FACULTY OF MEDICINE AND HEALTH SCIENCES

THE MECHANISM OF HIV-1 SENSING IN CD4$^+$ T CELLS:

THE ROLE OF NEWLY EXPRESSED VPR AND VPU

Nadia Elhemdaoui
Student number: 01811127

Promotor and mentor: Prof. Dr. Bruno Verhasselt
Tutor(s): Femke Van Gaever, Evelien Naessens
Department of Diagnostic Sciences

A dissertation submitted to Ghent University in partial fulfilment of the requirements for the degree of Master of Science in the Biomedical Sciences

Academic year: 2019- 2020

GHENT UNIVERSITY

FOREWORD

For my master, I had the luck to conduct research at the HIV lab in the Medical Research Building 2 at the UGhent. In the last two years, I experienced many things and I could not have achieved all of this without the help of many people. I would therefore like to take the opportunity to thank the following people:

First of all, I would like to thank my promotor Prof. Dr. Bruno Verhasselt for giving me the opportunity to conduct this thesis at his research group. Thank you for providing me with the needful guidance to complete the research and this thesis.

Next, I would like to express my sincere gratitude towards my tutor, Femke Van Gaever. You were always there to teach me techniques in the lab, gave me a lot of feedback and answered the overload of questions I had (I probably tested your patience multiple times…). I was so used to working together with you, that it was quite difficult for me to work independently when you left. However, even then, you would always support me, and gave me the necessary push when needed. Therefore, thank you for the talks and for your valuable insights and directions which helped me to complete this thesis! The same applies to Evelien Naessens. Evelien, your door was always open when I had questions, and you would always make time to assist me in the lab. At the beginning, I had little knowledge of cell culture, but with your tips and tricks I now have mastered it. I really appreciated your good advice, thank you!

Of course, I would like to thank PhD students Hannelore Hamerlinck, Laura Witjes, Rani Burm, Laura Corneillie and Laura Collignon for their talks and for creating a very pleasant working atmosphere. Also, a thanks to my friends and peer students of Immunology. In times of COVID-19, I didn't know what to do without your support. I am glad we could help each other and share our frustrations. At the end of the day, we are all in this together.

Finally, mom and dad, I would like to dedicate this last paragraph to you. You gave me the chance to study what I wanted, although this was quite an unfamiliar territory for you. In the last five years, you were always there in moments of frustration and tiredness when things weren't always going as expected. Therefore, thank you for your love, support and understanding, I could not have achieved all of this without you.

Thank you all for your support and take care.

PREAMBLE: COVID-19

On the 18th Of March, I went to the lab to check on my cells (lentiviral transduced SupT1 and Jurkat cell lines) that I had in culture. I had planned to maintain the cells in culture and use them for the next experiments which were planned to start on the 23rd of March. In the meanwhile, it was announced that all master students should stop all experimental activities, also the ones that were currently running. I therefore e-mailed my promotor and proposed to freeze the cells, in the hope to use them later on when the lockdown would end. I called the laboratory technician, Evelien Naessens, to explain me the freezing procedure. However, as I did not have access to the material, Evelien told me she would freeze them for me the next day.

Further experiments that were planned for thesis include: induction of the expression of Vpr and Vpu via doxycycline and evaluation via flow cytometry and Western Blot (**see 2.7**). Therefore, for the thesis, I could only discuss the results that I had obtained for the cloning of the lentiviral constructs and the viral titer determination.

TABLE OF CONTENTS

SUMMARY

Cyclic GMP-AMP synthase (cGAS) -mediated IFN-I response upon productive HIV-1 infection of activated $CD4^+$ T cells has been demonstrated recently. HIV-1 viral proteins Vpr and Vpu seem to regulate this IFN-I response in an opposite way and sensing only occurs upon *de novo* expression of Vpr. However, the mechanism behind this Vpr and Vpu effect remains to be clarified. In order to further evaluate the role of these proteins in the cGAS-mediated IFN-I response in $CD4^+$ T cells, Vpr and Vpu must be first introduced in the cells in the absence of HIV-1. We therefore constructed and evaluated inducible lentiviral vectors expressing Vpr/Vpu in a doxycycline-dependent manner. Vpr, Vpu and identification tag dNGFR were first amplified with PCR and Vpr and Vpu were then cloned in the MCS downstream of the P2A sequence of the all-in-one Tet-On backbone plasmid, which is a bicistronic expression system. Next, the constructed plasmids were used to clone dNGFR in the MCS upstream of the P2A sequence. Three inducible lentiviral vectors were generated: pCW57-dNGFR (control vector), pCW57-dNGFR-P2A-Vpr and pCW57-dNGFR-P2A-Vpu. The transduction efficiency of the produced lentiviruses was determined by assessment of viability in transduced SupT1 and Jurkat cells, and viral titers were assessed using the RT assay. Our results showed that Vpr, Vpu and dNGFR were successfully cloned in the all-in-one Tet-On inducible lentiviral plasmid. For the viability assay, we expected to have different observations for cell survival. However, for all three viruses in both cell lines the same observations were obtained. Further investigation of Vpr and Vpu expression in a doxycycline-dependent manner via flow cytometry and Western blotting is necessary.

1. INTRODUCTION

1.1. HIV-1: Human Immunodeficiency Virus type 1

Human Immunodeficiency Virus type 1 (HIV-1) is the causative agent of a disease called Acquired Immunodeficiency Syndrome (AIDS). Since its discovery in 1983, two years after the recognition of AIDS as a new disease, about 75 million people have become infected and 32 million people have died from AIDS-related illnesses [1].

The number of newly infected people and AIDS-related deaths are declining worldwide, owing to the better scientific understanding of HIV and the development of antiretroviral drugs. At the end of 2018, 37.9 million people globally were living with HIV, of which 770 000 died from AIDS-related illnesses, and 1.7 million were newly infected [1]. Most people living with HIV are now located in low- and middle- income countries, with an estimated 68% living in sub-Saharan Africa [2]. Nonetheless, despite many years of significant effort by the global health community to reduce new HIV infections, to increase access to treatment and to end AIDS-related deaths, the pace of progress is slowing down [2,3].

1.1.1. Classification

HIV-1 is a lentivirus from the *Retroviridae* family that is further subdivided in four distinct phylogenetic lineages: M, N, O and P [4,5]. Although members of all these groups are capable of causing $CD4^+$ T cell depletion and AIDS, they differ in prevalence. The HIV-1 group M viruses are responsible for more than 95% of all HIV-1 infections and therefore contributed to the global HIV pandemic and AIDS mortality/morbidity rates. Group O is much less prevalent than group M and represents less than 1% of global HIV-1 infections. It is largely restricted to West-African countries, whereas the other two groups are very rare, and cases have only been documented in a few individuals.

The global spread of HIV-1 group M led to the introduction of different group M lineages, called subtypes, each with their own geographic spread [4,6]. Group M viruses are currently classified into nine additional subtypes (A-D, F-H, J and K). Besides these groups and subtypes, HIV-1 can further diversify by a mechanism called “recombination”. In this process, more than one virus infects a cell and interactions between the two RNA strands of these viruses can take place. Recombination can then occur inter-group (between groups M and O), as well as inter- and intra-subtype (within group M), which generates new types of HIV-1, referred to as recombinants [7]. The genetic diversity of HIV-1 enables the virus to become more virulent to easily withstand the host’s immune responses and increase its survival.

In the late ‘80s, a genetically distinct virus with morphological similarities to HIV-1 was discovered. This virus, named HIV-2, could also cause AIDS and was mainly observed in West African patients [4,8]. However, HIV-2 is characterized by lower transmission rates and a less pathogenic course than HIV-1, with slow progression towards AIDS and clinical symptoms occurring at later time points [4,9]. The fact that HIV-2 has a completely distinct course of infection can be explained by its origin which differs from HIV-1 (**see also 1.1.2**).

1.1.2. Origin

Although the Human Immunodeficiency Viruses and AIDS gained worldwide attention in the early 1980s, the viruses had probably already spread from African non-human primates to humans throughout the 1900s [10]. The viruses are assumed to have emerged from Simian Immunodeficiency Viruses (SIVs), considering the genetical and phylogenetic resemblance [4,11]. HIV-1 and HIV-2 even seemed to be more closely related to different types of SIVs than to each other. HIV-1 arose from SIVcpz, a virus with the African chimpanzee (*Pan troglodytes*) as natural host, whereas HIV-2 evolved from SIVsm which is present in the Sooty Mangabey

monkey (*Cerocebus atys*). Hence, it is now assumed that the various HIV-1 lineages and HIV-2 are the result of independent direct cross-species transmissions from SIV to humans [4,11].

1.1.3. Structure and genomic organization of HIV-1

i. HIV-1 genome structure

Like all retroviruses, HIV-1 has a positive sense, single-stranded RNA (ssRNA) genome of which two copies are packed into the viral particles' core. Upon reverse transcription, HIV utilizes alternative splicing to obtain nine different gene products from a 9.7 kb long double-stranded (dsDNA) transcript, also known as the provirus [5,12]. The 5' and 3' ends of this proviral genome are flanked by a long terminal repeat (LTR) promoter region (**Figure 1.1**). Between the two LTR regions, there are genes encoding for polyproteins Gag, Pol and Env, regulatory proteins and accessory proteins.

The three major genes: *gag, pol* and *env*, which are present in all retroviruses, encode for three polyproteins that give rise to structural proteins [5,12,13]. The *gag* reading frame is followed by the *pol* reading frame and both are first transcribed into a polyprotein precursor: Gag (p55) and Gag-Pol (p160). Both polyproteins are further cleaved into proteins during maturation of the viral particle. The Gag polyprotein is processed by the viral protease into the following proteins: p17, p24, p7, p6 and two short peptides, SP1 and SP2. p17 (MA, matrix protein) associates with the inner part of the lipid bilayer which helps provide structural integrity to the viral particle and assists in incorporation of the *env* glycoprotein into the viral particle. p24 (capsid protein, CA) can assemble in a conical core structure to surround and protect the viral genome. p7 (nucleocapsid, NC) plays a role in protecting the viral RNA by non-specifically binding to it. p6 is a small protein present in the viral core that plays a role in viral budding from the host cell and in incorporating Vpr in the viral particle. The Gag polyprotein also gives rise to two other small proteins: spacer peptide (SP) 1 and 2. SP1 and SP2 are located between CA – NC and NC – p6, respectively.

Furthermore, three viral enzymes are encoded by the *pol* gene: protease (PR), reverse transcriptase (RT) and integrase (IN). PR is an enzyme that is required for cleaving polyproteins (*i.e.* Gag and Gag-pol), whereas RT and IN are required for the viral replication. RT supports the reverse transcription of viral RNA into proviral cDNA prior to viral integration into the host genome. Moreover, RT can also function as a ribonuclease (RNase H) to remove the original RNA template from the first DNA strand. [13]. IN, on the other hand, assists in the integration of the proviral DNA into the genomic DNA of an infected cell. In this process, IN functions as a 3' exonuclease, an endonuclease and a ligase.

Finally, the *env* reading frame, which follows the *pol* gene, will give rise to the precursor envelope glycoprotein gp160 that is further cleaved by cellular proteases into two glycoproteins: gp120 (surface protein, SU) and gp41 (transmembrane protein, TM). These glycoproteins are assembled into trimeric complexes which are inserted into the lipid bilayer of the virus. The gp120 spikes allow attachment of the virus to host cell receptors (*i.e*. CD4 and chemokine receptors) and regulate viral entry, whereas gp41 is involved in viral entry and fusion with the hosts' cell membrane.

In addition to *gag, pol* and *env*, the HIV-1 genome consists of six genes: *tat, rev, vif, vpr, vpu* and *nef* which encode for proteins with regulatory or immuno-modulatory functions [5,12,13]. Tat (transactivator protein) and Rev (RNA splicing-regulator) are regulatory proteins present in all retroviruses. They are necessary for the initiation of the viral replication. The viral accessory proteins Nef (negative regulating factor), Vif (viral infectivity factor), Vpr (viral protein r) and Vpu (viral protein u) on the other hand, are additional proteins which are not needed for viral replication *in vitro* but seem to be essential for the viral replication process and pathogenesis *in vivo* [14]. An overview of all the HIV-1 proteins and their functions is provided in **supplementary figure 1**. The role of Vpr and Vpu will be further discussed in the second section (**1.2**), where HIV-1 and host immune interactions are introduced.

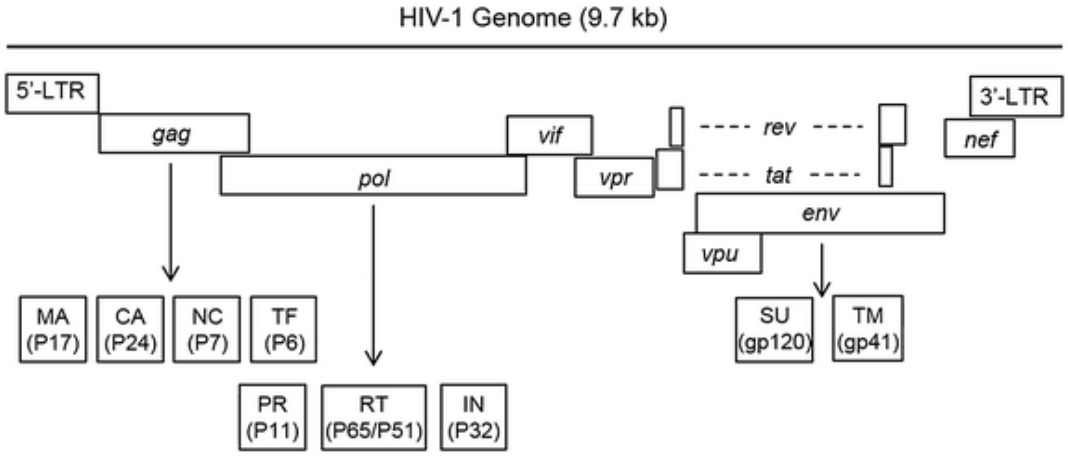


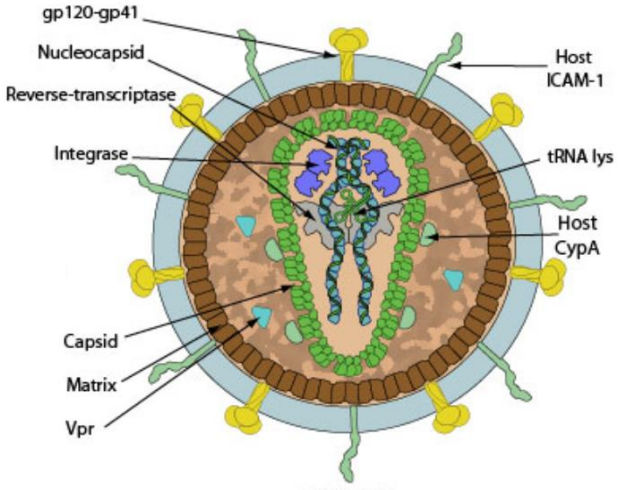
Figure 1.1 | Overview of the structure and organization of the HIV-1 genome. The size of the HIV-1 provirus is approximately 9.7 kb. Shown are the reading frames of the genes coding for structural and regulatory proteins. Arrows point to cleaved protein products, with the molecular weight of each protein indicated in parenthesis. Dashed lines represent RNA splicing. LTR, long terminal repeat; Gag, group-specific antigen; Pol, polymerase; MA, matrix protein; CA, capsid domain; NC, nucleocapsid; TF, trans-frame protein; PR, protease; RT, reverse transcriptase; IN, integrase; Env, envelope protein; SU, surface membrane protein; TM, trans-membrane protein; Vif, viral infectivity factor; Vpr, viral protein R; Vpu, viral protein U; Nef, negative regulatory factor; Rev, regulator of expression of viral proteins; Tat, trans-activator of transcription [15].

ii. HIV-1 particle structure

During the maturation process, the morphology of the viral particle completely changes. The immature "donut-shaped" virion undergoes structural changes to eventually become an infective mature virion [5,16].

The mature HIV-1 particles have a round shape that measures approximately 100 nm in diameter. They consist of an outer lipid envelope, a cone-shaped capsid and matrix proteins in between (**Figure 1.2**). The lipid envelope consists of a bilayer of cellular origin, with viral glycoproteins gp120 and gp41 attached to it. During viral budding, host proteins (*i.e.* ICAM-1) can get incorporated in this lipid bilayer. Underneath the viral lipid bilayer, a membrane is formed by matrix protein p17 which surrounds the inner core [16,17].

The cone-shaped core is assembled from the inner capsid protein p24 and encloses the nucleocapsid together with the viral genome. This nucleocapsid forms a hexameric structure that associates with the two RNA molecules and thus protects the RNA from digestion by nucleases [17].



Legend on next page

Figure on previous page

Figure 1.2 | Structure of a mature HIV-1 virion. During the replication cycle, the virion will undergo structural changes and become mature. The mature HIV-1 virion compromises the viral RNA that is tightly bound to nucleocapsid proteins and surrounded by enzymes needed in the viral replication cycle, such as reverse transcriptase, proteases, and integrase. The capsid surrounds the nucleocapsid that is bound to the viral genomic RNA. Adapted from ViralZone Swiss Institute of Bioinformatics (2015) [18].

iii. Difference with HIV-2

HIV-2 is structurally and genomically similar to HIV-1, but clearly differs from HIV-1 by the types of accessory genes it encodes. Instead of the vpu gene, HIV-2 harbors another gene called vpx (viral protein x) [5,19]. The encoded Vpx protein has some overlapping functions with Vpr and it is also present in some SIV strains (**see Figure 1.3**) [14]. Like Vpr, Vpx is packed into the virion and its expression enhances viral replication. The fact that HIV-2 is characterized by lower pathogenicity *in vivo* can be explained by the difference in this accessory protein [4,5,14].

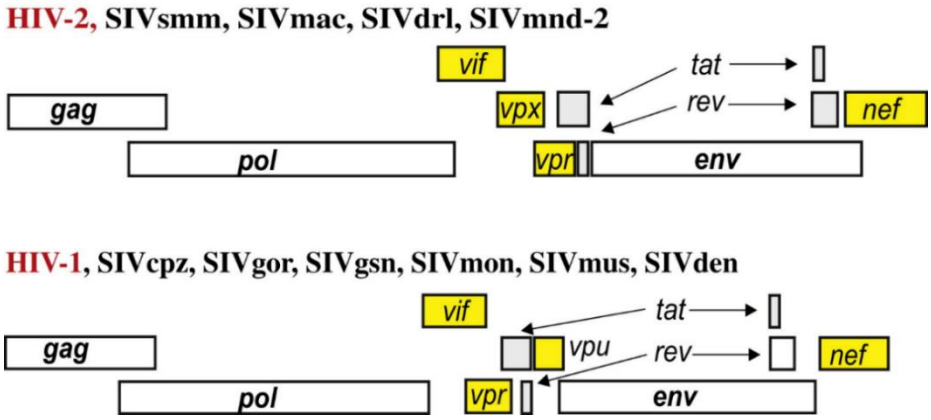


Figure 1.3 | Genomic organization of HIV-1, HIV-2 and some SIV lentiviruses
Adapted from Kirchhoff 2010 [14].

1.1.4. HIV-1 replication cycle

The HIV tropism is largely determined by surface receptors necessary for HIV-1 binding. HIV shows a high affinity for immune cells expressing the primary CD4 receptor as well as a secondary co-receptor CXCR4 or CCR5 for successful entry. HIV replication proceeds in a series of events that are displayed in **figure 1.4** and described in detail below.

i. Attachment and entry

HIV infection is initiated by binding of the viral gp120 subunit to the CD4 receptor of the target cell (step **1** on **Figure 1.4**). This attachment results in a conformational change which in turn enables subsequent binding to the chemokine co-receptor (CCR5 or CXCR4). Upon co-receptor binding, the fusion peptide of viral gp41 subunit inserts into the target cell membrane, allowing fusion (step **2**) of viral and cellular membranes and thus entry of the virus [18,20]. Once the virus has entered the host cell, its core will disassemble in order to release the viral cargo into the cytoplasm, known as “uncoating” (step **3**). Although this process has not been fully elucidated, it has been suggested that uncoating plays an important role in further processes in the replication cycle, such as reverse transcription, nuclear transport and nuclear import [21].

ii. Reverse transcription and integration

After uncoating, the viral ssRNA genome is transcribed into a dsDNA molecule by the viral RT released from the viral core (step **4**). Prior to entry in the nucleus, the viral DNA will associate with viral IN and other proteins of the viral core into a pre-integration complex (PIC). Once the virus enters the nucleus (step **5**), the viral dsDNA will integrate in the host’s genome (step **6**) [18,22].

iii. Virus protein synthesis

Following integration, the provirus acts as a template for viral gene expression and takes advantage of the host replication machinery to produce viral proteins. The host RNA polymerase II enzyme recognizes the long terminal repeat (LTR) located in the 5' region of the provirus and initiates transcription (step **7**) into viral (in)completely spliced and unspliced mRNAs. The viral mRNAs are then transported to the cytoplasm (step **8**) for translation into viral proteins (step **9**). Completely spliced mRNAs encode Rev, Tat and Nef accessory genes. The early expression of Rev is necessary to promote nuclear export of unspliced and incompletely spliced mRNAs. The unspliced full length mRNA will serve as genomic RNA to be packaged into virions or will be translated into gag and gag-pol polyproteins. The incompletely spliced mRNAs, on the other hand, are translated into Env that is further cleaved into SU and TM envelope proteins, and the accessory proteins Vif, Vpu and Vpr [13,18].

iv. Assembly, budding and maturation of the virion

The viral genomic RNA together with host proteins and the translated viral proteins assemble at the host cell membrane to form new virions (step **10**). The immature virions will be released from the host cell by budding through the cell membrane (step **11, 12**). During this process, the viral core will get enclosed by the host cell-derived lipid membrane. After budding, the viral protease will cleave Gag and Gag-Pol polyprotein precursors into structural and enzymatic viral proteins, which in turn will rearrange within the virion core and transform the immature virions into mature viral particles (step **13**) [18,23].

