfettini *et al.,* 2004).

When apoptosis is initiated in the Env positive cells, pre-apoptotic Env positive cells can fuse with healthy CD4 cells, thus forming syncytia that rapidly die through caspase-dependent apoptosis. This “contagious” apoptosis is exclusively observed when treated Env positve cells show a pre-apoptotic chromatin condensation (PACC), a phenomenon that precedes caspase activation and mitochondrial membrane permeabilization (MMP). PACC positive cells manifest other biochemical alterations such as DNA strand breaks and DNA foci with phosphorylated histone H2AX on Ser 135 and Chk2 on threonine 58. The transmission of lethal signals requires cytoplasmic (but not nuclear) fusion and leads first to PACC of the nuclei originating from the $CD4^+$ T-cells and then to full-blown apoptosis with MMP, caspase activation, nuclear pyknosis and chromatolysis (Prefettini *et al.,* 2004).

In conclusion, it is clear from the literature that the gp41 fragment may be a target for vaccines design but it may also play a specific role in the apoptosis of non-infected $CD4^+$ T-cells.

1.8 Hypothesis and Aims

Gp41 contains a sequence that can be used as a component for a safe vaccine which is antigenic and do not elicit autoimmunity or apoptosis.

Primary aims of the study:

1. To create an understanding of the degree to which HIV positive patients contain antibodies to various fragments of the gp41 fusion peptide.
2. To determine the correlation between the CD4 T-cell count and the antibodies to gp41 derived peptides in the sera of HIV positive patients.
3. To determine whether and how gp41 peptides are involved in CD4 T-cells apoptosis.

Chapter 2: Focuses on the PCR mediated synthesis of HIV-1 gp41 for functional expression in *E. coli* and *P. pastoris* expression host systems to obtain soluble protein.

Chapter 3: Describes the immunological characterization of synthetic HIV-1 gp41 by means of ELISA, western blotting and apoptotic studies.

Chapter 2

PCR mediated synthesis of the synthetic HIV-1 Subtype C gp41 gene and the expression of the synthetic gp41 gene in *E.coli* and *Pichia pastoris*

2.1 Functional regions of the gp41 gene

Gp41 plays a critical role in the fusion of the viral membrane and the host T lymphocytes and is also one of the major targets for the antiviral immune response in infected hosts. As indicated in Figure 2.1, HIV-1 gp41 is composed of three domains: an extracellular domain (ecto-domain), a transmembrane domain, and a cytoplasmic domain (endo-domain). The ecto-domain contains three major functional regions, namely the fusion peptide located on the N-terminus of gp41, followed by two 4-3 hydrophobic heptad repeats, designated N36 (N-terminal heptad repeat) and C34 (C-terminal heptad repeat). gp41(80) which will be investigated in this project is located on the ecto-domain of gp41.

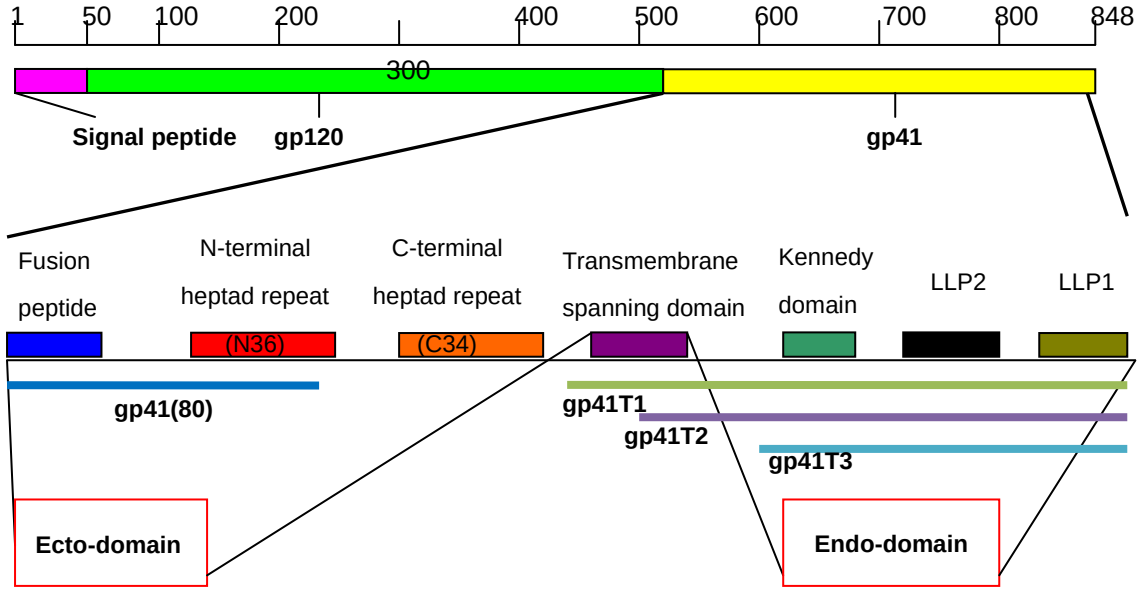

Figure 2.1: **Organization of the HIV-1 env gene product and functional domains of gp41.** The functional domains of gp41 are the following: Fusion peptide indicated in blue, N-terminal heptad repeat (N36) in red, C-terminal heptad repeat (C34) in orange, membrane spanning domain in purple, Kennedy domain in green, amphipathic "calmodulin-binding" domain LLP2 in black and amphipathic "calmodulin-binding" domain LLP1 in khaki (Chan *et al.,* 1997). The location of gp41(80), gp41T1, gp41T2 and gp41T3 which will be synthesized is also indicated.

The transmembrane domain is 22 amino acids in length and consists of 12 hydrophobic amino acids but also extended C-terminally with 3 hydrophobic amino acids. The long cytoplasmic domain (CPD) of the envelope glycoprotein is one of the unique characteristics of lentiviruses which distinguish them from other retroviruses. In HIV-1, the CPD of gp41 is approximately 150 amino acids (aa) in length and highly conserved (Day *et al.,* 2004, Wyma *et al.,* 2004). Despite numerous studies of HIV gp41 the full functional significance of the CPD remains unresolved. However, it plays an important role in the incorporation of envelope proteins into virions during maturation and modulates viral infection in a cell specific manner (Iwantani *et al.,* 2001, Piller *et al.,* 2000). The full-length CPD of gp41 has undoubted importance *in vivo*, because HIV engineered to contain C-terminal CPD truncations of more than 20 amino acids are almost invariably non-infectious.

The cytoplasmic domain consists of 3 major functional regions, namely the Kennedy domain located on the N-terminus of gp41. Gp41T1, gp41T2 and gp41T3 peptides are also located on the cytoplasmic domain of gp41 as indicated in Figure 2.1 and will be investigated in this study. The Kennedy domain is followed by two amphipathic calmodulin binding domains, designated LLP2 and LLP1. The Kennedy peptide, PRGPDRPEGIEEEGGERDRDRS, contains a conformationally dependent neutralizing epitope (ERDRD) and a linear non-neutralizing epitope (IEEE) (Cheung *et al.,* 2005). Therefore, the Kennedy peptide from the gp41 of HIV-1 can elicit non-neutralizing as well as neutralizing antibodies. The non-neutralizing IEEE epitope is immunodominant over the neutralizing ERDRD epitope (Cheung *et al.,* 2005).

Synthetic peptides corresponding to the LLP1 and LLP2 regions have been shown to bind to calmodulin (CaM) with high affinity in the presence of calcium (Miller *et al*., 1993; Srinivas *et al*., 1993). These amphipathic peptides inhibit CaM-regulated activation of bovine brain phosphodiesterase *in vitro*, and have also been shown to inhibit mitogen-induced lymphocyte activation, a property shared by CaM antagonists. These properties are not an artefact of peptides since purified HIV-1 gp160 as well as full-length gp41 binds to CaM and expression of HIV Env in cells results in CaM co-localization (Radding *et al*.,

1996). Moreover, while CaM binding by the C-terminal amphipathic helix is very sensitive to amino acid changes, the natural variation found in this region of Env leaves CaM-binding intact (Tencza *et al*., 1997). Since calmodulin is intimately involved in the metabolism and functioning of T-cells, these studies point to a novel mechanism of viral cytopathogenesis mediated by the interaction of the HIV TM protein with cellular CaM that could in turn uncouple critical cellular signal transduction pathways (Miller *et al*., 1993; Srinivas *et al*., 1993).

2.1.1 Functional regions of the synthethic gp41 ecto and cytoplasmic domain

1 2 Anchor Domain

REISNYTNIYRLLEDSQNQQEKQNEKDLLALDKWQNLWSWFDITNWLWYIKIFIMIVGGLIGLRIIFAI
1 36

3 Kennedy Domain

LSIVNRVRQGYSPLSFQTLTPSPRGPDRLGRIEEEGGEQDRDRSIRLVGGFLALAWDDLRSLCLFSYHR
95

LRDLILIATRVVELLGRSSLRGLQRGWEILKYLGSLVQYWGLELKKSAINLLDNITAIAVAEGTDRIIE

IIQRICRAIYHIPRRIRQGFEAALQ
221

Figure 2.2: **Synthesised gp41 sequences of HIV-1 subtype C (UG268A): 1-3:** 1: gp41T1 starts from the amino acid number 1 (purple R) up to amino acid 221 (purple Q). 2: gp41T2 starts from the amino acid number 36 (green N) up to amino acid 221 (purple Q). 3: gp41T3 starts from the amino acid number 95 (Yellow P) up to amino acid 221 (purple Q). **The Black bold underlined parts:** Glycosylation signals where NIT: N-linked glycosylation signals and NYT: CXCR4 N-linked glycosylation signal. **Blue Anchor Domain:** Consist of 12 hydrophobic amino acids and extended C-terminally with 3 hydrophobic amino acids. **Pink ALDKWQ:** Epitope for Ab, but there is no evidence that there is neutralization Ab formed against this epitope. **Light green Kennedy Domain:** Contains a conformational dependent neutralizing epitope EQDRD and a linear non-neutralizing IEEE epitope. **Red amino acids:** Are conserved amino acids and translation-specific cysteine residues (Cheung *et al.,* 2005, Kennedy *et al*., 1986).

The HIV-1 gp41 gene was truncated stepwise to produce three segments coding for three different peptides which vary in length namely 405 bp, 591 bp and 696 bp as indicated in Figure 2.2. The reason for this was to construct fragments of gp41 which do not contain the helices and which vary in length for the determination of the smallest antigenic peptide, which may be used as a vaccine target or even for chemotherapy in the treatment or prevention of HIV/AIDS.

2.1.2 Gene expression

2.1.2.1 Codon optimization for functional expression in *E.coli*

The expression of functional proteins in heterologous hosts is the cornerstone of modern biotechnology. Unfortunately, many proteins are difficult to express outside their original contexts. Proteins may contain expression limiting regulatory elements, come from organisms that use non-canonical nucleotide codes or from a gene rife with codons rarely used in the desired expression host. Improvements in the speed and efficiency of gene synthesis have rendered feasible complete gene redesign for maximum protein expression. For example, protein expression can improve dramatically when the codon frequency of the gene in the study is matched to that of the host expression system. Despite the many advantages of *E. coli*, high-level expression is not routinely achieved. The many reported causes preventing efficient heterologous protein production in *E. coli* include biased codon usage, gene product toxicity, solubility, mRNA secondary structure, and mRNA stability (Kane *et al*., 1995). In addition, rare codon gene expression can lead to translational errors as a result of ribosomal stalling at a position requiring incorporation of amino acids coupled to minor tRNAs, or even at sites requiring major tRNAs, but which are depleted because of over utilization of a particular amino acid (Kane *et al*., 1995). The mistranslational events related to rare tRNAs are observed as codon misreadings or processing errors and they manifest themselves as amino acid substitutions or frameshift events. Specifically, the rare arginine (AGG, AGA, CGG, and CGA), leucine (CUA), isoleucine (AUA) and proline codons (CCC) often lead to frameshift errors and ultimately to undesired products. Accordingly, several systems are now available to alleviate many of the problems through tight transcriptional regulation, or through heterologous expression of tRNAs to

circumvent problems with rare codon usage (Kane *et al*., 1995).

2.1.2.2 Expression systems for synthetic genes

In 1977 it was shown that a gene from a higher eukaryote could be expressed in a micro-organism, *E.coli*, to produce a biologically active protein (Kingsman *et al*., 1985). It appears, however, that *E.coli* may not be the most suitable host for the expression of all proteins. A disadvantage of the *E.coli* system is the necessity to renature the heterologous polypeptides, since in most cases this organism was found to be unable to produce the proteins in a properly folded, soluble form, i.e. not ending up as insoluble matter in inclusion bodies. Moreover, many eukaryotic proteins depend on post-translational modifications such as glycosylation, which prokaryotic hosts are unfit to perform (Gellissen *et al*., 1992). In many instances yeast became the preferred hosts for heterologous protein and gene expression. The best-characterized yeast species, *Saccharomyces cerevisiae*, was the first to be used for the successful production of eukaryotic proteins such as human α-interferon, hepatitis B surface antigens and enzymes like calf pro-chymosin and *Aspergillus* glucoamylase (Gellissen *et al*., 1992). Therefore, this enables researchers to perform functional complementation in yeast.

2.1.2.2.1 *Pichia pastoris*

P. pastoris was initially developed by Phillips Petroleum Company for the production of single cell protein and is now probably the non-conventional yeast in which the most recombinant proteins have been expressed. Since *P. pastoris* has no native plasmids, expression of the foreign gene was achieved by chromosomal integration, using integrative plasmids.

As methylotrophic yeast, *P. pastoris* can grow in methanol as the sole carbon and energy source. Key enzymes in this pathway may comprise up to 20-35% of the total intracellular protein content in methanol-grown cells. In this system, almost all the foreign genes have been expressed under control of the *P. pastoris* alcohol oxidase 1 (the first enzyme of the pathway) promoter. This promoter is tightly regulated by a carbon source-dependent repression/induction mechanism. Its expression is undetectable in cells grown on glucose or glycerol,

but is maximally induced during growth on methanol (Waterham *et al*., 1997).

In *P. pastoris* the foreign gene expression cassette is normally integrated in the host genome using *HIS4* as a selectable marker. For optimal genetic stability, the expression cassette is cleaved from the vector and transplanted into the chromosomal *AOX1* locus. This results in the site-specific eviction of the *AOX1* structural gene to create a strain that grows slowly in methanol (Mut^-). Alternatively, the vector can be linearized and targeted to integrate by non-disruptive insertion into *HIS4*, generating a strain that grows normally on methanol (Mut^+). For intracellular expression, it is preferable to use Mut^- cells because they have a lower level of alcohol oxidase protein (Sreekrishna *et al*., 1997).

P. *pastoris* was engineered for the expression of a variety of heterologous proteins like Type II human collagen (Ruottinen *et al*., 2008) to mention one.

2.1.2.2.2 Cell-free expression

As early as the 1950s, researchers realized that protein biosynthesis does not require the integrity of the cell and can continue after cell disruption. Eventually, Zamecnik and his colleagues (Littlefield *et al*., 1955) made a real cell-free ribosomal system of protein synthesis (translation) based on mitochondria-free cytoplasmic extracts of animal cells (rat liver). Soon after, cell-free translation systems based on bacterial (*Escherichia coli*) extracts were made by German and American groups (Schachtschabel and Zillig, 1959). Ribosomes in all of those systems, however, were programmed with endogenous mRNA: they were simply reading the messages to which they had been already attached at the time of cell disruption.

A revolutionary step in the development of cell-free translation system was the introduction of exogenous messages. This was first done in 1961 with a bacterial system (Nirenberg and Matthaei, 1961). Preincubation of the cell extract of *E.coli* at physiological temperature was sufficient to remove the endogenous mRNA from ribosomes. The vacant ribosomes in the extract were found to accept either exogenous natural mRNAs or synthetic polyribonucleotides as template for

polypeptide synthesis. Later, an efficient cell-free system for translation of exogenous messages was made from rabbit reticulocyte lysate: it was freed from endogenous mRNAs by pre-treatment with micrococcal Ca-dependent RNase (Pelham and Jackson, 1976). Another cell-free system, derived from wheat germ extract could be used for expression of exogenous messages without pre-treatment, because it had intrinsically low levels of endogenous mRNA (Roberts and Peterson, 1973). Thus, the *E.coli* extract, the rabbit reticulocyte lysate and the wheat germ extract became the three commonly used extracts and are now the best systems for cell-free protein synthesis.

2.1.3 Objective of the study

The three expressed proteins of the cytoplasmic tail of gp41 and the ecto-domain gp41(80) fusion protein are to be used in antigenicity studies for the determination of Abs directed to the specific gp41 peptides. Therefore, the following aspects were addressed.

- The gp41(80) peptide (80 amino acids in length) located on the ecto-domain of gp41 will be constructed. The gene coding for gp41(80) will be cloned into the pMAL-c2E vector system and expressed as a fusion peptide, which is coupled to a maltose binding protein (MBP-2).
- The gene coding for the cytoplasmic tail of gp41 will be truncated to produce gene products that will be expressed as three peptides of different lengths (405 bp, 591 bp, 696 bp) and the codon optimized peptide named 052741 which are based on the cytoplasmic tail 696bp peptide.
- These genes will be cloned into different vector systems namely pMAL-c2E, pET28a, pJC20 and pPIC9. The pMAL-c2E vector for the expression of a fusion peptide for purification reasons. The pET28a vector for the expression of a protein with an N-terminal His-tag to assist in protein purification. The pJC20 vector for the fact that it is known to produce increased levels of protein (*M.M. Botha,* 2006). The pPIC9 vector for expression of the protein in *P. pastoris* for the glycosylation of the protein.

2.2 Methods

2.2.1 Construction of pMAL-gp41(80)

2.2.1.1 PCR product preparation

The PCR reaction was done using a 5' primer (0113F), which contains a *Kpn*I site for an in-frame fusion with the maltose binding protein (MBP) in the vector pMAL-c2E (New England Biolabs). The 3' primer (0113R) was designed with an *Xba*I site. The primers are listed below, with the restriction sites in bold.

0113F: 5' GAG**GTACCT**AGAGCAGTGGGA 3'
0113R: 5' ATAGAAAGACACG**TCTAGA**GC 3'

The PCR product was digested with *Kpn*I and *Xba*I in a two-step digestion, first with *Kpn*I (Amersham) in the supplied buffer, then with *Xba*I in the supplied buffer. The restriction enzymes were subsequently heat-inactivated.

2.2.1.2 pBluescripts KS^+ preparation and construction of $pBKS^+$: HIV

pBluescripts KS^+ was used for cloning the inserts. Vector DNA was digested with *Kpn*I and *Xba*I in a two-step digestion. Ligation of $pBSK^+$ and the insert was set up using a 20:1 insert to vector ratio. The reaction was set up using Epicentre Technology's FastLink DNA ligation Kit and incubated at room temperature for 10 minutes. The reaction mixture was subsequently dried down, washed with 1 ml 70% ethanol and resuspended in dH_2O. The DNA was electroporated into *Escherichia coli cells* (SP GEN 001) and the transformants selected for on LB agar (1% Tryptone, 0.5% yeast extract, 1% NaCl at pH 7 and agar) supplemented with 50 µg/ml ampicillin, 80 µg/ml IPTG and 0.1 mM X-Gal plates.

White colonies from the ligation were PCR-screened using 0113F and 0113R as the primer pair. Plasmid isolation was performed on selected colonies appearing positive from the PCR screen, and the DNA digested with *SacI* and *KpnI* to check for the presence of an insert. Unligated $pBSK^+$ was used as the control. Vector only will give bands of 2856 bp and 102 bp, while the insert should give bands of 2859 bp and 260 bp.

2.2.1.3 Construction of pMAL-c2E:HIV

Plasmid DNA of pBSK$^+$: HIV80 was isolated. To isolate the insert, the plasmid DNA was digested with *Kpn*I and *Xba*I, in a two-step digestion as described previously. The digest was electrophoresed on a 1% agarose TAE gel to separate the insert from the pBKS$^+$ vector. The insert was purified from the agarose.

The HIV80 insert (*Kpn*I/*Xba*I) was ligated to pMAL-c2E (*Kpn*I/*Xba*I, CIAP), using the FastLink DNA Ligation Kit at an insert to vector ratio of ~50:1. The ligation reaction was incubated for 5 minutes at room temperature, dried down, washed with 1 ml 70% ethanol and electroporated into TB1 cells. Transformants were selected for on LB agar plate supplemented with 50 µg/ml ampicillin. LB agar plates supplemented with 50 µg/ml ampicillin, 80 µg/ml IPTG and 0.1 mM X-Gal plates were not used for the initial transformant selection, as the *tac* promotor on pMAL-c2E controlling the expression of the MBP fusion protein is very strong. Therefore, direct selection on IPTG could result in mutants, which either lost the fusion gene or no longer express it at high levels.

Transformants on the LB agar plate supplemented with 50 µg/ml ampicillin were toothpicked onto an LB agar plate supplemented with 50 µg/ml ampicillin as well as an LB agar plate supplemented with 50 µg/ml ampicillin, 80 µg/ml IPTG and 0.1 mM X-Gal. Blue-white selection on the IPTG, X-Gal plate was used and the corresponding colonies on the LB with ampicillin plate were selected for PCR screening as described previously.

Plasmids from possible positive colonies were isolated and treated with restriction enzymes to check for the presence of an insert. As previously, two-step *Kpn*I/*Xba*I digest was tedious to perform and the *Kpn*I/*Sac*I digest resulted in very small bands, a different screening digest was done using *Bgl*II and *Sal*I. For pMAL-c2E, this should result in 5822 + 829 bp; HIV80: 5822 +1056 bp.

2.2.1.4 Transformation of pMAL-c2E:HIV constructs into *Escherichia coli* ER2507

E. coli ER2507 (New England Biolabs #E4121S)) is a strain that has no chromosomally-encoded MBP. Its genotype is: F- *ara-14 leuB6 fhuA2 Δ(argF-lac) U169 lacY1 glnV44 galK2 rpsL20 xyl-5 mtl-5 Δ(malB) zjc:Tn5(Kan^R) Δ(mcrC-mrr)$_{HB101}$ = PR700 fhuA* = RR1 *Δ(malB) Δ(lac)U169 pro$^+$ fhuA.*

Plasmid DNA of pMAL-c2E: HIV80 was electroporated into electro-competent ER2507 cells, and transformants selected for on LB + Amp$_{50}$ plates. Colonies were PCR screened, and possible positives used for plasmid isolation and restriction digestion with *Bgl*II and *Sal*I.

2.2.2. Cultivation of transformants and subsequent purification of the constructs

2.2.2.1 Cultivation of transformants

pMAL-c2E:Hiv 80 was cultivated in 50 ml Luria-Bertani (LB) liquid media (1% Tryptone, 0.5% yeast extract, 1% NaCl at pH 7) supplemented with 100 µg/ml ampicillin and grown overnight at 37 ℃ with shaking (220 rpm). The overnight cultures were diluted by the addition of 250 ml LB liquid medium (supplemented with 100 µg/ml ampicillin) to 2.5 ml of the overnight culture. This was followed by incubation at 37 ℃ with shaking until an optical density of approximately 0.6 (logarithmic growth phase) at 600 nm was reached. Consequently, 0.238 mg/ml isopropyl-β-D-thiogalactopyranoside (IPTG) (pEQ Lab Biotechnology, GmbH) was added to induce protein expression. After the addition of IPTG the cultures were incubated at 37 ℃ for 16 hours with shaking (220 rpm).

2.2.2.2 Purification of the construct

Protein extraction from expression in pMAL-c2E vector and purification was done in the following way. The cells were harvested by centrifugation at 16,300 x g for 10 minutes at 4 ℃ (Sorvall Plus RC-5C, GSA rotor, Sorvall, UK) and the pellet frozen for overnight storage (-20 ℃). The cells were allowed to thaw on ice and resuspended in the appropriate buffer. The cells were disrupted by sonication for

20 minutes using a 50% on / off cycle. The sample was centrifuged at 9000x g for 30 minutes. The supernatant was retained and was subsequently loaded onto the amylose resin. Amylose resin is a composite amylose/agarose bead functioning as an affinity matrix used for the isolation of proteins fused to maltose-binding protein (New England Biolabs, Boston, USA). The affinity purification of the recombinant protein using 1 ml Amylose resin entailed the following steps. The resin was equilibrated with 5 ml Column buffer (20 mM Tris-HCl, 200 mM NaCl, 1 mM EDTA, pH 7.4). Of the cell free extract, 1 ml was loaded onto the Amylose resin. The mixture was incubated for 2 hours at 4 ℃ with constant shaking at 50 rpm. The Amylose resin was washed three times with 1 ml Column buffer. The fusion peptide was eluted from the amylase resin with 1 ml Elute buffer (20 mM Tris-HCl, 200 mM NaCl, 1 mM EDTA, 10 mM maltose, pH7.4). The mixture was incubated for 2 hours at 4 ℃ with constant shaking at 50 rpm. The fusion peptide was transferred to a clean container and stored at 4 ℃ until further analysis.

2.2.3 Oligonucleotides for the construction of gp41T1, gp41T2, gp41T3 and codon optimized 052741

Oligonucleotides were obtained from IDT (Coraville, USA) and were dissolved in TE buffer (10 mM Tris, 1 mM EDTA pH 8.0) at 37 ℃ to give a final stock solution concentration of 100 µM. The stock solutions were stored at -20 ℃. Working stocks of 2.5 µM were prepared. Primers were designed for a small 405 base pairs (bp) fragment, a 591 bp fragment and a 696 bp fragment on the cytoplasmic domain of gp41. The primer sets for the pMAL-c2E vector were designed as follows: the 5’ primers were designed to contain a *Kpn*I site, and the 3’ primer a *Bam*HI site, in order to facilitate directional cloning of the genes. For directional cloning of the genes the primer sets for the gp41 gene fragments in pET28a and pJC20 vectors were designed to contain a *Nde*I site on the 5’ end primer and a *Bam*HI site at the 3’ end primer. The primer sets for the pPIC9 vector were designed so that the 5’ primers contained an *Eco*RI site and the 3’ primers a *Not*I site, for directional cloning purposes. For the 052741 the primer set was designed to contain a *Nde*I site on the 5’ end primer and a *Xho*I site at the 3’ end primer. The primers are indicated in Table 2.1

Table 2.1: Forward and Reverse oligonucleotides used for the synthesis of different gene fragments.

Primer name*	Sequence (5' to 3')†	Length	T_m (℃)
gp41-T1 F mal	GAGGTACCTAATCTGTGGAGTTG	23	62.7
gp41 R	GCGGATCCTTATTGCAAAGCTG	22	62.6
gp41-T2 F mal	GTGGTACCTCTCGGAAGAATCG	22	64.5
gp41-T3 F mal	GAGGTACCTAGAGAAATTAGTA	22	57.1
gp41-T1N F	GCCATATGAATCTGTGGAGTTG	22	60.8
gp41-T2N F	GACATATGCTCGGAAGAATCG	21	60.6
gp41-T3N F	GCCATATGAGAGAAATTAGTA	21	54.7
gp41C9 F	GCGAATTCAGAGAAATTAGTAATTACACA	29	54.0
gp41C9 R	TCGCGGCCGCTTGCAAAGCTGC	22	68.1
052741 F	GCCATATGCGCGAAATCAGCAA	22	53.1
052741 R	CTCTCGAGCTGCAGCGCCGCTTC	23	62.3

†**Red: *Kpn*I cleavage site. Blue: *Bam*HI cleaving site. Green: *Nde*I cleaving site. Purple: *Eco*RI cleaving site. Light blue: *Not*I cleaving site. Orange: *Xho*I cleaving site.** * **Primers with mal were for cloning into pMAL-2cE vector. Primers with N were for cloning into pET28a and pJC20 vectors. Primers with C9 were for cloning into pPIC9 vector.**

The DNA concentration of the plasmids and the gel-purified PCR products were determined by ultraviolet absorbance (260 nm) with a A_{260} nm unit equivalent to 50 ng/µl double stranded and 33 ng/µl single stranded DNA. Alternatively, concentrations were also estimated by comparison of the fluorescence intensity of the sample bands to marker concentrations (MassRuler DNA Ladder Mix, Fermentas, USA).