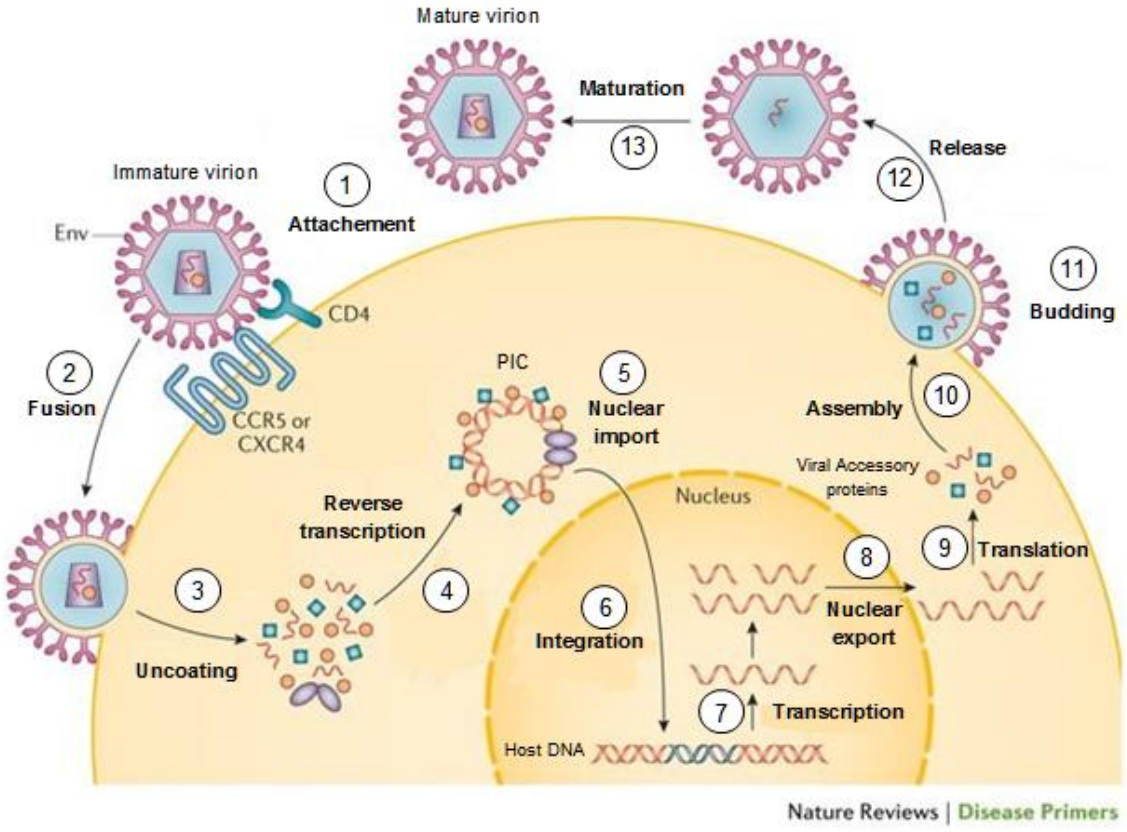


Figure 1.4 | Overview of the HIV-1 replication cycle. Different steps in HIV-1 replication process, starting from attachment of the virus to the target cell (step **1**) to the formation of new mature virions (step **13**). Adapted from Deeks *et al.* 2015 [24].

1.1.5. Pathogenesis of HIV-1 infection

i. Natural course of infection

The primary targets for HIV-1 infection and replication are activated **CD4$^+$ T cells**, while resting CD4$^+$ T cells remain non-permissive *in vitro* [25]. Activated CD4$^+$T cells typically die quickly as a result of the cytopathic effects of the virus or host immune responses. However, some activated CD4$^+$T cells may become infected and survive long enough to revert to resting CD4$^+$ memory T cells, which serve as a latent reservoir that remains invisible to the immune system. With time, this latency may be reversed and the CD4$^+$T cells can produce new HIV-1 virions [25,26]. **Macrophages** can also get infected by HIV-1 but seem to be more resistant to cytopathic effects of HIV-1 replication compared to activated CD4$^+$ T cells [25-27]. Being long-living cells, macrophages can contribute to the latent reservoir for HIV-1. **Dendritic cells (DCs)** express high levels of restriction factor SAMHD1 (Sterile alpha motif and HD-domain-containing protein 1) and are inefficiently infected. However, they may play an important role in the viral spread as they can capture the virus and promote trans-infection of neighboring CD4$^+$ T cells [25,28]. Besides binding to individual DCs, HIV-1 can also attach to the follicular dendritic network in lymphoid tissue, which contains follicular DCs that may provide a reservoir for continuing viral dissemination [24,29].

HIV-1 normally enters the body via intact mucous membranes, injured skin or mucosa and by percutaneous inoculation. Mucosal exposure includes mother-to-child transmission and sexual contact. Percutaneous inoculation, on the other hand, is mainly observed among drug users through needle sharing. However, the most common route of transmission is through sexual contact, hence AIDS is primarily a sexually transmitted disease. Following successful transmission, the natural course of HIV-1 infection is characterized by different stages (**Figure 1.5**) [5,24,30].

Upon infection of the first target cells (*e.g.* macrophages), there is an initial period of 1-3 weeks called **the eclipse phase**. This stage is characterized by undetectable HIV RNA levels in the plasma, spread to the local draining lymph nodes for transmission to CD4$^+$ T cells and direct innate immune responses (*e.g.* type I interferon (IFN-I)).

Once the virus disseminates to other lymphoid tissues (*e.g.* gut-associated lymphoid tissues), replication increases and results in a rise in HIV RNA levels and a decrease in CD4$^+$ T cell number. In this **acute phase**, flu-like symptoms such as, fever, joint pain, malaise, muscle aches may be observed. HIV-specific adaptive immune responses, including CD8$^+$ T cells and neutralizing antibodies will further increase to partially control the viral replication. As a result, there is an asymptomatic period (**clinical latency**) in which the immune system manages to partially recover the CD4$^+$ T cell concentration, and the viral load will reach a stable level (**set-point**). This clinical latency period varies from patient to patient and can last for many years. Persistent infection together with a chronic hyper-immune activation result in a continued destruction of HIV-1 target cells and hence a progressive loss of CD4$^+$ T cells [5,24].

After several years, the CD4$^+$ T cell levels drop below a critical limit (200 cells/µL) and the **AIDS phase** is reached. This stage is further characterized by an increase in HIV viral load, predisposition to opportunistic infections and (oncological) complications as a result of immune failure.

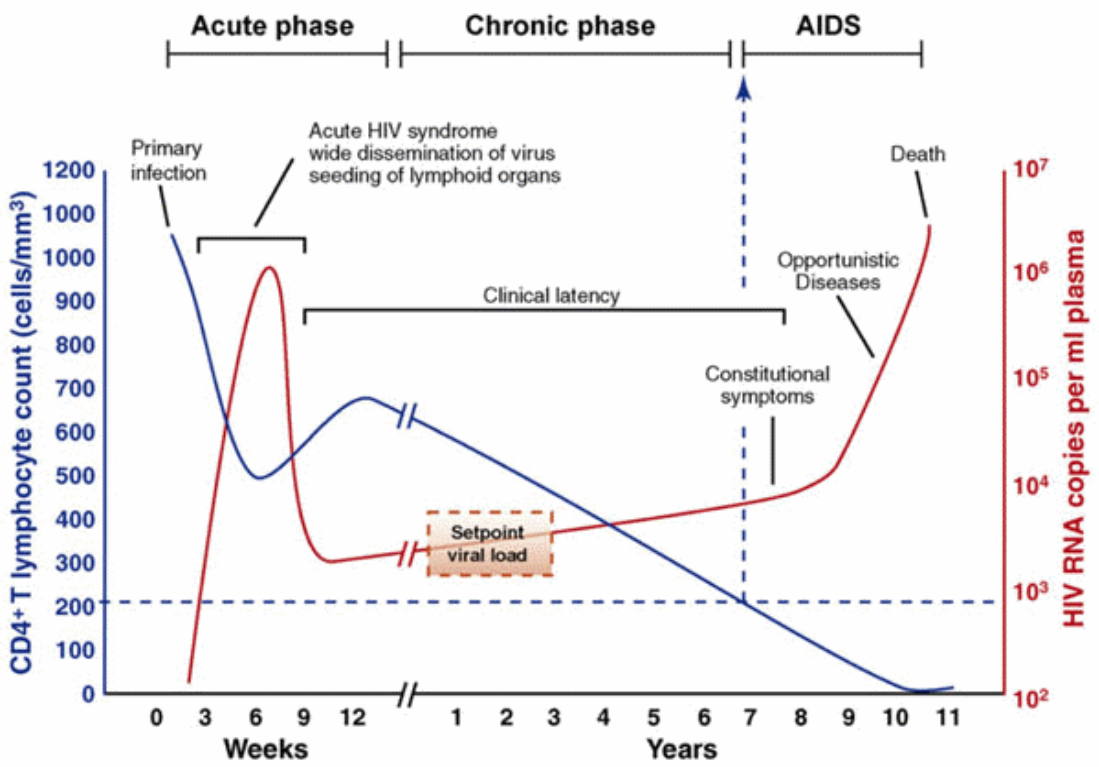


Figure 1.5 | Natural course of HIV-1 infection. The first weeks after HIV-1 infection, no clinical symptoms are observed (**eclipse phase**). The **acute phase** is associated with detectable HIV RNA levels (red line), reaching a peak a few weeks later, at which point the adaptive immune response results in partial control. A steady-state level (**set point**) of viral load is then established in the **clinical latency period.** Immune system exhaustion together with viral replication result in a progressively decline of the $CD4^+$ T cell levels (blue line) during the **chronic phase** which eventually leads to immune deficiency and **AIDS** [31].

ii. Host immune responses

Upon HIV-1 infection, when virus replication is still largely restricted to the mucosal tissues and draining lymph nodes (eclipse phase), increasing levels of acute-phase cytokines and chemokines are observed in the plasma. The innate immune responses are further triggered to inactivate the virus and reduce viral replication. Innate immune cells (*e.g.* NK cells, dendritic cells) rapidly secrete cytokines (*e.g*. interferons) which can have direct antiviral effects by inducing HIV-1 restriction factors (*e.g.* APOBEC3G, tetherin and SAMHD1) or which can further activate innate immune cells and adaptive immune responses (*e.g*. $CD8^+$ T cells and B cells). However, to enable its replication, HIV-1 must counteract the antiviral effects of these cellular restriction factors [30]. Although the activation of innate immune responses during infection may initially be protective, prolonged exposure to IFN-I in the chronic phase is detrimental and likely enhances chronic immune activation and disease progression [32,33]. The innate immune recognition of HIV-1 and the IFN-I response will be further discussed below (**see 1.2.**).

In contrast to the innate immune response that can react quickly against the virus, the adaptive immune system is much more antigen-specific and takes days to weeks. These HIV-specific responses include $CD8^+$ cytotoxic T lymphocytes (CTLs) and specific neutralizing antibodies, which will partially control the viral replication, subsequently leading to a decline of the viral load. Nevertheless, this adaptive immune response fails to durably control the virus. HIV-1 can rapidly mutate in order to escape from CTL and antibody responses. This eventually leads to a loss of virus control and a more chronic phase of infection characterized with ongoing destruction of the immune cells and loss of immune control [24,32,33].

1.1.6. Antiretroviral therapy

In the last two decades, AIDS has become a manageable disease due to the development of antiretroviral therapy (ART). ART is developed to block *de novo* infection of susceptible cells by interfering at various stages of the HIV life cycle, such as targeting viral enzymes (reverse transcriptase, protease or integrase), viral entry or fusion processes. In the absence of viral replication, the immune system can recover much of its lost function and AIDS is prevented [34].

Although ART manages to reduce the plasma viral load to undetectable levels, the virus almost rebounds within weeks when the treatment is stopped. Therefore, life-long treatment is required to suppress viral replication and prevent HIV transmission. Furthermore, once the viral DNA is successfully integrated into its target cell and latency takes place, ART is not able to eradicate these reservoirs. Even after several years of ART, many individuals fail to restore optimal immune function. As such, chronic inflammation together with ART is associated with comorbidities, such as cardiovascular disease, frailty, some cancers, liver dysfunction, renal dysfunction and cognitive decline. On top of that, due to the lack of widespread HIV testing, the expensive costs and the toxicities associated with long- term ART, the majority of the infected population is not on effective ART [24]. Recognizing these limitations, the search for a preventive vaccine or an HIV cure is ongoing.

1.2. Innate immune sensing in HIV-1 infection

The innate immune system is the first line of defense against pathogens and is therefore crucial in bacterial and viral infections. The innate immune system can recognize invading pathogens by the pattern of their surface components, called pathogen-associated molecular patterns (PAMPs). These innate immune responses are typically measured through the production of inflammatory cytokines and IFN-I. Although HIV-1 has unique PAMPs that can alert the human cells of invasion, as a member of the *Retroviridae*, it has developed efficient strategies to evade or counteract immune sensing and downstream responses [35]. Nevertheless, elevated production of IFN-I has been shown in HIV-1 patients, which implies the recognition of HIV-1 by the innate immune system during infection *in vivo* [36].

Over the last years, several pattern recognition receptors (PRRs) have been identified in HIV-1 infection of cells. Many HIV-1 products could serve as a potential PAMP, including viral genomic molecules present within the virions and nucleic acids produced during viral replication [35]. Based on their localization within the cell, two groups of PRRs can be distinguished: Toll-like receptors (TLRs) and cytosolic sensors. TLRs consist of a wide array of receptors with expression varying between cell types. While TLRs are located on endosomal or plasma membranes, cytosolic sensors detect viral RNA and DNA that is introduced to the cytosol during viral infection or that accumulates during replication. Activation of these cytoplasmic sensors is more complex than TLR activation, with interference of host cell restriction factors and initialization of complex signaling networks that eventually lead to an innate immune response. Since HIV-1 can infect different cell types, and since the expression of PRRs varies between cell types, sensing of HIV-1 and the outcome in immune responses is cell type dependent. The HIV-1 target cells typically express both entry receptor CD4 and co-receptor CCR5/CXCR4. Analysis of IFN-I induction *in vitro* has revealed multiple sensing mechanisms of HIV-1 infection in plasmacytoid dendritic cells (pDCs), monocyte-derived dendritic cells (MDDCs), monocyte-derived macrophages (MDMs), and $CD4^+$ T cells [35].

1.2.1. Type I IFN-responses in HIV-1 infection

Upon recognition of HIV-1 by PRRs, innate immune responses are induced resulting in the secretion of IFN-I. These cytokines will subsequently bind to the corresponding receptor IFNAR, which then leads to the activation of the JAK/STAT signaling pathway. Eventually, activation results in the expression of various interferon-stimulated genes (ISGs) which have various effects on the host cells, such as activation of antiviral restriction factors (*e.g.* APOBEC3G, TRIM5α, tetherin and SAMHD1) and further enhancement of innate and adaptive immune responses [37]. However, these IFN-I responses are rather complex during HIV-1 infection and seem to play a dual role. While IFN-I responses may indeed limit viral spread and replication in the early phase of infection, IFN-I contributes to hyper-immune activation and disease progression at a later stage. Since pDCs have been shown to produce large amounts of IFN-α through TLR7 activation in response to HIV-1 infection *in vitro*, and since high levels of IFN-α are detected in infected patients, it has been proposed that pDCs contribute to HIV-1 induced immunopathogenesis [38]. Paradoxically, when infection progresses, the amount of blood pDCs declines in HIV-1 infected patients, but high IFN-I levels are still observed [36]. Therefore, pDCs are likely the main source of IFN-α during acute infection, but when disease progresses, other cell types must be involved in the IFN-I response [37]. Although it remains to be determined what these cells are, cGAS-dependent IFN-I response in activated $CD4^+$ T cells has been observed, which indicates that they might be a potential source of persistent IFN-I in the chronic phase of HIV-1 infection [39].

Given the dual role of IFN-I during HIV-1 infection, both stimulation and inhibition of IFN-I responses have been tested in several studies [37]. Various results were obtained which can be linked to the different stages of HIV-1 infection and the immune status of the patients. Hence, different treatment approaches over the whole course of infection may be required. A better understanding of the sources of IFN-I and the corresponding signaling pathways during different stages of infection is thus crucial for the development of successful therapeutic strategies [37].

1.2.2. cGAS-dependent sensing in HIV-1 infection

Among all the recently identified DNA sensors detecting HIV-1, cGAS is widely recognized as the major sensor to mount antiviral responses in HIV-1 infected cells [40]. cGAS (cyclic GMP-AMP synthase) is a cytosolic DNA sensor which holds an enzymatic activity and will catalyze the cyclization of GTP and ATP to form cGAMP [41]. Upon activation of cGAS, cGAMP will subsequently bind to stimulator of interferon genes (STING), which in turn will recruit 2 kinases: TANK-binding kinase 1 (TBK1) and IκB kinase (IKK). IKK will phosphorylate IκBα, leading to its degradation and the activation of nuclear factor κB (NF-κB), whereas TBK1 phosphorylates IFN regulatory factor 3 (IRF3) which leads to its dimerization. Finally, dimerized IRF3 together with NF-κB translocate into the nucleus and act as transcription factors which are essential for the production of pro-inflammatory cytokines and IFN-I (**Figure 1.6**) [42]. Although cytosolic DNA has long been known to trigger innate immune responses, the activation of cGAS upon HIV-1 infection is rather complex. While cGAS is believed to only detect cytoplasmic DNA, activation of cGAS has so far only been observed after HIV-1 integration [40,43]. Therefore, the origin of the ligand which triggers cGAS is still a topic of debate, and several scenarios have been proposed. Given that cGAS is a DNA sensor, it might recognize cytosolic DNA which originates from the HIV-1 pre-integration complex or which is induced upon intracellular stress and mitochondrial stress [39,44]. This DNA might either originate from a previous infection or might be induced during a new infection event. Additionally, DNA damage induced by HIV-1 upon integration has also been suggested, as DNA damage can induce nuclear translocation of cGAS, which will subsequently bind self-DNA and activate cGAS-mediated signaling [40]. Alternatively, cGAS could sense newly expressed viral RNA or proteins in the presence of a co-receptor [39].

In the following section, an overview is provided of cGAS sensing that has been observed so far in HIV-1 infected primary target cells.

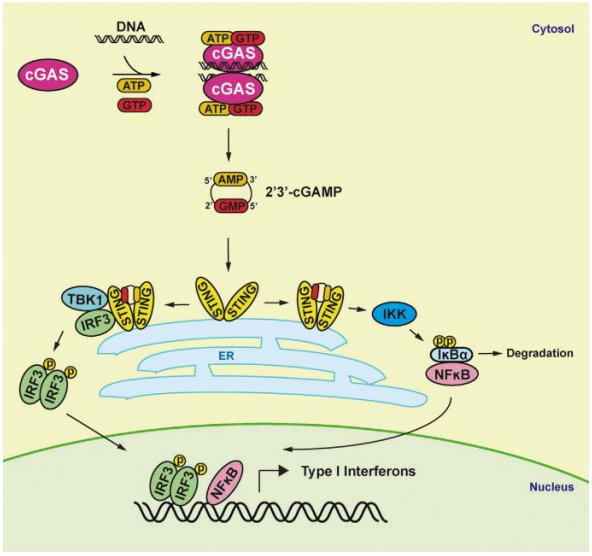


Figure 1.6 | cGAS-STING signaling pathway. Cytosolic DNA activates cGAS to form a dimeric cGAS-DNA complex which synthesizes cGAMP from ATP and GTP. cGAMP binds and activates STING at the endoplasmatic reticulum (ER), which will recruit kinases TBK1 and IKK and lead to downstream activation of IRF3 and NF-κB [42].

i. MDMs and MDDCs

Myeloid cells (macrophages and DCs) do also express CD4 and CCR5/CXCR4 co-receptors and are therefore target cells for HIV-1. To study HIV-1 infection of these cells *in vitro*, monocyte-derived dendritic cells (MDDCs) and monocyte-derived macrophages (MDMs) are used. However, when compared to $CD4^+$ T cells, they are less susceptible to HIV-1 infection. In contrast to HIV-1 infection, HIV-2 infection of myeloid cells is successful, which can be explained by the presence of small accessory protein Vpx. This viral protein can degrade host restriction factor SAMHD1, a cellular deoxynucleoside-triphosphate (dNTP) triphosphohydrolase, which is abundantly expressed in MDDCs and restricts HIV-1 infection in MDDCs. As such, to investigate HIV-1 immune sensing of myeloid cells *in vitro*, HIV-1 productive infection is increased by complementing HIV-1 with HIV-2 Vpx protein [43].