2.2.4 PCR synthesis of the gp41/052741 peptide

The *env* gene of HIV-1 subtype C known as BU/91/01/13 was kindly provided by Dr. R. Daniels as indicated in Figure 2.3.

```
REISNYTNIYRLLEDSQNQQEKQNEKDLLALDKWQNLWSWFDITNWLWYIKIFIMIVGG
                                        NLWSWFDITNWLWYIKIFIMIVGG

LIGLRIIFAILSIVNRVRQGYSPLSFQTLTPSPRGPDRLGRIEEEGGEQDRDRSIRLVG
LIGLRIIFAILSIVNRVRQGYSPLSFQTLTPSPRGPDRLGRIEEEGGEQDRDRSIRLVG
                                        PDRLGRIEEEGGEQDRDRSIRLVG

GFLALAWDDLRSLCLFSYHRLRDLILIATRVVELLGRSSLRGLQRGWEILKYLGSLVQY
GFLALAWDDLRSLCLFSYHRLRDLILIATRVVELLGRSSLRGLQRGWEILKYLGSLVQY
GFLALAWDDLRSLCLFSYHRLRDLILIATRVVELLGRSSLRGLQRGWEILKYLGSLVQY

WGLELKKSAINLLDNITAIAVAEGTDRIIEIIQRICRAIYHIPRRIRQGFEAALQ
WGLELKKSAINLLDNITAIAVAEGTDRIIEIIQRICRAIYHIPRRIRQGFEAALQ
WGLELKKSAINLLDNITAIAVAEGTDRIIEIIQRICRAIYHIPRRIRQGFEAALQ
```

Figure 2.3: **Amino acid sequence of the envelope gene of HIV-1 Subtype C known as BU/91/01/13 and the truncations made to the sequence.** The red amino acid sequence indicates the sequence of gp41T1. The green amino acid sequence indicates the sequence of gp41T2. The blue amino acid sequence indicates the sequence of gp41T3.

2.2.4.1 PCR synthesis of gp41 and 052741 genes

A temperature gradient PCR was set up for all the different constructs to determine the annealing temperatures. The PCR was carried out in a Mastercycler Personal 5332 Eppendorf Version 2.12.32. The PCR conditions for the individual genes are indicated in Table 2.2. The intended size of the PCR products (405 bp, 591 bp, 696 bp) was confirmed by agarose gel electrophoresis using a 1.5% agarose gel (section 2.2.5).

Table 2.2: PCR synthesis of gp41 and 052741 genes.

Genes	Vector	Anneal-ling temp.	Cycle conditions for all the gene constructs	PCR reaction for all the gene constructs
gp41	pET28	50 ℃	95 ℃ for 5 min was followed by 30 cycles of 95 ℃ for 1 min, Specific annealing temperature is indicated for each vector for 1 min, 72 ℃ for 1 min and 72 ℃ for 5 min	2 µl of template, 10 µl each of forward (2.5 µM) and reverse (2.5 µM) primers (Inqaba Biotech), 2 mM dNTPs and a proofreading Taq polymerase, TaKaRa ExTaq (TaKaRa Biomedicals, Shuzo, Japan)
052741	pET28	55 ℃		
gp41	pJC20	50 ℃		
052741	pJC20	55 ℃		
gp41	pMAL-c2E	50 ℃		
052741	pMAL-c2E	55 ℃		
gp41	pPIC9	60 ℃	95 ℃ for 5 min, was followed by 30 cycles of 95 ℃ for 30 sec, 60 ℃ for 30 min, 72 ℃ for 30 min and 72 ℃ for 2 min.	

2.2.5 Agarose gel electrophoresis and purification

2.2.5.1 Agarose gel electrophoresis and purification of the gp41 and 052741 genes

The correct size of the PCR products (405 bp, 591 bp, 696 bp) respectively were confirmed by agarose gel electrophoresis using 1.5% agarose (Promega, Wisconsin, USA) gel. TAE buffer (0.04 M Tris-Acetate, 1 mM EDTA, pH 8.0) was used for electrophoresis. Ethidium bromide, a DNA base intercalator (1.25 ng/ml) was included in the gel solution for DNA visualization. Samples were electrophoresed at 8 $V.cm^{-1}$ and the bands were visualized at 312 nm on a Bio-Rad Chemidoc Gel Documentation System (Bio-Rad Laboratories, USA), using the QuantityOne software. The PCR products were cleaned up (removel of

primer dimers, template and non-specific contaminants) using a High Pure PCR Purification Kit (Roche diagnostics, Germany). The restriction enzyme digestion was set up for the gp41/052741 constructs for cloning into the different expression vectors as indicated in Table 2.3.

Table 2.3: Restriction enzyme digests of gp41 and 052741 genes.

Gene	Vector	Enzyme	Reaction
gp41/052741	pET28	*Nde*I and *Bam*HI (10 U, Fermentas, USA)	40 µl of the DNA (20 ng/µl) was digested with 1 µl *Nde*I (10 U, Fermentas, USA) at 37 ℃ for 6 hours in the unique buffer for the enzyme followed with 1 µl *Bam*HI and digested at 37 ℃ for 15 hours supplemented with the unique buffer for the *Bam*HI enzyme.
gp41/052741	pJC20	*Nde*I and *Bam*HI (10 U, Fermentas, USA)	
gp41/052741	pMAL-c2E	*BamH*I and *Kpn*I (10 U, Fermentas, USA)	40 µl of the DNA (20 ng/µl) was digested with 1 µl *BamH*I (10 U, Fermentas) at 37 ℃ for 6 hours in its unique buffer, followed with 1 µl *Kpn*I at 37 ℃ for 15 hours in its unique buffer.
gp41	pPIC9	*Not*I and *Eco*RI (10U, Fermentas, USA)	40µl of the DNA (20 ng/µl) was digested with 1 µl *Not*I (10U, Fermentas) at 37 ℃ for 6 hours in its unique buffer, followed with 1 µl *Eco*RI at 37 ℃ for 15 hours in its unique buffer.

The digested products were separated by electrophoresis on a 1.5% agarose gel. The fragments were then purified from the agarose gel to remove any larger contaminating non-specific fragments.

2.2.6 Preparation of electrocompetent cells

2.2.6.1 Preparation of *E.coli* electrocompetent cells

Electrocompetent cells were prepared by the following method (Sambrook *et al*., 1989): A single colony of BL21(DE3) from *E. coli* (Novagen, Madison, Wisconsin, USA) was inoculated into 15 ml Luria-Bertani (LB) liquid media (1% Tryptone, 0.5% yeast extract, 1% NaCl at pH 7) and grown overnight at 37 ℃ with shaking (220 rpm). Fresh LB-Broth (250 ml) was inoculated with 1 ml of the overnight culture and grown at 37 ℃ with shaking (220 rpm) u ntil the cells reached early to

mid-log phase (OD at 600 nm of 0.3-0.6). To harvest the cells, they were transferred to two cold centrifuge tubes and pelleted at 4,000 x g in a Sorvall RC-5C Plus (Sorvall, UK) for 10 min at 4 ℃. All the s ubsequent steps were done at 4 ℃. After the supernatant was discarded, the cells were washed with 250 ml ice-cold water. The suspension was centrifuged at 4,000 x g in a Sorvall RC-5C Plus (Sorvall, UK) for 10 min. This washing step was repeated twice. After the final centrifugation step the supernatant was immediately removed from the loose pellets. The pellets were resuspended into 10 ml of ice-cold 10% glycerol and incubated on ice for 30 minutes. Cells were subsequently pelleted (3,000 x g for 10 min), the supernatant removed by vacuum suction and the pellets were resuspended in 800 µl of 10% ice-cold glycerol. This was divided into 90 µl aliquots and frozen at -70 ℃.

2.2.6.2 Preparation of *P. pastoris* electrocompetent cells

Electrocompetent cells were prepared by the following method (Scorer *et al*., 1994): A single colony of GS115 from *P. pastoris* (Invitrogen, Madison, Wisconsin, USA) was inoculated into 50 ml Yeast Extract Peptone Dextrose (YPD) liquid media (1% yeast extract, 2% peptone, 2% dextrose) and grown overnight at 30 ℃ with shaking (300 rpm). Fresh YP D media (250 ml) was inoculated with 0.5 ml of the overnight culture and grown at 30 ℃ with shaking (300 rpm) until the cells reached an optical density at 600 nm of 1.3-1.5. To harvest, the cells were transferred to two cold centrifuge tubes and centrifuged at 1,500 x g (Sorvall RC-5C Plus, Sorvall, UK) for 5 min at 4 ℃. All the subsequent steps were done at 4 ℃. After the supernatant was discarded, the cells were washed with 500 ml ice-cold water. The suspension was centrifuged at 1,500 x g (Sorvall RC-5C Plus, Sorvall, UK) for 5 min. After the supernatant was discarded, the cells were washed with 250 ml ice-cold water. The suspension was centrifuged at 1,500 x g (Sorvall RC-5C Plus, Sorvall, UK) for 5 min. After the supernatant was discarded, the cells were washed with 20 ml of ice-cold 1 M sorbitol. The suspension was centrifuged at 1,500 x g in a Sorvall RC-5C Plus (Sorvall, UK) for 5 min. After the final centrifugation step the supernatant was immediately aspirated from the loose pellets. The pellets were resuspended in ice-cold 1 M sorbitol to a final volume of approximately 1.5 ml. This was divided into 80 µl aliquots and frozen at -70 ℃.

2.2.7 Cloning of truncated gp41 genes and codon optimised genes (052741) into various vector systems

2.2.7.1 Cloning of truncated gp41 genes and codon optimised genes (052741) into pGem T-easy vector system

A ligation mixture was set up at a picomole insert:vector ratio of 20:1 using T4 DNA Ligase (Promega Corporation, USA) as per the manufacturer's instructions. The gp41 and 052741 inserts were ligated into pGem T-easy vector. The ligation was cleaned up by using the following method: the reaction mixture was dried *in vacuo* using a speedy vac for 20 min and the pellet was dissolved in 1 ml 70% ethanol. It was centrifuged at 4 ℃ at 10 000 x g f or 30 min and dried for 10 min. The ligation mixture was then transformed into electrocompetent XL10 gold cells or DH5α cells using electroporation. The electrocompetent XL10 gold cells or DH5α cells were thawed on ice. To 2 µl of plasmid DNA (pGEM T-easy gp41/052741), 100 µl of electrocompetent cells were added. This was then transferred to a pre-chilled electroporation cuvette and a pulse of 2500 V applied for 5 ms in a Bio-Rad MicroPulser electroporator (Bio-Rad Laboratories, Hercules, California) (Dower *et al*., 1988). LB liquid media (1 ml) was added directly after electroporation and the cells were incubated for 1 hour at 37 ℃ with shaking (180 rpm). The transformation mixture was plated on LB agar plates containing 100 µg/mL ampicillin, 80 µg/mL X-Gal and 0.1 mM isopropyl-β-D-thiogalactopyranoside (IPTG). White colonies on these plates were subsequently screened using PCR to identify positives containing the different HIV constructs.

2.2.7.2 Cloning the HIV gp41 and 052741 genes into pET28a, pJC20, pMAL-c2E and pPIC9 expression vector systems

Different expression vectors were used (Table 2.4): the IPTG inducible T7*lacUV5* promoter of the pET28a vector, pJC20 vector and the pMAL-c2E vector systems, the cell-free expression system (RTS-system, Roche Diagnostics) and the methanol inducible $^{3'}$*AOX1* (TT) promoter of the pPIC9 vector system (Invitrogen Technical Data, La Jolla, USA).

Table 2.4 Vector systems used for the recombinant expression of the synthetic gp41. Refer to Appendix E for the vector maps.

Vector	Antibiotic Resistance	Tag	Induction	Expression
pET28a	Kanamycin	N-terminal hexahistidine peptide tag	T7*lacUV5*	Intracellular
pJC20	Ampicillin	None	T7*lacUV5*	Intracellular
pMAL-c2E	Ampicillin	MBP	T7*lacUV5*	Intracellular
pPIC9	Ampicillin	none	*3'AOX1* (TT)	Secreted

The pET28a vector system (Novagen, EMD Biosciences, Germany) was chosen for the His_6-tagged expression and affinity purification, the pJC20 vector system (ATCC, Germany) was chosen because it is a high copy number vector and the pPIC9 vector system (Invitrogen Technical Data, La Jolla, USA) was chosen for the secreted expression in *Pichia pastoris* with the α-factor signal sequence.

Two different host systems were used: the *E.coli* system and the *Pichia pastoris* system. Different *E.coli* cell lines were used to determine which cell line gave optimal protein expression (Table 2.5). Cells containing the ply-S plasmid show increased stability of toxic gene expression by the system and are used when the protein produced might be toxic to the cell-line. The plys S plasmid encodes a small amount of T7 lysozyme, which binds to T7 RNA polymerase thereby inhibiting transcription and preventing cell death (Huang *et al.,* 1999). BL21 Star (DE3) cells were chosen for their high level of gene expression (Invitrogen Technical Data, La Jolla, USA). BL21 Rosetta cells were regarded beneficial due to the fact that many eukaryotic genes have codons that have low usage in *E.coli*, leading to low expression levels. To compensate for this the cell line contains a pRARE1 plasmid for higher expression levels of eukaryotic genes (Invitrogen Technical Data, La Jolla, USA). The ER2508 is a derivative of the *E.coli* RR1 strain that has been made *lon*$^-$. The *lon* mutation knocks out the major

ATP-dependent protease in the *E.coli* cytoplasm, thus improving the stability of many foreign proteins when expressed in *E.coli*. The *mal'e* gene, which encodes Maltose-Binding protein, is deleted as part of the *malB* deletion (New England Biolabs, Boston, USA). Only one *Pichia pastoris* strain (GS115) was used for the reason that these cells can yield both classes of transformants, histidine positive and methanol positive (His^+ Mut^+) and Histidine positive and methanol sensitive (His^+ Mut^S).

Table 2.5 Properties of *E.coli* and *Pichia pastoris* strains used as hosts for protein expression

Host strain	Expression host	Genotype	Antibiotic resistance
E.coli	BL21(DE3) (Invitrogen Technical Data, La Jolla, USA)	B strain F^- *dcm ompT* $hsdS_B$ (r^-_B m^-_B) *gal* λ (DE3) [plys S Cam^r]	None
E.coli	BL21 (DE3) plys S (Invitrogen Technical Data, La Jolla, USA)	B strain F^- *dcm ompT* $hsdS_B$ (r^-_B m^-_B) *gal* λ (DE3) [plys S Cam^r]	Chloramphenicol
E.coli	BL21 Star (DE3) (Invitrogen Technical Data, La Jolla, USA)	B strain F^- *ompT* $hsdS_B$ (r^-_B m^-_B) *gal dcm rne 131* (DE3)	None
E.coli	BL21 Rosetta (DE3) plys S (Invitrogen Technical Data, La Jolla, USA)	B strain F^- *dcm ompT* $hsdS_B$ (r^-_B m^-_B) *gal* λ (DE3) [plys S Cam^r]	Chloramphenicol
E.coli	ER2508 (New England Biolabs)	F^- *ara-14 leuB6 fhuA2 Δ(argF-lac)U169lacY1 lon::miniTn*10(Tet^R)*glnV44 galK2 rpsL20* (Str^R) *xyl-5 mtl-5 Δ (malB)zjc::Tn*5 (Kan^R) $Δ(mcrC\text{-}mrr)_{HB101}$	Kanamycin
Pichia pastoris	GS115 (Invitrogen Technical Data, La Jolla, USA)	*his4*	None

Restriction enzyme digestion mixture was set up for the gp41/052741 constructs: the plasmid DNA (40 µl; 20 ng/µl) was digested with 1 µl *Nde*I/*KpN*I/*EcoR*I (10 U, Fermentas, USA) at 37 ℃ for 6 hours and followed w ith 1 µl *BamH*I/*Not*I/*Xho*I at 37 ℃ for 15 hours. The digested products were sepa rated by electrophoresis on a 1.5% agarose gel. The fragment was then purified from the agarose gel to remove any larger contaminating non-specific fragments. Ligation reactions into pET28a, pJC20, pMLA-2ce and pPIC9 were set up at a picomole ratio of 20:1 insert:vector using T4 DNA Ligase (Promega Corporation, USA) as per the manufacturer's instructions. The ligations were then transformed into different electrocompetent cell lines as indicated in Table 2.5. To 2 µl of DNA (gp41/052741) 100 µl of electrocompetent cells were added. This was then transferred to a pre-chilled electroporation cuvette and a pulse of 2500 V applied for 5 ms in a Bio-Rad MicroPulser electroporator (Bio-Rad Laboratories, Hercules, California) (Dower *et al.,* 1988). LB liquid media (1 ml) was added directly after electroporation to the *E.coli* cell lines and the cells were incubated for 1 hour at 37 ℃ with shaking (180 rpm), and pla ted on LB solid media (1% noble agar, 0.5% yeast extract, 1% NaCl, pH 7.0) supplemented with 100 µg/ml ampicillin in the case of the pJC20 and pMAL-c2E vectors and 100 µg/ml kanamycin in the case of pET28a vector. The BL21 (DE3) plys and BL21 Rosetta transformations were plated on LB solid media supplemented with additional 34 µg/ml chloroamphenicol. Ice cold 1 M Sorbitol (1 ml) was added directly after electroporation to the *P. pastoris* cells. The transformations were plated on YNB solid media (1% noble agar, 0.5% yeast extract, 1% NaCl, pH 7.0) supplemented with 100 µg/ml ampicillin. Colonies on these plates were subsequently screened using PCR to identify positives containing the different gp41/052741 peptides.

2.2.8 Protein expression

2.2.8.1 Protein expression in pET28a, pJC20 and pMAL-c2E vectors

The systems used for protein expression were the pJC20, pET28a and pMAL-c2E vectors for the expression of recombinant genes in *E.coli*. Electrocompetent expression hosts (Section 2.2.4.1) were freshly transformed with the pJC20 or pET28a or pMAL-c2E plasmids containing gp41 or the codon optimized gp41 gene constructs and plated onto LB plates supplemented with 100 µg/ml

ampicillin in the case of the pJC20 and pMAL-c2E vectors or 100 µg/ml kanamycin in the case of the pET28a vector to select for the cells containing the gene. One positive colony was picked and grown for 16 hours in 50 ml LB liquid media supplemented with 100 µg/ml ampicillin, or 100 µg/ml kanamycin and incubated overnight at 37 ℃ with shaking (220 rpm) for population expression. The overnight cultures were diluted by the addition of 250 ml LB liquid medium (supplemented with 100 µg/ml ampicillin or 100 µg/ml kanamycin) to 2.5 ml of the overnight culture. This was followed by incubation at 37 ℃ with shaking until an optical density of approximately 0.6 (logarithmic growth phase) at 600 nm was reached. Consequently, 0.238 mg/ml isopropyl-β-D-thiogalactopyranoside (IPTG) (pEQ Lab Biotechnology, GmbH) was added to induce protein expression. After the addition of IPTG the cultures were incubated at 37 ℃ for 16 hours with shaking (220 rpm).

2.2.8.2 Protein expression in pPIC9 vectors

The pPIC9 vector is used to produce recombinant protein in a yeast system. This system was used to express the recombinant genes in *Pichia pastoris* under the control of T7 RNA polymerase (Clos and Brandau, 1994). Electrocompetent expression hosts (Section 2.2.4.2) were freshly transformed with the pPIC9 plasmid containing the gp41 gene construct and plated onto YNB plates supplemented with 100 µg/ml ampicillin to select for the cells containing the gene. One positive colony was picked and transferred into 50 ml BMGY liquid media (Buffered Glycerol-complex Medium) supplemented with 100 µg/ml ampicillin and incubated overnight at 30 ℃ with sh aking (350 rpm) until an optical density of approximately 2 (logarithmic growth phase) at 600 nm was reached. The overnight cultures were transferred to 250 ml BMMY liquid medium (Buffered Methanol-complex Medium) (supplemented with 100 µg/ml ampicillin). This was followed by incubation at 30 ℃ with shaki ng (350 rpm) for 17 hours.

2.2.9 Protein extraction

Protein extraction from expression in pET28a, pJC20 and pMAL-c2E vectors and purification was done using the following steps. The cells were harvested by centrifugation at 16,300 x g for 10 minutes at 4 ℃ (Sorvall Plus RC-5C, GSA rotor, Sorvall, UK) and the pellet frozen for overnight storage (-20 ℃). The cells

were allowed to thaw on ice and resuspended in the appropriate buffer. Complete Mini, EDTA-Free protease inhibitor cocktail tablet (Roche diagnostics, Germany) was added before sonication. The cells were sonicated for 10 cycles of 30 seconds pulsed sonication followed by 30 seconds incubation on ice water (Vibracell sonicator, output control 5, and duty cycle 50). The cell debris was removed by centrifugation at 12,100 x g for 10 minutes at 4 ℃ (Sorvall Plus RC-5C, rotor SS34, Sorvall, UK). The soluble protein-containing supernatant was transferred to a clean tube for overnight storage at 4 ℃ before purification commenced.

2.2.10 Protein purification

2.2.10.1 Affinity purification with His-bind Quick 900 Cartridges

His-bind Quick Cartridges (Novagen) are precharged with Ni^{2+} and use a large diameter cellulose matrix that has flow rates 5-50 times faster than agarose resins. The total binding capacity is around 2 mg. Using these formats, target proteins can be purified from crude lysates under native or denaturing conditions in as little as five minutes.The affinity purification of the recombinant protein using the 1 ml pre-packed His-bind Quick cartridge (Novagen, Madison, Wisconsin, USA) entailed the following steps: The cartridge was equilibrated with 6 ml Binding buffer (0.5 M NaCl, 20 mM Tris-HCl, 5 mM Imidazole pH 7.9). The cell free extract (5 ml) was loaded onto the cartridge. The cartridge was washed with 20 ml Binding buffer (followed with 10 ml Wash buffer (0.5 M NaCl, 60 mM imidazole, 20 mM Tris-HCl pH 7.9). The recombinant protein was eluted from the cartridge with 4 ml Elution buffer (1 M imidazole, 0.5 M NaCl, 20 mM Tris-HCl pH 7.9).

2.2.10.2 Affinity purification with Amylose resin

Amylose resin is a composite amylose/agarose bead functioning as an affinity matrix used for the isolation of proteins fused to maltose-binding protein (New England Biolabs, Boston, USA). The affinity purification of the recombinant protein using 1 ml Amylose resin entailed the following steps. The resin was equilibrated with 5 ml Column buffer (20 mM Tris-HCl, 200 mM NaCl, 1 mM EDTA, pH 7.4). Of the cell free extract, 1 ml was loaded onto the Amylose resin.

The mixture was incubated for 2 hours at 4 ℃ with constant shaking at 50 rpm. The Amylose resin was washed three times with 1 ml Column buffer. The fusion peptide was eluted from the Amylose resin with 1 ml Elute buffer (20 mM Tris-HCl, 200 mM NaCl, 1 mM EDTA, 10 mM maltose, pH 7.4). The fusion peptide was transferred to a clean container and stored at 4 ℃ until further analysis.

2.2.11 SDS-PAGE analysis

The Laemmli method for SDS (Sodium dodecyl-sulphate) poly-acrylamide gel electrophoresis (PAGE) was used to analyse the protein separation based on molecular mass (Laemmli, 1970). Gels were prepared as follows: 4% stacking gel (4% Bio-Rad Acrylamide-Bisacrylamide mix, 0.1% SDS, 0.05% ammonium persulphate, 0.1% TEMED, 0.05 M Tris-HCl, pH 6.8) and a 12% running gel (12% Bio-Rad Acrylamide-Bisacrylamide mix, 0.1% SDS, 0.05% ammonium persulphate, 0.1% TEMED, 0.375 M Tris-HCl, pH 8.8) with 10 wells. A 15 µl purified protein sample from each of the collected fractions was transferred to a clean microcentrifuge tube. Subsequently, an equal volume of denaturing buffer (1.2% SDS, 30% glycerol, 15% ß-mercaptoethanol, 0.18 mg/ml bromophenol blue, 0.15 M Tris, pH 6.8) was added. The protein samples were denatured at 90 ℃ for 5 minutes. Of each sample 10 µl was loaded onto the gel. Electrophoresis was performed in a 0.025 M Tris-HCl, 0.2 M Glycine buffer (pH 8.3) and separated at 200 V in a Bio-Rad Mini Protean 3 Electrophoresis system. Protein bands were visualized with Coomassie Blue G250 staining solution (0.1 g Coomassie Blue G250 in 40% methanol, 10% acetic acid, in H_2O), followed by a treatment with destaining solution (40% methanol, 10% acetic acid, in H_2O).

2.2.12 Protein concentration determination

Protein concentrations were determined according to the Folin-Lowry method (Lowry *et al.,* 1951), using calibration curves constructed with bovine serum albumin (BSA). A stock solution of 0.3 mg/ml was diluted to obtain a standard concentration range of 300, 240, 210, 180, 150, 120, 90, 60 and 30 µg/ml BSA protein. To each BSA standard reaction and protein sample, 300 µl of Solution ABC (20:1:1) was added (Solution A: 2% Na_2CO_3 in 1 M NaOH; Solution B: 1% $CuSO_4.H_2O$; Solution C: 2% Potassium tartrate). After a 15 minute incubation period, 900 µl of 10% 2 N Folin-Ciocalteau reagent (mixture of phosphotungstic

acid and phosphomolybdic acid in phenol) was added. The Copper (II) ion in alkaline solution B reacts with the protein to form complexes with functional groups of tyrosine, tryptophan and cysteine. These complexes react with the Folin-Ciocalteau reagent. The product becomes reduced to molybdenum/tungsten blue and can be detected colourimetrically by absorbance at 660 nm (Lowry *et al.,* 1951). Presence of strong acids or ammonium sulphate can interfere with the assay. After a 45 minute incubation period in the dark, the absorbance at 660 nm was read. After the standard curve for BSA was obtained, the slope and the y-intercept of the BSA standard curve were used to extrapolate the concentrations for the truncated gp41/052741 protein samples.

2.2.13 Western Blotting

Proteins were blotted onto PVDF (polyvinylidine-difluoride, Merck, Germany) membranes using a Mini Trans-Blot electrophoretic transfer cell (Bio-Rad) and 1x Transfer buffer (25 mM Tris-HCl, 192 mM glycine, 20% v/v methanol, pH 8.3) at 30 V for 16 hours. Membranes were blocked with 5% Blocking reagent (ECL plus western blotting Analysis System, Amersham Biosciences. UK) in TBST (50 mM Tris-HCl, 150 mM NaCl, 0.5% Tween-20, pH 7.4-7.6) overnight at 4 ℃. Blots were then incubated with the primary antibodies (Polyclonal antiboby 1577 from rabbit, directed against Kennedy domain of gp41 supplied by Dr. Hugh Brady) at a dilution of 1:6000 in TBST for 1 hour at room temperature. The membranes were rinsed twice with TBST followed by TBST wash for 15 minutes at room temperature repeated three times. The blots were then incubated with the secondary antibody (Horseradish peroxidase conjugated anti-rabbit or anti-mouse IgG whole molecule) (Amersham Biosciences, UK) at a dilution of 1: 15 000 in TBST for 1 hour at room temperature. The membranes were rinsed twice with TBST followed by TBST wash for 15 minutes at room temperature repeated three times. The antigen-antibody complex was detected with lumigen PS-3 detection reagent (Amersham Biosciences, UK) according to the manufacturer's instructions.