Infection of Vpx treated MDDCs with HIV-1 showed that sensing occurs through the cytosolic DNA sensor cGAS (**Figure 1.7 panel A**). However, production of IFN-β was only observed after integration and innate immune sensing was found to require the capsid and its interaction with cellular cyclophilin A (CypA). The incoming capsid likely "shields" the viral cDNA by forming complexes with CypA and prevents sensing by cGAS before integration. However, once the virus is integrated, newly expressed viral capsid interacts with CypA which, in turn, relieves the "shielding" and allows sensing of viral cDNA by cGAS [43]. In this way, post-integration sensing by cGAS can be circumvented by infection of Vpx treated MDDCs with mutant HIV-1 that is defective for CypA interaction. How this newly expressed HIV-1 capsid controls sensing of the viral cDNA by cGAS is not fully understood yet.

For MDMs, it has been proposed that HIV-1 sensing is similar to sensing in MDDCs, and cytosolic cDNA could serve as a ligand for cGAS only when SAMHD1 is counteracted by Vpx [43]. However, it appears that CypA has a different role in MDMs as compared to MDDCs. Infection of MDMs with HIV-1 that is defective for CypA - in the absence of Vpx - could induce cytosolic sensing and evoke an innate immune response (**Figure 1.7 panel B**).

Overall, although the mechanism of action is not fully understood yet, it is believed that prior to integration, cGAS-mediated immune sensing of HIV-1 infection in myeloid cells is evaded through host-viral interactions. The capsid seems to play a pivotal role in enabling cGAS-mediated immune activation, although its interaction with host CypA is virus strain- and cell type-specific. Once integration takes place, cGAS is activated, which in turn will activate a downstream signaling pathway for the secretion of IFN-β and the expression of ISGs. Of note, MDDCs and MDMs do also express TLR7 and TLR8 which can thus, in theory, recognize ssRNA HIV-1 upon infection. However, this type of sensing has not been reported yet.

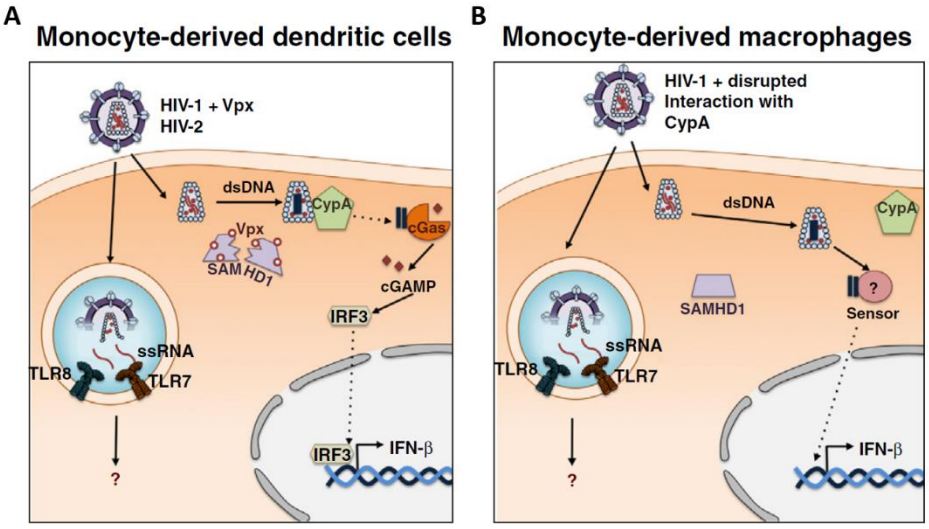


Figure 1.7 | Innate sensing of HIV-1 infection in target cells of the virus. (A) In MDDCs, infection by HIV-1 supplemented with Vpx or by HIV-2 leads to degradation of SAMHD1 and induces IFN-ß production. This requires the viral cDNA (dsDNA) and the cytosolic DNA sensor cGAS. However, sensing is also dependent on capsid-CypA interactions. **(B)** In MDMs, disrupting the interaction between the viral capsid and CypA leads to IFN-ß production in a viral DNA-dependent manner. The sensor is not known, but cGAS cannot be excluded. Question marks indicate possible innate sensing pathways. Adapted from Silvin and Manel 2015 [43].

ii. CD4⁺ T cells

Innate immune sensing mechanisms of HIV-1 in CD4⁺ T cells remain to be clarified. In fact, HIV-1 sensing and subsequent IFN-I production was considered to be absent in CD4⁺ T cells due to effective evasion strategies of the virus [40]. Recently, sensing of HIV-1 has been demonstrated for the first time in manipulated resting tonsil-derived CD4⁺ T cells. This type of sensing likely involves the DNA sensor IFI16 (IFN-Ɣ-inducible protein 16) which senses DNA products that were generated upon abortive infection of the cells. IFI16 subsequently activates the inflammasome, which in turn will lead to pyroptosis, a process that is characterized by cell death. It is still unclear how IFI16 leads to inflammasome activation, but this pathway may explain the mechanism by which CD4⁺ T cells are depleted in HIV-1 infection [43] .

As stated above, cGAS-mediated IFN-I response to HIV-1 infection was first only observed in Vpx treated MDDCs and MDMs. However, Vermeire *et al.* recently showed that HIV-1 can be sensed by cGAS in primary activated CD4⁺ T cells and can successfully mount an IFN-I response [39]. Since *integrase-/tat-* deficient or integrase inhibitor treated viruses strongly reduced IFN-I production, HIV-1 proviral integration seems to be required for the induction of an IFN-I response in activated CD4⁺ T cells. This is in line with what has been reported in myeloid cells (see above), and in this way, cGAS-mediated sensing may require *de novo* expression of HIV-1 proteins to induce an IFN-I response. Infection with HIV-1 with mutations in the capsid proteins did not affect the IFN-I levels in activated CD4⁺ T cells. This is in contrast to what has been observed in MDDCs (see above), which implicates that IFN-I induction in primary activated CD4⁺ T cells does not require newly expressed capsid proteins.

Contrarily, newly expressed HIV-1 accessory proteins Vpr and Vpu were shown to regulate this cGAS-mediated IFN-I response in a opposite manner. *De novo* expression of Vpr was shown to be required for potentiating the post-integration IFN-I response, as virion incorporated Vpr seemed to be insufficient for IFN-I induction. However, when compared to *integrase* deficient virus; *vpr*- and *vpu-vpr*-deleted viruses did not completely reduce IFN-I induction. This suggests that Vpr is likely involved in enhancement of the IFN-I response upon integration, but it is not the only factor that contributes. Newly expressed Vpu, on the other hand, was shown to counteract the IFN-I response. Viruses lacking Vpu induced higher IFN-I responses in $CD4^+$ T cells, and Vpu seems to exert its counteracting effect independent from its ability to bind host restriction factor tetherin (**see 1.2.3**). Overall, these findings demonstrate that HIV-1 can be sensed in its main target cells through cGAS and it is mediated by newly expressed viral proteins. HIV-1 infected $CD4^+$ T cells might thus be a potential source for IFN-I and in this way contribute to the detrimental high immune responses in the chronic phase of the disease.

1.2.3. Strategies by HIV-1 to modulate IFN-I response

Despite the fact that the early IFN-I response can limit viral replication and improve the course of disease, HIV-1 has developed numerous mechanisms to evade host immune responses to enable its replication. HIV-1 can counteract antiviral responses by inhibiting IFN-induced restriction factors -of which tetherin, APOBEC3G and TRIM5α are well studied- as well as by interfering with signaling pathways [14,37,40]. HIV-1 has acquired the four small accessory proteins Vif, Vpr, Vpu and Nef to antagonize antiviral host responses. In general, Vif, Vpu, and Vpr all seem to regulate immune responses by targeting host restriction factors, whereas Nef has numerous effects on infected cells which will eventually enhance viral dissemination and viral immune evasion (**supplementary figure 1**) [14]. In the context of this thesis, I will focus on Vpr and Vpu, and their ability to interfere with innate immune responses.

i. Functions of Vpr and Vpu

Vpu is an integral membrane phosphoprotein which is produced during the late stage of the HIV-1 replication cycle. Vpu is involved in the enhancement of virion release and CD4 degradation. The release of the virion from the plasma membrane is normally blocked by tetherin, a restriction factor induced by IFN-α. Vpu can bind to tetherin in a highly specific manner and can make it a target for polyubiquitylation and subsequent proteasomal degradation [14]. In this way, Vpu promotes the release of newly formed virions. Similar to its interaction with tetherin, Vpu can interact with newly synthesized CD4 receptor in the endoplasmic reticulum and can make it a target for the proteasome. Degradation/down-regulation of CD4 prevents the formation of gp120/CD4 complexes in virally infected cells, which in turn may facilitate virus release or enhance the incorporation of functional Env proteins in virions [14].

Furthermore, Vpr is a virion-incorporated protein which is also expressed in the late phase of HIV-1 infection and is known to induce cytopathic effects in cells. In contrast to Vpu, Vpr has been reported to affect several host cellular processes to promote HIV-1 replication, including cell cycle arrest, induction of cell death, DNA damage, activation of HIV-1 transcription and enhancement of nuclear import via PIC [5,14,19]. However, it remains unclear how alteration of these cellular processes by Vpr facilitates HIV-1 replication. Additionally, many studies support that Vpr, similar to Vpu, might induce proteasomal degradation of host factors as it seems to recruit multiple host proteins (*e.g*. cullin4A-DDB1-DCAF1 complex). Although some studies have reported that Vpr may target certain substrates (*e.g.* TET2, HLTF), it is not immediately clear how this would affect HIV-1 replication [14,40].

ii. The role of Vpr and Vpu in IFN-I response

Vpu and Vpr may exert immunosuppressive effects by interfering at the signaling level and in this way modulate IFN-I responses. It is believed that Vpu suppresses IFN-I expression through inhibition of cellular transcription factor NF-κB (either tetherin-mediated or not) or IRF3. As stated above, Vermeire *et al*. observed that Vpu counteracts the IFN-I response induced upon cGAS sensing in activated $CD4^+$ T cells. Moreover, suppression of IFN-I response by Vpu is independent of its ability to downregulate tetherin or to mediate IRF3 (*e.g.* by degradation). However, HIV-1 virus with a mutation in the Beta-transducin repeat-containing protein (β-TrCP) binding domain of Vpu failed to counteract IFN-I induction in $CD4^+$ T cells, which implies that Vpu likely uses cellular β-TrCP to counteract the cGAS-mediated IFN-I response. Since this is in line with previous reports that suggested that Vpu recruits β-TrCP to prevent translocation of NF-κB to the nucleus, and since dsDNA-mediated STING activation has been shown to also trigger the NF-κB pathway through TBK1, Vpu likely interferes with the cGAS-mediated NF-κB signaling to counteract the IFN-I response (**Figure 1.8**) [39]. Recently, Langer *et al.* used a global transcriptional profiling approach in primary $CD4^+$ T cells and confirmed the hypothesis that Vpu suppresses antiviral gene expression by inhibiting the activation of NF-κB. They could demonstrate that the capability to suppress NF-κB target genes is lost when cells are infected with HIV-1 variants lacking Vpu and released larger amounts of IFN-I and other cellular proteins compared to HIV-1 variants expressing Vpu. Furthermore, they showed that Vpu indeed inhibits immune responses through counteraction of NF-κB signaling, which in turn suppressed the expression of several restriction factors and the release of IFNs. However, the pathways which trigger NF-κB activation in HIV-1 infected $CD4^+$T cells remain to be determined [40].

While it is well established that Vpu counteracts the IFN-I response, the effects of Vpr on the IFN-I response are rather controversial. For example; IFN-I inhibiting effects have recently been observed in MDDCs and MDMs, which could be explained by the interaction of Vpr with TBK1 and hence the blocking of the downstream IRF3 signaling and IFN-I response [45]. In contrast, Zahoor *et al.* observed that HIV-1 Vpr potentiates IFN-I responses and downstream ISGs expression in both MDDCs and MDMs [46,47]. Furthermore, activation of NF-κB by Vpr has also been reported. Liu *et al.* demonstrated that Vpr associates with transforming growth factor-β-activated kinase 1 (TAK1) in order to enhance its phosphorylation and polyubiquitination, which consequently activates NF-κB (and AP-1) signaling [48]. Additionally, Vermeire *et al.* showed that in $CD4^+$ T cells, cGAS-mediated IFN-I induction upon post-integration is enhanced by Vpr [39]. How Vpr interferes with this cGAS-mediated IFN-I response remains to be determined, but it is most likely dependent on its interaction with the DDB1 and CUL4 associated factor 1 (DCAF1) to exert this potentiating effect. Moreover, as inhibition of TAK1 resulted in a decrease of the IFN-I response after HIV-1 infection, and as infection with *vpr*-defective HIV-1 virus did not affect the IFN-I induction, activation of NF-κB dependent immune activation by Vpr is plausible (**Figure 1.8**). Furthermore, they also checked the effect of Vpr on the IFN-I response in co-cultures of Vpx-treated MDDCs and MT4 cells infected with either *vpr*-deficient or WT HIV-1 viruses. The IFN-I production was here also strongly reduced in the absence of Vpr, indicating that Vpr potentiates IFN-I production in both activated $CD4^+$T cells and MDDCs.

Overall, the molecular mechanisms by which Vpr and Vpu regulate innate immune responses in different HIV-1 target cells are still a matter of debate and remain to be clarified.

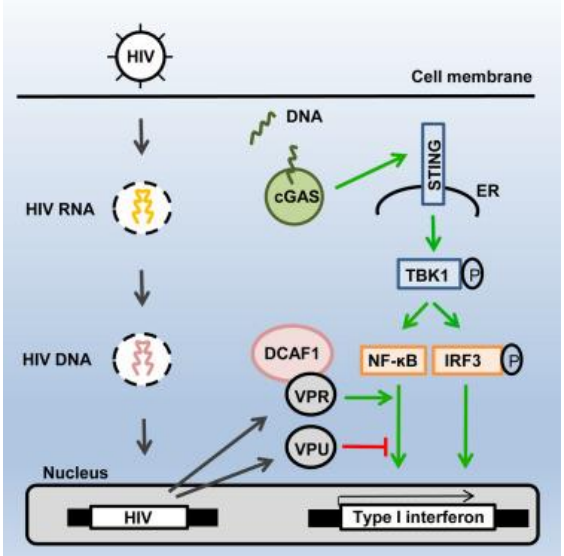


Figure 1.8| Schematic representation of cGAS-dependent, Vpu- and Vpr- regulated type I interferon responses in activated CD4⁺ T cells. Upon integration of HIV-1 in the host genome, cGAS is triggered and activates both the STING-TBK1-IRF3 signaling pathway and NF-κB. Nuclear translocation of IRF3 and NF-κB leads to the expression of type I interferons. Besides, the cGAS-dependent IFN-I response seems to be mediated by newly expressed viral products. Newly expressed Vpr was shown to potentiate IFN-I induction in a DCAF-1-dependent manner, whereas Vpu counteracted the IFN-I response [39].

1.3. Study design and objectives

As discussed above, HIV-1 has been shown to successfully mount an IFN-I response in primary activated CD4⁺ T cells through a post-integration mechanism. This requires the cytosolic DNA sensor cGAS and the downstream STING-TBK1-IRF3 signaling pathway. In addition, this IFN-I response seems to be regulated by two newly expressed HIV-1 accessory proteins: Vpr and Vpu. Interestingly, both viral proteins exert opposite effects on the cGAS-mediated IFN-I response: Vpr potentiates IFN-I induction, while Vpu counteracts this response. Given that the CD4⁺ T cells might contribute to the elevated IFN-I production and hyper-immune activation during the chronic disease stage, characterization of the mechanisms by which these newly expressed proteins mediate the IFN-I response will allow us to better understand the pathology of HIV-1 infection. This in turn may pave the way for the development of new therapeutic approaches. Furthermore, as cGAS-mediated IFN-response in primary activated CD4⁺ T cells only occurs upon integration, we can speculate that *de novo* expression of proteins is likely needed. As newly expressed Vpr and Vpu where shown to alter the cGAS-mediated IFN-response, we hypothesize that both viral proteins might interfere with factors in the cGAS-mediated IFN-I signaling pathway. To assess whether this hypothesis is true, the function of Vpu and Vpr must be further characterized. Vermeire *et al*. already evaluated the effect of Vpr and Vpu on the IFN-I response, by infecting CD4⁺ T cells with WT/*vpr*- *vpu*-deficient HIV-1 virus, or by infecting with *vpr*-deficient HIV-1 complemented with a WT Vpr protein[39]. To further evaluate the function of Vpr and Vpu, we will first investigate which effect these proteins have on IFN-I when introduced – in the absence of HIV-1 - in activated CD4⁺ T cells. As a first approach, this effect was studied by transduction of CD4⁺ T cells with retroviral vectors LZRS-IRES-Vpr-NGFR and LZRS-IRES-Vpu-NGFR. However, long-term expression of Vpr resulted in cytopathic effects in the cells. To overcome this, inducible retro- or lentiviral constructs are required to regulate the expression of the viral proteins in a dose and time dependent manner.

Therefore, the aim of this thesis was to construct and validate new inducible lentiviral vectors encoding either Vpr or Vpu, which will allow further research to evaluate the role of newly expressed Vpr and Vpu on the cGAS-mediated HIV-1 immune sensing in $CD4^+$ T cells. An inducible lentiviral system that can be used is the Tet-On inducible system, which is based on the addition of tetracycline/doxycycline (Dox) in the culture medium. This will, in turn, activate transcription of the gene of interest through binding of a mutant transactivator (rtTA) to the tet-responsive elements. In this way, the gene encoding Vpr or Vpu will be mostly inactive, and will only be turned on occasionally, which could prevent toxicity in the cells. We first generated lentiviral inducible constructs by introducing either Vpr and dNGFR, Vpu and dNGFR or dNGFR alone in the Tet-On inducible plasmid. Lentivirus was then produced, and the quality of the virus production was verified by the Reverse transcriptase (RT) assay. We also assessed the transduction efficiency of all three constructs by lentiviral transduction of SupT1 and Jurkat cell lines using the puromycin viability assay. Finally, in order to evaluate the Dox-dependent expression of Vpr and Vpu in Jurkat and SupT1 cells, flow cytometry and Western Blot experiments were prepared.

2. MATERIALS AND METHODS

All experiments were performed under supervision unless indicated otherwise.

2.1. Construction of inducible lentiviral vectors

We used the all-in-one doxycycline inducible lentiviral plasmid pCW57-MCS1-2A-MCS2 as a backbone for cloning of Vpr, Vpu and dNGFR (**Figure 2.1**). pCW57-MCS1-2A-MCS2 was a gift from Adam Karpf (Addgene plasmid #71782; http://n2t.net/addgene:71782; RRID:Addgene_71782)[49]. This plasmid allows bicistronic expression mediated by P2A. We first generated two constructs by cloning either Vpr or Vpu in the MCS downstream of the P2A sequence. The developed constructs were then used to introduce marker gene dNGFR upstream (MCS1) of the P2A sequence. We also generated a control plasmid by cloning only dNGFR upstream of P2A sequence in pCW57-MCS1-2A-MCS2.

Vpr (291 bp) and *Vpu* (246 bp) (**supplementary figure 2**) were first amplified by PCR from the template plasmids that were previously constructed at our lab. Retroviral LZRS-Vpr-IRES-dNGFR and LZRS-Vpu-IRES-eGFP template plasmids were used respectively. PCR was performed using Thermoscientific Phusion High Fidelity DNA polymerase[50] and tailed primers to add restriction sites for *MluI* and *BamHI* and random nucleotides. The PCR program consisted of an initial denaturation step of 30 seconds (s) at 98 °C, 35 amplification cycles of 10 s denaturation at 98 °C, 30 s annealing at 63 °C and 15 s elongation at 72 °C, followed by a final elongation step of 2 minutes (min) at 72 °C. Primer sequences are depicted in **Table 1,** details for the PCR reaction can be found in **supplementary figure 3.** In a next step, lengths of the amplified fragments were checked on 1.5% agarose gel and PCR products were purified with the QIAquick PCR Purification Kit (QIAGEN) according to the manufacturer's instructions. Both the PCR products and the lentiviral backbone pCW57-MCS1-2A-MCS2 were digested by High-Fidelity (HF) restriction enzymes *MluI* and *BamHI* (NEB, New England Bio Labs, USA) for 2 hours (h) at 37 °C to create overhangs, and the restriction products were subsequently purified with the QIAquick PCR Purification Kit (QIAGEN). The purified products were then ligated (insert to vector ratio of 3:1) in the *MluI* and *BamHI* sites of the digested backbone using the T4 DNA ligase (NEB, New England Bio Labs, #MO2O2M) at room temperature (RT) for 3 h. The generated plasmids were then introduced into DH5alpha competent cells via heat shock treatment (at 42 °C for 2 min), grown for 1 h at 37 °C while shaking (225 rpm), and then transferred to LB agar plates for overnight incubation at 37 °C in the presence of ampicillin. The next morning, single colonies were picked for further growing (at 37 °C, 225 rpm, 16 h) to increase the plasmid yield. Plasmids were then purified using the QIAprep Spin, Miniprep Kit (QIAGEN, #27106) according to the manufacturer's instructions, verified by digestion with the HF restriction enzymes (New England Bio Labs, USA) and subjected to 1.5% agarose gel electrophoresis for separation. The generated constructs were termed pCW57-Vpr and pCW57-Vpu and were also verified by Sanger sequencing (Eurofins genomics, Germany) and Nucleotide BLAST(ncbi). The sequencing primers are listed in **supplementary figure 4.**

Marker gene *dNGFR* (830 bp) (**supplementary figure 2**) was amplified by PCR from the retroviral template plasmid LZRS-Vpr-IRES-dNGFR previously constructed at our lab. PCR was performed as described above using tailed primers with random nucleotides and restriction sites for *NheI* and *AgeI*. However, primer annealing temperatures and elongation time were adjusted depending on the primer and template length, respectively. Primer sequences are also listed in **Table 1,** details on the PCR reaction and cycling conditions can be found in **supplementary figure 3** and **5**, respectively. In order to maintain the reading frame with P2A and the downstream gene, an additional bp was added in the reverse primer. In this way, an 831 bp dNGFR fragment was formed upon PCR amplification, in which the last codon (AGC) will be translated into serine, which is the next amino acid that is formed in the non-truncated NGFR.