2.2.14 RTS 100 cell-free protein expression

2.2.14.1 RTS 100 *E.coli* HY reaction

The codon optimised gp41 sequence for *E.coli* was used as template in the RTS 100 *E.coli* HY reaction for cell-free protein expression. The reaction mixture was prepared according to the manufacturer's guidelines (Roche Diagnostics). The mixture was added to 0.5 µg of linear template to a final volume of 50 µl. The reaction mixture was incubated at 30 ℃ for 6 hours .

2.2.14.2 Acetone precipitation of proteins

Ice cold acetone (50 µl) was added to 5 µl of the reaction mixture and incubated on ice for 5 minutes. The mixture was centrifuged at 10,000 x *g* in a microcentrifuge at 4 ℃. The supernatant was discar ded and the pellet was air-dried for 10 min. The pellet was resuspended in 20 µl of SDS-PAGE sample buffer and heated for 5 minutes at 95 ℃. The sampl es were analysed by means of a SDS-PAGE gel (section 2.2.11).

2.3 Results

The gp41 gene was truncated into three gene fragments which varied in length namely gp41T3 which is 405 bp in length, gp41T2 which is 591 bp in length and gp41T1 which is 696 bp in length. Gp 41T1 also underwent codon optimization and is called 052741 which is also 696 bp in length.

Different tags were used for expression and purification purposes, which will influence the size of the expressed protein as indicated in Figure 2.4

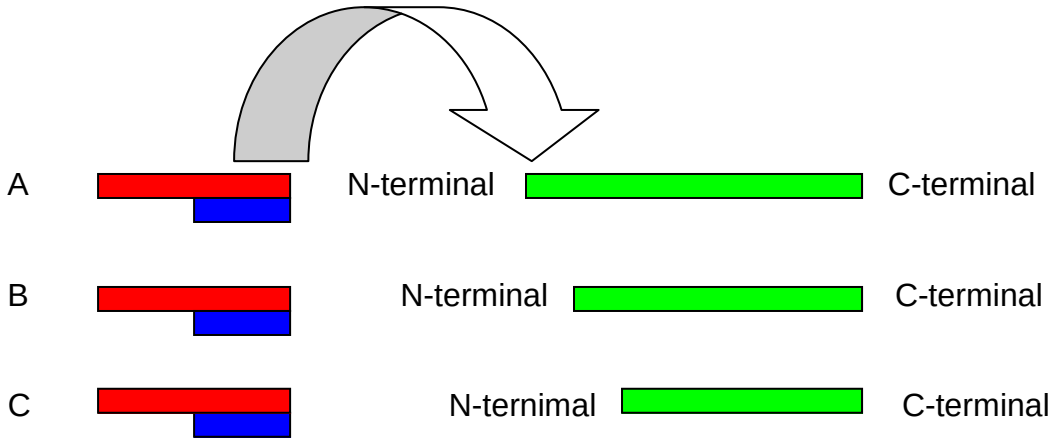

Figure 2.4: **Fragment sizes of truncated gp41 peptides with different tags: In A, B and C** red represents the N-terminal MBP-2 tag which is 42.5 kDa and blue represents the N-terminal 6xHis-tag. **A** represents gp41T1 which is 27.8 kDa and indicated in green. **B** represents gp41T2 which is 22.6 kDa and indicated in green. **C** represents gp41T3 which is 15.3 kDa and indicated in green.

Comparison of the codon frequencies between HIV-1 gp41 and *E.coli* identified points of codon bias, which were eliminated in the design of the codon optimized gene. GeneArt® converted the native HIV-1 gene codons (gp41T1 gene) to the preferred *E.coli* codons (052741 gene). The alignment of the synthetic (052741) and native (gp41T1) genes are shown in Figure 2.5.

```
gp41T3N    ...ATGAGAGAAATTAGTAATTACACAAACACAATATACA      37
052741     CATATGCGCGAAATCAGCAACTATACCAACACCATCTATC      40
Consensus     atg g gaaat ag aa ta ac aacac at ta

gp41T3N    GGTTGCTTGAAGACTCGCAAAACCAGCAGGAACAAAATGA      77
052741     GCCTGCTGGAAGATAGCCAGAATCAGCAGGAACAGAACGA      80
Consensus g  tgct gaaga    ca aa cagcaggaaca aa ga

gp41T3N    AAAAGATTTACTAGCATTGGACAAGTGGCAAAATCTGTGG     117
052741     AAAAGATCTGCTGGCGCTGGATAAATGGCAGAATCTGTGG     120
Consensus aaaagat t ct gc  tgga aa tggca aatctgtgg

gp41T3N    AGTTGGTTTGACATAACAAATTGGCTGTGGTATATAAAAA     157
052741     AGCTGGTTCGATATTACCAACTGGCTGTGGTACATCAAAA     160
Consensus ag tggtt ga at ac aa tggctgtggta at aaaa

gp41T3N    TATTCATAATGATAGTAGGGGGCTTGATAGGTTTAAGAAT     197
052741     TCTTCATCATGATTGTGGGCGGTCTGATTGGCCTGCGTAT     200
Consensus t ttcat atgat gt gg gg  tgat gg  t  g at

gp41T3N    AATTTTTGCTATACTCTCTATAGTGAATAGAGTTAGGCAG     237
052741     CATTTTTGCCATCCTGAGCATTGTGAACCGTGTTCGTCAG     240
Consensus  atttttgc at ct    at gtgaa  g gtt g cag

gp41T3N    GGATACTCACCCTTATCGTTTCAGACCCTTACCCCAAGCC     277
052741     GGTTATAGCCCGCTGTCTTTTCAGACCCTGACCCCGAGCC     280
Consensus gg ta    cc  t tc tttcagaccct acccc agcc

gp41T3N    CGAGGGGACCCGACAGGCTCGGAAGAATCGAAGAAGAAGG     317
052741     CGCGTGGTCCGGATCGTCTGGGTCGTATTGAAGAAGAAGG     320
Consensus cg g gg cc ga  g ct gg  g at gaagaagaagg

gp41T3N    TGGAGAGCAAGACAGAGACAGATCCATTCGATTAGTGGGC     357
052741     CGGCGAACAGGATCGTGATCGTAGCATTCGTCTGGTGGGT     360
Consensus  gg ga ca ga  g ga  g   cattcg  t gtggg

gp41T3N    GGATTCTTAGCACTTGCCTGGGACGATCTACGGAGCCTGT     397
052741     GGTTTTCTGGCGCTGGCGTGGGATGATCTGCGTAGCCTGT     400
Consensus gg tt  t gc ct gc tggga gatct cg agcctgt

gp41T3N    GCCTCTTCAGCTACCACCGATTGAGAGACTTAATATTGAT     437
052741     GCCTGTTTAGCTATCATCGTCTGCGCGATCTGATTCTGAT     440
Consensus gcct tt agcta ca cg  tg g ga  t at  tgat

gp41T3N    TGCAACGAGAGTGGTGGAGCTTCTGGGACGCAGCAGTCTC     477
052741     TGCCACCCGTGTTGTTGAACTGCTGGGTCGTAGCAGCCTG     480
Consensus tgc ac  g gt gt ga ct ctggg cg agcag ct

gp41T3N    AGGGGACTACAGAGGGGGTGGGAAATCCTTAAGTATCTGG     517
052741     CGTGGTCTGCAGCGTGGTTGGGAAATCCTGAAATATCTGG     520
Consensus  g gg ct cag g gg tgggaaatcct aa tatctgg

gp41T3N    GAAGTCTTGTGCAGTATTGGGGTCTAGAGCTAAAAAAGAG     557
052741     GCAGCCTGGTTCAGTATTGGGGCCTGGAACTGAAAAAAAG     560
Consensus g ag ct gt cagtattgggg ct ga ct aaaaa ag

gp41T3N    TGCTATTTAATCTGCTTAATATCACAGCAATAGCAGTAGCT     597
052741     CGCCATCAACCTGCTGAACATTACCGCCATTGCCGTGGCG     600
Consensus  gc at aa ctgct aa at ac gc at gc gt gc

gp41T3N    GAAGGAACAGATAGGATTATAGAAATCATACAAGAATTT     637
052741     GAAGGCACCGATCGTATCATCGAAATCATCCAGCGTATTT     640
Consensus gaagg ac gat g at at gaaatcat ca  g attt

gp41T3N    GTAGAGCTATCTATCACATACCTAGAAGAATAAGACAGGG     677
052741     GCCGTGCCATTTATCATATTCCGCGTCGTATTCGCCAGGG     680
Consensus g  g gc at tatca at cc  g  g at  g caggg

gp41T3N    CTTTGAAGCAGCTTTGCAATAA......                  699
052741     TTTTGAAGCGGCGCTGCAGTAAGGATCC                 708
Consensus  tttgaagc gc  tgca taa
```

Figure 2.5: **Alignment of the synthetic (052741) and native (gp41T3) genes:** The black parts indicate 100% homology between the two genes. The blue parts indicate 50-75% homology between the two genes.

2.3.1 Cloning of synthetic gp41 genes for expression in *E. coli*

2.3.1.1 Restriction enzyme digestion of pMAL-gp41 constructs

The truncated gp41 genes (gp41T1, gp41T2, gp41T3 and 052741) were sub-cloned into pGem T-easy, followed by enzyme digestion and ligation into pMAL-c2E. These plasmids were then used to transform ER 2507 *E.coli* cells via electroporation. The presence of the plasmid in the cells was confirmed by antibiotic selection with 100 µg/ml ampicillin, as the pMAL-c2E plasmid carries an ampicillin resistance marker gene. After plasmid isolation (peqGOLD plasmid Miniprep Kit 1, Optima Scientific), the presence of the gp41 gene in the plasmid was confirmed with restriction enzyme digestion using *Kpn*I and *Bam*HI. The gel in Figure 2.6 A shows the bands (Lane 2-5). Indicated in Figure 2.6 B is the enzyme digest of pMAL-gp41(80) (lane 5) to indicated the presence of the gp41gene in the plasmid, The plasmid was digested with *Kpn*I and *Xba*I in a two-step digestion, first with *Kpn*I (Amersham) in L buffer, then with *Xba*I in M buffer. The restriction enzymes were subsequently heat-inactivated.

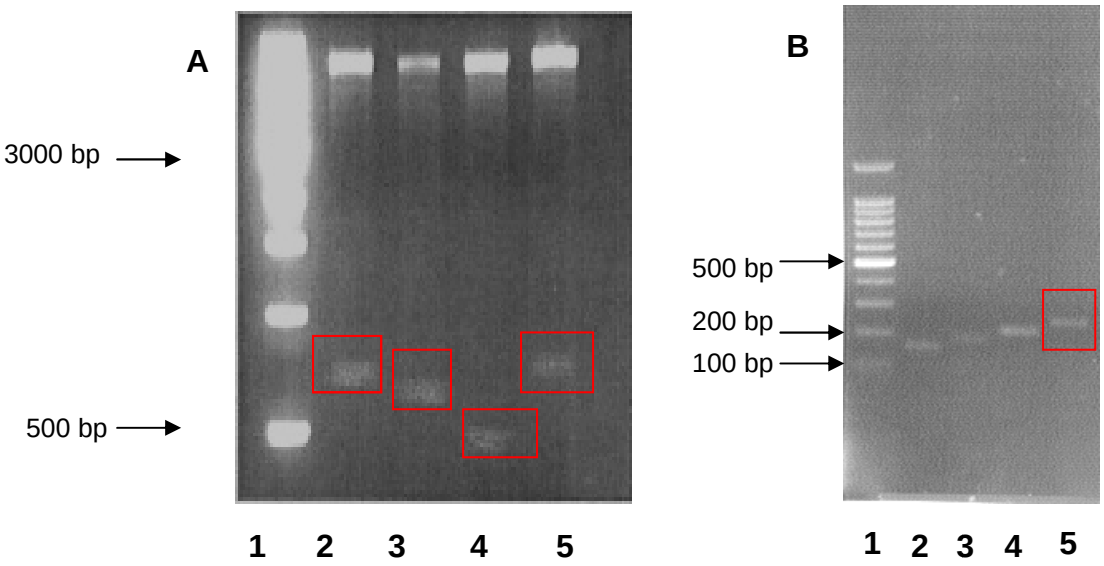

Figure 2.6: ***Kpn*I and *Bam*HI restriction analysis of the pMAL-gp41 constructs.** In Figure A: Lane 1: 1 kb DNA Ladder (New England Biolabs, Boston, USA). Lane 2: Restriction digestion of the pMAL-gp41T1 plasmid. Lane 3: Restriction digestion of the pMAL-gp41T2 plasmid. Lane 4: Restriction digestion of the pMAL-gp41T3 plasmid. Lane 5: Restriction digestion of the 052741 plasmid. The 3000 bp and 500 bp bands of the 1 kb DNA Ladder are indicated. In Figure B: Lane 1: 100 bp ladder (Promega). Lane 2-4: Truncated gp41(80). Lane 5: Restriction digestion of pMAL-gp41(80). The 500 bp, 200 bp and 100 bp bands of the 100 bp DNA Ladder are indicated.

In Figure 2.6A the bigger band is of size 6642 bp (Lane 2-5) that represents the pMAL-c2E plasmid back-bone and the smaller bands in the red blocks respectively indicates the gp41 genes with sizes of 405 bp (Lane 4) for gp41T3, 591 bp (Lane 3) for gp41T2 and 696 bp for both gp41T1(Lane 2) and 052741 (Lane 5). This was an indication that the ER 2507 *E.coli* cells were transformed with the pMAL-c2E plasmid. From Figure 2.6B the gp41 gene with a size of 260 bp (Lane 5) is indicated in the red block.

2.3.1.2 Restriction enzyme digestion of pET28a-gp41/052741 constructs

After sub-cloning of the truncated gp41 genes (gp41T1, gp41T2, gp41T3 and 052741 into pGem T-easy, enzyme digestion and ligation of the constructs into pET28a the plasmid was used to transform BL21 (DE3), BL21 (DE3) plys, BL21 Star (DE3) and BL21 Rosetta *E.coli* cells via electroporation. The presence of the plasmid in the cells was confirmed by antibiotic selection with 100 µg/ml kanamycin as the pET28a plasmid carries a kanamycin resistance marker gene. After alkaline lysis and plasmid isolation (peqGOLD plasmid Miniprep Kit 1, Optima Scientific), the presence of the gp41 gene in the plasmid was confirmed with restriction enzyme digestion using *Nde*I and *Bam*HI/*Xho*I. The gel in Figure 2.7 shows the results.

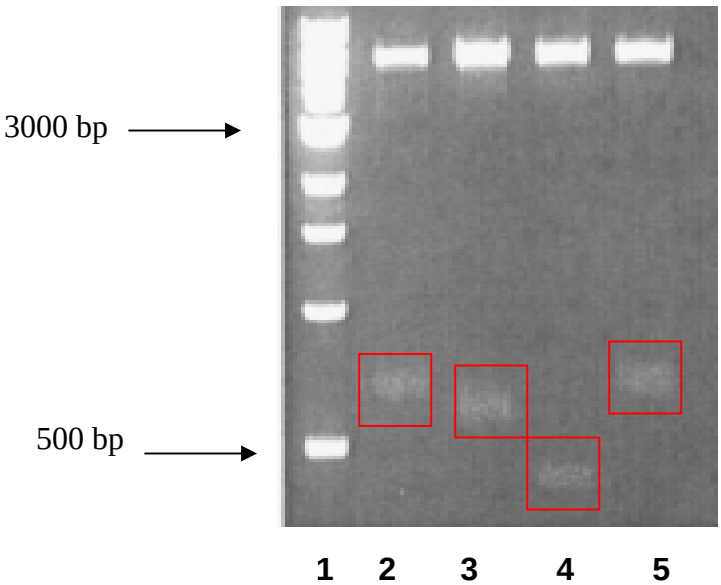

Figure 2.7: ***Nde*I and *Bam*HI/*Xho*I restriction analysis of the pET28a-gp41/052741 constructs.** Lane 1: 1 kb DNA Ladder (New England Biolabs, Boston, USA). Lane 2: Restriction digestion of the gp41T1 plasmid. Lane 3: Restriction digestion of the gp41T2 plasmid. Lane 4: Restriction digestion of the gp41T3 plasmid. Lane 5: Restriction digestion of the 052741 plasmid.The 3000 bp and 500 bp bands of the 1 kb DNA Ladder are indicated.

The bigger band of size 5392 bp (Lane 2-5) represents the pET28a plasmid back-bone and the smaller bands respectively indicate the gp41 genes with sizes of 405 bp (Lane 4) for gp41T3, 591 (Lane 3) bp for gp41T2 and 696 bp (Lane 2) for gp41T1 and 052741. This was an indication that the BL21 (DE3), BL21 (DE3) plys, BL21 Star (DE3), BL21 Rosetta *E. coli* cells were transformed with the pET28a plasmid.

2.3.1.3 Restriction enzyme digestion of pJC20 052741 constructs

After sub-cloning of the 052741 gp41 gene into pGem T-easy, enzyme digestion and ligation of the constructs into pJC20, this plasmid was used to transform BL21 (DE3) *E. coli* cells via electroporation. The presence of the plasmid in the cells was confirmed by antibiotic selection with 100 µg/ml ampicillin as the pJC20 plasmid carries an ampicillin resistance marker gene. After alkaline lysis and plasmid isolation, the presence of the gp41 gene in the plasmid was confirmed with restriction enzyme digestion using *Nde*I and *Xho*I. The gel in Figure 2.8 shows the bands (Lane 2).

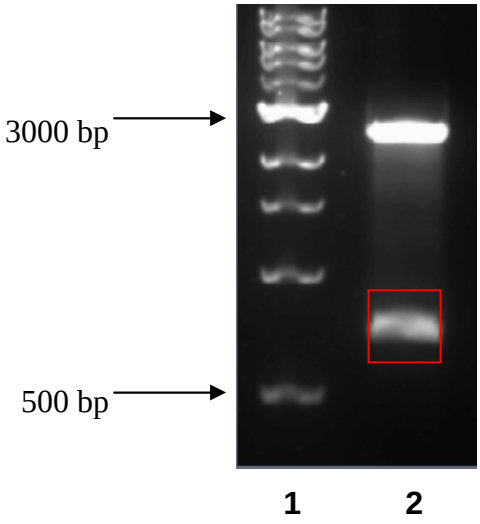

Figure 2.8: ***Nde*I and *Xho*I restriction analysis of the pJC20 052741 constructs.** Lane 1: 1 kb DNA Ladder (New England Biolabs, Boston, USA). Lane 2: Restriction digestion of the pJC20-052741 plasmid. The 3000 bp and 500 bp bands of the 1 kb DNA Ladder are indicated.

The bigger band is of size 2345 bp (Lane 2) that represents the pJC20 plasmid back-bone and the smaller band indicates the 052741 gene with a size of 696 (Lane 2) bp. This was an indication that the BL21 (DE3) *E.coli* cells were transformed with the pJC20 plasmid.

2.3.2 Protein expression of synthetic gp41 genes in *E. coli*

Initial attempts to express the truncated fragments of the HIV gp41 gene as well as the codon-optimized HIV gp41 genes in pET vectors did not succeed. The codon optimized gp41T1 gene namely 052741 was used as template for the cell-free protein expression system. The RTS 100 *E. coli* HY protein expression system was used and the protein was analysed on a 12% SDS-PAGE gel with no success.

2.3.2.1 Protein expression of pMAL-gp41 fusion proteins

The truncated pieces of the HIV gp41 gene namely gp41T1, gp41T2, gp41T3 were cloned into the pMAL-c2E expression vector and *E .coli* ER2508 cells were transformed. Protein expression took place at 37 ℃. The analysis in Figure 2.9A revealed that only expression of the fusion protein pMAL-gp41T3 (Lane 4-5) took place which is the smallest fragment of all. The HIV gene fragment gp41(80) was cloned into the pMAL-c2E expression vector and *E .coli* ER2508 cells were transformed. Protein expression took place at 37 ℃. The analysis in Figure 2.9B revealed the successful expression of the fusion protein pMAL-gp41(80) (Lane 2-3).

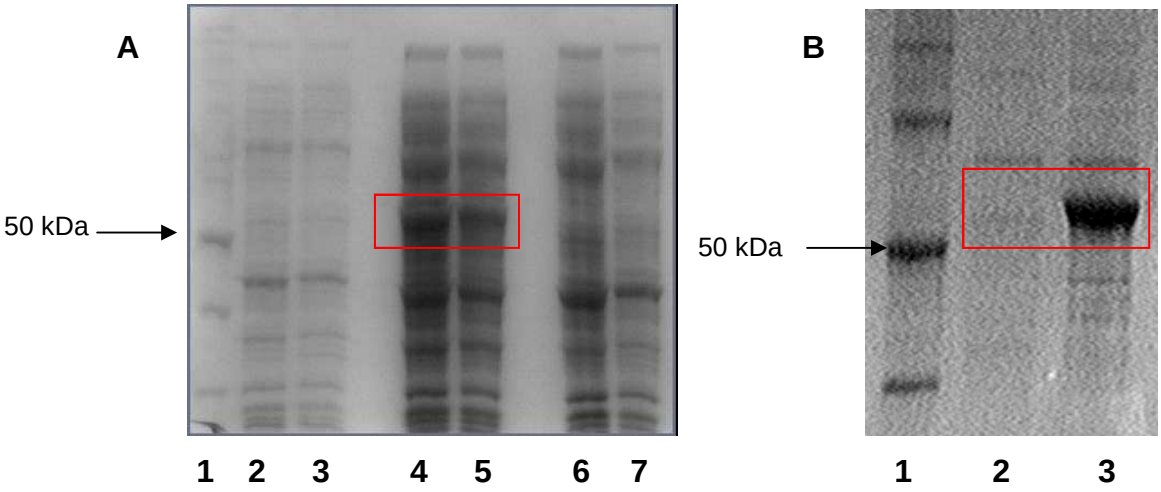

Figure 2.9: **SDS-PAGE analysis of protein expression of different pMAL-gp41 fusion proteins at 37 ℃.** In Figure A: Lane 1: PageRuler Protein Ladder. Lane 2: Total protein content of pMAL-gp41T1 (10 µl). Lane 3: Total protein content of pMAL-gp41T1 (5 µl). Lane 4: Total protein content of pMAL-gp41T3 (10 µl). Lane 5: Total protein content of pMAL-gp41T3 (5 µl). Lane 6: Total protein content of pMAL-gp41T2 (10 µl). Lane 7: Total protein content of pMAL-gp41T2 (5 µl). pMAL-gp41T3 is indicated in red as a band at 57.8 kDa. In Figure B: Lane 1: PageRuler Protein Ladder. Lane 2: Total protein content of pMAL-gp41(80) (1 µl). Lane 3: Total protein content of pMAL-gp41(80) (5 µl). pMAL-gp41(80) is indicated in red as a band at 54 kDa.

2.3.2.2 Protein purification of pMAL-gp41 fusion proteins

After the presence of the pMAL-gp41T3 protein had been confirmed on a 8.5% SDS-PAGE gel, the soluble fraction was used in the Amylose batch purification. From the 8.5% SDS-PAGE gel in Figure 2.10A it was evident that the protein purification was successful, since there was an intense band at 57.8 kDa (Lane 2-5). The purified protein was dialyzed overnight in Column buffer (20 mM Tris-HCl, 25 mM NaCl, pH 8.0) before using it in the assay. After the presence of the pMAL-gp41(80) protein had been confirmed on a 12% SDS-PAGE gel, the soluble fraction was used in the Amylose batch purification. From the 12% SDS-PAGE gel in Figure 2.10B it was evident that the protein purification was successful, since the band appeared at 54 kDa (Lane 3).

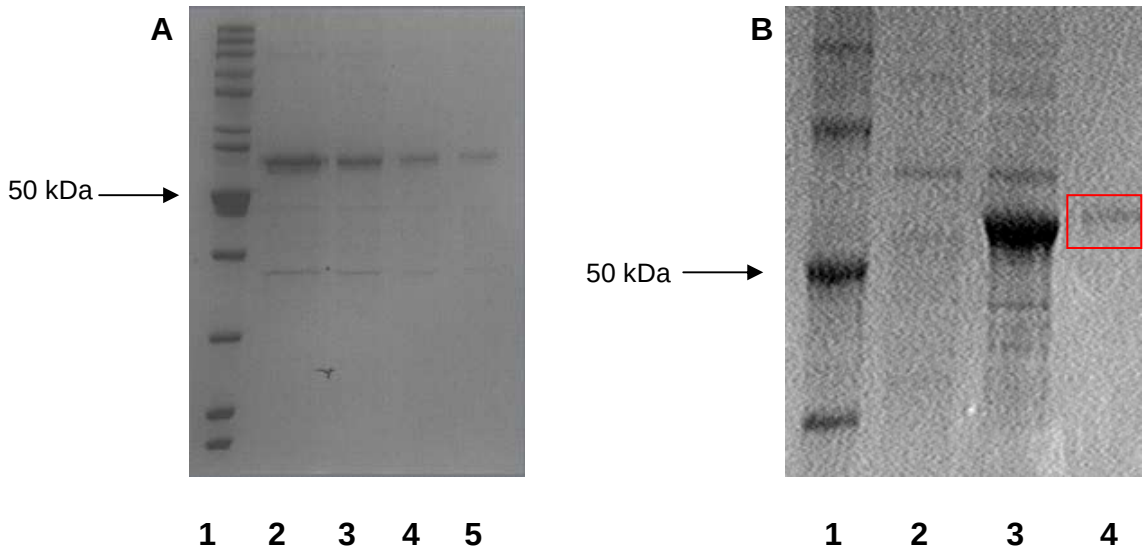

Figure 2.10: **SDS-PAGE analysis of affinity protein purification with amylose resin batch purification.** In Figure A: Lane 1: PageRuler Protein Ladder. Lane 2: pMAL-gp41T3 (4 µg/ml). Lane 3: pMAL-gp41T3 (2 µg/ml). Lane 4: pMAL-gp41T3 (1 µg/ml). Lane 5: pMAL-gp41T3 (0.5 µg/ml). The pMAL-gp41T3 band can be seen at 57.8 kDa. In Figure B: Lane 1: PageRuler Protein Ladder. Lane 2: Total protein content of pMAL-gp41(80) (1 µl). Lane 3: Total protein content of pMAL-gp41(80) (5 µl).Lane 4; pMAL-gp41(80). pMAL-gp41(80) is indicated in red as a band at 54 kDa.

In Figure 2.10A a contaminating band just below 40 kDa can be seen (Lane 2-5). When the purified pMAL-gp41T3 and MBP-2 were analysed on the same SDS-PAGE the following was illustrated in Figure 2.11: the size of the MBP-2 is 43 kDa (Lane 1-4) and our contaminating band is just below 40 kDa (Lane 6-9). Due

to the small difference between the sizes of the two proteins the possibility is there that the contaminating band might be MBP-2. The formation of this contaminating band could be due to protein degradation during protein extraction and purification.