Next, the PCR product was checked on 1.5% agarose gel, excised from the gel and subsequently purified with the Wizard® SV Gel and PCR Clean-Up kit according to the manufacturer's instructions [51]. The purified dNGFR PCR product, the pCW57-MCS1-2A-MCS2 backbone and the previously constructed pCW57-Vpr and pCW57-Vpu plasmids were all digested by HF restriction enzymes *NheI* and *AgeI* (New England Bio Labs, USA) and purified with QIAquick PCR Purification Kit (QIAGEN). The ligation, transformation of competent cells and verification of the generated constructs were performed as described above for Vpr and Vpu cloning. The final lentiviral constructs were termed pCW57-dNGFR-P2A-Vpr, pCW57-dNGFR-P2A-Vpu and pCW57-dNGFR (control vector). All three constructs were also verified by DNA sequencing (Eurofins Genomics, Germany) and Nucleotide BLAST(ncbi).The sequencing primers are listed in **supplementary figure 4.**

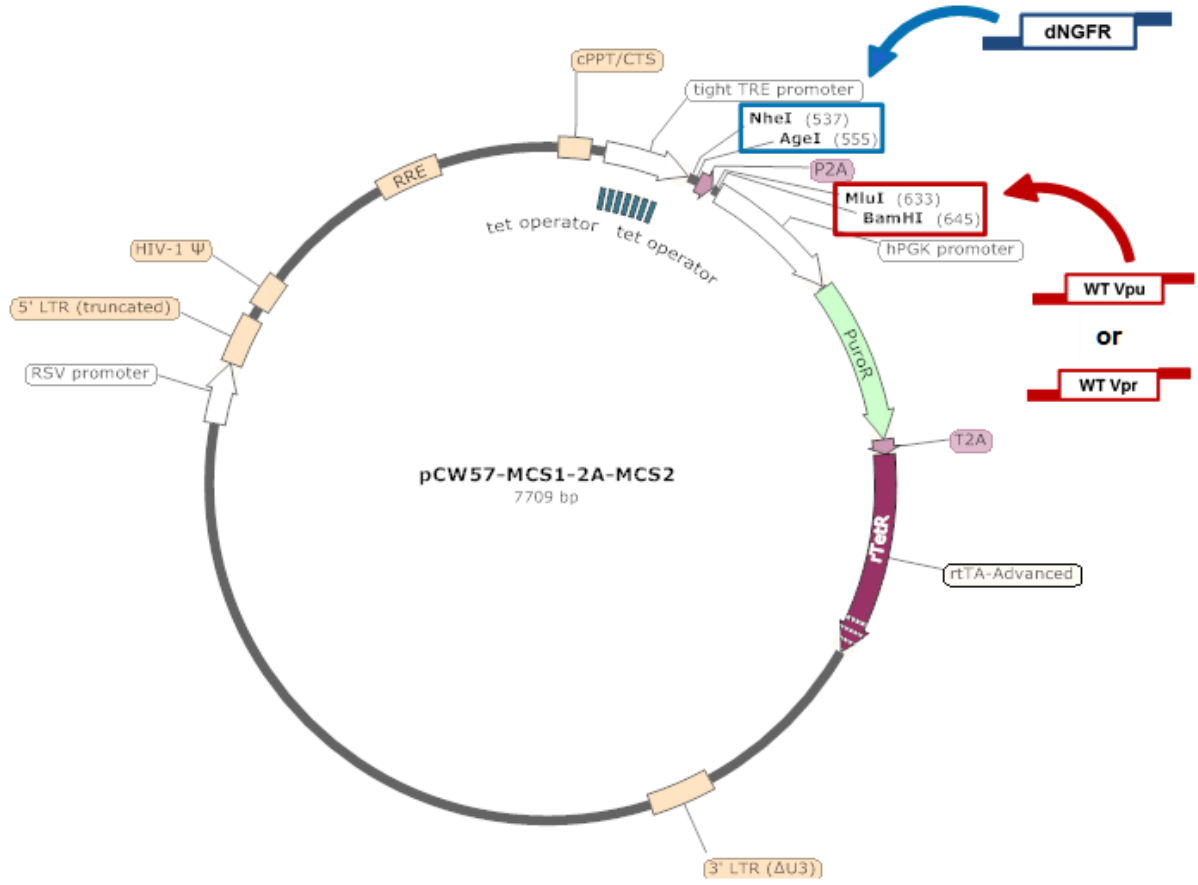


Figure 2.1| Schematic illustration of the molecular cloning procedure of Vpr, Vpu and dNGFR the all-in-one Tet-On inducible lentiviral backbone pCW57-MCS1-2A-MCS2. After PCR amplification with tailed primers, restriction sites were added to wild-type (WT) Vpr/Vpu and dNGFR for cloning into the backbone plasmid.

Table 1 | Primer design for PCR amplification of Vpu, Vpr and dNGFR*.

Primer ID	Gene	Direction	Primer sequence (5' – 3')
MluI-Vpr-Fwd	*vpr*	forward	*AAAGTT*ACGCGTATGGAACAAGCCCCAGAAG
BamHI-Vpr-Rev	*vpr*	reverse	*AAAGTT*GGATCCCTAGGATCTACTGGCTCCATTTC
MluI-Vpu-Fwd	*vpu*	forward	*AAAGTT*ACGCGTATGCAACCTATAATAGTAGCAATAGTAGC
BamHI-Vpu-Rev	*vpu*	reverse	*AAAGTT*GGATCCCTACAGATCATCAATATCCCAAGGAG
NheI-dNGFR-Fwd	*dngfr*	forward	*AAAGTT*GCTAGCATGGGGGCAGGTGCCACC
AgeI-dNGFR-Rev	*dngfr*	reverse	*AAAGTT*ACCGGTGCTGTTCCACCTCTTGAAGGCTATGTAGG

*All primers were designed in-house with SnapGene software. Tail part of primers is underlined: random nucleotides are shown in italic, additional restriction sites (ACGCGT = *MluI*, GGATCC = *BamHI,* GCTAGC= *NheI*, ACCGGT= *AgeI*) are shown in red, and additional nucleotides are in green.

2.2. Cell culture and thawing of cells

293T cells were used for the production of lentivirus, and Jurkat E6.1 and SupT1 cells were used for transductions with lentiviral supernatant. Furthermore, transduced SupT1 cells (TRP-EF1A-LINK-IRES-NGFR-WPRE) and transduced Jurkat cells (LZRS-LIN-I-NGFR) were used as control stable cell lines to assess the activity of monoclonal phycoerythrin (PE)-conjugated antibody nerve growth factor receptor (NGFR) low-affinity receptor (anti-NGFR-PE, clone ME20.4, Chromaprobe, Maryland Heights, MO) via flow cytometry. All cell lines were cultured at 37 °C in a 7 % CO_2 humidified atmosphere in Iscove modified Dulbecco complete medium (IMDMc): IMDM (Invitrogen, Merelbeke, Belgium) supplemented with 10% heat-inactivated fetal calf serum (FCS, Hyclone, Perbio) and 1% L-glutamine (Invitrogen).

All frozen cells were stored in liquid nitrogen and were quickly thawed in a 37 °C warm water bath. The cells were then stepwise diluted in cold IMDMc, centrifuged (1500 rpm, 5 min ,4 °C) and subsequently resuspended in warm IMDMc. All cell lines were maintained in suitable culture flasks and were passaged routinely until they were used for experiments. Cell counting was done manually by Trypan Blue staining (Invitrogen) using a Bürker counting chamber.

2.3. Production of lentiviral supernatants.

For the production of replication-deficient lentiviral supernatant from pCW57-dNGFR-Vpr, pCW57-dNGFR-Vpu or pCW57-dNGFR, 293T cells were seeded 24 h before transfection in 6-wells plates at a density of 400,000 cells per well. Transfection with the lentiviral plasmid of interest together with the MISSION® Lentiviral Packaging Mix (lentiviral packaging plasmid and the VSV envelope plasmid, Sigma-Aldrich) was carried out using FuGENE® HD Transfection Reagent (Promega, Leiden, The Netherlands) following the protocol recommended by the manufacturer. Medium (IMDMc) was refreshed 24 h after transfection and viral supernatant was harvested 48 h after transfection, centrifugated at 2000 rpm for 10 min to purify the supernatant from remaining cells. The viral supernatants were stored at -80 °C until use.

2.4. Viral titer determination: Reverse transcriptase activity assay

This assay was performed by Evelien Naessens.

To check the quality of the lentiviral production, we quantified the reverse transcriptase (RT) activity in the lentiviral supernatants with the one-step SG-PERT assay as described earlier by our research group [52]. Briefly, lentiviral supernatants were first lysed and were added to a reaction mix of exogenous RNA template (bacteriophage MS2 RNA), MS2 specific primers and a SYBR Green I qPCR mastermix. The activity of the RT in the lentiviral supernatants was then assessed by quantification of the synthesized cDNA using qPCR and subsequent calculation (via Excel) of the titer from the obtained Cq values. The titers are typically expressed as RT activity units (mU) per milliliter (mL) (mU RT/mL), although the equivalent p24 value (ng/mL) can also be calculated.

2.5. Puromycin viability assay

A serial dilution (1:10,1:50,1:250,1:1250,1:6250,1:31250,1:156250 and 1:781250) of supernatants of each stock of lentivirus was first made. SupT1 and Jurkat cells were then plated at 50,000 cells per well in 50 µL IMDMc in a 96-well flat bottom plate and transduced with 50 µL lentiviral supernatant in the presence of polybrene (8µg/mL, Sigma-Aldrich). Cells were subsequently spinoculated for 30 min, 2300 rpm at 32 °C. After 24 h, cells were spinoculated (1500 rpm, 5 min, 4 °C), medium was aspirated, and fresh IMDMc was added. Puromycin selection (2 µg/mL, Sigma-Aldrich) was started 72 h after transduction.

Given that at high dilutions none of the cells become infected (because little to no particles are present), more wells were taken for the highest dilutions. For each dilution, 1 to 2 wells with transduced cells without puromycin were included as a positive control for cell survival. Wells with non-transduced Jurkat and SupT1 cells were also included to control for puromycin toxicity. The percentage (%) of transduced wells was determined after 9 days of puromycin selection by manually recording the number of wells for each virus dilution in which cell death was observed. All calculations were performed in Microsoft Office Excel 2016.

To obtain stable cell lines, positively selected SupT1 and Jurkat cells from the viability assay were further maintained in culture for about 14 days, alternating between addition and removal of puromycin (2 µg/mL).

2.6. Flow cytometry

Flow cytometry was performed to assess the activity of human monoclonal anti-NGFR-PE. Transduced control cell lines (TRP-EF1A-LINK-IRES-NGFR-WPRE TD SupT1 and LZRS-LIN-I-NGFR TD Jurkat) and non-transduced Jurkat and SupT1 cells were each plated in 96-well plate at a density of 200,000 cells per well, stained with anti-NGFR-PE for 30 min at 4 °C and analyzed for NGFR expression on a MACSQuant® flow cytometer (Miltenyi Biotec). Analysis of data was performed with the MACSQuantify™ v2.8 software.

For gating, non-transduced SupT1 and Jurkat cells were used. Cells were selected based on forward (FSC) vs side scatter (SSC) (green) and the viable cells (red) were then gated by means of propidium-iodide (negative). An additional gating (blue) was then performed on the viable cells to obtain the $NGFR^+$ cells (shown in **Figure 2.2**). For both transduced SupT1 and Jurkat cell lines, NGFR expression was measured in 5 samples.

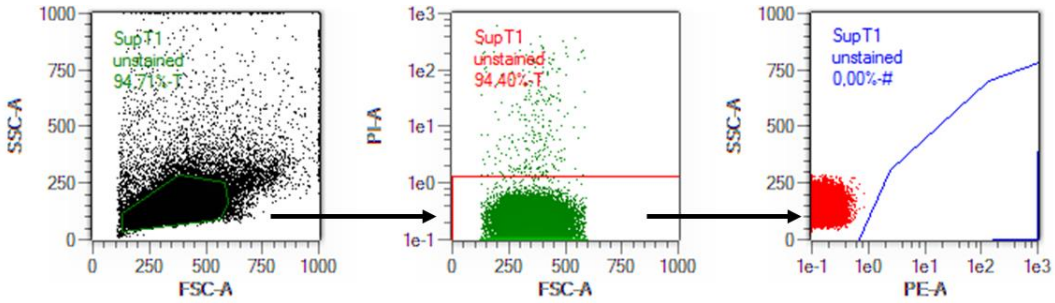


Figure 2.2 | Gating strategy applied on non-transduced control cells (shown here: SupT1 cells)

2.7. Future experiments

2.7.1. Doxycycline induction

In order to determine whether the produced inducible lentiviral vectors are indeed expressing Vpr, Vpu and dNGFR, functional assays must be performed. Given that dNGFR is simultaneously expressed with Vpr or Vpu, it can be used as a marker to identify Vpr/Vpu expressing cells by flow cytometry. We first want to evaluate the expression of dNGFR upon doxycycline (Dox) induction by using flow cytometry. The idea is to plate pCW57-dNGFR, pCW57-dNGFR-P2A-Vpr and pCW57-dNGFR-P2A-Vpu transduced SupT1 and Jurkat cells (stable cell lines) at 200,000/ well in a 96-well flat bottom plate in IMDMc containing various concentrations of Dox (0.1, 1 and 10 µg/mL), and keep the cells at 37 °C in an atmosphere of 7% CO_2 (**Figure 2.3**). For transduced cells, an un unstained sample (Dox(-), anti-NGFR-PE (-)) must be included together with a negative control for cell death caused by the transduction process (Dox (-), anti-NGFR-PE(+)). Non-transduced cells in the presence of Dox can also be included as controls to distinguish toxicity (if present) caused by HIV Vpr/Vpu proteins from toxicity caused by the (high) Dox concentration. After 24h and 48h, the expression of dNGFR will be determined by flow cytometry using anti-NGFR-PE staining. In this way, we can assess whether dNGFR (and thus Vpr and Vpu) are indeed expressed only after Dox addition.

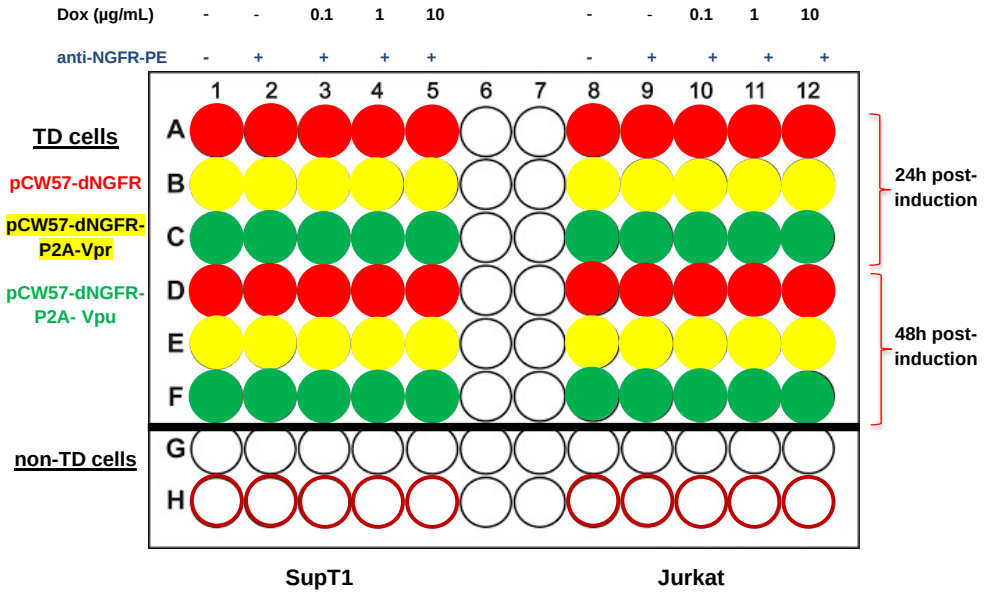


Figure 2.3| Experiment outline: induction of Vpr, Vpu and dNGFR expression with doxycycline. Wells with transduced (TD) cells are depicted in different colors (transduced with: pCW57-dNGFR= red, pCW57-dNGFR-P2A-Vpr = yellow or pCW57-dNGFR-P2A-Vpu=green). Non-TD cells are shown in row H. Anti-NGFR-PE = staining with antibody targeting dNGFR. An unstained sample (Dox(-),anti-NGFR-PE (-)), negative controls (Dox (-), anti-NGFR-PE (+)) and non-transduced Jurkat and SupT1 cells are included.

2.7.2. Western blot

To ensure that WT Vpr and Vpu are expressed by the inducible lentiviral vectors, we also planned to perform a Western Blot. Similar to what has been described for the flow cytometry analysis, both transduced Jurkat and SupT1 cells will be treated with several concentrations of Dox. The cells will then be lysed after 24h and 48h, and the expression of WT HIV-1 Vpr and Vpu will then be detected using primary antibodies.

3. RESULTS

3.1. Construction of Lentiviral Vectors expressing Vpr or Vpu.

To establish inducible expression of HIV accessory proteins Vpr or Vpu, we used the all-in-one doxycycline-inducible lentiviral vector plasmid pCW57-MCS1-2A-MCS2. This plasmid encodes 2A self-cleaving peptide sequences (P2A and T2A), which are virus derived elements that can be used to establish polycistronic mRNA in eukaryotes. The P2A sequence in this all-in-one lentiviral backbone can be used to clone two genes of interest that are under the control of an inducible promoter (TetO), whereas the T2A sequence allows the continuous expression of rtTA and the puromycin resistance gene by using an independent promoter (hPGK). As such, the lentiviral constructs continuously express the puromycin resistance gene, which allows selection of efficiently transduced cells. The cloning strategies we used for generating our Vpr and Vpu constructs are summarized in **Figure 3.1**.

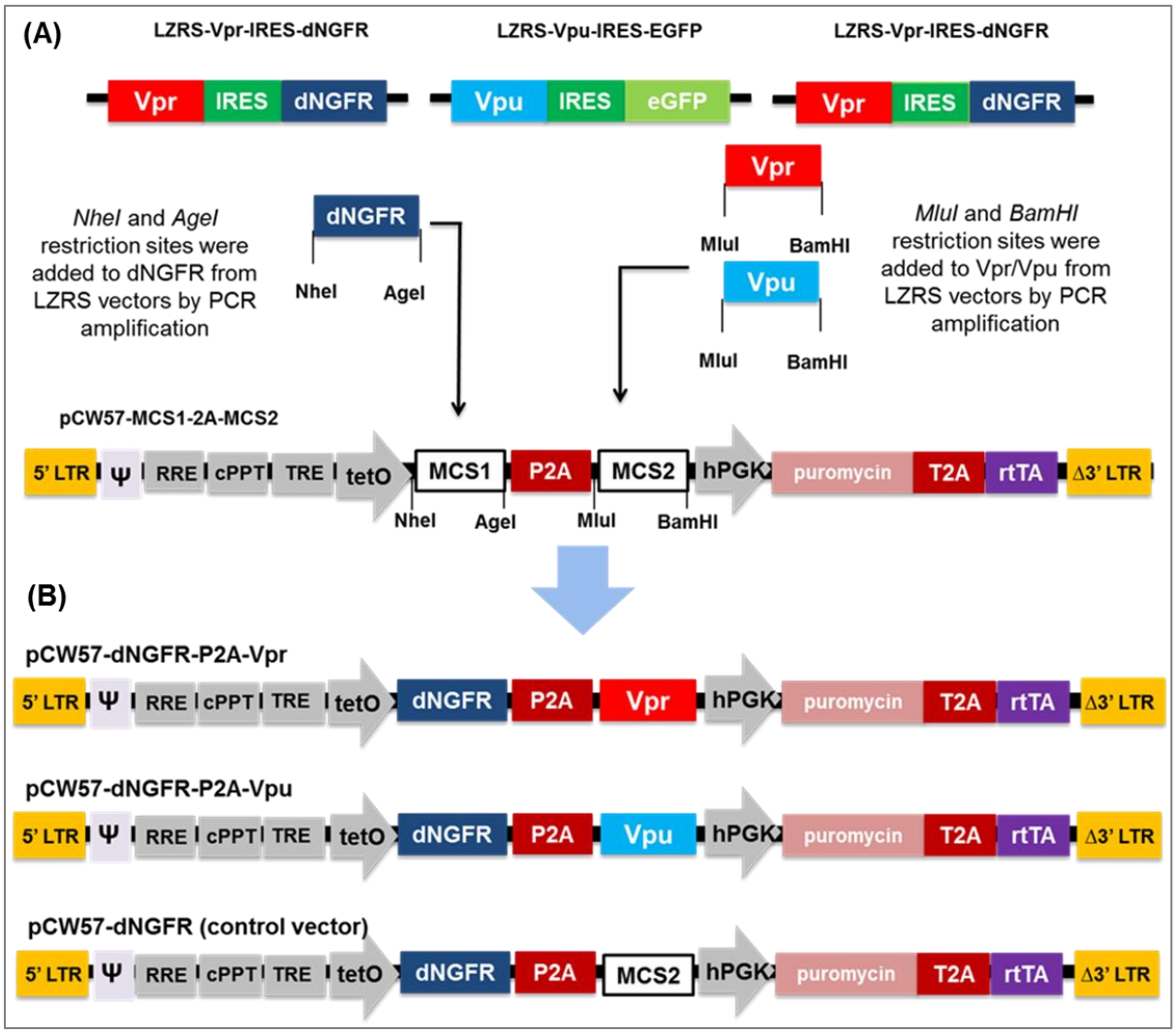


Figure 3.1| Schematic representation of the cloning steps used to make inducible lentiviral vectors expressing either Vpr and dNGFR, Vpu and dNGFR or dNGFR alone. (A) Vpr/Vpu and dNGFR were amplified from retroviral plasmids (LZRS) and inserted downstream and upstream of the P2A sequence of pCW57-MCS1-2A-MCS2, respectively. **(B)** The final constructs were termed pCW57-dNGFR-P2A-Vpr, pCW57-dNGFR-P2A-Vpu and pCW57-dNGFR (empty vector). Abbreviations: 5'LTR, truncated 5' long terminal repeat from HIV-1; Ψ, packaging signal HIV-1; RRE, HIV Rev response element; TRE, Tet-responsive elements; cPPT, central polypurine tract; tetO, Tet-operator (inducible promoter); hPGK, human phosphoglycerate kinase 1 promoter; rtTA, improved tetracycline-controlled transactivator; Δ3'LTR, self-inactivating truncated 3' long terminal repeat from HIV-1.

3.1.1. Cloning of Vpr and Vpu

WT Vpr (291 bp) and Vpu (246 bp) from retroviral constructs were amplified by PCR using primers with overhangs containing restriction enzyme sites. Amplification products were subsequently checked on 1.5% agarose gel. We expected to see fragments of 303 bp and 258 bp for Vpr and Vpu, respectively. Both fragments with Vpr (**Figure 3.2A**) and Vpu (**Figure 3.2B**) were indeed present in the PCR product sample.

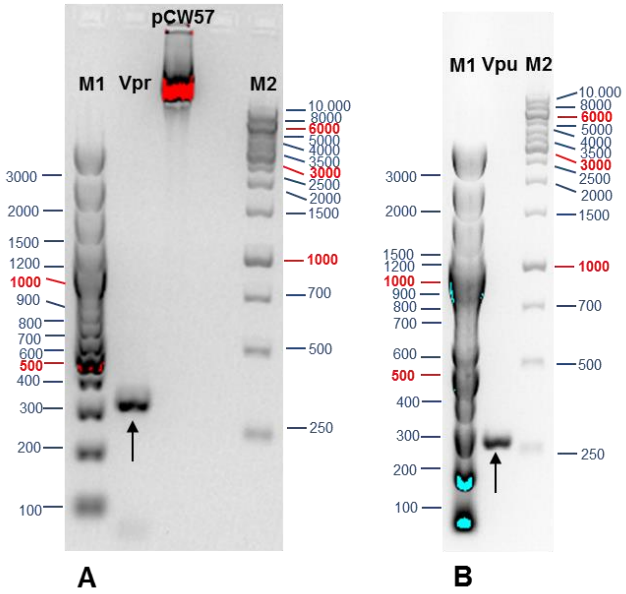
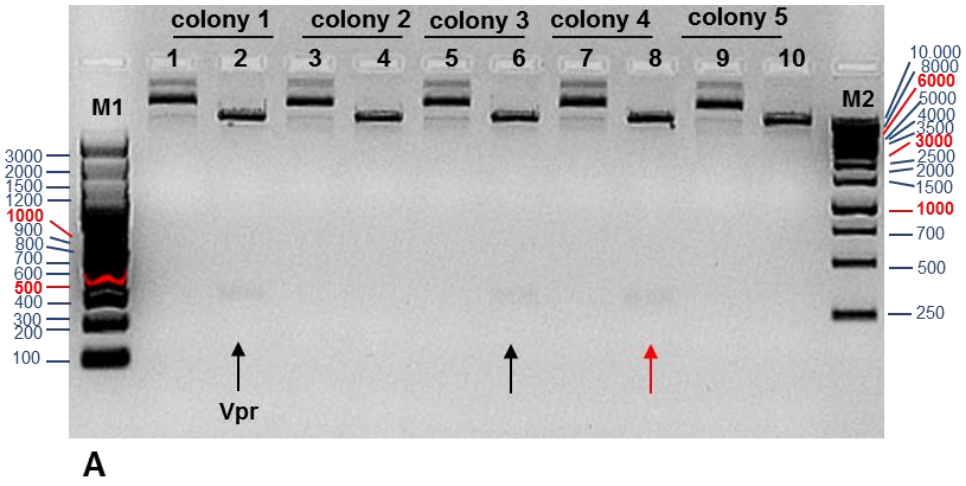


Figure 3.2| Vpr and Vpu cloning: analysis of PCR products on 1.5% agarose gel electrophoresis. Restriction sites were added to **(A)** Vpr (291 bp) and **(B)** Vpu (246 bp) via PCR amplification and PCR products were subsequently analyzed on 1.5% agarose gel. pCW7-MCS1-2A-MCS2 plasmid (7709 bp) was also checked on 1.5% agarose gel. Lane M1: 100 bp DNA ladder, lane M2 :1kb DNA ladder. Arrows indicate fragment of interest.

We then ligated Vpr/Vpu in the *MluI* and *BamH*I restriction sites (in MCS2) of the pCW57-MCS1-2A-MCS2 plasmid to yield two plasmids we termed pCW57-Vpr and pCW57-Vpu. Upon transformation of competent *E. coli* with the cloned plasmids, five single colonies were picked, and screened for the 291bp (Vpr) and 246 bp (Vpu) fragments by restriction enzyme digestion and gel electrophoresis. Three of the five recombinant colonies were positive for the Vpr (**Figure 3.3A**) or Vpu (**Figure 3.3B**) fragment and were confirmed with DNA sequencing (see **supplementary figure 6** for sequence analysis). Colony 4 for both pCW57-Vpu and pCW57-Vpr was eventually selected for the subsequent cloning steps.



Legend on next page

Figure 3.3A on previous page

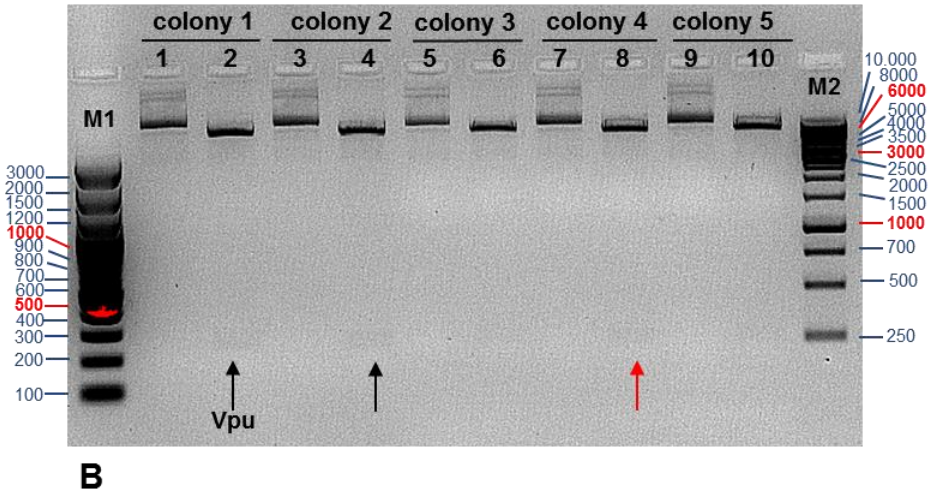
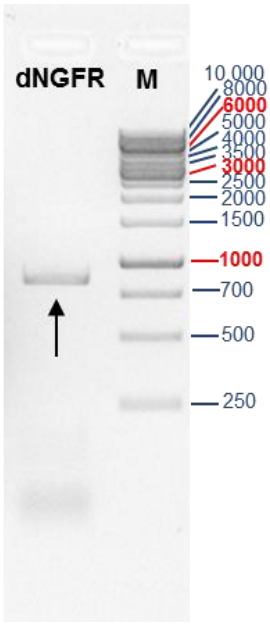


Figure 3.3| Vpr and Vpu cloning: analysis of restriction digest products on 1.5% agarose gel electrophoresis. (A) Restriction digest of pCW57-Vpr recombinant colonies with *MluI* and *BamHI*. Lanes 1,3, 5, 7 and 9 are uncut controls for each colony. Only lanes 2, 6 and 8 show a fragment of 291 bp (Vpr). **(B)** Restriction digest of colonies with pCW57-Vpu with *MluI* and *BamHI*. Lanes 1, 3, 5, 7 and 9 are uncut controls for each colony. Only lanes 2, 4 and 8 show a fragment of 246 bp (Vpu). Lane M1: 100 bp DNA ladder, lane M2 :1kb DNA ladder. Red arrows indicate selected colony.

3.1.2. Cloning of dNGFR

In order to allow the identification of Vpr/Vpu expressing cells later on, we also added a marker gene in our construct. We PCR-amplified dNGFR (830 bp) from a retroviral vector template plasmid and cloned it upstream (in MCS1) of the P2A sequence. In this way, dNGFR will be simultaneously expressed with Vpr/Vpu upon induction with doxycycline. Given the characteristics of P2A constructs, it is important that (1) both genes are in frame with the P2A sequence and that (2) no STOP codon is formed before the P2A sequence, which would further prevent the translation of the P2A sequence as well as the downstream protein. For this reason, besides the *NheI* and *AgeI* restriction sites, we also added an extra cytosine in the reverse primer. In this way, the last codon (AGC) of dNGFR will be translated into a serine. The PCR product was checked on 1.5% agarose gel (**Figure 3.4**), the amplified dNGFR fragment of 843 bp was retrieved, purified and cut with *NheI* and *AgeI.*



Legend on next page

Figure on previous page

Figure 3.4| dNGFR cloning: analysis on 1.5% agarose gel electrophoresis dNGFR was amplified via PCR and subsequently analyzed on 1.5% agarose gel. The amplified dNGFR fragment of 843 bp was retrieved. Lane M :1kb DNA ladder. Arrow indicates fragment of interest.

Next, we ligated the purified dNGFR fragment in the *NheI* and *AgeI* restriction sites of pCW57-Vpr, pCW57-Vpu and pCW7-MCS1-2A-MCS2 to yield the desired lentiviral plasmids: pCW57-dNGFR-P2A-Vpr, pCW57-dNGFR-P2A-Vpu and pCW7-dNGFR. After transformation of competent cells, single colonies were picked and grown. Plasmids were then purified and screened for the dNGFR fragment (831 bp) by restriction enzyme digestion and gel electrophoresis (**Figure 3.5**). For pCW57-dNGFR, all five colonies were positive for dNGFR, and for pCW57-dNGFR-P2A-Vpr four colonies were positive. For pCW57-dNGFR-P2A-Vpu, however, only 1 colony was positive for the dNGFR fragment. Positive colonies (lanes 1,2,5,7,8,9,13 on **Figure 3.5**) were also confirmed with DNA sequencing (see **supplementary figure 7** for sequence analysis). Colonies 2, 2 and 4 (lanes 2,7 and 13) of pCW57-dNGFR, pCW57-dNGFR-P2A-Vpr and pCW57-dNGFR-P2A-Vpu, respectively, were eventually selected as final plasmids for lentivirus production.

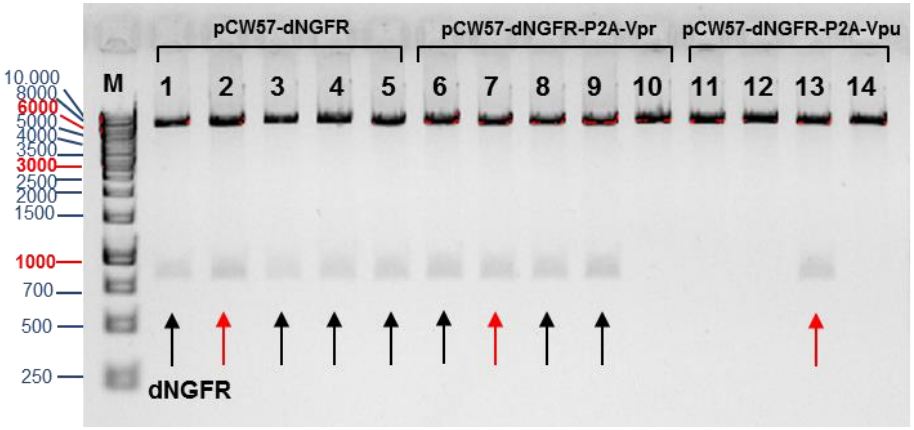


Figure 3.5| dNGFR cloning: analysis of restriction digest products on 1.5% agarose gel electrophoresis. Restriction digest with *NheI* and *AgeI* on selected colonies. Lanes 1-5: colonies with pCW57-dNGFR, lanes 6-10: colonies with pCW57-dNGFR-P2A-Vpr and lanes 11-14: colonies with pCW57-dNGFR-P2A-Vpu. Lanes 1-9 and lane 13 show a fragment of 831 bp (dNGFR). Lane M, 1 kb DNA ladder. Red arrows indicate selected colonies for later experiments.

3.2. Puromycin viability assay results

To estimate the transducing capacity of the produced lentiviral vectors, we assessed the viability of SupT1 and Jurkat cells upon transduction with serially diluted lentivirus. Cells that have been successfully transduced, will obtain the puromycin resistance gene and will survive in the presence of puromycin. When the wells of the lowest dilutions (or non-diluted) do not survive, we can state that the produced lentivirus has a low transduction capacity in that cell line. In contrast, when survival is observed in almost all wells, even in the highest dilutions, we can assume that the virus has a high transduction efficiency.

Both SupT1 and Jurkat cells were transduced with serially diluted lentiviral supernatant (1:10,1:50,1:250,1:1250,1:6250,1:31250,1:156250 and 1:781250) and selected with puromycin. After 9 days of selection, cell death was observed under the microscope and manually recorded for each well (**Figure 3.6**). These observations were then used to calculate the percentage (%) of transduced wells (wells with survival/total number of wells). Control wells with non-transduced Jurkat and SupT1 cells were also included to check for puromycin toxicity. We expected that the puromycin would kill non-transduced cells. Indeed, these control wells did show cell death when treated with puromycin (data not shown). As a positive control for cell survival, we included also for each dilutions some wells with transduced Jurkat and SupT1

cells where puromycin was not added. Since we used two cell lines and three different lentiviral viruses, we expected to obtain different observations for cell survival.

However, for all lentiviruses, in both transduced cell lines, the same observations for cell death were obtained. Cell survival was high in the lowest dilutions, but starting from dilution 1:1250, all cells showed signs of cell death (**Table 2**).

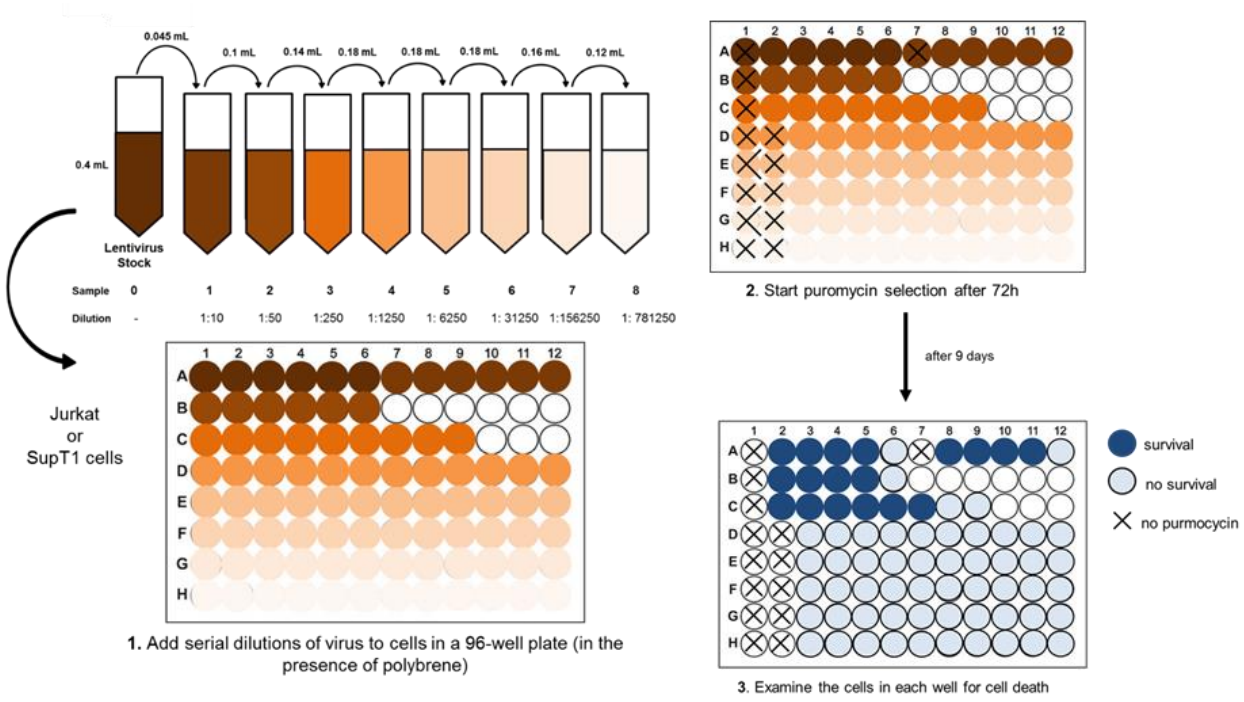


Figure 3.6 | Puromycin viability assay for pCW57-dNGFR-P2A-Vpr, pCW57-dNGFR-P2A-Vpu and pCW57-dNGFR lentiviruses. A diagram illustrating the steps involved in the viability assay in both Jurkat and SupT1 cell lines. The virus dilution samples were added in the following wells: A1-A7: virus stock, A8-A12: 1:10 dilution, B1-B6: 1:50 dilution, C1-C9: 1:250 dilution, row D: dilution 1:1250 dilution, row E: dilution 1:6250, row F: dilution 1:31250, row G:dilution 1:156250 and row H: dilution 1:781250. X indicates control wells with transduced cells in the absence of puromycin, whereas wells B6-B12 and C10-C12 were left blank. The number of wells with survival /no survival were recorded after 9 days of selection with puromycin.

Table 2 | The percentage (%) of transduced wells per dilution after 9 days of puromycin selection.

Dilution	# transduced wells *	% transduced wells
0	4 of 5	80
1/10	4 of 5	80
1/50	4 of 5	80
1/250	6 of 8	75
1/1250	0 of 10	0
1/6250	0 of 10	0
1/31250	0 of 10	0
1/156250	0 of 10	0
1/781250	0 of 10	0

*transduced wells = wells with transduced cells that did not show cell death

In order to obtain stable transduced Jurkat and SupT1 cell lines for the doxycycline induction assay, cells of the lowest dilutions showing survival were pooled together and were further maintained in culture. The cells were then intermittently selected with puromycin to ensure that the population of cells is not taken over by non-transduced cells. The cells did grow rapidly, and they were passaged routinely upon reaching confluency.

3.3. Physical titers of the produced lentiviruses

Physical titers are a measure for the total number of viral particles in a sample and are usually based on the quantification of a viral protein (*e.g.* HIV p24 protein, retroviral RT activity) or viral nucleic acid with molecular assays such as ELISA and quantitative real-time PCR [53]. We therefore assessed the physical titer by using the SYBR Green I-qPCR-based Product-Enhanced Reverse Transcriptase (SG-PERT) assay to measure the RT activity in all produced lentiviral supernatants.