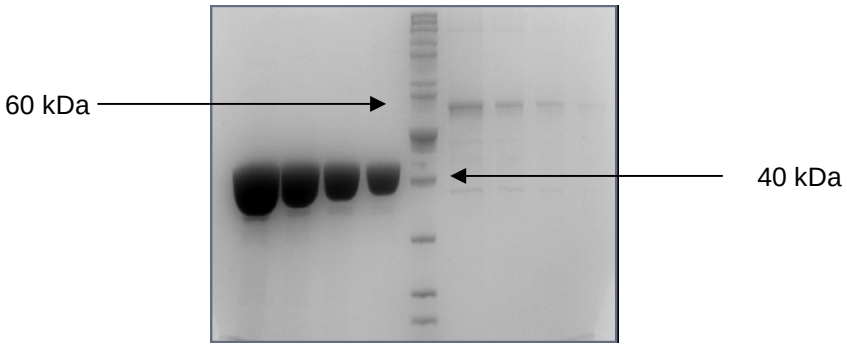

Figure 2.11: **SDS-PAGE analysis of affinity protein purification with amylose resin batch purification.** Lane 1: MBP-2 (20 µg/ml). Lane 2: MBP-2 (10 µg/ml). Lane 3: MBP-2 (5 µg/ml). Lane 4: MBP-2 (2.5 µg/ml). Lane 5: PageRuler Protein Ladder. Lane 6: pMAL-gp41T3 (4 µg/ml). Lane 7: pMAL-gp41T3 (2 µg/ml). Lane 8: pMAL-gp41T3 (1 µg/ml). Lane 9: pMAL-gp41T3 (0.5 µg/ml).

2.3.2.3 Protein concentration determination of pMAL-gp41T3 fusion protein

A bovine serum albumin standard curve for protein concentration was produced and quantified by the Folin-Lowry method (Lowry *et al.,* 1951). The standard curve shown in Figure 2.12 has an R-square value of 0.9913; this is an indicator of how well the linear regression model fits the data. By reading from the linear part of the standard curve the protein concentration of pMAL-gp41T3 was determined to be 1 mg/ml.

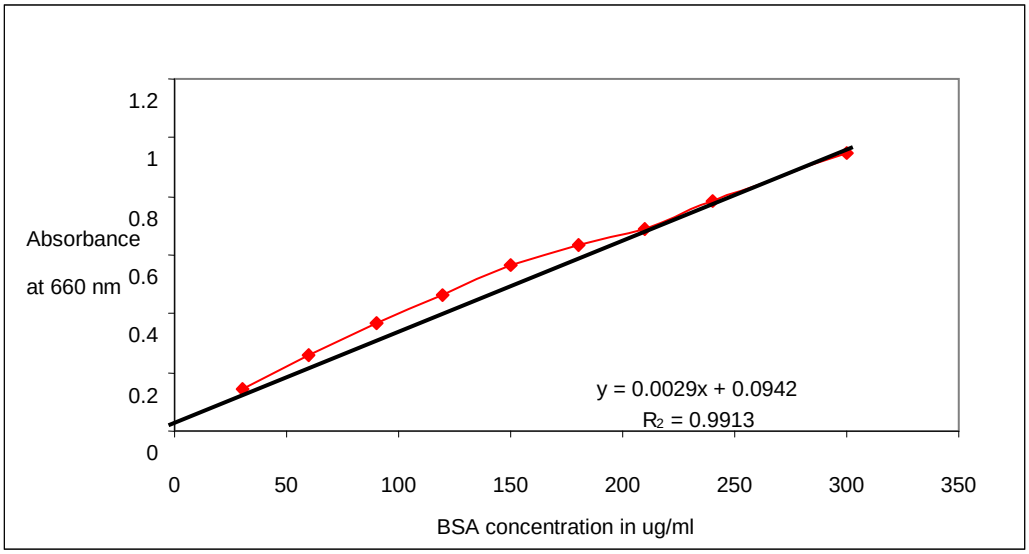

Figure 2.12: **Standard curve for protein concentration determination constructed by the Folin-Lowry assay.** Red line: Bovine Serum Albumin (BSA) protein standards. Black line: Linear regression of the data provided by BSA protein standards.

2.3.2.4 Protein identification by means of western blotting of pMAL-gp41T3 fusion protein

After the presence of the fusion protein pMAL-gp41T3 (Lane 4) was confirmed on a 8.5% SDS-PAGE gel (Figure 2.13), the protein was identified by means of western blotting by using a polyclonal antibody directed to gp41. From the Western blot in Figure 2.14 it was evident that the protein expression and purification of the fusion protein pMAL-gp41T3 (Lane 4) was successful due to the presence of the 57.8 kDa band. A comparison between the Western blot in Figure 2.14 and the 8.5% SDS-PAGE gel in Figure 2.13 (duplicate gels) shows that the purification of only the smallest fusion protein pMAL-gp41T3 was achieved and it could be detected with a polyclonal antibody directed towards it by western blotting.

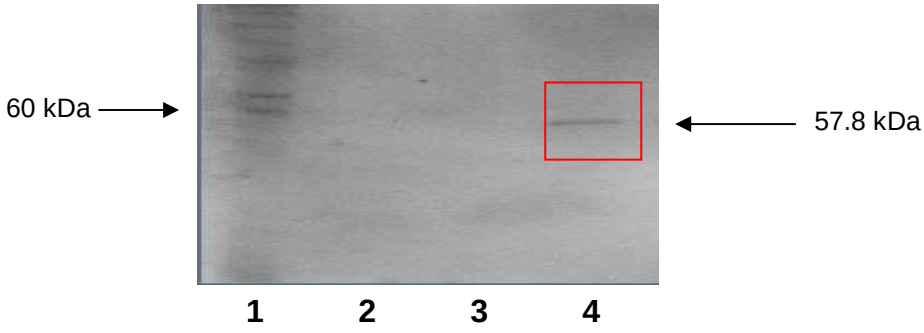

Figure 2.13: **SDS-PAGE analysis of affinity protein purification of different pMAL-gp41 fusion proteins with Amylose resin.** Lane 1: PageRuler Protein Ladder. Lane 2: purified pMAL-gp41T1. Lane 3: purified pMAL-gp41T2. Lane 4: purified pMAL-gp41T3. The 60 kDa band of the molecular marker is indicated and pMAL-gp41T3 is indicated in the red square as a band of 57.8 kDa.

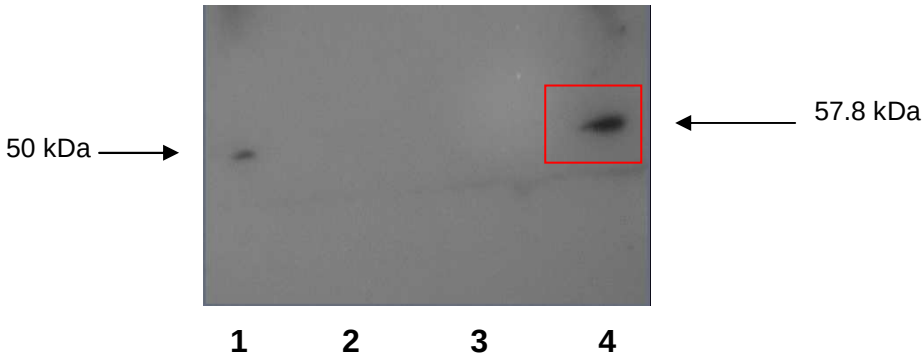

Figure 2.14: **Western Blot analysis of affinity protein purification of different pMAL-gp41 fusion proteins with Amylose resin.** Lane 1: PageRuler Protein Ladder. Lane 2: purified pMAL-gp41T1. Lane 3: purified pMAL-gp41 T2. Lane 4: purified pMAL-gp41T3. The 50 kDa band of the molecular marker is indicated and pMAL-gp41T3 is indicated in the red square as a band of 57.8 kDa.

Protein expression of the gp41T3 gene was obtained in the ER2507 cells as a fusion protein with maltose binding protein. Protein concentrations of 1 mg/ml pMAL-gp41T3 culture were observed. The identity of the pMAL-gp41T3 proteins was confirmed by western blotting by making use of a polyclonal antibody directed to gp41(Kennedy domain). Although gp41T3 is the smallest fragment cloned this protein can still be used for antigenicity studies. One of the reasons why the bigger fragments were not expressed might be due to the anchor domain which is located just before the start of the smaller fragment as indicated in Figure 2.2.

The synthetic truncated fragments of the HIV gp41 gene as well as the codon optimized truncated HIV gp41 fragments were cloned into pET28a and pJC20 vectors and transformed into different *E. coli* strains for expression, with little success. The codon optimized 052741 fragment cloned into pJC20 and transformed into BL21 DE3 was expressed in the insoluble fraction and solubulized with 1mM Tris-HCl pH 7.5 buffer. The results are not shown.

2.3.3 Cloning of synthetic gp41 genes for expression in *P. pastoris*

2.3.3.1 Restriction enzyme digestion of the pPIC9-gp41T1 construct

After sub-cloning of gp41T1 into pGem T-easy, enzyme digestion and ligation of the construct into pPIC9 the plasmid was transformed into GS115 *P. pastoris* cells via electroporation. The presence of the plasmid in the cells was confirmed by antibiotic selection with 100 µg/ml ampicillin as the pPIC9 plasmid carries an ampicillin resistance marker gene. After alkaline lysis and plasmid isolation (peqGOLD plasmid Miniprep Kit 1, Optima Scientific), the presence of the gp41 gene in the plasmid was confirmed with restriction enzyme digestion using *Eco*RI and *Not*I. The gel in Figure 2.15 shows the bands (Lane 2).

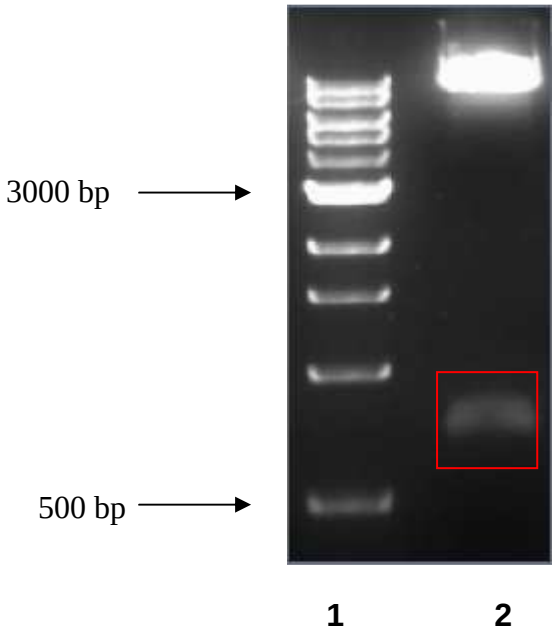

Figure 2.15: ***Eco*RI and *Not*I restriction analysis of the pPIC9-gp41T1 construct.** Lane 1: 1 kb DNA Ladder (New England Biolabs, Boston, USA). Lane 2: Restriction digestion of the gp41T1 plasmid. The 3000 bp and 500 bp bands of the 1 kb DNA Ladder are indicated.

Indicated in Figure 2.15 the bigger band is of size 8000 bp (Lane 2) which represents the pPIC9 plasmid back-bone and the smaller band indicated in the red block (Lane 2) is the gp41T1 gene with a size of 696 bp. This was an indication that the GS115 *P. pastoris* cells were transformed with the pPIC9 plasmid.

2.3.3.2 Protein expression of the pPIC9 construct

The smallest fragment of the gp41 gene (gp41T3) was expressed and purified in *E. coli;* hence there was no need to express it in *P. pastoris*. The expression of gp41T2 was unsuccessful in both *E. coli* and *P.pastoris* therefore, only the expression of gp41T1 will be described belowe.

The gp41T1 (full length peptide) was cloned into pPIC9a secreted expression vector and *P. pastoris* (GS115) cells were transformed. From the 12% SDS-PAGE in Figure 2.16 it was clear that secreted expression did take place. Secreted expression of gp41 T1 took place seventeen hours after induction at 30 ℃ as indicated by the 27.8 kDa protein band (Lane 2-6) in Figure 2.16.

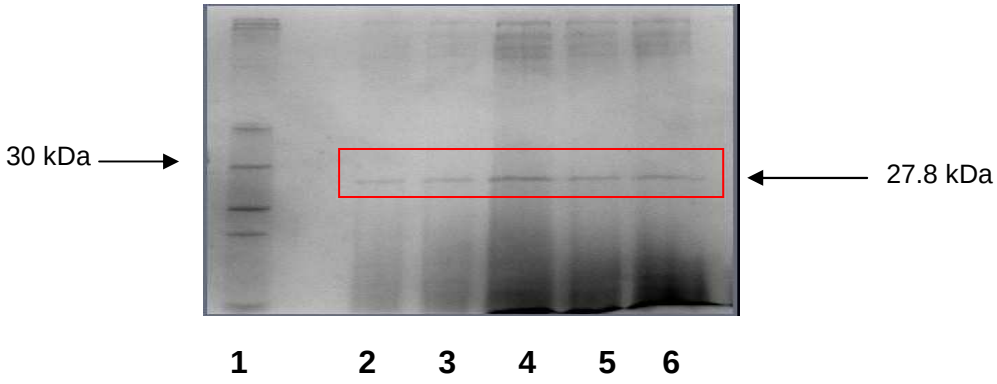

Figure 2.16: **SDS-PAGE analysis of protein expression of gp41T1 in *P. pastoris* at 30 ℃ after 17 hours.** Lane 1: PageRuler Protein Ladder. Lane 2: Secreted gp41T1 (3 µl). Lane 3: Secreted gp41T1 (5 µl). Lane 4: Secreted gp41T1 (7 µl). Lane 5: Secreted gp41T1 (10 µl). Lane 6: Secreted gp41T1 (13 µl). gp41T1 indicated as a band at 27.8 kDa.

2.3.3.3 Protein identification by means of western blotting of the pPIC9a construct

After the presence of the gp41T1 protein (Lane 3-4) had been confirmed on a 12% SDS-PAGE gel Figure 2.17, the protein was identified by means of western blotting using Polyclonal antiboby 1577 from rabbit, directed against Kennedy domain of gp41 supplied by Dr. Hugh Brady. From the western blot in Figure 2.18 it was evident that secreted protein expression did take place. Protein expression was determined by a time study where samples were taken every two hours for the first 14 hours and then taken every hour for the next eight hours. These samples were analyzed on a 12% SDS-PAGE gel (complete gel not shown here). Optimal protein expression did take place after 17 hours of incubation at 30 ℃. A comparison between the weste rn blot in Figure 2.18 and the 12% SDS-PAGE gel in Figure 2.17 (duplicate gels) shows that the gp41T1 protein was expressed at 17 hours after induction (Lane 3-4).

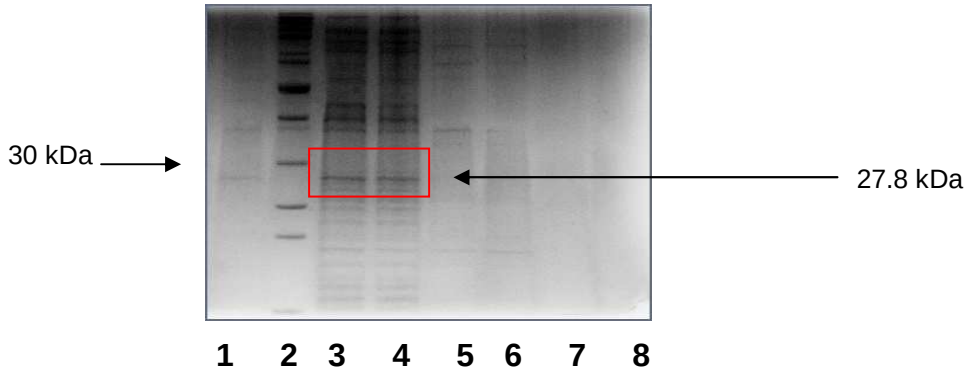

Figure 2.17: **SDS-PAGE analysis of protein expression of gp41T1 in *P. pastoris* at different time intervals.** Lane 1: gp41T1 17 hours after induction. Lane 2: PageRuler Protein Ladder. Lane 3 and 4: gp41T1 17 hours after induction. Lane 5: gp41T1 19 hours after induction. Lane 6: gp41T1 21 hours after induction. Lane 7 and 8: Negative control (pPIC9a transformed into GS115). gp41T1 is indicated in the red square as a band of 27.8 kDa.

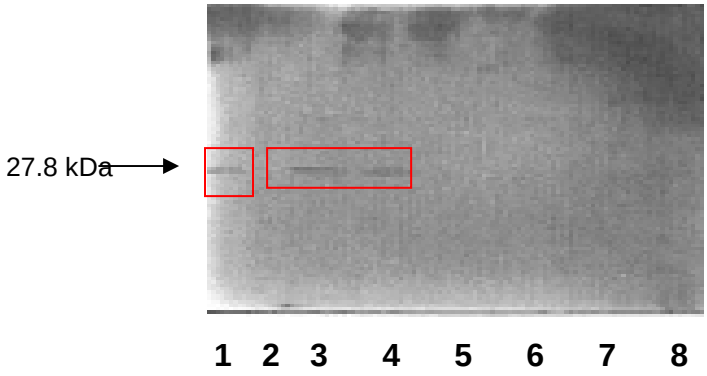

Figure 2.18: **Western Blot analysis protein expression of gp41T1 in *P. pastoris* at different time intervals.** Lane 1, 3 and 4: gp41T1 17 hours after induction. Lane 2: PageRuler Protein Ladder. Lane 5: gp41T1 19 hours after induction. Lane 6: gp41T1 21 hours after induction. Lane 7 and 8: Negative control (pPIC9a transformed into GS115). gp41T1 indicated as a band at 27.8 kDa.

Optimal protein expression of the gp41T1 gene was obtained in the GS115 strain of *P. pastoris*. During time studies it was determined that the protein was optimally expressed 17 hours after induction. Protein expression was confirmed by means of western blotting.

2.4 Discussion

During this study the HIV-1 gp41 gene was truncated into three gene fragments which vary in length namely 405 bp, 591 bp and 696 bp. The truncations were done to determine the smallest antigenic peptide of the gp41 endo-domain, which will indicate that the peptide or peptides present an antigenic epitope. This was compared with an ecto-domain gp41 peptide, gp41(80).

Functional expression of the genes in *E. coli* and *P. pastoris* was achieved. The 405 bp (gp41T3) peptide was expressed as a fusion protein with maltose binding protein in *E. coli*. The 260 bp (gp41T(80)) peptide was expressed as a fusion protein with maltose binding protein in *E. coli*. The 696 bp (gp41T1) peptide was expressed using *P. pastoris* as expression host.

The synthetic truncated fragments of the HIV gp41 gene as well as the codon optimized truncated HIV gp41 fragments were cloned into pET28a and pJC20 vectors and transformed into different *E. coli* strains for expression, with little success. The codon optimized 052741 fragment cloned into pJC20 and transformed into BL21 DE3 was expressed in the insoluble fragment and solubulized with 1 mM Tris-HCl pH 7.5 buffer.

Protein expression of the gp41T1 gene (696 bp) was obtained in the GS115 *P. pastoris* cells with a protein concentration of 0.6 mg/ml. Protein concentrations of 1 mg pMAL-gp41T3/ml culture were observed when the peptide was expressed as a fusion protein in *E. coli* ER2507 cells.

The identity of the gp41T1 and the pMAL-gp41T3 proteins was confirmed by western blotting by making use of a polyclonal antibody directed to the Kennedy domain of gp41.

One can therefore, conclude that cloning and protein expression of pMAL-gp41(80), pMAL-gp41T3 and gp41T1 was achieved. In Chapter 3 the antigenic and apoptotic properties of the three expressed proteins will be determined.

Chapter 3

Immunological characterization of synthetic gp41

3.1 Introduction

The involvement of the gp120-41 complex in viral infectivity and the presence of epitopes that are recognized by all arms of the immune system make the glycoproteins major targets for vaccine design. Individuals infected with HIV-1 develop a neutralizing antibody response, which is directed against epitopes of the gp120 and gp41 surface glycoproteins. Two major seroreactivity regions of gp120, the V3 loop region and the CD4 binding site, correlate closely with virus neutralizing ability and a few more are located on the transmembrane protein, gp41. Although the bulk of gp41 might be expected to be hidden from the humoral immune system within the Env oligomeric complex, several regions of this molecule are recognized by antibodies from individuals infected with HIV (Wyatt *et al*., 1998; Eckert *et al*., 2008).

The focus of this study excluded gp120 for the following reasons: Retroviruses are known to share structural epitope domains with host-cell proteins which are described as molecular mimicry and has been suggested as a fundamental mechanism underlying the pathogenicity of auto-immune disease. gp120 and MHC II on the CD4 T-cells of the human immune system do have consensus sequence similarity. This may lead to the formation of Ab against gp120 but also against MHC II which will lead to the destruction of the human immune response (Tishkoff *et al.,* 2000; Veljkovic *et al.,* 2005).

3.1.1 Immunodominant epitopes of gp41

3.1.1.1 Cluster I

The first of the immunodominant epitopes of gp41 is often referred to as Cluster I which extends from around residue 588 to residue 615. The immunodominant epitope is centred around a seven-residue peptide (CSGKLIC) and contains an intramolecular disulfide bond forming a disulfide bonded loop from residues 607

to 613 which is located within the central region of the gp41 ecto-domain (Gnann *et al*., 1987b; Gnann *et al*., 1987c). NMR studies showed that this domain maintains a preference for a folded conformation, including a type I reverse turn about the residues SGKL. The presence of the disulfide bond is integral to the formation of the structure of the loop in solution. Mutations (Cys to Ser) which eliminate the possibility of loop formation resulted in the failure of antibody binding to this peptide (Oldstone *et al*., 1991).

Almost all sera from infected individuals contain antibodies that react with this epitope (Eberle *et al*., 1997) but they are not neutralizing and do not block infection by HIV. On the contrary, monoclonal antibodies reactive with this region have been reported to enhance infection through a complement-dependent mechanism (Robinson *et al*., 1990a; Robinson *et al*., 1990b).

3.1.1.2 Cluster 2

In addition to this primary immunodominant epitope (Cluster 1), a second ecto-domain region, located proximal to the membrane-spanning region of gp41, is also recognized by sera from infected individuals (Horal *et al*., 1991). While a majority of antibodies reactive with this region do not neutralize HIV infection, one human monoclonal antibody (2F5) has been identified that neutralizes a variety of laboratory strains and clinical isolates of HIV-1. The amino acid sequence ELDKWA (residues 671–676) has been defined as the epitope recognized by this antibody (Muster *et al*., 1993). Sequence analysis of a variety of primary isolates suggests that the major determinant of MAb 2F5 binding corresponds to the amino acid sequence LDKW, since naturally occurring and *in vitro* selected neutralization-resistant viruses contain changes in the D and K positions of the ELDKWA motif (Purtscher *et al*., 1996). While 2F5 recognizes a putatively linear determinant present on short peptides, it seems likely that the epitope reflects a specific conformation of the contributing amino acids, since attempts to elicit neutralizing antibodies to a synthetic peptide encompassing this region or to the epitope presented on the hepatitis B surface antigen have been unsuccessful (Eckhart *et al*., 1996). However, insertion of this epitope into the loop of antigenic site B of the influenza virus hemagglutinin did elicit ELDKWA-specific, neutralizing immunoglobulins in antisera of mice (Muster *et al*., 1994).

The mechanism of neutralization is not understood, although Neurath and colleagues have reported multiple effects of binding the 2F5 monoclonal antibody to virions that are suggestive of post-binding conformational changes in Env (Neurath *et al*., 1995). It is also relevant that the 2F5 epitope overlaps the amino acid sequence of a peptide derived from gp41 that is a potent inhibitor of fusion between the virus and the host cell and a tryptophan-rich region which plays a critical role in fusion and infectivity (Vishwanathan *et al.,* 2008).

3.1.1.3 Cluster 3

An enigma, as far as anti-gp41 neutralizing antibodies are concerned, has been the fact that monoclonal antibodies raised against a hydrophilic peptide corresponding to residues 743–760 (DRPEGIEEEGGERDRDRS) of the cytoplasmic domain could neutralize a variety of HIV-1 isolates. Both monoclonal antibodies and polyclonal antisera raised against this epitope have neutralizing potential (Dalgleish *et al*., 1988; Evans *et al*., 1989). There is currently no evidence that this region of the cytoplasmic domain can be translocated to the outer surface of the cell and so how such antibodies can neutralize has remained a puzzle. Recent experiments, however, indicate that a monoclonal antibody to this epitope can bind with equal efficiency to infected and uninfected CD4 positive cells (Sattentau *et al*., 1995), suggesting that the epitope is also present on a cell surface protein. Since it is clear that the lipid envelope of HIV incorporates a large number of cell surface proteins (Arthur *et al*., 1992), it is possible that neutralization by antibodies to this epitope is mediated through one of these cell-derived molecules. The transmembrane protein also elicits activity from the cellular arm of the immune response. This study will focus on the Kennedy domain (residue 743 – 760) which is located in Cluster 3 which is described above. The reason for the focus on this domain is its known status as an epitope for generation of protective antibodies and its homology to TNF. I address the hypothesis that this seemingly protective antigen domain might be responsible for the apoptosis of non-infected T-cells.

3.1.2 Antigen presentation

Immune cells present antigenic fragments to T-cells. The most common and well known mechanism for antigen presentation is by means of MHC I and MHC II molecules. Despite the structural similarity between the two classes they differ in many aspects. MHC I is recognized by CD8 cytotoxic T lymphocytes (CTL), whereas MHC II is recognized by CD4 helper T-cells. The type of peptide which binds to the two classes also differs from each other. In the case of MHC I it will bind shorter peptides which are normally between 8 and 10 amino acid residues where both the N- and C-terminal ends of the peptide will bind into the binding groove of MHC I. MHC II can bind peptides between 9 to 22 residues where both N- and C-terminal ends extend beyond the binding grove (Stern *et al*., 1994).

The way B and T lymphocytes recognize lipid antigens differs greatly. B-cells use membrane-bound antibodies that interact with the hydrophilic heads of glycolipid antigens, i.e. they recognize exposed shapes of the molecules protruding from a hydrophobic inner structure. T-cells instead use the TCR αβ or γδ and recognize complexes formed by lipids associated with dedicated antigen-presenting molecules, which belong to the CD1 family. In humans the CD1 genes encode 5 proteins, CD1a, b, c, d and e. CD1 antigen presenting cells are a lineage of antigen presenting cells with the ability to present lipids and glycolipids to T-cells. The presentation mechanism of lipids has physical requirements dictated by the very nature of lipids. As lipids are poorly soluble in water they are always associated with membranes or with lipid-transfer proteins (LTP) in tissues and biological fluids. This emphasizes the importance of lipid-binding proteins in lipid uptake by antigen-presenting cells (APC), in lipid transport inside cells and in lipid handling until CD1 loading. The different CD1 isoforms have different lipid-binding specificity with some redundancy in function. This has a direct impact on the repertoire of lipid antigens that stimulate T-cells. The intracellular compart-mentalization of lipids may facilitate or hinder their capacity to form complexes with CD1 molecules, thus directly contributing to lipid immunogenicity. The structure of CD1 molecules, their trafficking capacity as well as the structures of immunogenic lipids have been extensively described in a number of reviews (Sugita *et al.,* 2000; Kasmar *et al.,* 2009) and therefore, will not be discussed here.

3.1.3 Antibody directed against the immunodominant epitopes of gp41

Despite the induction of an intense virus-specific antibody (Ab) response during HIV infection, neutralizing Abs able to inhibit infection of a broad range of primary isolates are hardly ever detected in sera and plasma from HIV infected individuals (Eckert *et al*., 2008).

Neutralisation by antibody is, for a number of viruses, an *in vitro* correlate for protection *in vivo*. For HIV-1 this is controversial. However, the induction of a potent anti-HIV neutralising antibody response remains one of the principal goals in vaccine development. A good knowledge of the fundamental mechanisms underlying the neutralizing process would help direct research towards suitable vaccine immunogens. The primary determinant of HIV neutralisation appears to be antibody affinity for the trimeric envelope glycoprotein spike on the virion, suggesting that epitope-specific effects are secondary and implying a single, dominant mechanism of neutralisation. A gp41 specific antibody neutralises by interfering with post-attachment steps leading to virus membrane fusion (Eckert *et al*., 2008).