After production of lentiviral particles by co-transfection of 293T cells with Tet-On inducible lentiviral construct of interest and envelope and packaging plasmids, cell-free supernatant was harvested 48h post-transfection. The quality of the produced lentiviral supernatants was then quantified using the SG-PERT assay. The determined RT activity titers (and equivalent p24 titers) for pCW57-dNGFR-P2A-Vpr, pCW57-dNGFR-P2A-Vpu and pCW57-dNGFR are indicated in **Table 3.**

Table 3 | RT activity and corresponding HIV-1 p24 protein concentration in lentiviral supernatant.

Lentivirus	mU RT /mL	ng p24 /mL
pCW57-dNGFR-P2A-Vpr	5928.21	1067.08
pCW57-dNGFR- P2A-Vpu	4647.04	836.47
pCW57-dNGFR	4706.98	847.26

3.4. Flow cytometry analysis of dNGFR expressing control cell lines

Our final goal was to determine whether the produced inducible lentiviral vectors indeed express Vpr/Vpu in a doxycycline-dependent manner (**see 2.7**). Since dNGFR is simultaneously expressed with Vpr/Vpu upon induction with Dox, expression levels of Vpr and Vpu can be characterized by staining with anti-NGFR-PE and subsequent flow cytometry analysis. However, to interpret these flow cytometry data in a correct manner, positive controls are needed. We therefore assessed the expression of NGFR (%) in retroviral/lentiviral transduced Jurkat and SupT1 cell lines by anti-NGFR-PE staining and flow cytometry (**Figure 3.7**). For each control cell line, 5 wells (samples) with cells were measured for NGFR expression. The average $NGFR^+$ expression of these 5 wells can then be used as control sample in further experiments.

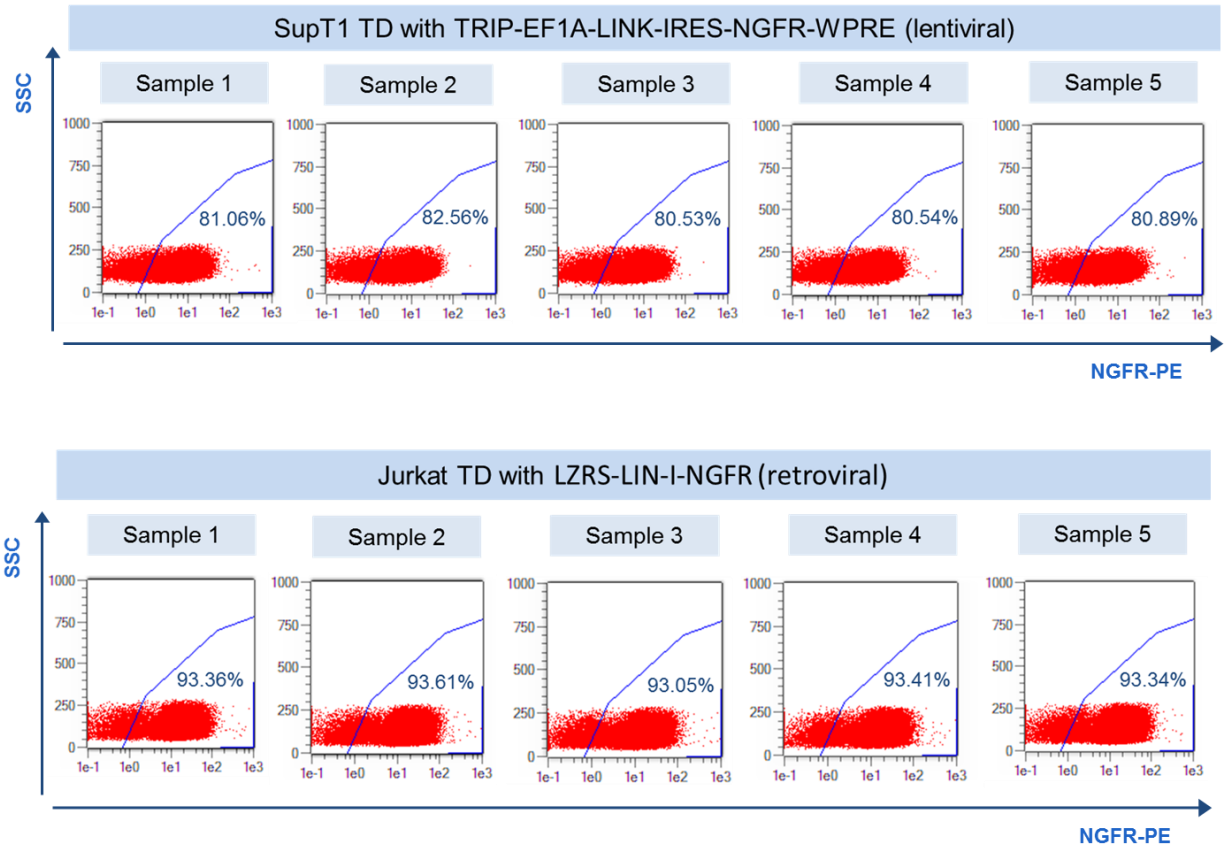


Figure 3.7 | Flow cytometric data of control samples. dNGFR expression in transduced SupT1 (TRIP-EF1A-LINK-IRES-NGFR-WPR) and Jurkat (LZRS-LIN-I-NGFR) cells were measured in 5 samples by flow cytometry after staining with PE-labelled antibody. Cells were first gated on cells of interest (lymphocytes) and viable cells. Additional gating was performed on NGFR$^+$ cells. Dot plots show side scatter (SSC) versus NGFR-PE. Numbers indicate percentage of NGFR+ cells among viable cells.

4. DISCUSSION AND CONCLUSION

cGAS-mediated IFN-response upon HIV-1 infection of primary activated $CD4^+$ T cells has recently been demonstrated by our research group [39]. This type of sensing seems to only occur upon integration, and newly expressed Vpr and Vpu were shown to respectively potentiate and counteract this IFN-response. It has been suggested that the ability of Vpr to potentiate the post-integration IFN-I response might be a result from its ability to interact with host factors and other processes which have a favorable outcome for the virus (*e.g*. NF-κB activation, G2 arrest) [39]. Similar to Vpr, Vpu might interfere with the cGAS-mediated NF-κB activation to counteract the IFN-I response. To fully understand the mechanism by which Vpr and Vpu exert these effects, Vpr and Vpu must be introduced in $CD4^+$ T cells in the absence of HIV-1. However, since constitutive expression of Vpr results in cytotoxicity, an inducible expression system is required to control its expression. The primary objective of the work described in this thesis was therefore to develop an inducible lentiviral expression system for Vpr, as well as for Vpu.

We wanted to use an all-in-one Tet-On inducible lentiviral expression system which contains all inducible expression components and allows co-expression of the gene of interest (Vpr/Vpu) and surface marker dNGFR. The advantage of this system is that it shows rapid gene expression kinetics as compared to the Tet-Off system and it allows activation of gene expression by the addition - instead of removal - of doxycycline [53]. In addition, we also wanted to have an antibiotic resistance gene to assess transduction efficiency. To achieve all this, we used the pCW57-MCS1-P2A-MCS2 plasmid, which consists of all elements required for Tet-On inducible expression. It has two MCS separated by the P2A sequence for co-expression, as well as the puromycin resistant gene. We cloned Vpr or Vpu in the MCS downstream of the P2A sequence, whereas the surface marker dNGFR was cloned in the MCS upstream of P2A. Three plasmid constructs were successfully produced: pCW57-dNGFR (control vector), pCW57-dNGFR-P2A-Vpr and pCW57-dNGFR-P2A-Vpu.

The aforementioned 2A peptides are small in size and, compared to other strategies for multi-gene co-expression such as internal ribosome entry sites (IRES), they lead to relatively high levels of expression of downstream protein. As we wanted to use dNGFR as a measure for Vpr/Vpu expression, equivalent levels of expression are required. We therefore used a P2A containing inducible expression system to allow expression of Vpr/Vpu and dNGFR in equimolar levels. By placing a P2A sequence between dNGFR and Vpr/Vpu, a polyprotein is first formed which is then “cleaved” into two individual proteins. This “cleavage” mechanism is achieved through ribosomal skipping of the peptide bond between glycine and proline in the C-terminal side of the P2A peptide, which results in two “cleaved” proteins [54]. Some studies have reported “cleavage” efficiencies of different 2A peptides in a bicistronic setting, and both P2A and T2A have been shown to have very high cleavage efficiency compared to other 2A peptides [55-57].

Although the 2A system opens opportunities for larger genes with comparable expression levels, it may involve extra optimization. The 2A-mediated skipping event might fail and result in undesired outcomes [58]. It is possible that (1) there is successful skipping but ribosomes fall-off and discontinued translation, which results in only the protein upstream of 2A, or (2) there can be unsuccessful skipping and continued translation resulting in a fusion protein. In terms of protein expression, Liu *et al.* aimed to characterize the effect of gene position on protein expression level in polycistronic 2A constructs and to determine whether different 2A systems affect the level of expressed proteins [58]. By comparing Green Fluorescent Protein (GFP) intensity among the different 2A bicistronic constructs in different cell types, they could show that the protein expression at the second gene position was the greatest in the T2A constructs and slightly less in the P2A construct.

They also determined the efficiency of 2A-mediated “cleavage” and the protein expression level of GFP at the second gene position. Quantification of flow cytometry and western blotting results revealed that the ratio of uncleaved to cleaved protein at the first gene position is ~0.1 in both P2A and T2A constructs. Additionally, the ratio of expression of cleaved protein at the second gene position (GFP) to cleaved protein at the first gene position was equivalently high in P2A and T2A constructs. Their flow cytometry and western blotting results do also suggest that, in a bicistronic setting, the ratio of occurrence of ribosome skipping and recommencement of translation: ribosome skipping and fall-off and discontinued translation: ribosome read-through is about 30% to 60% to 10%, respectively. This indicates that the ‘cleavage’ product N-terminal of 2A is produced in excess when compared to the product C-terminal of 2A. Overall, their findings suggest that the protein expression at the second gene position decreases by ~70% when compared to the first gene position, and that T2A leads to the highest level of protein expression at the second gene position compared to other 2As.

Contrarily, other studies have demonstrated that 2A containing vectors can mediate efficient and comparable expression of the proteins [56,59,60]. Therefore, the conflicting data that has been reported on 2A peptides cleavage efficiencies and gene expression levels might result from differences in experimental conditions (*e.g.* cell line or *in vivo* model that is used). Altogether, it will be necessary to assess the efficiency of 2A-mediated “cleavage” of downstream Vpr and Vpu, and to further evaluate the ratio of protein expression between dNGFR and Vpu/Vpr, as different levels of protein expression will lead to misinterpretation of the expression of Vpr/Vpu.

It should also be noted that the protein upstream of the P2A (dNGFR) is still attached to the complete P2A peptide except for the C-terminal proline, and the protein downstream of the 2A (Vpr or Vpu) is attached to a proline at the N-terminus. These remaining P2A residues might cause adverse effects on the upstream and downstream proteins. For example, N-terminal acetylation is believed to affect protein stability, degradation and function. By placing a proline at the first (or second) position of a protein, post-translational N-terminal acetylation is prevented [61]. However, in the context of Vpr, unpublished data from our research group showed that N-terminus conservation seems to be important for sufficient Vpr protein levels in infected cells and virions, but Nt-acetylation did not really affect its function or degradation [62]. In this regard, additional experiments will be required to further characterize the impact of these residues on the expressed proteins, with a focus on the functionality of Vpr and Vpu. For example, Vpr/Vpu expression levels can be compared with expression levels of other vectors which constitutively express Vpr/Vpu. If then Vpr/Vpu protein levels seem to be affected, another approach must be used. Alternatively, the P2A-residues can also be removed. Fang *et al.* showed that the N-terminal-P2A residues can be removed through insertion of a furin cleavage site in front of the P2A sequence, whereas the C-terminal proline was removed by inserting a signal peptide sequence at the N-terminus of the second protein [63]. However, this approach was shown for secreted proteins, since furin is mainly localized within the Golgi apparatus.

Furthermore, for all three produced lentiviruses, we assessed the RT activity as a measure for the physical viral titer by using the SG-PERT assay, and high RT values were obtained for all produced lentiviruses (5928.21, 4647.04 and 4706.98 mU/RT for pCW57-NGFR-P2A-Vpr, pCW57-NGFR-P2A-Vpu and pCW57-NGFR-resepctively). We also determined the transduction capacity of the lentiviruses after transducing cells with a limiting dilution series of the lentiviruses. We used SupT1 and Jurkat cells, which are both T lymphocyte cell lines and translate thus well to the target cell line (primary human $CD4^+$ T cells). As transduction and gene expression are highly dependent on the cell type, and as we generated three different lentiviruses, we expected to obtain different observations for cell death. However, the same observations were obtained for all viruses in both cell lines. These results can be explained by the fact that determination of cell death is quite subjective, and wells were scored for cell death/survival based on the color of the medium and by means of a microscope.

Based on the obtained results, it is difficult to draw conclusions from the viability assay. Therefore, the viability assay for all three lentiviruses needs to be reevaluated. Alternatively, as these lentiviral vectors do also express an inducible identification tag (dNGFR), the transduction efficiency can also be determined in transduced cells after selection with puromycin and upon doxycycline induction. Based on the dNGFR expression, the viral titer (transducing units/ml) can then be calculated using flow cytometry. This might be a reliable alternative approach to evaluate the transduction efficiency of the produced lentiviruses.

In order to allow further evaluation of the inducible expression Vpr, Vpu and dNGFR, positively transduced cells are required. We therefore developed stable SupT1 and Jurkat cell lines expressing Vpr/Vpu and dNGFR in a doxycycline-controlled manner. Additional assays such as flow cytometry and Western Blot assays are required to evaluate dose-dependent response and ensure expression of Vpr and Vpu upon doxycycline induction.

In conclusion, we state that inducible lentiviral vectors expressing Vpr, Vpu or dNGFR were successfully constructed. Nevertheless, further evaluation of protein expression and functionality is required.

5. REFERENCE LIST

1 UNAIDS. *Global HIV & AIDS statistics — 2019 fact sheet*, <https://www.unaids.org/en/resources/fact-sheet> (2019).

2 Avert. *Global HIV and AIDS statistics*, <https://www.avert.org/global-hiv-and-aids-statistics> (2019).

3 HIV.gov. *Global Statistics*, <https://www.hiv.gov/hiv-basics/overview/data-and-trends/global-statistics> (2019).

4 Sharp, P. M. & Hahn, B. H. Origins of HIV and the AIDS pandemic. *Cold Spring Harbor perspectives in medicine* **1**, a006841, doi:10.1101/cshperspect.a006841 (2011).

5 German Advisory Committee Blood, S. A. o. P. T. b. B. Human Immunodeficiency Virus (HIV). *Transfusion medicine and hemotherapy : offizielles Organ der Deutschen Gesellschaft fur Transfusionsmedizin und Immunhamatologie* **43**, 203-222, doi:10.1159/000445852 (2016).

6 Hemelaar, J. Implications of HIV diversity for the HIV-1 pandemic. *The Journal of infection* **66**, 391-400, doi:10.1016/j.jinf.2012.10.026 (2013).

7 Lau, K. A. & Wong, J. J. Current Trends of HIV Recombination Worldwide. *Infectious disease reports* **5**, e4, doi:10.4081/idr.2013.s1.e4 (2013).

8 Clavel, F. *et al.* Isolation of a new human retrovirus from West African patients with AIDS. *Science (New York, N.Y.)* **233**, 343-346, doi:10.1126/science.2425430 (1986).

9 Campbell-Yesufu, O. T. & Gandhi, R. T. Update on human immunodeficiency virus (HIV)-2 infection. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **52**, 780-787, doi:10.1093/cid/ciq248 (2011).

10 Faria, N. R. *et al.* HIV epidemiology. The early spread and epidemic ignition of HIV-1 in human populations. *Science (New York, N.Y.)* **346**, 56-61, doi:10.1126/science.1256739 (2014).

11 Keele, B. F. *et al.* Chimpanzee reservoirs of pandemic and nonpandemic HIV-1. *Science (New York, N.Y.)* **313**, 523-526, doi:10.1126/science.1126531 (2006).

12 Frankel, A. D. & Young, J. A. T. HIV-1: Fifteen proteins and an RNA. *Annu Rev Biochem* **67**, 1-25, doi:DOI 10.1146/annurev.biochem.67.1.1 (1998).

13 HIV InSite UCSF. *Structure, Expression, and Regulation of the HIV Genome*, <https://www.hiv.lanl.gov/content/sequence/HIV/MAP/landmark.html> (2000).

14 Kirchhoff, F. Immune evasion and counteraction of restriction factors by HIV-1 and other primate lentiviruses. *Cell host & microbe* **8**, 55-67, doi:10.1016/j.chom.2010.06.004 (2010).

15 Nkeze, J., Li, L., Benko, Z., Li, G. & Zhao, R. Y. Molecular characterization of HIV-1 genome in fission yeast Schizosaccharomyces pombe. *Cell & bioscience* **5**, 47, doi:10.1186/s13578-015-0037-7 (2015).

16 Sakuragi, J. Morphogenesis of the Infectious HIV-1 Virion. *Frontiers in microbiology* **2**, 242, doi:10.3389/fmicb.2011.00242 (2011).

17 IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. *HUMAN IMMUNODEFICIENCY VIRUS-1 - Biological gents*, <https://www.ncbi.nlm.nih.gov/books/NBK304351/> (2012).

18 ViralZone Swiss Institute of Bioinformatics. *Human immunodeficiency virus 1*, <https://viralzone.expasy.org/7?outline=all_by_species> (2015).

19 Abfalterer W., A. G., Fischer W., Funkhouser B., Gaschen B.,Hraber P, Lo C., Macke J., Szinger J.J., Thurmond J., Yoon H. ,and Zhang M. *HIV Sequence Compendium 2008*, <https://www.hiv.lanl.gov/content/sequence/HIV/COMPENDIUM/2008/frontmatter.pdf> (2008).

20 Wilen, C. B., Tilton, J. C. & Doms, R. W. HIV: cell binding and entry. *Cold Spring Harbor perspectives in medicine* **2**, doi:10.1101/cshperspect.a006866 (2012).

21 Ambrose, Z. & Aiken, C. HIV-1 uncoating: connection to nuclear entry and regulation by host proteins. *Virology* **454-455**, 371-379, doi:10.1016/j.virol.2014.02.004 (2014).

22 Hughes, S. H. Reverse Transcription of Retroviruses and LTR Retrotransposons. *Microbiology spectrum* **3**, MDNA3-0027-2014, doi:10.1128/microbiolspec.MDNA3-0027-2014 (2015).

23 Freed, E. O. HIV-1 assembly, release and maturation. *Nature reviews. Microbiology* **13**, 484-496, doi:10.1038/nrmicro3490 (2015).

24 Deeks, S. G., Overbaugh, J., Phillips, A. & Buchbinder, S. HIV infection. *Nature reviews. Disease primers* **1**, 15035, doi:10.1038/nrdp.2015.35 (2015).

25 Pan, X., Baldauf, H. M., Keppler, O. T. & Fackler, O. T. Restrictions to HIV-1 replication in resting CD4+ T lymphocytes. *Cell research* **23**, 876-885, doi:10.1038/cr.2013.74 (2013).

26 Siliciano, R. F. & Greene, W. C. HIV latency. *Cold Spring Harbor perspectives in medicine* **1**, a007096, doi:10.1101/cshperspect.a007096 (2011).

27 Koppensteiner, H., Brack-Werner, R. & Schindler, M. Macrophages and their relevance in Human Immunodeficiency Virus Type I infection. *Retrovirology* **9**, 82, doi:10.1186/1742-4690-9-82 (2012).
28 Laguette, N. *et al.* SAMHD1 is the dendritic- and myeloid-cell-specific HIV-1 restriction factor counteracted by Vpx. *Nature* **474**, 654-657, doi:10.1038/nature10117 (2011).
29 Wu, L. & KewalRamani, V. N. Dendritic-cell interactions with HIV: infection and viral dissemination. *Nature reviews. Immunology* **6**, 859-868, doi:10.1038/nri1960 (2006).
30 Killian, M. S. & Levy, J. A. HIV/AIDS: 30 years of progress and future challenges. *European journal of immunology* **41**, 3401-3411, doi:10.1002/eji.201142082 (2011).
31 An, P. & Winkler, C. A. Host genes associated with HIV/AIDS: advances in gene discovery. *Trends in genetics : TIG* **26**, 119-131, doi:10.1016/j.tig.2010.01.002 (2010).
32 Levy, J. A. Virus-host interactions in HIV pathogenesis: directions for therapy. *Advances in dental research* **23**, 13-18, doi:10.1177/0022034511398874 (2011).
33 McMichael, A. J., Borrow, P., Tomaras, G. D., Goonetilleke, N. & Haynes, B. F. The immune response during acute HIV-1 infection: clues for vaccine development. *Nature reviews. Immunology* **10**, 11-23, doi:10.1038/nri2674 (2010).
34 Arts, E. J. & Hazuda, D. J. HIV-1 antiretroviral drug therapy. *Cold Spring Harbor perspectives in medicine* **2**, a007161, doi:10.1101/cshperspect.a007161 (2012).
35 Iwasaki, A. Innate immune recognition of HIV-1. *Immunity* **37**, 389-398, doi:10.1016/j.immuni.2012.08.011 (2012).
36 Hardy, G. A. *et al.* Interferon-alpha is the primary plasma type-I IFN in HIV-1 infection and correlates with immune activation and disease markers. *PloS one* **8**, e56527, doi:10.1371/journal.pone.0056527 (2013).
37 Sivro, A., Su, R. C., Plummer, F. A. & Ball, T. B. Interferon responses in HIV infection: from protection to disease. *AIDS reviews* **16**, 43-51 (2014).
38 Lepelley, A. *et al.* Innate sensing of HIV-infected cells. *PLoS pathogens* **7**, e1001284, doi:10.1371/journal.ppat.1001284 (2011).
39 Vermeire, J. *et al.* HIV Triggers a cGAS-Dependent, Vpu- and Vpr-Regulated Type I Interferon Response in CD4(+) T Cells. *Cell reports* **17**, 413-424, doi:10.1016/j.celrep.2016.09.023 (2016).
40 Yin, X. *et al.* Sensor Sensibility-HIV-1 and the Innate Immune Response. *Cells* **9**, doi:10.3390/cells9010254 (2020).
41 Gao, D. *et al.* Cyclic GMP-AMP synthase is an innate immune sensor of HIV and other retroviruses. *Science (New York, N.Y.)* **341**, 903-906, doi:10.1126/science.1240933 (2013).
42 Cai, X., Chiu, Y. H. & Chen, Z. J. The cGAS-cGAMP-STING pathway of cytosolic DNA sensing and signaling. *Molecular cell* **54**, 289-296, doi:10.1016/j.molcel.2014.03.040 (2014).
43 Silvin, A. & Manel, N. Innate immune sensing of HIV infection. *Current opinion in immunology* **32**, 54-60, doi:10.1016/j.coi.2014.12.003 (2015).
44 Krapp, C., Jonsson, K. & Jakobsen, M. R. STING dependent sensing - Does HIV actually care? *Cytokine & growth factor reviews* **40**, 68-76, doi:10.1016/j.cytogfr.2018.03.002 (2018).
45 Harman, A. N. *et al.* HIV Blocks Interferon Induction in Human Dendritic Cells and Macrophages by Dysregulation of TBK1. *Journal of virology* **89**, 6575-6584, doi:10.1128/JVI.00889-15 (2015).
46 Zahoor, M. A., Xue, G., Sato, H. & Aida, Y. Genome-wide transcriptional profiling reveals that HIV-1 Vpr differentially regulates interferon-stimulated genes in human monocyte-derived dendritic cells. *Virus research* **208**, 156-163, doi:10.1016/j.virusres.2015.06.017 (2015).
47 Zahoor, M. A. *et al.* HIV-1 Vpr induces interferon-stimulated genes in human monocyte-derived macrophages. *PloS one* **9**, e106418, doi:10.1371/journal.pone.0106418 (2014).
48 Liu, R. *et al.* HIV-1 Vpr stimulates NF-kappaB and AP-1 signaling by activating TAK1. *Retrovirology* **11**, 45, doi:10.1186/1742-4690-11-45 (2014).
49 Barger, C. J., Branick, C., Chee, L. & Karpf, A. R. Pan-Cancer Analyses Reveal Genomic Features of FOXM1 Overexpression in Cancer. *Cancers (Basel)* **11**, doi:10.3390/cancers11020251 (2019).
50 Thermo Scientific. *Phusion™ High-Fidelity DNA Polymerase (2 U/µL)*, <https://www.thermofisher.com/order/catalog/product/F530L?SID=fr-phusion-1#/F530L?SID=fr-phusion-1> (2019).
51 Promega. *Wizard® SV Gel and PCR Clean-Up System protocol*, <https://be.promega.com/resources/protocols/technical-bulletins/101/wizard-sv-gel-and-pcr-cleanup-system-protocol/> (2019).
52 Vermeire, J. *et al.* Quantification of reverse transcriptase activity by real-time PCR as a fast and accurate method for titration of HIV, lenti- and retroviral vectors. *PLoS One* **7**, e50859, doi:10.1371/journal.pone.0050859 (2012).

53 Meghan Rego. *Tips for Titering Your Lentiviral Preps*, <https://blog.addgene.org/tips-for-titering-your-lentiviral-preps> (2016).
54 Donnelly, M. *e. a.* Analysis of the aphthovirus 2A/2B polyprotein 'cleavage' mechanism indicates not a proteolytic reaction, but a novel translational effect: a putative ribosomal 'skip'. *J Gen Virol.* **82**, 1013-1025, doi:10.1099/0022-1317-82-5-1013 (2001).
55 Kim, J. H. *et al.* High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice. *PLoS One* **6**, e18556, doi:10.1371/journal.pone.0018556 (2011).
56 Szymczak, A. *et al.* Correction of multi-gene deficiency in vivo using a single 'self-cleaving' 2A peptide–based retroviral vector. *Nat Biotechnol* **22**, 589–594, doi:https://doi.org/10.1038/nbt957 (2004).
57 Donnelly, M. *et al.* The 'cleavage' activities of foot-and-mouth disease virus 2A site-directed mutants and naturally occurring '2A-like' sequences. *J Gen Virol.* **82**, 1027-1041, doi:10.1099/0022-1317-82-5-1027 (2001).
58 Liu, Z. *et al.* Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector. *Sci Rep* **7**, 2193, doi:10.1038/s41598-017-02460-2 (2017).
59 Ibrahimi, A. *et al.* Highly efficient multicistronic lentiviral vectors with peptide 2A sequences. *Hum Gene Ther* **20**, 845-860, doi:10.1089/hum.2008.188 (2009).
60 Richard, W. D., Adam, J. R., Gregory T. M. and Barry G. Expression of Multiple Transgenes from a Single Construct Using Viral 2A Peptides in Drosophila. *PLoS One* **9**, e100637, doi:10.1371/journal.pone.0100637 (2014).
61 Ree, R., Varland, S. & Arnesen, T. Spotlight on protein N-terminal acetylation. *Exp Mol Med* **50**, 1-13, doi:10.1038/s12276-018-0116-z (2018).
62 Van Nuffel *et al.* Vpr content of HIV-1 virions determines infection of resting peripheral blood CD4+ lymphocytes. *Retrovirology* **10**, doi:https://doi.org/10.1186/1742-4690-10-S1-P92 (2013).
63 Fang, J. *et al.* Stable antibody expression at therapeutic levels using the 2A peptide. *Nat Biotechnol.* **23**, 584–590, doi:https://doi.org/10.1038/nbt1087 (2005).

6. POSTER

GHENT UNIVERSITY

The mechanism of HIV-1 sensing in CD4⁺ T cells: the role of newly expressed Vpr and Vpu

Nadia Elhemdaoui, Evelien Naessens, Femke Van Gaever, and Bruno Verhasselt

Department of Diagnostic Sciences, Ghent University, 9000 Ghent, Belgium

Background

Human immunodeficiency virus type 1 (HIV-1) infection of innate immune cells activates downstream pathways with an antiviral type I interferon (IFN-I) response as a result. This HIV-1 sensing occurs through the activation of many pathogen PRRs, including the cytosolic DNA sensor cGAS.

cGAS signaling has been demonstrated in HIV-1 infected activated CD4⁺T cells by Vermeire *et al.* (2016). Moreover, HIV-1 provirus integration is required to trigger a cGAS-mediated IFN-I response in CD4⁺T cells, and viral proteins Vpr and Vpu seem to stimulate and counteract this IFN-I response respectively. However, how these newly expressed Vpr and Vpu proteins result in a distinct IFN-I response remains unclear.

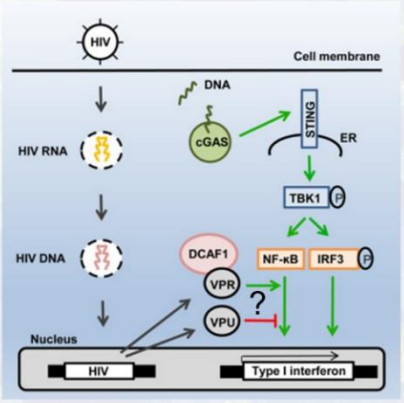
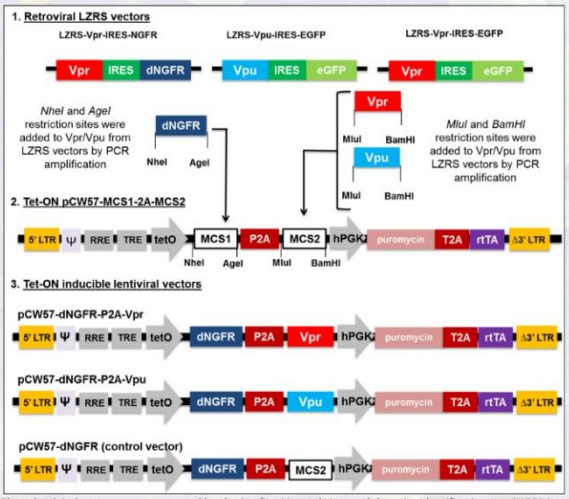


Fig. 1 | Schematic representation of cGAS dependent, Vpu- and Vpr- regulated type I interferon response in CD4⁺T cells.
Vermeire J, Roesch F, Sauter D, Rua R, Hotter D, Van Nuffel A, Vanderstraeten H, Naessens E, Iannucci V, Landi A, *et al.* (2016) HIV Triggers a cGAS-Dependent, Vpu- and Vpr-Regulated Type I Interferon Response in CD4(+) T Cells. *Cell Rep* **17**: 413-424

Methods

To better understand the role of Vpr and Vpu in the cGAS dependent immune sensing of HIV-1 in CD4⁺T cells, we constructed Tet-ON inducible lentiviral constructs in which both Vpr and Vpu are regulated, in a dose and time dependent manner by tetracycline (Tet) / doxycycline (DOX).



Three lentiviral vectors were generated by cloning first Vpr and Vpu, and then the identification tag dNGFR into the Tet-ON plasmid pCW57-MCS1-2A-MCS2 (Addgene 71782)

Contact
Nadia.Elhemdaoui@Ugent.be
Universiteit Gent
@ugent
Ghent University

Results

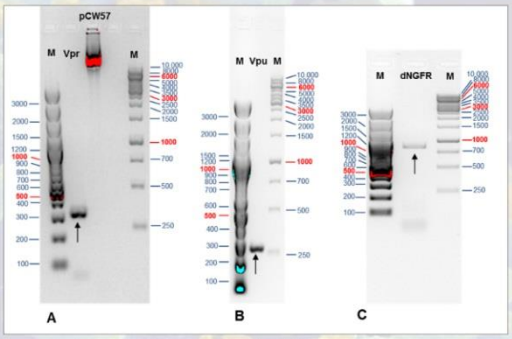


Fig. 3 | Analysis of PCR products (insert) on 1.5% agarose gel electrophoresis
Restriction sites were added to Vpr, Vpu and dNGFR via PCR amplification.
The PCR amplicons were purified afterwards and analyzed on 1.5% agarose gel. **(A)** Vpr (291 bp) and Tet-ON plasmid pCW57 (7709 bp), **(B)** Vpu (246 bp) and **(C)** dNGFR (830 bp) with lane M: DNA ladder.

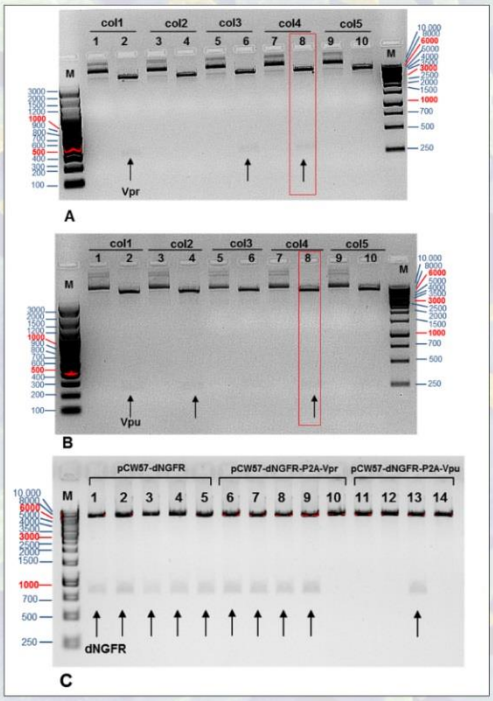


Fig. 4 | Restriction digestion of cloned plasmids and analysis on 1.5% agarose gel electrophoresis. Upon transformation of competent *E. coli* with the cloned plasmids, five colonies were picked for restriction enzyme digestion and analysis on 1.5% agarose gel. **(A)** Restriction digest of pCW57-Vpr plasmids with *MluI* and *BamHI* . Lanes 1,3, 5, 7 and 10 are uncut controls for each colony. Only lanes 2, 6 and 8 show a fragment of 291 bp (Vpr). **(B)** Restriction digest of pCW57-P2A-Vpr plasmids with *MluI* and *BamHI*. Lanes 1,3, 5, 7 and 10 are uncut controls for each colony. Only lanes 2, 4 and 8 show a fragment of 246 bp (Vpu). **(C)** dNGFR was cloned into pCW57-Vpr and -Vpu plasmids from colony 4 (red squares) and restriction digest of cloned pCW57 plasmids with *NheI* and *AgeI* was performed. Lanes 1-5: colonies with pCW57-dNGFR , lanes 6-10: pCW57-dNGFR-P2A-Vpr and lane 11-14 : pCW57-dNGFR-P2A-Vpu. Lanes 1-9 and lane 13 show a fragment of 831 bp (dNGFR). Lane M, DNA ladder.

Conclusion and future directions

- These data show that Vpr and Vpu, together with the identification marker dNGFR were succesfully cloned within the Tet-ON inducible expression vectors.
- After lentivirus production in 293T cells, the constructed Tet-ON lentiviral vectors will be validated in control cell lines (Jurkat and SupT1 cells) and CD4⁺T cells.
- Quantification of the IFN-I in HIV-1 infected CD4⁺ T cells expressing Vpu and Vpr upon lentiviral transduction will be subject of further investigation.

7. ADDENDUM

Supplementary Information

Supplementary figure 1 | Overview of the viral proteins of HIV-1 and their function.

Class	Gene	Size (kDA)	Protein	Function(s)	Localization	Expression
Viral structural proteins and enzymes	*gag*	Pr55Gag	Gag polyprotein	Precursor of inner structural proteins	Virion	Late
		p17	Matrix protein (MA)	Myristoylated protein that forms the inner membrane layer to provide stability of the viral particle	Virion	Late
		p24	Capsid protein (CA)	Formation of the cone-shaped capsid to protect the viral genome	Virion	Late
		p7	Nucleocaspid-protein (NC)	Formation of nucleocapsid by non-specifically binding to viral RNA	Virion	Late
		p6	Nucleocaspid-protein	Involved in viral particle release (budding) from the host cell, as well as incorporation of the Vpr protein in the viral particle	Virion	Late
	pol	Pr160 GagPol	Gag-pol polyprotein	Precursor of the viral enzymes	Virion	Late
		p11	Protease (PR)	Proteolytic cleavage of Gag and Gag-Pol precursor proteins to release structural proteins and viral enzymes (maturation)	Virion	Late
		p51, p66	Reverse transcriptase (RT)	Transcription of HIV ssRNA in proviral ssDNA and dsDNA synthesis (polymerase activity)	Virion	Late
		p15	RNase H	Degradation of viral RNA in the viral RNA/DNA replication complex	Virion	Late
		p32	Integrase (IN)	Integration of proviral DNA into the host genome	Virion	Late
	env	PrGp160	gp160 precursor	Precursor of the viral envelope proteins that is cleaved by cellular proteases	Plasma membrane, virion envelope	Late

		gp120	Surface (SU) glycoprotein	Attachment of virus to the target cell	Plasma membrane, virion envelope	Late
		gp41	Transmembrane (TM) glycoprotein	Bound to gp120, fusion of viral envelope and host cell membrane	Plasma membrane, virion envelope	Late
Viral regulatory proteins	*tat*	p14	Transactivator protein (Tat)	Transcription factor that binds to the TAR RNA element to activate transcription initiation and elongation from the LTR promoter	Nucleus	Early
	rev	p19	RNA splicing regulator (Rev)	Phosphoprotein that binds to RRE for promoting the nuclear export of non-spliced and partially spliced mRNA	Nucleus and cytoplasm	Early
Viral accessory proteins	*nef*	p27	Negative regulating factor (Nef)	Myristoylated protein that enhances viral replication and infectivity, and protection of infected cells from immune responses (*e.g.* by downregulating CD4 and MHC-I cell surface markers).	Plasma membrane, cytoplasm	Early
	vif	p23	Viral infectivity factor (Vif)	Preventing the action of cellular APOBEC-3 restriction factors present in the host (promotes virus infectivity)	Cytoplasm and virion	Late
	vpr	p14	Viral protein R (Vpr)	Important role in viral replication by promoting nuclear import of PIC, and has also effect on the cell cycle (G2 arrest)	Virion nucleus	Late
	vpu	p16	Viral protein U (Vpu)	Phosphoprotein that enhances virion release from the plasma membrane of infected cells and is involved in CD4 degradation	Integral membrane protein	Late

Abbreviations: RRE, HIV Rev response element; TAR, trans-activation response element; APOBEC-3G, apolipoprotein B mRNA-editing enzyme catalytic polypeptide 3; PIC, Pre-integration complex. Adapted from German Advisory Committee Blood SAoPTbB 2016 and HIV Sequence Compendium 2008 [5,19].

Supplementary figure 2 | Vpr, Vpu and dNGFR sequences retrieved from retroviral template plasmids.

WT Vpr sequence (291 bp) from LZRS-Vpr-IRES-NGFR
5'ATGGAACAAGCCCCAGAAGACCAAGGGCCACAGAGGGAGCCATACAATGAATGGACACTAGAGCTTTT AGAGGAACTTAAGAGTGAAGCTGTTAGACATTTTCCTAGGATATGGCTCCATAACTTAGGACAACATATCT ATGAAACTTACGGGGATACTTGGGCAGGAGTGGAAGCCATAATAAGAATTCTGCAACAACTGCTGTTTATC CATTTCAGAATTGGGTGTCGACATAGCAGAATAGGCGTTACTCGACAGAGGAGAGCAAGAAATGGAGCCA GTAGATCC**TAG** **3'**
WT Vpu sequence (246 bp) from LZRS-Vpu-IRES eGFP
5'ATGCAACCTATAATAGTAGCAATAGTAGCATTAGTAGTAGCAATAATAATAGCAATAGTTGTGTGGTCCAT AGTAATCATAGAATATAGGAAAATATTAAGACAAAGAAAAATAGACAGGTTAATTGATAGACTAATAGAAAG AGCAGAAGACAGTGGCAATGAGAGTGAAGGAGAAGTATCAGCACTTGTGGAGATGGGGGTGGAAATGGG GCACCATGCTCCTTGGGATATTGATGATCTG**TAG** **3'**
dNGFR sequence (830 bp) from LZRS-Vpr-IRES-NGFR
5'ATGGGGGCAGGTGCCACCGGCCGCGCCATGGACGGGCCGCGCCTGCTGCTGTTGCTGCTTCTGGGG GTGTCCCTTGGAGGTGCCAAGGAGGCATGCCCCACAGGCCTGTACACACACAGCGGTGAGTGCTGCAAA GCCTGCAACCTGGGCGAGGGTGTGGCCCAGCCTTGTGGAGCCAACCAGACCGTGTGTGAGCCCTGCCT GGACAGCGTGACGTTCTCCGACGTGGTGAGCGCGACCGAGCCGTGCAAGCCGTGCACCGAGTGCGTGG GGCTCCAGAGCATGTCGGCGCCGTGCGTGGAGGCCGACGACGCCGTGTGCCGCTGCGCCTACGGCTAC TACCAGGATGAGACGACTGGGCGCTGCGAGGCGTGCCGCGTGTGCGAGGCGGGCTCGGGCCTCGTGTT CTCCTGCCAGGACAAGCAGAACACCGTGTGCGAGGAGTGCCCCGACGGCACGTATTCCGACGAGGCCA ACCACGTGGACCCGTGCCTGCCCTGCACCGTGTGCGAGGACACCGAGCGCCAGCTCCGCGAGTGCACA CGCTGGGCCGACGCCGAGTGCGAGGAGATCCCTGGCCGTTGGATTACACGGTCCACACCCCCAGAGGG CTCGGACAGCACAGCCCCCAGCACCCAGGAGCCTGAGGCACCTCCAGAACAAGACCTCATAGCCAGCAC GGTGGCAGGTGTGGTGACCACAGTGATGGGCAGCTCCCAGCCCGTGGTGACCCGAGGCACCACCGACA ACCTCATCCCTGTCTATTGCTCCATCCTGGCTGCTGTGGTTGTGGGCCTTGTGGCCTACATAGCCTTCAA GAGGTGGAACAG **3'**

Start (ATG) and stop (TAG) codons are marked in red

Supplementary figure 3 | Reagentia for Vpr, Vpu and dNGFR amplification with PCR.