This study does not only focus on the identification of a specific epitope which can be used for vaccine development, but also on the apoptotic properties of specific domains of gp41. In particular, the gp41T1 peptide and the pMAL-gp41T3 fusion peptide which are both located on the C-terminal of gp41 and the pMAL-gp41(80) fusion peptide which is located on the N-terminal of gp41 is investigated to contribute to the general understanding of the biology of HIV/AIDS.

3.1.4 $CD4^+$ T-cells killing during HIV/AIDS infection

During HIV/AIDS infection both infected and non-infected $CD4^+$ T-cells are killed. The mechanism or mechanisms of $CD4^+$ T-cell depletion is still an unresolved issue in HIV/AIDS research, but since 1991 T-cell apoptosis has been proposed as one (Ameisen and Capron, 1991; Cotton *et al.,* 1997). Several mechanisms have been proposed to explain the accelerated apoptosis of infected and uninfected cells in HIV patients: (i) the direct role of HIV specific proteins, (ii) activation induced cell death, (iii) direct infection of T-cells, (iv) autologous cell-

mediated killing of uninfected T-cells and (v) disregulation of cytokine/chemokine production (Phenix and Badley, 2002). Due to the observation that the degree of cell loss largely exceeds the number of infected cells, HIV-1-induced apoptosis in bystander T-cells is a likely theoretical reason for the depletion of T-cells in HIV patients. Several HIV-1 proteins like HIV envelope glycoproteins (Env), Tat, Vpr, Nef, Vpu and the protease can induce apoptosis in T-cells. The impact of this process is not fully understood, but cumulative data demonstrated that Env fulfils an important role in uninfected T-cell death (Laurent-Crawford *et al.,* 1993; Ohnimus *et al.,* 1997; Heinkelein *et al.,* 1998; Roshal *et al.,* 2001). One question is whether $CD4^+$ T-cell death is due to the immunological consequences of viral infection or whether it is due to the release of viral peptides that act as T-cell apoptosis mediators (Zhou *et al.,* 2005). If by induced apoptosis, the elucidation of the mechanism by which it occurs will contribute to the better understanding of the basic biology of HIV/AIDS.

3.1.4.1 Tumor necrosis factor

Tumor necrosis factor (TNF) is the prototypic member of an emerging family of cytokines that function as prominent mediators of immune regulation and the inflammatory response. The apparent involvement of TNF in septic shock, autoimmune disorders, and graft-versus-host disease is well documented (reviewed by Revel and Schattner, 1987; Cerami and Beutler, 1988). A unique feature of this family of ligands is the ability of some members to induce directly the apoptotic death of chronically activated T-cells and B-cells (Daniel and Krammer, 1994; Alderson *et al.,* 1995). Likewise, TNF has been shown to induce apoptosis in normal thymocytes under appropriate conditions (Hernandez-Caselles and Stutman, 1993). Further, peripheral T-cells from HIV infected individuals have been shown to be much more sensitive to Fas-mediated apoptosis than uninfected controls (Katsikis *et al*., 1995).

In this study, it was deemed important to determine both the humoral antigenicity and apoptotic nature of the mentioned gp41 viral peptides.

3.1.5 Aims

In order to characterize the specific domains of gp41 for its potential towards AIDS prevention and therapy, the following aspects are addressed:

- Ability to induce antibodies in HIV infected human patients
- Correlation between antibody binding signal in ELISA with AIDS progression (CD4 count) of the patient
- Ability to induce apoptosis in non-infected T lymphocytes

An ideal and commercially exploitable outcome of the research would be the finding that a particular part of the gp41 cytoplasmic domain or ecto-domain induces antibodies in patients that correlate positively with CD4 count (indicative of protection) and displays little apoptotic effect (not toxic). This ideal outcome serves as the hypothesis to be tested experimentally:

Hypothesis: gp41 peptides pMAL-gp41T3, pMAL-gp41(80) or gp41T1 induces neutralizing antibodies with minimal apoptotic side-effect.

3.2 Materials and Methods

3.2.1 Immunological characterization

Serum samples were obtained from 44 HIV-infected individuals by Dr Gunter Schleicher of the Department of Pulmonology at the University of the Witwatersrand, South Africa. These serum samples formed part of a bigger serum collection that was used in a different study (Schleicher *et al*., 2002)

The blood was drawn from patients in sterile Vacutainer tubes (with brown lids, Aquila Health Care, Pinegowrie, SA) by Dr. Gunter Schleicher at the Helen Joseph Hospital, Auckland Park, Johannesburg. These tubes were silicone coated and without any anticoagulants. Samples were stored and transported to the laboratory at 3-5 ºC. After the blood clotted, serum was removed from the blood clot with plastic pipettes. The serum samples were then centrifuged in an Eppendorf centrifuge (326 x g, 5 min, 4 ºC). This was done to remove any red blood cells that were still in the serum. All serum from one patient were then combined in a 10 ml tube and after gentle mixing samples were aliquoted in 500 µl portions into 1.8 ml cryo tubes (NUNC Brand products, Nalge Nunc international, Denmark) and stored at -70 ºC until use.

Samples were thawed and then γ-irradiated (30 Grates, Pretoria Academic Hospital) as an additional safety precaution. Both γ-irradiated and non-irradiated sera were analyzed in ELISA to confirm that antibody activity was not affected by γ-irradiation (data not shown).

The patient samples which were included in this study were selected according to the following criteria: The HIV status of the patients as HIV positive or HIV negative as well as the CD4 T-cell counts.

3.2.1.1 Precipitation of immune complexes

Serum samples were tested at 1:20 serum dilution. A mixture of 50 µl undiluted HIV patient sera and 50 µl of 8% PEG 8000 (8% w/v Polyethelyne glycol 8000, 0.01 M PBS, pH 7.4) was prepared to precipitate the immune complexes and IgG in patient serum by incubation at 4 ºC overnight. The mixtures were centrifuged at 4 ºC for 30 minutes at 4500 x g and the supernatants discarded. The precipitates were washed two times with 100 µl 4% PEG 8000 (4% w/v Polyethelyne glycol 8000, 0.01 M PBS, pH 7.4) and centrifuged at 4 ºC for 30 minutes at 4500 x g. The wash buffer (4% w/v Polyethelyne glycol 8000, 0.01 M PBS, pH 7.4) was completely removed. The precipitate was dissolved in 50 µl 1x PBS (pH 7.4) and 25 µl 0.2 M Glycine HCl (pH 2.8) and left on ice for 15 minutes to release the antibodies from the complexes. The precipitate was neutralised with 12.5 µl 1 M K_2HPO_4. This prevented the antibodies that were just dissociated to be damaged by the lower pH (Chia *et al.,* 1979). H_2O (12.5 µl) and 0.5% Casein/PBS (900 µl) were added to the samples to obtain a final volume of 1 ml and a 1:20 dilution equivalent of the original serum. All samples were kept on ice throughout the whole experiment.

3.2.1.2 Enzyme-linked Immunosorbant Sandwich Assay (ELISA)

The wells of a flat bottom 96 wells plate (Bibby Sterilin Ltd., Sterilab, UK) were coated with 50 µl of the pMAL-gp41T3 or pMAL-gp41(80) fusion peptides obtained after affinity chromatography (3 µg per well) and incubated at 4 ºC for 16 hours. For the negative control the wells were also coated with the two peptides and negative HIV patient sera was used in the next step. The fluid was flicked out and the plates were blocked with 400 µl blocking buffer (0.01 M PBS, 0.5% casein, pH 7.4) for 2 hours at room temperature. The plates were washed once with blocking buffer and aspirated. Plates were incubated with 50 µl of the precipitated immune complex samples (reconstituted to the equivalent of 1:20 dilution of serum in blocking buffer) for 60 minutes at room temperature. Plates were washed four times in blocking buffer and aspirated. The plates were incubated with an appropriate anti-IgG (whole molecule) peroxidase conjugate (Amersham Biosciences, UK) at a 1:10 000 dilution. After a second washing step, 50 µl developing buffer (10 ml citrate, 10 mg OPD and 8 mg H_2O_2, pH 4.5) were added and the reaction monitored using the SLT 340 ATC photometer with

a 450 nm measuring filter and a 690 nm reference filter.

3.2.2 Flow cytometry analysis

Purified pMAL-gp41T3 at a concentration of 1 mg/ml was sent to Dr. Hugh J.M Brady at the University College London for apoptotic function determination by means of fetal thymic organ cultures (FTOCs) and flow cytometry.

The flow cytometer was equipped with a single 488 nm argon laser and 530 nm, 585 nm and > 650 nm filters for analysis in FL1, FL2 and FL3 channels respectively.

The FTOCs cells were treated with the pMAL-gp41T3 compound for 16 hours at doses of 1, 0.1 and 0.01 µg/ml. The FTOC cells were aliquotted at 1 x10^5 cells per well of a V-bottomed 96-well plate. The cells were stained with DRAQ5 DNA dye (Sigma-Aldrich Chemicals Co.) to quantify the apoptosis. The cells were washed once with PBS - 0.1% w/v NaN_3 and resuspended in 200 µl of binding buffer (10 mM *N*-hydroxyethylpiperazine-*N*'-2-ethanesulfonate, pH 7.4; 150 mM NaCl; 5 mM KCl; 1 mM MgCl; 1.8 mM $CaCl_2$). Annexin V (FITC labelled) obtained from Pharmingen was added to a final concentration of 1 µg/ml and incubated for 15 min at 37 ºC in the dark. The cells were analyzed immediately with a flow cytometer detecting annexin V staining in the appropriate FL channel (FL1 for FITC) on a logarithmic scale. Apoptotic cells will show increased binding of annexin V.

3.3 Results

3.3.1 Antigenicity of pMAL-gp41T3

After purification and protein determination of pMAL-gp41T3 the antigenicity of this fusion peptide and its MBP-2 fusion partner alone was tested against HIV positive patients by means of ELISA with 12 randomly chosen HIV positive patient sera. The negative control consisted of 5 HIV negative patient serums that were pooled. The results in Figure 3.1 revealed that fusion protein (pMAL-gp41T3), but not its fusion partner MBP-2, showed antigenicity when it was tested on the ELISA assay. This finding confirmed that the fusion of gp41T3 to MBP-2 did not destroy the former's antigenicity and that human patient antibodies present in HIV positive patient sera did not cross-react with the MBP-2 fusion partner.

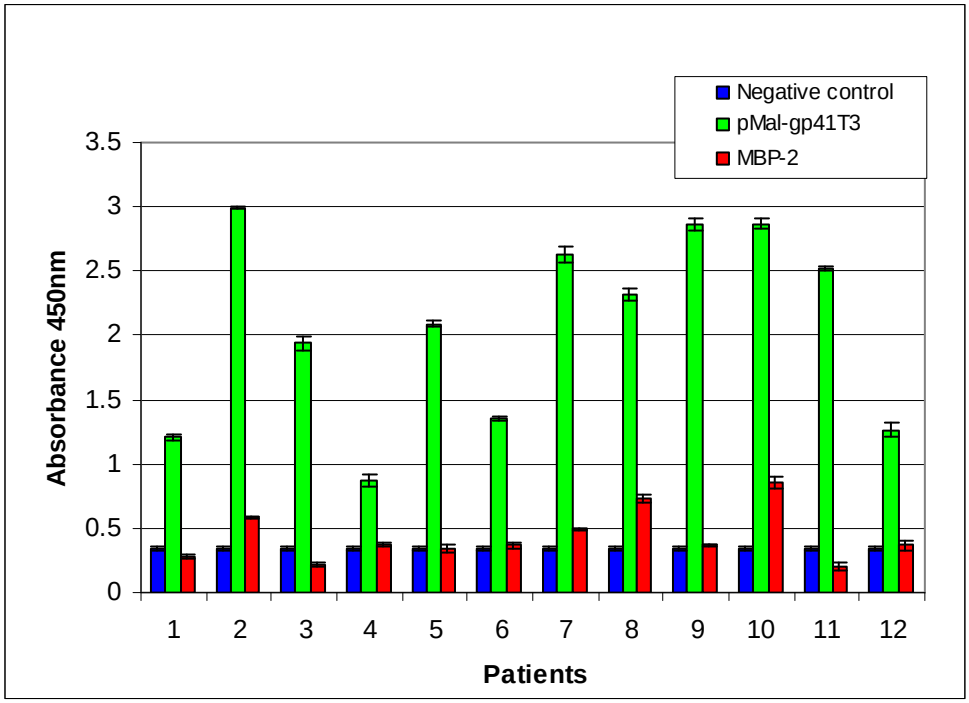

Figure 3.1: **Antigenicity of pMAL-gp41T3 and MBP-2 towards HIV positive patient sera determined by ELISA:** Twelve random HIV positive patient sera samples were tested against both pMAL-gp41T3 and MPB-2. A negative control consisting of 5 HIV negative patient serums which were pooled was also tested against both pMAL-gp41T3 and MPB-2. The blue bars indicate the HIV negative sera control. The green bars indicate the twelve different HIV positive sera samples tested against pMAL-gp41T3. The red bars indicate the twelve different HIV positive sera samples tested against MBP-2. The standard deviation calculated from 3 repeats is shown by the error bars.

In other words: the specificity of the antibodies is only for the HIV peptide component of the fusion protein and not for the MBP-2. After the specificity of the fusion protein was confirmed by testing both the fusion protein and the MBP-2, the antigenicity of pMAL-gp41T3 was determined by testing all the available HIV positive patient sera in the collection as well as the negative control. Figure 3.2 reveals that pMAL-gp41T3 showed universal antigenicity to all the patients tested with the ELISA assay. The experiment turns out in favour of an immune active peptide in respect of antibody elicitation.

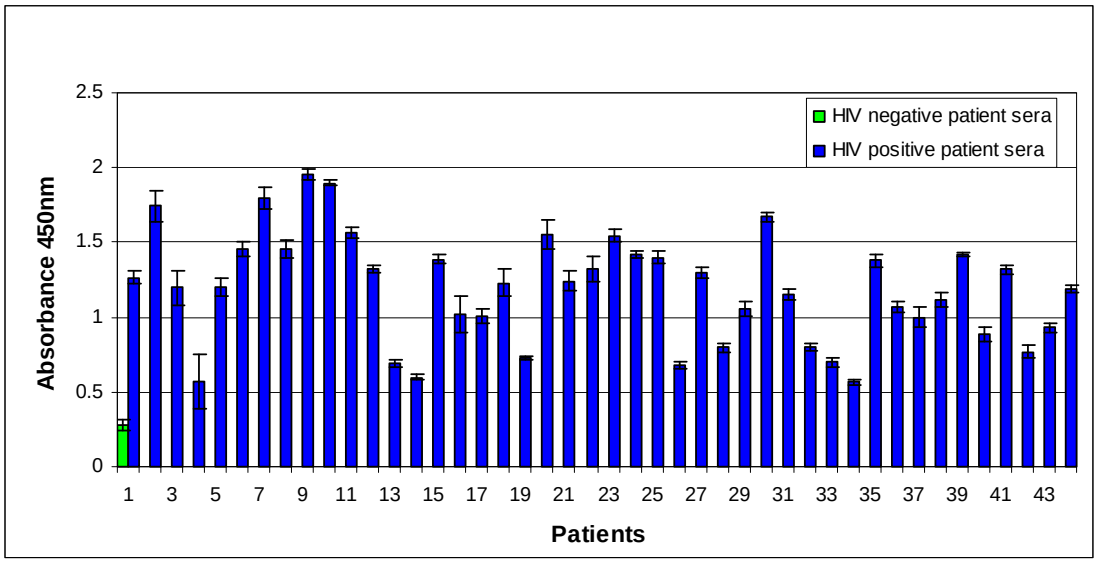

Figure 3.2: **Antigenicity of pMAL-gp41T3 towards HIV positive patient sera and HIV negative patient sera determined by ELISA:** The blue bars indicate the HIV positive patient sera response towards pMAL-gp41T3. The green bar indicates the HIV negative patient sera response towards pMAL-gp41T3. The population size used in the study consists of 44 HIV positive patient sera, which were compared to 7 HIV negative patient sera that were pooled. The standard deviations for 3 repeats are shown by the error bars.

3.3.2 Antigenicity of pMAL-gp41(80)

The antigenicity of pMAL-gp41(80) was determined by testing 38 HIV positive patient sera as well as the negative control. The negative control consisted of 5 HIV negative patient sera that were pooled. Figure 3.3 reveals that pMAL-gp41(80) showed antigenicity to all the HIV positive patient sera tested as compared to that of the pooled HIV negative sera. The experiment turns out in favour of an immune active peptide in respect of antibody elicitation.

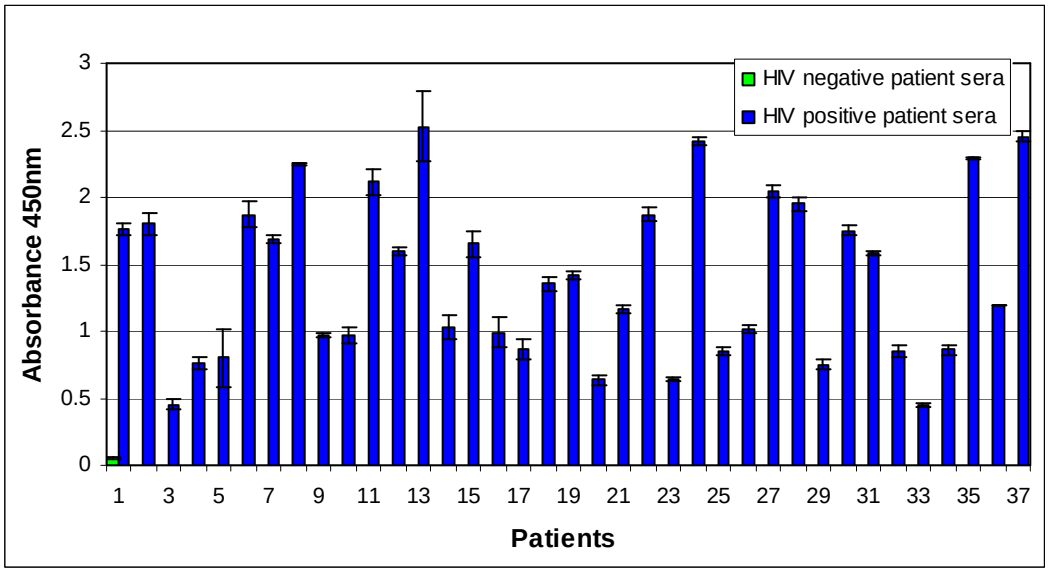

Figure 3.3: **Antigenicity of pMAL-gp41(80) towards HIV positive patient sera and HIV negative patient sera determined by ELISA:** The blue bars indicate the HIV positive patient sera response towards pMAL-gp41(80). The green bar indicates the HIV positive patient sera response towards pMAL-gp41(80). The population size used in the study consists of 37 HIV positive patient sera, which were compared to 7 HIV negative patient sera that were pooled. The standard deviations calculated from 3 repeat measurements are shown by the error bars.

3.3.3 Correlation of the antigenicities of pMAL-gp41(80) and pMAL-gp41T3 to the CD4 T-cell counts of HIV positive patients

It was expected that the antibody binding signal to the HIV peptides will increase as the HIV infection progresses, due to the immune stimulatory effect of increased antigen concentration, until AIDS sets in that destroys the antibody response. Therefore, a negative correlation was expected between the $CD4^+$ T-cell count of the patients and the antibody binding signal. Alternatively, a neutralizing antibody would be expected to reduce its binding capacity as it fails to fight off the viral infection, thereby showing a positive correlation between CD4 count and antibody binding signal. The antibody binding signals were determined using the ELISA assay as outlined in 3.2.1.2 of the methods. The $CD4^+$ T-cell counts of the patients are presented in Appendix A.

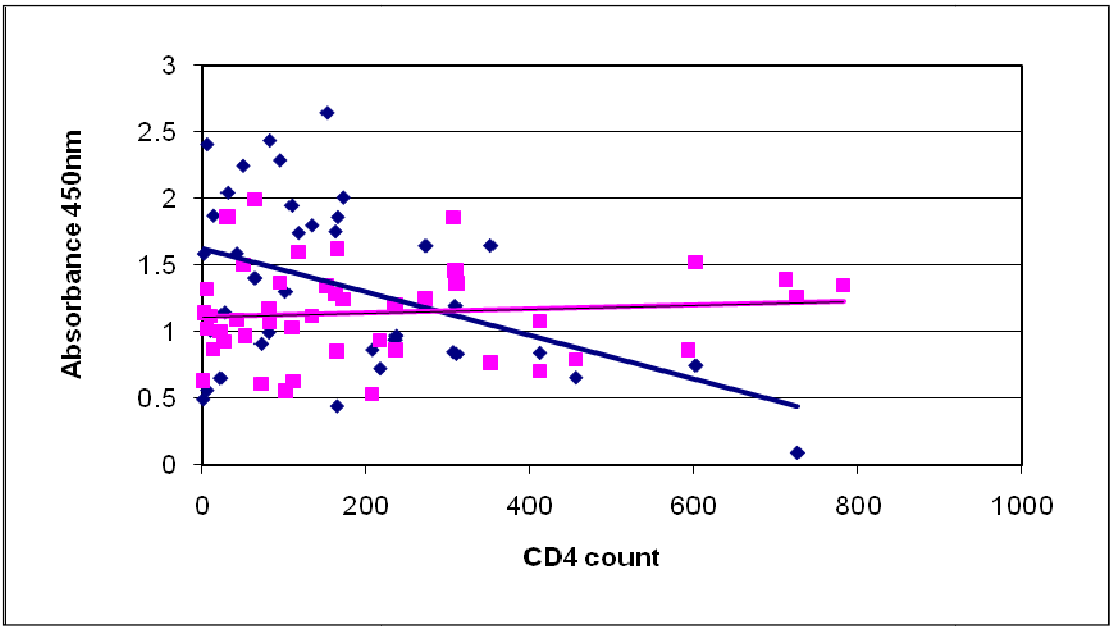

Figure 3.4: **The absorbance values of pMAL-gp41(80) and pMAL-gp41T3 in relation to patient CD4 T-cell count.** The relationships between the absorbance values of pMAL-gp41T3 and the CD4 count is indicated by the purple squares. The correlated trend line is indicated by the purple line with a correlation coefficient of r = 0.02717. The relationship between the absorbance values of pMAL-gp41(80) and the CD4 count is indicated by the blue diamonds. The correlated trend line is indicated in blue with a correlation coefficient of r = -0.2088.

From the results obtained in Figure 3.4 the following conclusions were made: The comparison between the two trend lines of pMAL-gp41T3 and pMAL-gp41(80) indicates that there is a different relationship between the absorbance values and the CD4 counts of the two sets of data. Correlation coefficient can be used to quantify the relationship between two sets of data. The pMAL-gp41(80) results (blue diamonds) in Figure 3.4 show that there is a negative correlation between the antibody binding signals and the CD4 count in the case of the pMAL-gp41(80) peptide as indicated by the correlation coefficient of -0.2088. In contrast, pMAL-gp41T3 (purple squares) in Figure 3.4 has no correlation between the CD4 count and the antibody binding with a correlation coefficient of 0.02717. This shows that the production of antibodies to pMAL-gp41T3 and pMAL-gp41(80) might be regulated differently by antigen presentation in the patient. Both peptides proved to be antigenic and may therefore, be considered as potential vaccine candidates. Gp41-80 stimulates antibody production and/or antigen binding affinity as HIV infection proceeds towards AIDS, while gp41T3 maintains the antibody levels and/or antigen binding affinity irrespective of how

low CD4 counts become. In the case of gp41(80), the antibody binding signal may simply indicate that antibody activity is increasingly induced as more viral antigen is released into the circulation. This is typical of an adaptive antibody response to higher concentrations of antigen that is mediated by $CD4^+$ T-cell help. For the reason that the absorbance values against pMAL-gp41(80) correlates with AIDS, this can also represent the pattern of a pathological anti-self Ab which may contribute to the transmission from HIV infection to AIDS. If this is the case, it will be very dangerous to use gp41(80) as a vaccine. However, gp41T3 induces an antibody response that is independent of antigen concentration and the state of progression of the HIV infection. This is a peculiar response that may find explanation by a different mode of antibody induction for example, by means of CD1 presentation to NKT (Cohen *et al.,* 2009; Kasmar *et al.,* 2009) cells. Antibodies to viral peptides that are not affected by the degree of infection with virus are ideal diagnostic markers for HIV infection.

3.3.4 Protein hydrophobicity determination of pMAL-gp41T3 and pMAL-gp41(80)

The most well known method for antigen presentation is the presentation of antigens via the MHC-molecules on antigen presenting cells (APC). The antigenic peptides are presented by the MHC class I or II molecules and recognized by CD8 or CD4 T-cells respectively. In 1994, another class of antigen presenting cells was identified, namely CD1. CD1 is a family of antigen-presenting proteins that are involved in presenting lipid and glycolipid proteins to CD4/CD8 double negative T-cells (Bendelac *et al.,* 2001). The hypothesis of CD1 antigen presentation stipulates that CD1 has evolved to present lipids and glycolipids to T-cells. The model currently used is that fatty acid tails of glycolipid or lipidated proteins, or the hydrophobic parts of the peptides themselves will fit into the hydrophobic pockets of the binding site of CD1 and the polar part of the peptide protrude out of the binding site and bind to the TCR.

The hydrophobicity plots for both the peptides were determined to see if there is a chance that one of the two peptides can be regulated by the CD1 presenting pathway, while the other could be the normal MHC-II presenting pathway. This would be an explanation for the different results obtained when the CD4 count

was correlated with the antibody binding signals. Kyte-Doolittle is a widely applied scale for delineating hydrophobic character of a protein. The hydrophobicity plots were constructed by making use of DNAMAN (Lynnon Biosoft, Canada). Peptide stretches with values above 0 are hydrophobic in character while stretches with values below 0 are hydrophilic in nature. Figure 3.5 and Figure 3.6 represent the hyrophobicity of pMAL-gp41T3 and pMAL-gp41(80) respectively.

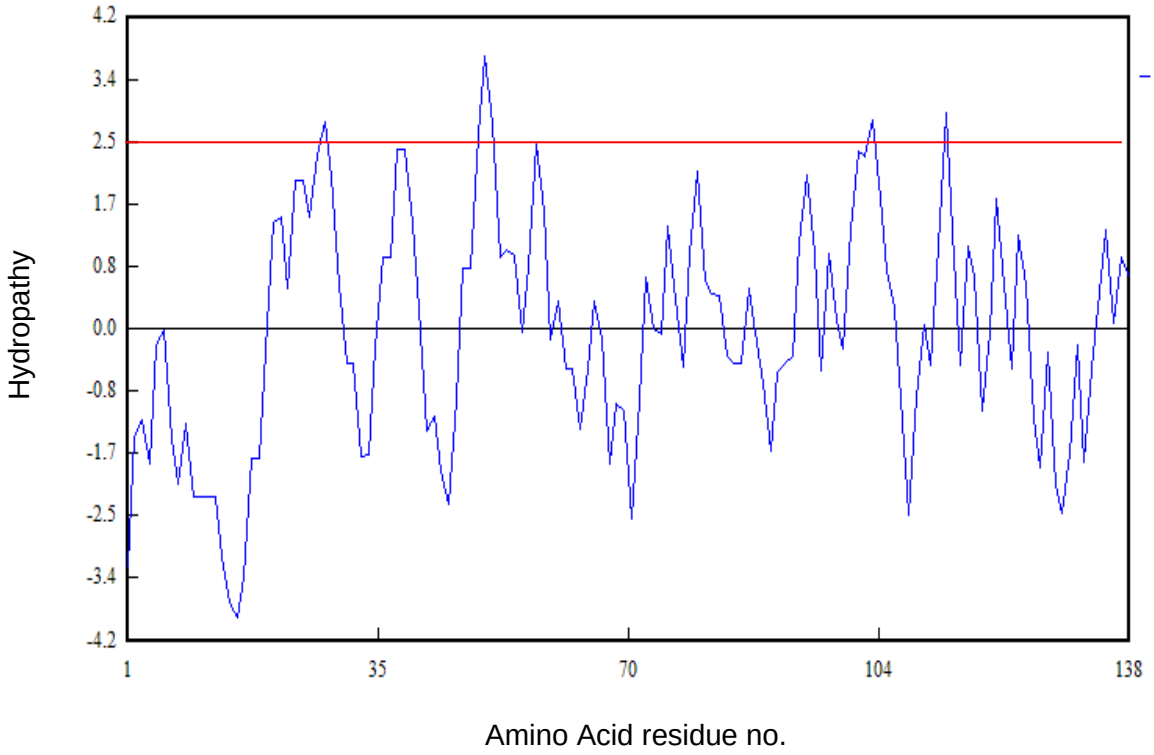

Figure 3.5: **Kyte-Doolittle hydrophobicity plot of gp41T3:** The regions with values above 0 indicate the hydrophobic character of the specific amino acids of the peptide while the regions with a value under 0 indicate the hydrophilic character around the specific amino acids of the peptide. The red line indicates the hydrophobic part on the peptide with a value of 2.5 and higher. A sample window of five amino acid residues was used in the construction of the plot.

From the two hydrophobicity plots indicated by Figure 3.5 and 3.6 it is clear that there are large regions on both the pMAL-gp41T3 and pMAL-gp41(80) peptides that are hydrophobic. The hydrophobic regions on the two peptide fragments were compared. The red line indicated in both the figures shows the regions of hydrophobicity at 2.5 or higher for the specific peptides. From Figure 3.5 it can been seen that gp41T3 exhibits four regions with a hydrophobicity index value

above 2.5, as opposed to only one in the case of pMAL-gp41(80), indicated in Figure 3.6. The degree of hydrophobicity of one peak in particular in pMAL-gp41T3 is huge, in comparison to that found in pMAL-gp41(80).

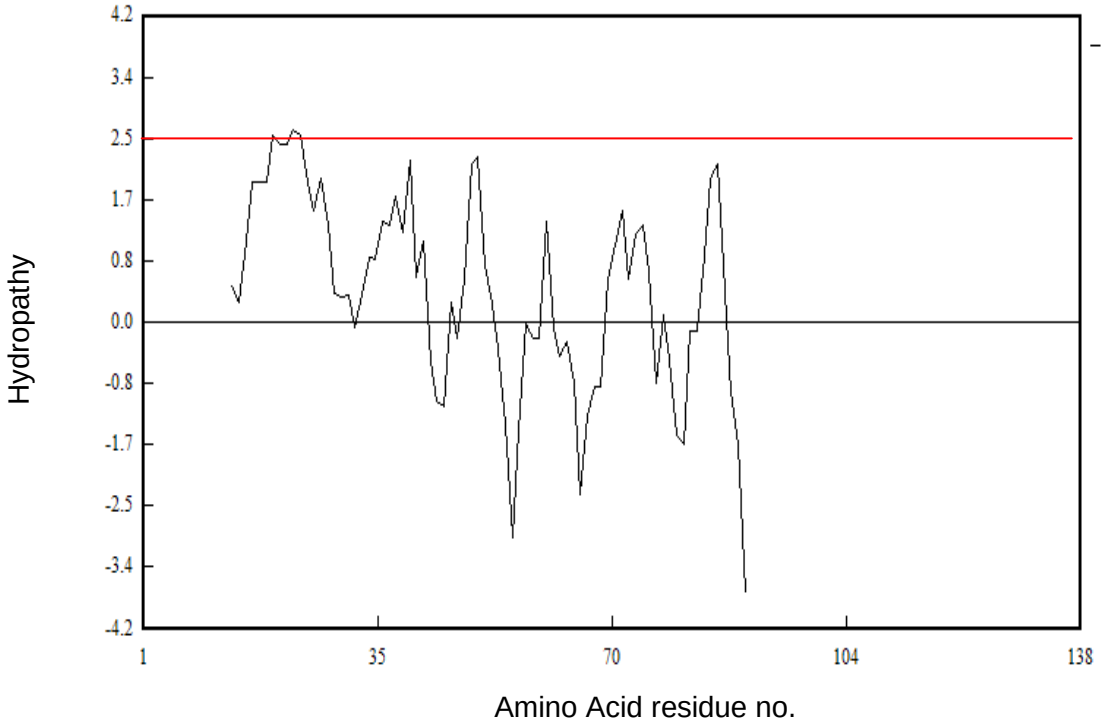

Figure 3.6: **Kyte-Doolittle hydrophobicity plot of gp41(80):** The regions with values above 0 indicate the hydrophobic character of the specific amino acids of the peptide while the regions with a value under 0 indicate the hydrophilic character of the specific amino acids of the peptide. The red line indicates the hydrophobic part on the peptide with a value of 2.5 and higher. A sample window of five amino acid residues was used in the construction of the plot.

This may explain why the antibody titre towards pMAL-gp41T3 is independent of the patient CD4 T-cell count in Figure 3.4, if pMAL-gp41T3 is presented to T-cells via CD1, instead of MHCII presentation. When a large portion of a peptide is hydrophobic, the peptide may be presented by APC on CD1 to CD4/8 double negative T-cells, which may help the humoral antibody response to come about even in the absence of any CD4 T helper cells (Sugita *et al.,* 2000; Kasmar *et al.,* 2009). The pMAL-gp41T3 peptide might therefore, be presented as a “lipid-like” peptide by CD1 on APC, while the pMAL-gp41(80) peptide may be presented via MHCII. For consideration as a vaccine or immunotherapy, it is important that the immune response should not be paralysed by HIV infection. As the pMAL-

gp41(80) peptide appeared to be CD4 dependent in its ability to induce antibodies, gp41T3 may be the preferred antigen to evoke antibodies that are stable and predictable under all conditions of $CD4^+$ T-cell demise.

3.3.5 Antigenicity of gp41T1 expressed in *Pichia pastoris*

After induction and protein concentration determination the antigenicity of gp41T1 was tested against all 44 HIV positive patient by means of ELISA. The negative control consisted of five pooled HIV negative patient sera's. From Figure 3.7 one can conclude that gp41T1 exhibits no significant antigenicity against the HIV positive patient sera. Reasons for this may be due to the larger length of the peptide in comparison to gp41T3, enabling it to fold in such away that the epitope is not presented properly. Alternatively, it may be that the structure of the peptide is not stable due to its length. Based on the results obtained from Figure 3.7 it was decided not to use the gp41T1 peptide in further studies seeing that it did not show any antigenicity. The pMAL-gp41T3 was therefore, used in further studies as it showed the best potential for exploitation as a vaccine or an immunotherapeutic.

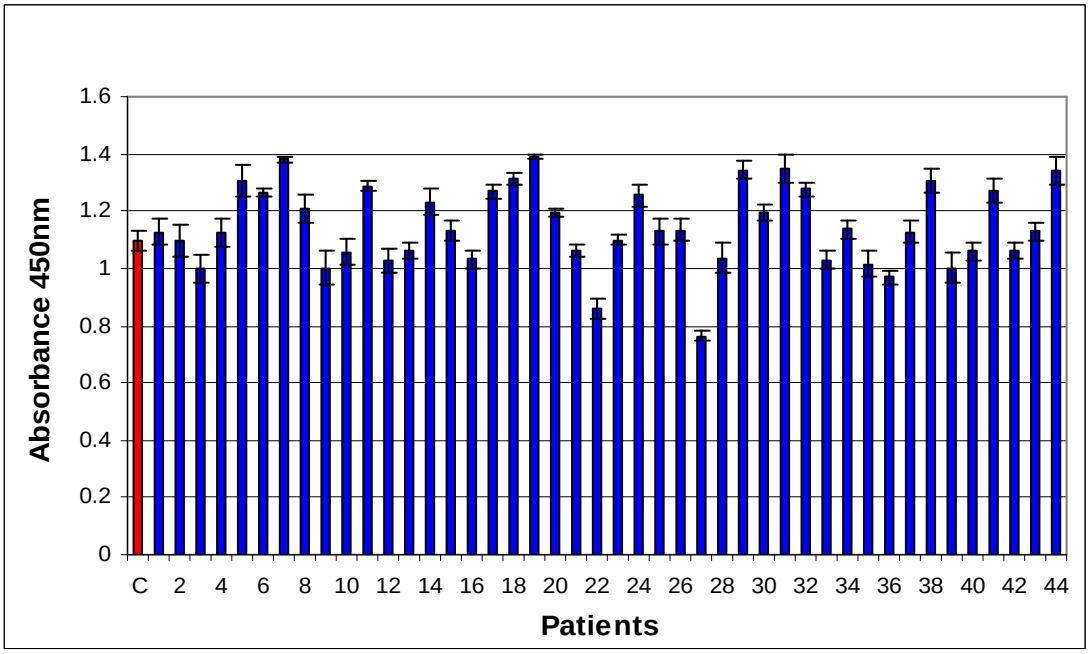

Figure 3.7: **Antigenicity of gp41T1 towards HIV positive and HIV negative patient sera:** The blue bars indicate the 44 HIV positive patient sera responses towards gp41T1. The red bar indicates the HIV negative patient sera response towards gp41T1. The negative patient sera consisted of 7 negative patient sera that were pooled. The standard deviations among 3 repeats are shown by the error bars.

3.3.6 Pro-apoptotic function determination of pMAL-gp41T3

To determine the possible unwanted side effects of gp41 applied as a vaccine or immunotherapy and provide better understanding of the general biology of HIV/AIDS the apoptotic functionality of the peptide was investigated. During HIV infection both infected and non-infected $CD4^+$ T-cells die due to an unknown mechanism, of which viral peptide induced apoptosis is not excluded (Moanna *et al.,* 2005; Vieillard *et al.,* 2008). Should this mechanism become unravelled, it may explain how HIV-infection destroys the immune system.

Plasma membrane phospholipid asymmetry is lost in cells undergoing apoptosis, leading to the exposure of phosphatidylserine (PS) to the outer leaflet of the membrane where it can be detected using fluorescein isothiocyanate (FITC) labelled PS-binding anticoagulant annexin V.

Figure 3.8 shows the pro-apoptotic effect of gp41. These experiments were done in fetal thymic organ cultures (FTOCs) which are extremely sensitive for measuring T-cell apoptosis. The experiment was done by Dr. Hugh Brady, using Trail (a member of the TNF family, able to activate cell death pathways) as the positive control for apoptosis. It was the extracellular domain of mouse Trail (aa 99-291) fused at the N-terminus to a His_6-tag and a linker peptide (produced by Alexis Biochemicals in *E. coli*) that was used. Trail induces apoptosis at concentrations of >10 ng/ml as tested on human tumour cells. The enhancer in the study was a cross-linker that improves Trail action as Trail multimers are more effective at triggering Trail receptors. The negative control is just the medium without any Trail or enhancer. The experiment was repeated three times under the same conditions to produce the results indicated in Figure 3.8.

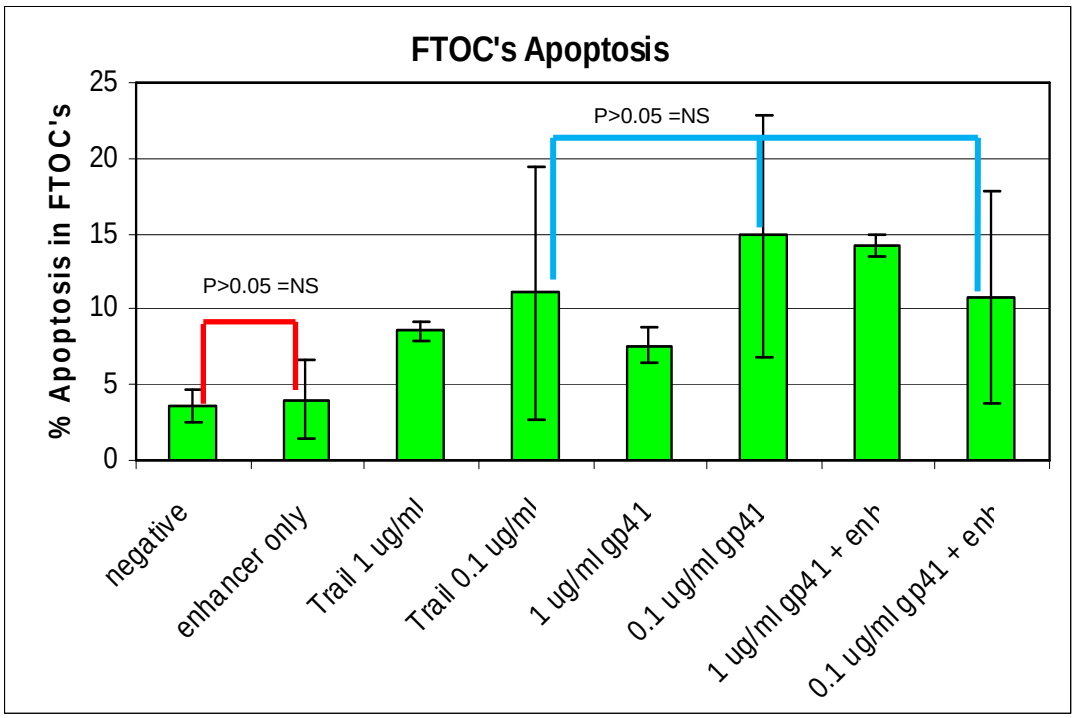

Figure 3.8: **Pro-apoptotic function of pMAL-gp41T3 in FTOCs:** Negative control indicated by the first bar is the media without enhancer and Trail. The enhancer is indicated in the second bar and enhances the signal by acting as a cross-linker. Trail indicated in the third and fourth bars is a member of the TNF family and is able to activate cell pathways that lead to cell death and was used in the experiment as the positive control. The standard deviations among 3 repeats are shown by the error bars. The red line indicates the (lack of) statistical significance between the negative control and the enhancer. The blue line indicates the (lack of) statistical significance between the 0.1 µg/ml samples.

The unpaired t-test with a two tail P-value was used to determine the statistical significance of the result where $P<0.05$ indicates statistical significance between sample groups. As indicated in Figure 3.8 there is no statistical significance between the 0.1 µg/ml samples. When the percentage apoptosis caused by negative control (3.6%), gp41 (0.1 µg/ml) (14.8%) and gp41 (0.1 µg/ml) + enhancer (10.8%) were compared the following could be concluded: gp41 induce (11.2%) more apoptosis compared to the negative control but with ($P>0.05$) and therefore, the 11.2% is not significant, with the addition of the enhancer the apoptotic activity increased with 7.2% but with ($P>0.1531$) therefore, no significant % of apoptosis. The negative control and the enhancer only are not significantly different with $P= 0.8091$. When the percentage apoptosis caused by

negative control (3.6%), gp41 (1 µg/ml) (7.6%) and gp41 (1 µg/ml) + enhancer (14.3%) were compared the following could be concluded: gp41 induced more apoptosis (4%) than the negative control ($P<0.05$), with the addition of the enhancer the apoptotic activity increased with 10.7% ($P<0.0001$). 1 µg/ml gp41 plus enhancer is 39.8% more apoptotic than Trail 1 µg/ml ($P<0.001$).

From the results of the preliminary test it is clear that the fusion peptide does have apoptotic characteristics. This preliminary indication of the apoptotic activity of gp41T3 calls for care in its use as a vaccine or immunotherapy, but may contribute to the better understanding of the functionality of g41 in the bigger picture of the biology of HIV/AIDS.

Sequence alignment between gp41T3 and several TNF's (TNF-α; FasL, Trail, APRIL, TRANCE and VEGI) is indicated in Figure 3.9. Figure 3.9 shows that there are homology boxes found in gp41T3 and several other TNFs. These homology boxes between the sequences may explain the apoptotic nature of gp41T3.

```
TNFA-TRNK       .......TPSDKPVAHVVANPQAEGQLQWLNRRANALLAN      33
FASL-TRNK       .......KKELRKVAHLTGKSNSRSMPLEWEDTYGIVLLS      33
TRAIL-TRNK      ......RSNTLSSPNSKNEKALGRKINSWESSRSGHSFLS      34
APRIL-TRNK      .....KKQHSVLHLVPINATSKDDSDVTEVMWQPALRRGR      35
TRANCE-TRNK     KLEAQPFAHLTINATDIPSGSHKVSLSSWYHDR.GWGKIS      39
VEGI-TRNK       ............QTPTQHFKNQFPALHWEHELGLAFTKN       27
GP41-TRNK       .NQQEQNEKDLLALDKWQNLWSWFDITNWLWYIKIFIMIV      39
Consensus

TNFA-TRNK       GVELRDNQLVVPSEGL...YLIYSQVLFKGQGCP......      64
FASL-TRNK       GVKYKKGGLVINETGL...YFVYSKVYFRGQSCN......      64
TRAIL-TRNK      NLHLRNGELVIHEKGF...YYIYSQTYFRFQEEIKENT..      69
APRIL-TRNK      GLQAQGYGVRIQDAGV...YLLYSQVLFQDVTFT......      66
TRANCE-TRNK     NMTFSNGKLIVNQDGF...YYLYANICFRHHETSGDLA..      74
VEGI-TRNK       RMNYTNKFLLIPESGD...YFIYSQVTFRGMTSECSEIRQ      64
GP41-TRNK       GGLIGLRIIFAILSIVNRVRQGYSPLSFQTLTPSPRGPDR      79
Consensus                              y    f

TNFA-TRNK       ........................STHVLLTHTISRIAVS      80
FASL-TRNK       ..........................NLPLSHKVYMRNSK      78
TRAIL-TRNK      ........................KNDKQMVQYIYKYTS.      84
APRIL-TRNK      ............................MGQVVSREGQG      77
TRANCE-TRNK     ........................TEYLQLMVYVTKTSIK      90
VEGI-TRNK       AG.....................RPNKPDSITVVITKVTDS      84
GP41-TRNK       LGRIEEEGGEQDRDRSIRLVGGFLALAWDDLRSLCLFSYH     119
Consensus

TNFA-TRNK       YQTKVNLLSAIKSPCQRETPEGAEAKPWYEPIYLGGVFQL     120
FASL-TRNK       YPQDLVMMEG.KMMSYCTTGQ.....MWARSSYLGAVFNL     112
TRAIL-TRNK      YPDPILLMKSARNSCWSKDAE.....YGLYSIYQGGIFEL     119
APRIL-TRNK      RQETLFRCIRSMPSHPDRAYN.........SCYSAGVFHL     108
TRANCE-TRNK     IPSSHTLMKGGSTKYWSGNSE.....FHFYSINVGGFFKL     125
VEGI-TRNK       YPEPTQLLMG.TKSVCEVGS......NWFQPIYLGAMFSL     117
GP41-TRNK       RLRDLILIATRVVELLGRSSLRGLQRGWEILKYLGSLVQY     159
Consensus

TNFA-TRNK       EKGDRLSAEINRPDYLDFAESGQVYFGIIAL.........     151
FASL-TRNK       TSADHLYVNVSELSLVNFE.ESQTFFGLYKL.........     142
TRAIL-TRNK      KENDRIFVSVTNEHLIDMD.HEASFFGAFLVG........     150
APRIL-TRNK      HQGDILSVIIPRARAKLNLSPHGTFLGFVKL.........     139
TRANCE-TRNK     RSGEEISIEVSNPSLLDPD.QDATYFGAFKVRDID.....     159
VEGI-TRNK       QEGDKLMVNVSDISLVDYTKEDKTFFGAFLL.........     148
GP41-TRNK       WGLELKKSAINLLNITAIAVAEGTDRIIEIIQRICRAIYH     199
Consensus
```

Figure 3.9: **Sequence alignment of gp41T3 with some TNFs:** Homology boxes are indicated between several TNFs and gp41T3. Black represents 100% homology, pink represents >75% homology, blue represent >50% homology and yellow represent >33% homology.

3.4 Discussion

In the search for a vaccine against AIDS, the focus was originally directed against the viral surface proteins, of which gp120 was most prominent (Jones *et al.,* 1995). Several investigators predicted problems with the use of gp120 as a vaccine, including the possibility that it may give rise to auto-immune antibodies that may pose a health risk to the vaccinated individual (Tishkoff and Hunt, 2000). For this purpose, the current study focused on the viral gp41 protein as vaccine target, despite the fact that it is not directly exposed on the viral surface. Here the vaccine potential of gp41 was investigated by characterizing its cytoplasmic domain in terms of its ability to induce antibodies in HIV infected human patients, how that correlated with AIDS progression in terms of the CD4 count of the patients and whether and how its peptide fragments induced apoptosis in non-infected thymocytes.

There are reports of neutralizing and non-neutralizing antibodies directed towards the Kennedy domain (Dalgleish *et al*., 1988; Evans *et al*., 1989). In this study we confirmed, using ELISA, the antigenicity of a pMAL-gp41T3 fusion protein of 57 kDa originating from the Kennedy domain located on the C-terminal of gp41, located in Cluster 3. Positive and negative patient sera were analysed and antibody activity was found that was directed towards the pMAL-gp41T3 fusion peptide. The antigenicity of the pMAL-gp41(80) fusion protein, which is 80 amino acids in length and located on the N-terminal of gp41 was also confirmed by means of the ELISA test. Both these fusion proteins showed high levels of antigenicity.

When the antigenicity of the full length gp41T1 peptide located on the C-terminal of gp41 was determined using the same patient group, no antigenicity directed to gp41T1 was observed. The increased number of amino acids could possibly alter the folding of the peptide causing the T3 antigenic epitope to become hidden (Aleanzi *et al.,* 1996).

AIDS is characterized by a gradual depletion of $CD4^+$ T-cells from the blood, which correlates with declining immunocompetency. The mechanism by which HIV causes depletion of the $CD4^+$ T-cells in infected individuals remains

unknown (Dunham *et al.,* 2005). Numerous theories have been proposed, but none can fully explain all of the events observed to occur in patients. During the progression of the disease there is a gradual depletion of the $CD4^+$ T-cells and an increase in HIV RNA copies in plasma.

When working towards an immunotherapeutic or vaccine solution against AIDS, it is important to know that $CD4^+$ T-cells cannot be relied upon to support immunity in the patient. It is also important to exclude the potential danger of inducing auto-immunity to $CD4^+$ T-cells when using viral proteins or peptides in sub-component vaccines or immunotherapies. Both these aspects were addressed in this chapter to test the hypothesis that a particular part of the gp41 cytoplasmic domain induces antibodies in patients that correlate positively with CD4 count (indicative of protection) and displays little apoptotic effect (not toxic).

Three parts of the gp41 cytoplasmic domain were tested for antigenicity, of which pMAL-gp41T3 and pMAL-gp41(80), located on the C-terminal and N-terminal respectively, showed antigenicity to Ab from HIV positive patient sera. Neither of the two peptide antigens correlated positively with the $CD4^+$ T-cell count of the HIV positive patient sera, thereby arguing against their use as vaccine or immunotherapy components according to the stated hypothesis. In the case of the pMAL-gp41(80) fusion protein the correlation between the $CD4^+$ T-cell count and the antibody binding signal followed a negative correlation. This result may be due to the following: With high CD4 counts, the patients are still only mildly infected with HIV and specific anti-gp41 antibody induction is low. As HIV infection progresses, more viral antigen is released into the circulation to stimulate more specific antibody production. In the case of the pMAL-gp41T3 fusion protein there was no correlation between the antibody binding signal and the patient CD4 count. From the results one can conclude that the Abs directed to pMAL-gp41T3 are maximally induced already at an early state of HIV infection and are not sensitive for further induction by increased antigen release in the body. The hydrophobicity of pMAL-gp41T3 was found to be high compared to pMAL-gp41(80), providing a possible argument for why the two fusion peptides reacted so differently. It has been shown by Kasmar (Kasmar *et al.,* 2009) that high levels of hydrophobicity of peptides may lead to their presentation on CD1

as a “lipid like” peptide to CD4/CD8 negative T-cells or NKT cells with a very different outcome on Ab regulation (Lang and Glatman-Freedman, 2006). This may also explain why no correlation between the antibody binding signal and the patient CD4 count was found.

Whereas the antibody response to gp41(80) is clearly unable to prevent the progression of HIV infected patients to AIDS, the role of the antibodies to gp41T3 is more enigmatic. Gp41T3’s independence of the CD4 T-cell count may make it a potential beneficial medicinal peptide, as its antibody-inducing effect appears not to be influenced by the HIV-induced paralysis of the host immune response. However, antibodies and other forms of immunity induced by the CD4/CD8 double negative T-cell population (NKT cells) are often involved in auto-immunity (Strober *et al.,* 1996). The possible auto-immune side-effects of exposure of the body to gp41T3 remain to be determined.

If the induction of antibodies to gp41T3 is beneficial, then it is also important to determine whether the peptide is safe to use. To test the hypothesis that gp41 itself might play a role in the killing of $CD4^+$ T-cells, one needs to investigate the possible mechanisms of cell death that it may induce.

In our studies it is shown that pMAL-gp41T3 shows an apoptotic effect on FTOCs. In the studies done by Micoli it was shown that when point mutations are introduced into HIV-1 gp160, it reduced apoptosis and calmodulin binding without affecting viral replication. The point mutation was introduced in the calmodulin binding domain. Due to this mutation the apoptosis effect in HIV/AIDS was reduced (Micoli *et al.,* 2006). This domain is also included in the pMAL-gp41T3 construct and therefore, our opservation of the apoptotic properties of gp41T3 is confirmed by the studies of Micoli. To further confirm that this peptide might be responsible for apoptosis of T-cells, the sequence alignment of gp41 and some TNFs which are responsible for apoptosis were compared and showed homology to several TNFs, which are known to induce apoptosis.

Chapter 4

Concluding Discussion

HIV is one of the first retroviruses to infect humans (Gallo, 2005). This may be the reason why it became so challenging to attempt to control this epidemic. Despite more than two decades of huge research investment, a protective vaccine against HIV still remains elusive. It is widely accepted that such a vaccine will require the elicitation of both T-cell mediated immunity and a broadly neutralizing antibody (bNAb) response (Johnston *et al.,* 2007; Walker *et al.,* 2009). It became evident that studies which focus on understanding the biology, biochemistry and structural biology of HIV as well as the interactions between viral components and new drug candidates will be required to determine an effective strategy to combat HIV/AIDS.

To learn more about possible ways to generate a vaccine against AIDS, we focused on the gp41 envelope protein of the virus. The hypothesis to be tested was that gp41 contains an element that can be used as a component for a safe vaccine that is antigenic and does not elicit autoimmunity or cause apoptosis. This research project was therefore, undertaken to improve the understanding of the immunological properties of gp41 by following its antibody footprints in HIV positive patients and establish whether gp41 could induce apoptosis of uninfected mammalian thymocytes.

During this study the ecto- and endo-domains of HIV-1 gp41 were studied. The gene encoding the endo-domain was truncated into three gene fragments, all of which contain the Kennedy domain located on the C-terminal of gp41 indicated in Figure 4.1. These gene fragments vary in length, namely the 405 bp gp41T3 fragment, 591 bp gp41T2 fragment and the 696 bp gp41T1 fragment. As indicated in Figure 4.1 the 80 amino acid fragment of the ecto-domain, which is located on the N-terminal of gp41 was also cloned and named gp41(80).

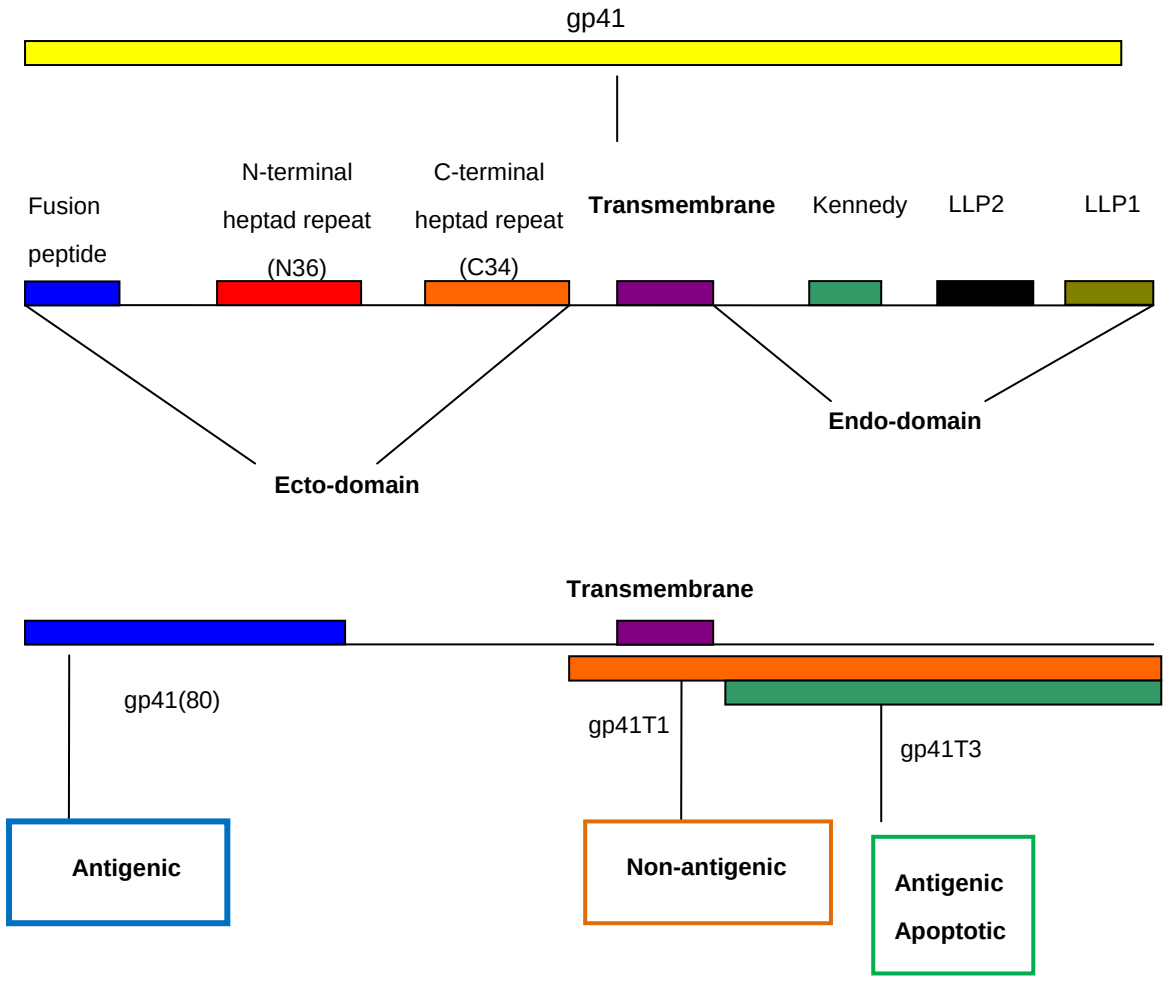

Figure 4.1: **Functional domains of gp41 and the location of gp41(80), gp41T1 and gp41T3 on the gp41 gene.** The location of gp41(80) is indicated by the blue box located on the Ecto-domain of gp41. The location of gp41T1, which includes the transmembrane spanning domain, is shown by the orange box while gp41T3 is indicated by the green box. The antigenicity and the apoptotic functions of these peptides are also indicated.

The pMAL-gp41T3, pMAL-gp41(80) and gp41T1 constructs were successfully expressed in either *E. coli* or *P. pastoris.* The pMAL-gp41T3 and pMAL-gp41(80) constructs were successfully purified by affinity purification while gp41T1 underwent secreted expression.

From literature one can conclude that gp41 might be of some interest for vaccine development because of the antigenic neutralizing epitope sites located on gp41 (Grag *et al.,* 2008). Some of these binding sites for neutralizing antibodies namely 2F5 and 4E10 are indicated in Figure 4.2 as well as the sequence of the fusion inhibitor T20 (Enfuvirtide). The work done by Grag and co-workers does not cover the peptide region expressed and purified by us as indicated in Figure 4.2.

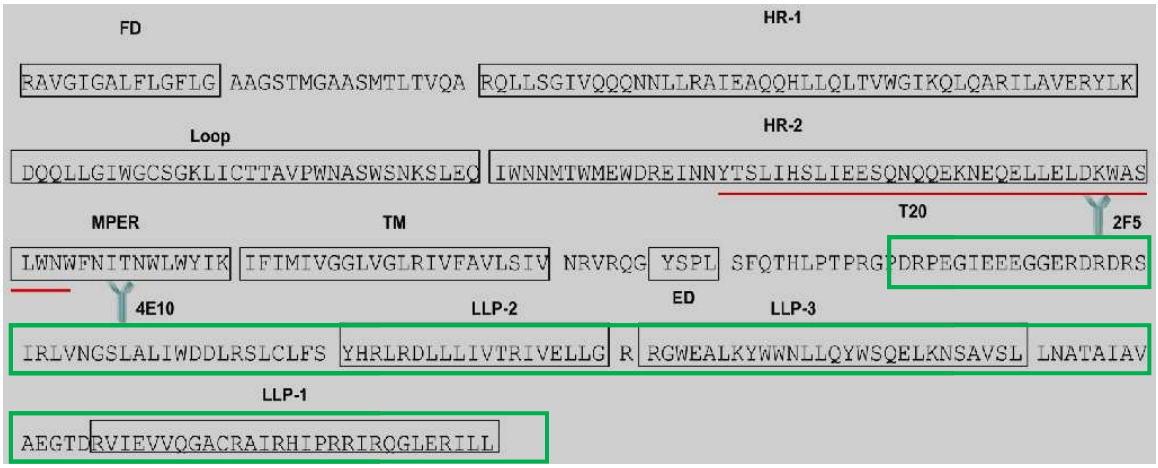

Figure 4.2: **Sequence of HIV-1 gp41 showing different domains.** FD (Fusion Domain), HR (Heptad Repeat), MPER (Membrane Proximal Ecto-domain Region), TM (Transmembrane Domain), ED (Endocytosis Domain), LLP (Lentiviral Lytic Peptide). Binding sites for neutralizing antibodies 2F5 and 4E10 are also shown along with the sequence of fusion inhibitor T20. Gp41T3 is indicated by the green box (Garg *et al.,* 2008).

The antigenicity of the different construct peptides was determined by making use of ELISA. Both the pMAL-gp41T3 and the pMAL-gp41(80) peptides proved to be antigenic towards HIV positive patient sera, but in the case of gp41T1 no antigenicity was found. Gp41(80) showed increasing antibody-antigen binding activity as HIV infection progressed towards AIDS, while gp41T3 showed constant antibody-antigen binding activity irrespective of the decline in CD4 counts. From a vaccine perspective the following can be concluded: In the case of pMAL-gp41(80), where antibody binding activity increased with a decrease in the patient CD4 count, the antibodies may have broadened their specificity range with increased concentration of circulating antigen. This would represent the

normal case of immunization protocol, where higher concentrations of antigen tend to saturate the B cell surface antibodies even of low affinity antigen binders, making for a lower average affinity of antibody binding to antigen, but providing a wider range of antibodies with specificity towards gp41(80). The Jerne plaque assay for determining the number of B cells responding to antigen in a spleen biopsy would have given a more accurate assessment of the phenomenon, but is not ethical to do in humans. Therefore, pMAL-gp41(80) might not be ideal for the development of a vaccine, because it is unable to prevent the progression of HIV infected patients to AIDS, despite broadening its specificity. From the ELISA data obtained for pMAL-gp41T3 there was no correlation between the CD4 count and the antibody-antigen binding activity. Therefore, irrespective of the decline in the CD4 counts of the patients there is a constant antibody-antigen binding activity. It was also indicated that the hydrophobicity of pMAL-gp41T3 was higher than for pMAL-gp41(80). Kasmar and co-workers illustrated that high levels of hydrophobicity of peptides may lead to their presentation on CD1 as a “lipid like” peptide to CD4/CD8 negative T-cell or NKT cells (Kasmar *et al.,* 2009). Therefore, the potential for this peptide to elicit auto-immunity is there. The reason for this is: antibodies and other forms of immunity induced by the CD4/CD8 double negative T-cell population (NKT cells) are often involved in auto-immunity.

Many viruses have, as part of their invasion strategy, the ability to modulate the apoptotic pathways of the host. It is counter-intuitive that such simple organisms would be efficient at regulating this, a most crucial pathway within the host, given the relative complexity of the host cells. Yet, viruses have the potential to initiate or maintain the onset of programmed cell death through the manipulation of a variety of key apoptotic proteins. It is not until recently that the depth at which viruses exploit the apoptotic pathways of their host was appreciated (Hay and Kannourakis 2002). A proper understanding may provide a novel opportunity for future therapeutic ventures.

Experimental prelimary evidence was provided here that the pMAL-gp41T3 construct has apoptotic properties. This was corroborated by the sequence alignment of gp41 and some TNFs, namely tumour necrosis factor alpha (TNFα), Fas ligand (FasL), TRAIL, APRIL, TRANCE and VEGI, showing considerable homology. These peptides are known to induce apoptosis in a wide variety of cells.

To conclude, gp41 may be considered as a vaccine target due to its antigenicity shown in this study, but there are some reservations to its efficacy and safety. The ecto-domain seems to elicit antibodies that do not prevent the progression to AIDS, while the gp41 endo-domain elicits an antibody response that is stably maintained, but exhibits apoptotic function. Even if no vaccine can be made from this, the antibody response to the endo-domain of gp41 was found not to be perturbed by the progression towards AIDS. It is therefore, another ideal surrogate marker for HIV infection that may be useful in diagnosis. The future prospects for this study includes an in depth apoptotic study of this peptide. It is evident from the study that both pMAL-gp41(80) and pMAL-gp41T3 are not ideal new vaccine targets, but their immune properties shed some new light on how HIV may establish itself to produce AIDS. This knowledge may be exploited to steer immunity away from AIDS, eg by selective suppression of pathological auto-immune antibody activity by idiotypic vaccines or suppression of peptide induced apoptosis of CD4 T-cells by pharmaceutical or biological means.

References

Alderson, M.R., Tough, T.W., Davis-Smith, T., Braddy, S., Falk, B., Schooley, K.A., Goodwin, R.G., Smith, C.A., Ramsdell, F. and Lynch, D.H. (1995). Fas ligand mediates activation-induced cell death in human T lymphocytes. *Journal of Experimental Medicine* **181**, 71-77.

Aleanzi, M., Perotti, N., Montaner, A., Lottersberger, J., Tonarelli, G. and Marcipar, J. (1996). Epitope contiguous chains and antibody recognition in HIV-1 synthetic peptide antigens. *Journal of Molecular Recognition* **9**, 631-638.

Allen, J.S. (1985). Major glycoprotein antigens that induce antibodies in AIDS patients are encoded by HTLV-III. *Science* **228**, 1091-1094.

Ameisen, J.C. and Capron, A. (1991). Cell dysfunction and depletion in AIDS: the programmed cell death hypotheses. *Immunology Today* **12**, 102-105.

Andreau, K., Castedo, M., Perfettini, J.L., Roumier, T., Pichart, E., Souquere, S., Vivet, S., Larochette, N. and Kroemer, G. (2004). Pre-apoptotic chromatin condensation upstream of the mitochondrial checkpoint. *Journal of Biological Chemistry* **279**, 55937-55945.

Artenstein, A.W., Van Cott, T.C., Mascola, J.R., Carr, J.K., Hegerich, P.A. and Gaywee, J. (1995). Dual infection with human immunodeficiency virus type 1 of distinct envelope subtypes in humans. *Journal of Infectious Diseases* **171**, 805-810.

Arthur, L.O., Bess, J.W., Sowder, R.C.I., Benveniste, R.E., Mann, D.L., Chermann, J.C. and Henderson, L.E. (1992). Cellular proteins bound to immunodeficiency virus: implications for pathogenesis and vaccines. *Science* **258**, 1935-1938.

Badley, A.D., Pilon, A.A., Landay, A. and Lynch, D.H. (2000). Mechanisms of HIV-associated lymphocyte apoptosis. *Blood* **96**, 2951-2964.

Barnett, S.W., Lu, S., Srivastava, I., Cherpelis, S., Gettie, A., Blanchard, J., Wang, S., Mboudjeka, I., Leung, L., Lian, Y., Fong, A., Buckner, C., Ly, A., Hilt, S., Ulmer, J., Wild, C.T., Mascola, J.R. and Stamatatos L. (2001). The ability of an oligomeric human immunodeficiency virus type 1 (HIV-1) envelope antigen to elicit neutralizing antibodies against primary HIV-1 isolates is improved following partial deletion of the second hypervariable region. *Journal of Virology* **75**, 5526-5540.

Bendelac, A., Bonneville, M. and Kearney, J.F. (2001). Autoreactivity by design: innate B and T lymphocytes. *Nature Reviews Immunology* **1**, 177-186.

Blanco, J., Barretina, J., Ferri, K.F., Jacotot, E., Gutierrez, A., Cabrera, C., Kroemer, G., Clotet, B. and Este, J.A. (2003). Cell-surface-expressed HIV-1 envelope induces the death of CD4 T cells during GP41-mediated hemifusion-like events. *Virology* **305**, 318-329.

Borrow, P., Lewicki, H. and Wei, X. (1997). Antiviral pressure exerted by HIV-1-specific CTLs during primary infection demonstrated by rapid selection of CTL escape virus. *Nature Medicine* **3**, 205-211.

Botha, M.M. (2006). Structure-Based inhibitor design and validation: Application to *Plasmodium falciparum* Glutathione S-transferase. *University of Pretoria.*

Bouhlal, H., Galon, J., Kazatchkine, M.D., Fridman, W.H., Sautès-Fridman, C. and Cavaillon, N.H. (2001). Soluble CD16 Inhibits CR3 (CD11b/CD18)-Mediated Infection of Monocytes/Macrophages by Opsonized Primary R5 HIV-1. *Journal of Immunology* **166**, 3377-3383.

Bowerman, B., Brown, P.O., Bishop, J.M. and Varmus, H.E. (1989). A nucleoprotein complex mediates the integration of retroviral DNA. *Genes Diversity* **4**, 469-478.

Bukrinsky, M.I., Haggerty, S., Dempsey, M.P., Sharova, N., Adzhubei, A., Spitz, L., Lewis, P., Goldfarb, D., Emerman, M. and Stevenson, M. (1993). A nuclear localization signal within HIV-1 matrix protein that governs infection of non-dividing cells. *Nature* **365**, 666-669.

Camaur, D. and Trono, D. (1996). Characterization of human immunodeficiency virus type 1 Vif particle incorporation. *Journal of Virology.* **70**, 6106-6111.

Cerami, A. and Beutler, B. (1988). The role of cachectin/TNF in endotoxic shock and cachexia. *Immunology Today* **9**, pp. 28-31.

Chan, D.C., Fass, D., Berger, J.M. and Kim, P.S. (1997). Core structure of gp41 from the HIV envelope glycoprotein. *Cell* **89**, 263-273.

Chan, D.C. and Kim, P.S. (1998). HIV entry and its inhibition. *Cell* **93**, 681-684.

Chan, D.C., Chutkowski, C.T. and Kim, P.S. (1998). Evidence that a prominent cavity in the coiled coil of HIV type 1 gp41 is an attractive drug target. *Proceedings of the National Academy of Science USA* **95**, 15613-15617.

Cheung, L., McLain, L., Hollier, M.J., Reading, S.A. and Dimmock, N.J. (2005). Part of the C-terminal tail of the envelope gp41 transmembrane glycoprotein of human immunodeficiency virus type 1 is exposed on the surface of infected cells and is involved in virus-mediated cell fusion. *Journal of General Virology* **86**, 131-138.

Chia, D., Barnett, E.V., Yamagata, J., Knutson, D., Restivo, C. and Furst, D. (1979). Quantitation and characterization of soluble immune complexes precipitated from sera by polyethylene glycol (PEG). *Clinical and Experimental Immunology* **37**(3), 399–407.

Clos, J. and Brandau, S. (1994). pJC20 and pJC40 - Two high-copy number vectors for T7 RNA polymerase-dependent expression of recombinant genes in *Escherichia coli*. *Protein expression and purification* **5**, 133-137.

Coffin, J., Haase, A., Levy, J.A., Montagnier, L., Oroszlan, S., Teich, N., Temin, H., Toyoshima, K., Varmus, H., Vogt, P. and Weiss, R. (1986). Human immunodeficiency viruses. *Science* **232**, 697.

Cohen, N.R., Garg, S. and Brenner, M.B. (2009). Antigen Presentation by CD1: Lipids, T Cells, and NKT Cells in Microbial Immunity. *Advances in Immunology* **102**, Chapter 1.

Connors, M., Kovacs, J.A. and Krevat, S. (1997). HIV infection induces changes in CD4 T-cell phonotype and depletions within the CD4 T-cell repertoire that are not immediately restored by antiviral or immune-based therapies. *Nature Medicine* **3**, 533-540.

Cotton, M.F., Ikle, D.N., Rapaport, E.L., Marschner, S., Tseng, P.O., Kurrle, R. and Finkel, T.H. (1997). Apoptosis of CD4+ and CD8+ T cells isolated immediately ex vivo correlates with disease severity in human immunodeficiency virus type 1 infection. *Pediatric Research* **42**, 656-664.

Courcoul, M., Patience, C., Rey, F., Blanc, D., Harmache, A., Sire, J., Vigne, R. and Spire, B. (1995). Peripheral blood mononuclear cells produce normal amounts of defective vif- human immunodeficiency virus type 1 particles which are restricted for the preretrotranscriptional steps. *Journal of Virology* **69**, 9358-9365.

Crise, B., Buonocore, L. and Rose, J.K. (1990). CD4 is retained in the endoplasmic reticulum by the human immunodeficiency virus type 1 glycoprotein precursor. *Journal of Virology* **64**, 5585-5593.

Dalgleish, A.G., Chanh, T.C., Kennedy, R.C., Kanda, P., Clapham, P.R. and Weiss, R.A. (1988). Neutralization of diverse HIV-1 strains by monoclonal antibodies raised against a gp41 synthetic peptide. *Virology* **165**, 209-215.

Daniel, P.T. and Krammer, P.H. (1994). Activation induces sensitivity toward APO-1 (CD95) mediated apoptosis in human B cells. *Journal of Immunology* **152**, 5624-5632.

Day, J.R., Munk, C. and Guatelli, J.C. (2004). The membrane-proximal tyrosine-based sorting signal of human immunodeficiency virus type I gp41 is required for optimal viral infectivity. *Journal of Virology* **78**, 1069-1079.

Delves, P.J and Roitt, I.M., (2000). The Immune System. *New England Journal of Medicine* **343**, 108-117.

Dower, W.J., Miller, J.F. and Ragsdale, C.W. (1988). High effiency transformation of *E.coli* by high voltage electroporation. *Nucleic Acids Research* **16** (13), 6127-6145.

Dunham, R., Moanna, A., Paiardini, M. and Silvestri, G. (2005). CD4 T-cell depletion in HIV infection: Killed by Friendly Fire? *Current HIV/AIDS Reports* **2**, 16-23.

Earl, P.L., Moss, B. and Doms, R.W. (1991). Folding, interaction with GRP78-BiP, assembly, and transport of the human immunodeficiency virus type 1 envelope protein. *Journal of Virology* **65**, 2047-2055.

Eberle, J., Loussert-Ajaka, I., Brust, S., Zekeng, L., Hauser, P.H., Kaptue, L., Knapp, S., Damond, F., Saragosti, S., Simon, F. and Gurtler, L.G. (1997). Diversity of the immunodominant epitope of gp41 of HIV-1 subtype O and its validity for antibody detection. *Journal of Virology Methods* **67**, 85-91.

Eckert, D.M., Shi, Y., Kim, S., Welch, B.D., Kang, E., Poff, E.S. and Michael S. Kay, M.S. (2008). Characterization of the steric defense of the HIV-1 gp41 N-trimer region. *Protein Science* **17**, 2091-2100.

Eckhart, L., Raffelsberger, W., Ferko, B., Klima, A., Purtscher, M., Katinger, H. and Ruker, F. (1996). Immunogenic presentation of a conserved gp41 epitope of human immunodeficiency virus type 1 on recombinant surface antigen of hepatitis B virus. *Journal of General Virology* **77**, 2001-2008.

Emerman, M. and Malim, M.H. (1998). HIV-1 regulatory /accessory genes: keys to unraveling viral and host cell biology. *Science* **280**, 1880-1884.

Evans, D.J., McKeating, J., Meredith, J.M., Burke, K.L., Katrak, K., John, A., Ferguson, M., Minor, P.D., Weiss, R.A. and Almond, J.W. (1989). An engineered poliovirus chimaera elicits broadly reactive HIV-1 neutralizing antibodies. *Nature* **339**, 385-388.

Ferri, K.F., Jacotot, E., Geuskens, M. and Kroemer, G. (2000b). Apoptosis and karyogamy in syncytia induced by HIV-1 ENV/CD4 interaction. *Cell Death and Differentiation* **7**, 1137-1139.

Foti, M., Mangasarian, A., Piguet, V., Lew, D.P., Krause, K.H., Trono, D. and Carpentier, J.L. (1997). Nef-mediated clathrin-coated pit formation. *Journal of Cell Biology* **139**, 37-47.

Fouchier, R.A., Simon, J.H., Jaffe, A.B. and Malim, M.H. (1996). Human immunodeficiency virus type 1 Vif does not influence expression or virion incorporation of gag-, pol- and env-encoded proteins. *Journal of Virology* **70**, 8263-8269.

Fouchier, R.A.M., Meyer, B.E., Simon, J.H.M., Fischer, U. and Malim, M.H. (1997). HIV-1 infection of non-dividing cells: evidence that the amino-terminal basic region of the viral matrix protein is important for Gag processing but not for post-entry nuclear import. *European Molecular Biology Organization Journal* **16**, 4531-4539.

Franke, K.E., Hui, E., Yuan, H. and Luban, J. (1994b). Specific incorporation of cyclophilin A into HIV-1 virions. *Nature* **24**, 359-362.

Freed, E.O. and Martin, A.M. (1995). Virion incorporation of envelope glycoproteins with long but not short cytoplasmic tails is blocked by specific, single amino acid substitutions in the human immunodeficiency virus type 1 matrix. *Journal of Virology* **69**, 1984-1989.

Freed, E.O., Englund, G. and Martin, A.M. (1995). Role of the basic domain of human immunodeficiency virus type 1 matrix in macrophage infection. *Journal of Virology* **69**, 3949-3954.

Gallo, R.C. (2005). History of the discoveries of the first human retroviruses: HTLV-1 and HTLV-2. *Oncogene* **24**, 5926-5930.

Garg, H. and Blumenthal R. (2008). The role of HIV gp41 mediated fusion/hemifusion in Bystander apoptosis. *Cellular and Molecular Life Science* **65**, 3134-3144.

Gellissen, G., Weydemann, U., Strasser, A.W.M., Piontek, M., Janowicz, Z.A. and Hollenberg, C.P. (1992). Progress in developing methyltrphic yeasts as expression systems. *TIBTECH* **10**, 413-417.

Gnann, J.W., Nelson, J.A. and Oldstone, M.B. (1987b). Fine mapping of an immunodominant domain in the transmembrane glycoprotein of human immunodeficiency virus. *Journal of Virology* **61**, 2639-2641.

Gnann, J.W., Schwimmbeck, P.L., Nelson, J.A., Truax, A.B. and Oldstone, M.B. (1987c). Diagnosis of AIDS by using a 12 amino acid peptide representing an immunodominant epitope of the human immunodeficiency virus. *Journal of Infectious Diseases* **156**, 261-267.

Golding, H. (1988). Identification of homologous regions in human immunedeficiency virus I gp41 and human MHC class II beta 1 domain. I. Monoclonal antibodies against the gp41-derived peptide and patients' sera react with native HLA class II antigens, suggesting a role for autoimmunity in the pathogenesis of acquired immune deficiency syndrome. *Journal of Experimental Medicine* **167**, 914-923.

Goulder, P.J., Edwards, A. and Phillips, R.E. (1997). Identification of a novel HLA-A24-restricted cytotoxic T-lymphocyte epitope within HIV-1 Nef. *AIDS* **11**, 1883-1884.

Greene, W.C. (1993). The molecular biology of HIV-1 and AIDS. *Contemporary Topics in Infectious Diseases* **10**,161-179.

Hay, S. and Kannourakis, G. (2002). A time to Kill: viral manipulation of the cell death program. *Journal of General Virology* **83**, 1547-1564.

Heinkelein, M., Sopper, S. and Jassoy, C. (1998). Contact of human immunodeficiency virus type 1-infected and uninfected $CD4^+$ T lymphocytes is highly cytolytic for both cells. *Journal of Virology* **69**, 6925-6931.

Hérnandez-Caselles, T. and Stutman, O. (1993). Immune functions of tumor necrosis factor. I. Tumor necrosis factor induces apoptosis of mouse thymocytes and can also stimulate or inhibit IL-6-induced proliferation depending on the concentration of mitogenic costimulation. *Journal of Immunology* **151**, 3999-4012.

Horal, P., Svennerholm, B., Jeansson, S., Rymo, L., Hall, W.W. and Vahlne, A. (1991). Continuous epitopes of the human immunodeficiency virus type 1 (HIV-1) transmembrane glycoprotein and reactivity of human sera to synthetic peptides representing various HIV-1 isolates. *Journal of Virology* **65**, 2718-2723.

Hoxie, J.A., Alpers, J.D., Rackowski, J.L., Huebner, K., Haggarty, B.S., Cedarbaum, A.J. and Reed, J.C. (1986). Alterations in T4 (CD4) protein and mRNA synthesis in cells infected with HIV. *Science* **234**, 1123-1127.

Huang, J., Villemain, J., Padilla, R. and Sousa, R. (1999). Mechanisms by which T7 lysozyme specifically regulates T7 RNA polymerase during different phases of transcription. *Journal of Molecular Biology* **293** (3), 457-475.

Iwantani, Y., Ueno, T., Nishimura, A., Zhang, X., Hattori, T., Ishimoto, A., Ito, M. and Sakai, H. (2001). Modification of virus infectivity by cytoplasmic tail of HIV-1 TM protein. *Virus Research* **74**, 75-87.

Johnston, M.I. and Fauci, A.S. (2007). An HIV vaccine-evolving concepts. *New England Journal of Medicine* **356**, 2073-2081.

Jones, D.H., McBride, B.W., Roff, M.A. and Farrar, H.F. (1995). Efficient purification and rigorous characterisation of a recombinant gp120 for HIV vaccine studies. *Vaccine* **13**, 991-999.

Jowett, J.B.M., Planelles, V., Poon, B., Shah, N.P., Chen, M.L. and Chen, I. S.Y. (1995). The human immunodeficiency virus type 1 vpr gene arrests infected T cells in the G2. M phase of the cell cycle. *Journal of Virology* **69**, 6304-6313.

Kane, J.F. (1995). Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coil. Current Opinion in Biotechnology* **6**, 494-500.

Kasmar, A., Van Rhijn, I. and Moody, D.B. (2009). The evolved functions of CD1 during infection. *Current Opinion in Immunology* **21**, 397-403.

Katsikis, P.D., Wunderlich, E.S., Smith, C.A. and Herzenberg, L.A. (1995). Fas Antigen Stimulation Induces Marked Apoptosis of T Lymphocytes in Human Immunodeficiency Virus-infected Individuals. *Experimental Mededicine* **181**, 2029-2036.

Klasse, P.J. and Moore, J.P. (2004). Is there enough gp120 in the body fluids of HIV-1-infected individuals to have biologically significant effects? *Virology* **2004**, 1-8.

Kennedy, R.C., Henkel, R.D., Pauletti, D., Allan, J. S., Lee, T.H., Essex, M. and Dreesman, G.R. (1986). Antiserum to a synthetic peptide recognizes the HTLV-III envelope glycoprotein. *Science* **231**, 1556-1559.

Kingsman, S.M., Kingsman, A.J., Dobson, M.J., Mellor, J. and Roberts, N. (1985). Heterologous gene expression in *Saccharomyees cerevisiae*. *Biotechnology and genetic engineering reviews* **3**, 377-416.

Köhler, G. and Milstein, C. (1975). Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* **256**, 495-497.

Kolchinsky, P., Kiprilov, E., Bartley, P., Rubinstein, R. and Sodroski, J. (2001). Loss of a single N-linked glycan allows CD4-independent human immuno-deficiency virus type 1 infection by altering the position of the gp120 V1/V2 variable loops. *Journal of Virology* **75** (7), 3435-3443.

Korber, B., Muldoon, M., Theiler, J., Gao, F., Gupta, R., Lapedes, A., Hahn, B.H., Wolinsky, S. and Bhattacharya, T. (2000). Timing the ancestor of the HIV-1 pandemic strains. *Science* **288**, 1789-1796.

Krammer, P.H., Behrmann, .I, Daniel, P., Dhein, J. and Debatin, K.M. (1994). Regulation of apoptosis in the immune system. *Current Opinion in Immunology* **2**, 279-289.

Kwong, P.D., Wyatt, R., Robinson, J., Sweet, R.W., Sodroski, J. and Hendrickson, W.A. (1998). Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* **393**, 648-659.

Lang, M.L. and Glatman-Freedman, A. (2006). Do CD1-restricted T cells contribute to antibody-mediated immunity against *Mycobacterium tuberculosis*? *Infection and Immunity* **74**, 803-809.

Laurent-Crawford, A.G., Krust, B., Riviere, Y., Desgranges, C., Muller, S., Kieny, M.P., Dauguet, C. and Hovanessian, A.G. (1993). Membrane expression of HIV envelope glycoprotein triggers apoptosis in CD4 cells. *AIDS Research and Human Retroviruses* **9**, 761-773.

Leist, M. and Jaattela, M. (2001). Four deaths and a funeral: from caspases to alternative mechanisms. *Nature Reviews Molecular and Cellular Biology* **2**, 589-598.

Levy, J.A., Hoffmann, A.D., Kramer, S.M., Landis, J.A., Shimabukuro, J.M. and Oshiro, L.S. (1984). Recovery of AIDS-associated retroviruses from patients with AIDS or AIDS-related conditions and from clinically healthy individuals. *Journal of Infectious Diseases* **152,** 734-738.

Lifson, J.D., Feinberg, M.B., Reyes, G.R., Rabin, L., Bonapour, B., Chakrabarti, S., Moss, B., Wong-Staal, F., Steimer, K.S. and Engleman E.G. (1986a). Induction of CD4- dependent cell fusion by the HTLV-III/LAV envelope glycoprotein. *Nature* **323**, 725-728.

Littlefield, J.W., Keller, E.B., Gross, J. and Zamecnik, P.C. (1995). Studies on cytoplasmic ribonucleoprotein particles from the liver of the rat. *Journal of Biological Chemistry* **217**, 111-123.

Liu, H., Wu, X., Newman, M., Shaw, G.M., Hahn, B.H. and Kappes, J.C. (1995). The Vif protein of human and simian immunodeficiency viruses is packaged into virions and associates with viral core structures. *Journal of Virology* **69**, 7630-7638.

Lowry, O.H., Rosebrough, N.J., Farr, A.L. and Randall, R.J. (1951). Protein measurement with Folin phenol reagent. *Journal of Biological Chemistry* **193** (1), 265-275.

Lu, M., Blacklow, S.C. and Kim, P.S. (1995). A trimeric structural domain of the HIV-1 transmembrane glycoprotein. *Nature Structural Biolology* **2**, 1075-1080.

Margottin, F., Bour, S.P., Durand, H., Selig, L., Benichou, S., Richard, V., Thomas, D., Strebel, K. and Benarous, R. (1998). A novel human WD protein, hbeta TrCp that interacts with the HIV-1 Vpu connects CD4 to the ER degradation pathway through an F-box motif. *Molecular Cell* **1**, 565-574.

MC Michael, A. (1998). T cell responses and viral escape. *Cell* **93**, 673-676.

Micoli, K.J., Mamaeva, O., Piller, S.C., Barker, J.L., Pan, G., Hunter, E. and McDonald, J.M. (2006). Point mutations in the C-terminus of HIV-1 gp160 reduce apoptosis and calmodulin binding without affecting viral replication. *Virology* **344**, 468-479

Miller, M.A., Mietzner, T.A., Cloyd, M.W., Robey, W.G. and Montelaro, R.C. (1993). Identification of a calmodulin-binding and inhibitory peptide domain in the HIV-1 transmembrane glycoprotein. *AIDS Research and Human Retroviruses* **9**, 1057-1066.

Miller, M.D., Farnet, C.M. and Bushman, F.D. (1997). Human immunodeficiency virus type 1 preintegration complexes: studies of organization and composition. *Journal of Virology* **71**, 5382-5390.

Moanna, A., Dunham, R., Paiardini, M. and Silvestri, G. (2005). CD4+ T-cell depletion in HIV infection: killed by friendly fire? *Current HIV/AIDS Reports* **2**, 16-23.

Montagnier, L. (2002). A History of HIV. *Science* **298**, 1727-1728.

Moore, J. P. (1997). Coreceptors: implications for HIV pathogenesis and therapy. *Science* **276**, 51-52.

Muster, T., Steindl, F., Purtscher, M., Trkola, A., Klima, A., Himmler, G., Ruker, F. and Katinger, H. (1993). A conserved neutralizing epitope on gp41 of human immunodeficiency virus type 1. *Journal of Virology* **67**, 6642-6647.

Muster, T., Guinea, R., Trkola, A., Purtscher, M., Klima, A., Steindl, F., Palese, P. and Katinger, H. (1994). Cross-neutralizing activity against divergent human immunodeficiency virus type 1 isolates induced by the gp41 sequence ELDKWAS. *Journal of Virology* **68**, 4031-4034.

Neurath, A.R., Strick, N., Lin, K. and Jiang, S. (1995). Multifaceted consequences of anti-gp41 monoclonal antibody 2F5 binding to HIV type 1 virions. *AIDS Research and Human Retroviruses* **11**, 687-696.

Nie, Z., Bergeron, D., Subbramanian, R.A., Yao, X.J., Checroune, F., Rougeau, N. and Cohen, E.A. (1998). The putative alpha helix 2 of human immunodeficiency virus type-1 Vpr contains a determinant which is responsible for the nuclear translocation of proviral DNA in growth-arrested cells. *Journal of Virology* **72**, 4104-4115.

Nirenberg, M.W. and Matthaei, J.H. (1961). The dependence of cell-free protein synthesis in *E.coli* upon naturally occurring or synthetic polynucleotides. *Proceedings of the National Academy of Sciences USA* **47**, 1588-1602.

Ohnimus, H., Heinkelein, M. and Jassoy, C. (1997). Apoptotic cell death upon contact of CD4+ T lymphocytes with HIV glycoprotein-expressing cells is mediated by caspases but bypasses CD95 (Fas/Apo-1) and TNF receptor 1. *Journal of Immunology* **159**, 5246-5252.

Ohno, M., Fornerod, M. and Mattaj, I.W. (1998). Nucleocytoplasmic transport: the last 200 nanometers. *Cell* **92**, 327-336.

Oldstone, M.B., Tishon, A., Lewicki, H., Dyson, H.J., Feher, V.A., Assa-Munt, N. and Wright, P.E. (1991). Mapping the anatomy of the immunodominant domain of the human immunodeficiency virus gp41 transmembrane protein: peptide conformation analysis using monoclonal antibodies and proton nuclear magnetic resonance spectroscopy. *Journal of Virology* **65**, 1727-1734.

Pantaleo, G. and Fauci, A. (1995). Apoptosis in HIV infection. *Nature Medicine* **1**, 118-120.

Pantaleo, G., Soudeyns, H., Demarest, J.F., Vaccarezza, M., Graziosi, C., Paolucci, S., Daucher, M.B., Cohen, O.J., Denis, F., Biddison, W.E., Sekaly, R.P. and Fauci, A.S. (1997). Accumulation of human immunodeficiency virus-specific cytotoxic T lymphocytes away from the predominant site of virus replication during primary infection. *European Journal of Immunology* **27**, 3166-3173.

Pantaleo, G. (1997). How immune-based interventions can change HIV therapy. *Nature Medicine* **5**, 483-486.

Patke, C.L. and Shearer, W.T. (2000). gp120- and TNF-α-induced modulation of human B cell function: Proliferation, cyclic AMP generation, Ig production, and B-cell receptor expression. *Journal of Allergy and Clinical Immunology* **105**, 975-982.

Pelham, H.R.B. and Jackson, R.J. (1976). An efficient mRNA-dependent translation system from reticulocyte lysates. *European Journal of Biochemistry.* **67**, 247-256.

Perfettini, J.L., Roumier, T., Castedo, M., Larochette, N., Boya, P., Reynal, B., Lazar, V., Ciccosanti, F., Nardacci, R., Penninger, J., Piacentini, M. and Kroemer, G. (2004b). NF-kB and p53 are the dominant apoptosis-inducing transcription factors elicited by the HIV-1 envelope. *Journal of Experimental Medicine* **199**, 629-640.

Phenix, B.N. and Badley, A.D. (2002). Influence of mitochondrial control of apoptosis on the pathogenesis, complications and treatment of HIV infection. *Biochimistry* **84**, 251-264.

Piller, S.C., Dubay, J.W., Derdeyn, C.A. and Hunter, E. (2000). Mutational analysis of conserved domains within the cytoplasmic tail of gp41 from human immunodeficiency virus type 1: effects on glycoprotein incorporation and infectivity. *Journal of Virology* **74**, 11717-11723.

Plata, F., Autran, B., Martins, L.P., Wain-Hobson, S., Raphael, M. and Mayaud, C. (1987). AIDS virus-specific cytotoxic T lymphocytes in lung disorders. *Nature* **328**, 348-351.

Pollard, V.W. and Malim, M.H. (1998). The HIV-1 Rev protein. *Annual Review Microbiology* **52**, 491-532.

Popov, S., Rexach, M., Zybarth, G., Reiling, N., Lee, M.A., Ratner, L., Lane, C.M., Moore, M.S., Blobel, G. and Bukrinsky, M. (1998). Viral protein R regulates nuclear import of the HIV-1 pre-integration complex. *European Molecular Biology Organization Journal* **17**, 909-917.

Purtscher, M., Trkola, A., Grassauer, A., Schulz, P.M., Klima, A., Dopper, S., Gruber, G., Buchacher, A., Muster, T. and Katinger, H. (1996). Restricted antigenic variability of the epitope recognized by the neutralizing gp41 antibody 2F5. *AIDS* **10**, 587-593.

Radding, W., Pan, Z.Q., Hunter, E., Johnston, P., Williams, J.P. and McDonald, J.M. (1996). Expression of HIV-1 envelope glycoprotein alters cellular calmodulin. *Biochemical and Biophysical Research Communications* **218**, 192-197.

Re, F., Braaten, D., Franke, E.K. and Luban, J. (1995). Human immunodeficiency virus type 1 Vpr arrests the cell cycle in G2 by inhibiting the activation of p34-cdc2-cyclin B. *Journal of Virology* **69**, 6859-6864.

Reil, H., Bukovsky, A.A., Gelderblom, H.R. and Gottlinger, H.G. (1998). Efficient HIV-1 replication can occur in the absence of the viral matrix protein. *European Molecular Biology Organization Journal* **17**, 2699-2708.

Reines, D., Conaway, J.W. and Conaway, R.C. (1996). The RNA polymerase II general elongation factors. *Biochemical Science* **9**, 351-355.

Reisinger, E.C., Vogetseder, W., Berzow, D., Kofler, D., Bitterlich, G., Lehr, H.A., Wachter, H. and Dierich, P.M. (1990). Complement-mediated enhancement of HIV-1 infection of monoblastoid cell line U937. *AIDS* **10**, 961-966.

Requejo, H.I.Z. (2006). Worldwide molecular epidemiology of HIV. Rev Saúde Pública **40**(2), 331-45.

Revel, M. and Schattner, A. (1987). Interferons: cytokines in autoimmunity. *Ciba Foundation symposium* **129**, 223-233.

Rizzuto, C. and Sodroski, J. (2000). Fine definition of a conserved CCR5-binding region on the human immunodeficiency virus type 1 glycoprotein 120. *AIDS Research and Human Retroviruses* **16** (8), 741-749.

Roberts, B.E. and Paterson, B.M. (1973). Efficient translation of tobacco mosaic virus RNA and rabbit globin 9S RNA in a cell-free system from commercial wheat germ. *Proceeding of the Natlional Academy of Science USA* **70**, 2330-2334.

Robinson, W., Kawamura, T., Gorny, M.K., Lake, D., Xu, J.Y., Matsumoto, Y., Sugano, T., Masuho, Y., Mitchell, W.M. and Hersh, E. (1990a). Human monoclonal antibodies to the human immunodeficiency virus type 1 (HIV-1) transmembrane glycoprotein gp41 enhance HIV-1 infection in vitro. *Proceeding of the Natlional Academy of Science USA* **87**, 3185-3289.

Robinson, W., Kawamura, T., Lake, D., Masuho, Y., Mitchell, W.M. and Hersh, E.M. (1990b). Antibodies to the primary immunodominant domain of human immunodeficiency virus type 1 (HIV-1) glycoprotein gp41 enhance HIV-1 infection in vitro. *Journal of Virology* **64**, 5301-5305.

Rogel, M.E., Wu, L.I. and Emerman, M. (1995). The human immunodeficiency virus type 1 vpr gene prevents cell proliferation during chronic infection. *Journal of Virology* **69**, 882-888.

Roggero, R., Robert-Hebmann, V., Harrington, S., Roland, J., Verne, L., Jaleco S., Devaux, C. and Biard-Piechaczyk, M. (2001). Binding of human immunodeficiency virus type 1 gp120 to CXCR4 induces mitochondrial transmembrane depolarization and cytochrome c-mediated apoptosis independently of Fas signaling. *Journal of Virology* **75**, 7637-7650.

Roitt, I., Brostoff, J. and Male, D. (2002). Immunology, sixth edition, London, Elsevier Science Limited.

Roshal, M., Zhu, Y. and Planelles, V. (2001). Apoptosis in AIDS. *Apoptosis* **6**, 103-116.

Ruottinen, M., Bollok, M., Kögler, M., Neubauer, A., Krause, M., Hämäläinen, E.R., Myllyharju, J., Vasala, A. and Neubauer, P. (2008). Improved production of human type II procollagen in the yeast *Pichia pastoris* in shake flasks by a wireless-controlled fed-batch system. *Biomedcentral Biotechnol*ogy **8**, 33.

Salzwedel, K., West, J.T. and Hunter, E. (1999). A Conserved Tryptophan-Rich Motif in the Membrane-Proximal Region of the Human Immunodeficiency Virus Type 1 gp41 Ecto-domain Is Important for Env-Mediated Fusion and Virus Infectivity. *Journal of Virology*, 2469-2480.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989). Molecular `Cloning: A laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Sattentau, Q.J. and Moore, J.P. (1993). The role of CD4 in HIV binding and entry. *Philosophical Transactions of the Royal Society B: Biological Science*. **342**, 59-66.

Sattentau, Q.J., Moore, J.P., Vignaux, F., Traincard, F. and Poignard, P. (1993). Conformational changes induced in the envelope glycoproteins of the human and simian immunodeficiency viruses by soluble receptor binding. *Journal of Virology* **67**, 7383-7393.

Sattentau, Q.J., Zolla-Pazner, S. and Poignard, P. (1995). Epitope exposure on functional, oligomeric HIV-1 gp41 molecules. *Virology* **206**, 713-717.

Schachtschabed, I. and Zillig, W. (1959). Untersuchen zur Biosynthese der Proteine, I. aer den Einbau 14C-markierter Aminosauren ins Protein zellfreier Nucleoproteid- Enzym-Systeme aus *Escherichia coli.* *Physiological Chemistry* **314**, 262-270.

Schubert, U., AntoÂn, L.C., Bacik, J., Cox, J.H., Bour, S., Bennick, J.R., Orlowski, M., Strebel, K. and Yewdell, J.W. (1998). CD4 glycoprotein degradation induced by human immunodeficiency virus type I Vpu protein requires the function of proteosomes and the ubiquitin-conjugating pathway. *Journal of Virology* **72**, 2280-2288.

Scorer, C.A., Clare, J.J. and McCombie, W.R. (1994). Rapid selection using G418 of high copy number transformants of *Pichia pastoris* for high-level foreign gene expression. *BioTechnology* **12**, 181-184.

Shi, Y. (2001). A structural view of mitochondria-mediated apoptosis. *Nature Structural Biology* **8**, 394-401.

Simon, J.H., Miller, D.L., Fouchier, R.A., Soares, M.A., Peden, K.W. and Malim, M.H. (1998). The regulation of primate immunodeficiency virus infectivity by Vif is cell species restricted: a role for Vif in determining virus host range and cross-species transmission. *European Molecular Biology Organization Journal* **17**, 1259-1267.

Sova, P., Van Ranst, M., Gupta, P., Balachandran, R., Chao, W., Itescu, S., McKinley, G. and Volsky, D.J. (1995). Conservation of an intact vif gene in vitro and in vivo. *Journal of Virology* **69**, 2557-2564.

Sreekrishna, K., Brankamp, R.G., Kropp, K.E., Blankenship, D.T., Tsay, J., Smith, P.L., Wierschke, J.D., Subramaniam, A. and Birkenberger, L.A. (1997). Strategies for optimal synthesis and secretion of heterologous proteins in the methylotrophic yeast *Pichia pastoris*. *Gene* **190**, 55-62.

Srinivas, S.K., Srinivas, R.V., Anantharamaiah, G.M., Compans, R.W. and Segrest, J.P. (1993). Cytosolic domain of the human immunodeficiency virus envelope glycoproteins binds to calmodulin and inhibits calmodulin-regulated proteins. *Journal of Biological Chemistry* **268**, 22895-22899.

Starcich, B.R., Hahn, B.H., Shaw, G.M., McNeely, P.D., Modrow, S., Wolf, H., Parks, E.S. and Parks, W.P. (1986). Identification and characterization of conserved and variable regions in the envelope gene of HTLV**-**III**/**LAV, the retrovirus of AIDS. *Cell* **45** (5), 637-648.

Stern, L.J. and Wiley, D.C. (1994). Antigen peptide binding by class I and class II histocompatibility proteins. *Structure* **2**, 245-250.

Strober, S., Cheng, L., Zeng, D., Palathumpat, R., Dejbakhsh-Jones, S., Huie, P. and Sibley R. (1996). Double Negative T Cells Which Promote Tolerance Induction and Regulate Autoimmunity. *Tmmunvhgicat Reviews* **149**, 217-230.

Sugita, M., Peters, P.J. and Brenner M.B. (2000). Pathways for lipid antigen presentation by Cd1 molecules: Nowhere for intracellular pathogens to hide. *Traffic* **1**, 295-300.

Sullivan, N., Sun, Y., Li, J., Hofmann, W., and Sodroski, J. (1995). Replicative function and neutralization sensitivity of envelope glycoproteins from primary and T-cell line-passaged human immunodeficiency virus type 1 isolates. *Journal of Virology* **69**, 4413-4422.

Sylwester, A., Murphy, S., Shutt, D. and Soll, D.R. (1997). HIV-induced T cell syncytia are self-perpetuating ant the primary cause of T cell death in culture. *Journal of Immunology* **158**, 3996-4007.

Sytwu, H.K., Liblau, R.S. and McDevitt, H.O. (1996). The role of Fas/Apo-1 (CD95) and TNF in antigen-induced programmed cell death in T cell receptor transgenic mice. *Immunity* **5***,* 17-30*.*

Tencza, S.B., Mietzner, T.A. and Montelaro, R.C. (1997). Calmodulin-binding function of LLP segments from theHIVtype 1 transmembrane protein is conserved among natural sequence variants. *AIDS Research and Human Retroviruses* **13**, 263-269.

Terai, C., Kornbluth, R.S., Pauza, C.D., Richman, D.D. and Carson, D.A. (1991). Apoptosis as a mechanism of cell death in cultured T lymphoblasts acutely infected with HIV-1. *Journal of Clinical investigate* **5**, 1710-1715.

Tilly, J.L. (2001). Commuting the death sentence: how oocytes strive to survive. *Nature Reviews: Molecular Cell Biology* **2**, 838-848.

Tishkoff, G.H. and Hunt, L.T. (2000). Unexpected molecular mimicry among peptides MHC Class II, Blood-Clotting Factor X and HIV-1 Envelope Glycoprotein Gp120. *Thrombosis Research* **98**, 343-346.

Turner, B.G. and Summers, F.S. (1999). Structural Biology of HIV. *Journal of Molecular Biology* **285**, 1-32.

Twu, C., Liu, N.Q., Popik, W., Bukrinsky, M., Sayre, J., Roberts, J., Rania, S., Bramhandam, V., Roos, K.P., MacLellan, W.R. and Fiala, M. (2002). Cardiomyocytes undergo apoptosis in human immunodeficiency virus cardiomyopathy through mitochondrion- and death receptor-controlled pathways. *Proceedings of the National Academy of Science USA* **99**, 14386-14391.

Van Parijs, L., Ibraghimov, A. and Abbas, A.K. (1996). The roles of costimulation and Fas in T cell apoptosis and peripheral tolerance. *Immunity* **4**, 321-328.

Veljkovic, N. (2005). Molecular mimicry of HIV gp120: Possible implications on prevention and Therapy of AIDS. *Archive of Oncology* **13,** 126-130.

Vieillard, V., Le Grand, R., Dausset, J. and Debre, P. (2008a). Vaccine strategy against AIDS: An HIV gp41 peptide immunization prevents NKp44L expression and CD4 T cell depletion in SHIV-infected macaques. *Proceedings of the National Academy of Science USA* **105** (6), 2100-2104.

Vishwanathan, S.A. and Hunter, E. (2008). Importance of the membrane-perturbing properties of the membrane-proximal external region of human immunodeficiency virus type 1 gp41 to viral fusion. *Journal of Virology* **82**, 5118-5126.

Vodicka, M.A., Koepp, D.M., Silver, P.A. and Emerman, M. (1998). HIV-1 Vpr interacts with the nuclear transport pathway to promote macrophage infection. *Genes Diversity* **12**, 175-185.

von Schwedler, U., Song, J., Aiken, C. and Trono, D. (1993). Vif is crucial for human immunodeficiency virus type 1 proviral DNA synthesis in infected cells. *Journal of Virology* **67**, 4945-4955.

von Schwedler, U., Kornbluth, R.S. and Trono, D. (1994). The nuclear localization signal of the matrix protein of human immunodeficiency virus type 1 allows the establishment of infection in macrophages and quiescent t lymphocytes. *Proceedings of the National Academy of Science USA* **91**, 6992-6996.

Walker, C.M and Levy, J.A. (1989). A diffusible lymphokine produced by $CD8^+$ T lymphocytes suppresses HIV replication. *Immunology* **66**, 628-630.

Walker, L.M., Phogat, S.K., Chan-Hui, P., Wagner, D. Phung, P., Goss, J.L., Wrin, T., Simek, M.D., Fling, S., Mitcham, J.L., Lehrman, J.K., Priddy, F.H., Olsen, O.A., Frey, S.M. and Hammond, P.W. (2009). Broad and Potent Neutralizing Antibodies from an African Donor Reveal a New HIV-1 Vaccine Target. *Science* **326**, 285-292.

Waterham, H.R., Russell, K.A., Vries, Y. and Cregg, J.M. (1997). Peroxisomal targeting, import, and assembly of alcohol oxidase in *Pichia pastoris*. *Journal of Cellular Biology* **139**, 1419-1431.

Watkins, D.I., Burton, D.R., Kallas, E.G., Moore, J.P. and Koff, W.C. (2008). Nonhuman primate models and the failure of the Merck HIV-1 vaccine in humans. *Nature Medecine* **14**, 617-621.

Wei, P., Garber, M.E., Fang, S.M., Fischer, W.H. and Jones, K.A. (1998). A novel CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its high-affinity, loop-specific binding to TAR RNA. *Cell* **92**, 451-462.

Wiegers, K., Rutter, G., Kottler, H., Tessmer, U., Hohenberg, H. and Krausslich, H.G. (1998). Sequential steps in human immunodeficiency virus particle maturation revealed by alterations of individual Gag polyprotein cleavage sites. *Journal of Virology* **72**, 2846-2854.

Wyatt, R. and Sodroski, J. (1998). The HIV-1 envelope glycoproteins: fusogens, antigens and immunogens. *Science* **280**, 1884-1888.

Wyatt, R., Kwong, P.D., Desjardins, E., Sweet, R.W., Robinson, J., Hendrickson, W.A. and Sodroski, J.G. (1998). The antigenic structure of the HIV gp120 envelope glycoprotein. *Nature* **393**, 705-711.

Wyma, D.J., Jiang, J., Shi, J., Zhou, J., Lineberger, J.E. and Miller, M.D. (2004). Coupling of human immunodeficiency virus type 1 fusion to virion maturation: a novel role of the gp41 cytoplasmic tail. *Journal of Virology* **78**, 3429-3435.

Zhivotovsky, B. and Kroemer, G. (2004). Apoptosis and genomic instability. *Nature Reviews Molecular and Cellular Biology* **5**, 752-762.

Zhou, Y., Zhang, H., Siliciano, J.D. and Siliciano, R.F. (2005). Kinetics of human immunodeficiency virus type 1 decay following entry into resting $CD4^+$ T cells. *Journal of* Virology **79**, 2199-2210.

Zhu, T. and Korber, B. (1998). An African HIV-1 sequence from 1959 and implications for the origin of the epidemic*. Nature* **391**, 594-597.

Appendix A

List of patient CD4 counts

Patient number	CD4 count	Patient number	CD4 count
2/HPTP-2	112	79/HPTN-27	594
9/HPTP-4	54	80/HPTN-26	713
11/HPTP-6	82	82/HPTN-28	28
12/HPTP-5	273	84/HPTN-21	166
15/HPTP-7	236	85/HPTN-22	23
18/HPTP-10	102	86/HPTN-24	7
19/HPTP-9	412	87/HPTN-23	307
20/HPTP-8	65	97/HPTN-18	32
23/HPTP-11	456	98/HPTN-19	110
46/HPTP-29	163	99/HPTN-20	602
47/HPTP-22	135	100/HPTN-118	118
48/HPTP-31	2	101/HPTN-12	3
52/HPTP-35	218	102/HPTN-13	782
55/HPTP-33	7	103/HPTN-14	311
56/HPTP-34	15	104/HPTN-15	726
59/HPTP-28	352	105/HPTN-16	165
61/HPTP-30	51	106/HPTN-17	208
68/HPTP-40	237	110/HPTN-1	96
69/HPTP-36	73	111/HPTN-2	309
70/HPTP-37	173	112/HPTN-3	83
71/HPTP-38	43		
76/HPTP-15	153		

HPTP: HIV positive, TB positive

HPTN: HIV positive, TB negative