Component	Volume (µL)
Nuclease free H_2O	Add to 50 µL
5X phusion HF buffer (Thermoscientific)	10 µL
50 mM $MgCl_2$ solution (Thermoscientific)	1 µL
10 mM dNTPs (Amersham)	1 µL
Forward primer* (ordered at IDT)	2.5 µL
Reverse primer* (ordered at IDT)	2.5 µL
Template DNA (in-house)**	1µL
Phusion DNA polymerase (Thermoscientific)	0.5 µL
Total Volume	**50 µL**

**Primers were first diluted to 10µM working stock from storage stocks at 100uM. ** Template DNA was first diluted to 20 ng/µL working stock from storage stock. All dilutions were performed in Nuclease free H_2O.*

Supplementary figure 4 | Primers for DNA sequencing to verify cloning of Vpr, Vpu and dNGFR in pCW57-MCS1-2A-MCS2 backbone plasmid. *

Primers for pCW57 -Vpr and pCW57-Vpu constructs			
Primer ID	**Gene**	**Direction**	**Primer sequence (5' – 3')**
LNCX-Fwd	*CMV promoter***	forward	AGCTCGTTTAGTGAACCGTC
hPGK-Rev	*hPGK promoter*	reverse	GAGGGTACTAGTGAGACG
Primers for pCW57 -dNGFR- Vpr, pCW57-dNGFR- Vpu and pCW57-dNGFR constructs			
Primer ID	**Gene**	**Direction**	**Primer sequence (5' – 3')**
LNCX-Fwd	*CMV promoter **	forward	AGCTCGTTTAGTGAACCGTC
Vpr-Rev	*vpr*	reverse	AAAGTTGGATCCCTAGGATCTACTGGCTCCATTTC
Vpu-Rev	*vpu*	reverse	AAAGTTGGATCCCTACAGATCATCAATATCCCAAGGAG
AgeI-dNGFR-Rev	*dngfr*	reverse	AAAGTTACCGGTGCTGTTCCACCTCTTGAAGGCTATGTAGG

** All primers were designed in-house with SnapGene software tool, sequencing of the constructs was performed by Eurofins Generomics, Germany. **The Human CMV promoter is last part of the inducible tight TRE promoter.*

Supplementary figure 5 | PCR program for dNGFR amplification with tailed primers.

PCR step	Temperature	Time	Cycles
Initial denaturation	98 °C	30 s	1
Denaturation	98 °C	10 s	35
Annealing (dependent on length primers)	69 °C	30 s	
Elongation	72 °C	25 s	
Final extension	72 °C	2 min	1
Hold	12 °C	∞	

Differences with PCR program for Vpu and Vpr amplification are indicated in red.
Abbreviations: s= seconds, min= minutes

Supplementary figure 6 |Sequence analysis of pCW57-Vpr and pCW57- Vpu by using NCBI blast.

With: RE sites *P2A sequence* VPU/VPR sequence

pCW57-VPR colony 4

LNCX-Fwd

GGCGGAATGGCTAGCGAATTCGTCGACACCGGGTTCCGGA***GCCACGAACTTCTCTCTGTTAAAGCAAGCA GGAGACGTGGAAGAAAACCCCGGTCCT***AGGCTGCAGACGCGTATGGAACAAGCCCCAGAAGACCAAGGG
CCACAGAGGGAGCCATACAATGAATGGACACTAGAGCTTTTAGAGGAACTTAAGAGTGAAGCTGTTAGACAT
TTTCCTAGGATATGGCTCCATAACTTAGGACAACATATCTATGAAACTTACGGGGATACTTGGGCAGGAGTG
GAAGCCATAATAAGAATTCTGCAACAACTGCTGTTTATCCATTTCAGAATTGGGTGTCGACATAGCAGAATAG
GCGTTACTCGACAGAGGAGAGCAAGAAATGGAGCCAGTAGATCCTAGGGATCCGGGGTTGGGGTTGCGCC
TTTTCCAAGGCAGCCCTGGGTTTGCGCAGGGACGCGGCTGCTCTGGGCGTGGTTCCGGGAAACGCAGCG
GCGCCGACCCTGGGTCTCGCACATTCTTCACGTCCGTTCGCAGCGTCACCCGGATCTTCGCCGCTACCCTT
GTGGGCCCCCCGGCGACGCTTCCTGCTCCGCCCCTAAGTCGGGAAGGTTCCTTGCGGTTCGCGGCGTGC
CGGACGTGACAAACGGAAGCCGCACGTCTCACTAGTACCCTCGCAGACGGACAGCGCCAGGGAGCAATGG
CAGCGCGCCGACCGCGATGGGCTGTGGCCAATAGCGGCTGCTCAGCAGGGCGCGCCGAGAGCAGCGGC
CGGGAAGGGGCGGTGCGGGAGGCGGGGTGTGGGGCGGTAGTGTGGGCCCTGTTCCTGCCCGCGCGGTG
TTCCGCATTCTGCAAGCCTCCGGAGCGCACGTCGGCAGTCGGCTCCCTCGTTGACCGAATCACCGACCTCT
CTCCCCAGCAATTCACCATGACCGAGTACAAGCCCACGGTGCGCCTCGCCACCCGCGACGACGTCCCCAG

GGCCGTACGCACCCTCGCCGCCGCGTTCGCCGACTACCCCGCCACGCGCCACACCGTCGATCCGGACCG
CCACATCGAGCGGGTCACCGAGCTGCAAGAACTCTTCCTCACGCGCGTCGGGCTCGACATCGGCAAGGTG
TGGGTCGCGGACGACGGCGCCGCGGTGGCGGTCTGGACCACGCCGGAGAACGTCAAACCGGGGGCGGT
GTTCGCCGAGATCGGCCCGCGCATGGCCCAAATTGAACGGTTCCCGGCTGGCCGCCCAGCAACAGATGGA
AAGGCCTCCTGGCGCCCCACCGGCCCAAGGAACCCGCGGGTTTCCGGGCCACCGTCGGGGTTTCGCCCC
AACCCCAGGGGAAAGGGTTTGGGCCACCCCCCCCCTTCCTCCCCCGAAAAGGAAGGCC

hPGK-Rev
GGGCTTGCGTCGGCCGCCGCGAACCGCAGGGAACCTTCCCGACTTAGGGGCGGAGCAGGAAGCGTCGCC
GGGGGGCCCACAAGGGTAGCGGCGAAGATCCGGGTGACGCTGCGAACGGACGTGAAGAATGTGCGAGAC
CCAGGGTCGGCGCCGCTGCGTTTCCCGGAACCACGCCCAGAGCAGCCGCGTCCCTGCGCAAACCCAGGG
CTGCCTTGGAAAAGGCGCAACCCCAACCCCGGATCCCTAGGATCTACTGGCTCCATTTCTTGCTCTCCTCT
GTCGAGTAACGCCTATTCTGCTATGTCGACACCCAATTCTGAAATGGATAAACAGCAGTTGTTGCAGAATTC
TTATTATGGCTTCCACTCCTGCCCAAGTATCCCCGTAAGTTTCATAGATATGTTGTCCTAAGTTATGGAGCCA
TATCCTAGGAAAATGTCTAACAGCTTCACTCTTAAGTTCCTCTAAAAGCTCTAGTGTCCATTCATTGTATGGC
TCCCTCTGTGGCCCTTGGTCTTCTGGGGCTTGTTCCATACGCGTCTGCAGCCT***AGGACCGGGGTTTTCTTC***
CACGTCTCCTGCTTGCTTTAACAGAGAGAAGTTCGTGGCTCCGGAACCGGTGTCGACGAATTCGCTAGCC
AATTCTCCAGGCGATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGGCCTCCCACCGTACACGCCTACC
TCGACATACGTTCTCTATCACTGATAGGGAGTAAACTCGACATACGTTCTCTATCACTGATAGGGATAAACTC
GACATACGTTCTCTATCACTGATAGGGAGTAAACTCGACATACGTTCTCTATCACTGATAGGGAGTAAACTC
GACATACGTTCTCTATCACTGATAGGGAGTAAACTCGACATCGTTCTCTATCACTGATAGGGAGTAAACTCG
ACATACGTTCTCTATCACTGATAGGGAGTAAACTCGACATATCGATTCGCGGCCAAAGTGGATCTCTGCTGT
CCCTGTAATAAACCCGAAAATTTTGAATTTTTGTAATTTGTTTTTGGAATTCTTAAGTTGGATGTCGGTGGCTA
TTATGTCTACTATTCTTTCCCCTGCACTGTACCCCCCAATCCCCCCTTTTCTTTTAAAATTGTGGATGAATACT
GCCATTTGTCTCGAGGTCGAGAATTGTCCCCTCGGGGTTGGGAAGGTGGGTCTGAAACGATAATGGTGAAA
TTCCCTGCCTAACTCCTTTCCCTAAAGAAAAGTACAGCAAAAACTATTCTTAAACCTACCAGCCTCCTACTAT
ATTATTGAAAAATTTTTAATCCCCCGCCCCTTTTTTTGGTAACCATTTCCAAAATTCCCCTTTTTTATTTCAAAA
ATTTTTGTTCTTTTTTTTTTGTGGGGTGTGGCAATCTCATATAGAAGGGTGTAAAAAGTGGTAAATTTGTATAT
TTTTCCGCCCCCCCTCCCCCCGCCGGGGAGGTACCAACTGTTTCCCAAAAATTTTTACCACCCCACCCCC

pCW57-VPU colony 4

LNCX-Fwd
GGGCAGGGATGGCTAGCGAATTCGTCGACACCGGGTTCCGGA***GCCACGAACTTCTCTCTGTTAAAGCAAG***
CAGGAGACGTGGAAGAAAACCCCGGTCCTAGGCTGCAGACGCGTATGCAACCTATAATAGTAGCAATAGT
AGCATTAGTAGTAGCAATAATAATAGCAATAGTTGTGTGGTCCATAGTAATCATAGAATATAGGAAAATATTA
AGACAAAGAAAAATAGACAGGTTAATTGATAGACTAATAGAAAGAGCAGAAGACAGTGGCAATGAGAGTGAA
GGAGAAGTATCAGCACTTGTGGAGATGGGGGTGGAAATGGGGCACCATGCTCCTTGGGATATTGATGATCT
GTAGGGATCCGGGGTTGGGGTTGCGCCTTTTCCAAGGCAGCCCTGGGTTTGCGCAGGGACGCGGCTGCT
CTGGGCGTGGTTCCGGGAAACGCAGCGGCGCCGACCCTGGGTCTCGCACATTCTTCACGTCCGTTCGCAG
CGTCACCCGGATCTTCGCCGCTACCCTTGTGGGCCCCCCGGCGACGCTTCCTGCTCCGCCCCTAAGTCGG
GAAGGTTCCTTGCGGTTCGCGGCGTGCCGGACGTGACAAACGGAAGCCGCACGTCTCACTAGTACCCTCG
CAGACGGACAGCGCCAGGGAGCAATGGCAGCGCGCCGACCGCGATGGGCTGTGGCCAATAGCGGCTGCT
CAGCAGGGCGCGCCGAGAGCAGCGGCCGGGAAGGGGCGGTGCGGGAGGCGGGGTGTGGGGCGGTAGT
GTGGGCCCTGTTCCTGCCCGCGCGGTGTTCCGCATTCTGCAAGCCTCCGGAGCGCACGTCGGCAGTCGG
CTCCCTCGTTGACCGAATCACCGACCTCTCTCCCCAGCAATTCACCATGACCGAGTACAAGCCCACGGTGC
GCCTCGCCACCCGCGACGACGTCCCCAGGGCCGTACGCACCCTCGCCGCCGCGTTCGCCGACTACCCCG
CCACGCGCCACACCGTCGATCCGGACCGCCACATCGAGCGGGTCACCGAGCTGCAAGAACTCTTCCTCAC
GCGCGTCGGGCTCGACATCGGCAAGGTGTGGGTCGCGGACGACGGCGCCGCGGTGGCGGTCTGGACCA
CGCCGGAGAACGTCCAAACGGGGGCGGTGTTCGCCGAAATCGGCCCCCGCATGGCCGAATTTAACGGTTC
CCGGCTGGCCCCCAGCAACAAATGGAAGGCTCCTGGCCCCCAACGGCCCAAGGAACCCGGGGGTTCTGG
CCACGTCGGGTTCCCCCCGACCCAAGGGAAGGGTTGGGACGCCCTCTGGCCCCCGAAGTGGGGGGGCCA
AA

hPGK-Rev
GGCGGTTGCGTCCGGCCGCCGCGAACCGCAAGGAACCTTCCCGACTTAGGGGCGGAGCAGGAAGCGTCG
CCGGGGGGCCCACAAGGGTAGCGGCGAAGATCCGGGTGACGCTGCGAACGGACGTGAAGAATGTGCGAG
ACCCAGGGTCGGCGCCGCTGCGTTTCCCGGAACCACGCCCAGAGCAGCCGCGTCCCTGCGCAAACCCAG
GGCTGCCTTGGAAAAGGCGCAACCCCAACCCCGGATCCCTACAGATCATCAATATCCCAAGGAGCATGGTG
CCCCATTTCCACCCCCATCTCCACAAGTGCTGATACTTCTCCTTCACTCTCATTGCCACTGTCTTCTGCTCTT
TCTATTAGTCTATCAATTAACCTGTCTATTTTTCTTTGTCTTAATATTTTCCTATATTCTATGATTACTATGGAC
CACACAACTATTGCTATTATTATTGCTACTACTAATGCTACTATTGCTACTATTATAGGTTGCATACGCGTCTG
CAGCCT***AGGACCGGGGTTTTCTTCCACGTCTCCTGCTTGCTTTAACAGAGAGAAGTTCGTGGC***TCCGGAA
CCGGTGTCGACGAATTCGCTAGCCAATTCTCCAGGCGATCTGACGGTTCACTAAACGAGCTCTGCTTATATA
GGCCTCCCACCGTACACGCCTACCTCGACATACGTTCTCTATCACTGATAGGGAGTAAACTCGACATACGTT
CTCTATCACTGATAGGGATAAACTCGACATACGTTCTCTATCACTGATAGGGAGTAAACTCGACATACGTTCT
CTATCACTGATAGGGAGTAAACTCGACATACGTTCTCTATCACTGATAGGGAGTAAACTCGACATCGTTCTC

TATCACTGATAGGGAGTAAACTCGACATACGTTCTCTATCACTGATAGGGAGTAAACTCGACATATCGATTC
GCGGCCAAAGTGGATCTCTGCTGTCCCTGTAATAAACCCGAAAATTTTGAATTTTTGTAATTTGTTTTTGTAAT
TCTTTAGTTTGTATGTCTGTTGCTATTATGTCTACTATTCTTTCCCCTGCACTGTACCCCCCAATCCCCCCTTT
TCTTTTAAAATTGTGGATGAATACTGCCATTTGTCTCGAGGTCGAGAATTGTCCCCTCGGGGTTGGGAAGGT
GGGTCTGAAAACGATAATGGGTGAATATCCCTGCCTAACTCA

The sequence of both colonies is correct: Vpr and Vpu are present in the colonies.

Supplementary figure 7 |Sequence analysis of pCW57-dNGFR, pCW57-dNGFR-P2A-Vpr and pCW57-dNGFR-P2A-Vpu by using NCBI blast.

With: RE sites NGFR sequence *P2A sequence* VPU/VPR sequence reading mistakes

pCW57-dNGFR colony 2

LNCX-Fwd

GGCCGGAATGGCTAGCATGGGGGGCAGGTGCCACCGGCCGCGCCATGGACGGGCCGCGCCTGCTGCTG
TTGCTGCTTCTGGGGGTGTCCCTTGGAGGTGCCAAGGAGGCATGCCCCACAGGCCTGTACACACACAGCG
GTGAGTGCTGCAAAGCCTGCAACCTGGGCGAGGGTGTGGCCCAGCCTTGTGGAGCCAACCAGACCGTGTG
TGAGCCCTGCCTGGACAGCGTGACGTTCTCCGACGTGGTGAGCGCGACCGAGCCGTGCAAGCCGTGCAC
CGAGTGCGTGGGGCTCCAGAGCATGTCGGCGCCGTGCGTGGAGGCCGACGACGCCGTGTGCCGCTGCG
CCTACGGCTACTACCAGGATGAGACGACTGGGCGCTGCGAGGCGTGCCGCGTGTGCGAGGCGGGCTCGG
GCCTCGTGTTCTCCTGCCAGGACAAGCAGAACACCGTGTGCGAGGAGTGCCCCGACGGCACGTATTCCGA
CGAGGCCAACCACGTGGACCCGTGCCTGCCCTGCACCGTGTGCGAGGACACCGAGCGCCAGCTCCGCGA
GTGCACACGCTGGGCCGACGCCGAGTGCGAGGAGATCCCTGGCCGTTGGATTACACGGTCCACACCCCCA
GAGGGCTCGGACAGCACAGCCCCCAGCACCCAGGAGCCTGAGGCACCTCCAGAACAAGACCTCATAGCCA
GCACGGTGGCAGGTGTGGTGACCACAGTGATGGGCAGCTCCCAGCCCGTGGTGACCCGAGGCACCACCG
ACAACCTCATCCCTGTCTATTGCTCCATCCTGGCTGCTGTGGTTGTGGGCCTTGTGGCCTACATAGCCTTCA
AGAGGTGGAACAGCACCGGTTCCGGA***GCCACGAACTTCTCTCTGTTAAAGCAAGCAGGAGACGTGGAAG***
AAAACCCCGGTCCTAGGCTGCAGACGCGTTGTACAGGATCCGGGGGGTTGGGGTTGCGCCTTTTCCAAGGC
AGCCCTGGGTTTGCGCAGGGACGCGGCTGCTCTGGGCGTGGTTCCGGGAAACGCAGCGGCGCCGACCCT
GGGTCTCGCACATTCTTCACGTCCGTTCGCAGCGTCACCCGGATCTTCGCCGCTACCCTTGTGGGCCCCCC
GGCGACGCTTCCTGCTCCGCCCCTAAGTCGGGAAAGGTTCCTTGCCGGTTCGCGCCCTGCCCGGACGTGA
A

AgeI-dNGFR-Rev

CCCCACAAGGGCCCACAACCACAGCAGCCAGGATGGAGCAATAGACAGGGATGAGGTTGTCGGTGGTGCC
TCGGGTCACCACGGGCTGGGAGCTGCCCATCACTGTGGTCACCACACCTGCCACCGTGCTGGCTATGAGG
TCTTGTTCTGGAGGTGCCTCAGGCTCCTGGGTGCTGGGGGCTGTGCTGTCCGAGCCCTCTGGGGGTGTGG
ACCGTGTAATCCAACGGCCAGGGATCTCCTCGCACTCGGCGTCGGCCCAGCGTGTGCACTCGCGGAGCTG
GCGCTCGGTGTCCTCGCACACGGTGCAGGGCAGGCACGGGTCCACGTGGTTGGCCTCGTCGGAATACGT
GCCGTCGGGGCACTCCTCGCACACGGTGTTCTGCTTGTCCTGGCAGGAGAACACGAGGCCCGAGCCCGC
CTCGCACACGCGGCACGCCTCGCAGCGCCCAGTCGTCTCATCCTGGTAGTAGCCGTAGGCGCAGCGGCA
CACGGCGTCGTCGGCCTCCACGCACGGCGCCGACATGCTCTGGAGCCCCACGCACTCGGTGCACGGCTT
GCACGGCTCGGTCGCGCTCACCACGTCGGAGAACGTCACGCTGTCCAGGCAGGGCTCACACACGGTCTG
GTTGGCTCCACAAGGCTGGGCCACACCCTCGCCCAGGTTGCAGGCTTTGCAGCACTCACCGCTGTGTGTG
TACAGGCCTGTGGGGCATGCCTCCTTGGCACCTCCAAGGGACACCCCCAGAAGCAGCAACAGCAGCAGGC
GCGGCCCGTCCATGGCGCGGCCGGTGGCACCTGCCCCCATGCTAGCCAATTCTCCAGGCGATCTGACGGT
TCACTAAACGAGCTCTGCTTATATAGGCCTCCCACCGTACACGCCTACCTCGACATACGTTCTCTATCACTG
ATAGGGAGTAAACTCGACATACGTTCTCTATCACTGATAGGGATAAACTCGACATACGTTCTCTATCACTGAT
AGGGAGTAAACTCGACATACGTTCTCTATCACTGATAGGGAGTAAACTCGACATACGTTCTCTATCACTGAT
AGGGAGTAAACTCGACATCGTTCTCTATCACTGATAGGGAGTAAACTCGACATACGTTCTCTATCACTGATA
GGGAGTAAACTCGACATATCGAATTCGCGGCCCAAAGTGGATCCCTGCTGTCCCTGTAATAAACACCGAAA
A

The sequence is correct, as the extra G in the forward sequence is not present as an extra C in the reverse sequence (it is a reading mistake).

pCW57-dNGFR-P2A-VPR colony 2

LNCX-Fwd

GGCCGGAATGGCTAGCATGGGGGGCAGGTGCCACCGGCCGCGCCATGGACGGGCCGCGCCTGCTGCTG
TTGCTGCTTCTGGGGGTGTCCCTTGGAGGTGCCAAGGAGGCATGCCCCACAGGCCTGTACACACACAGCG
GTGAGTGCTGCAAAGCCTGCAACCTGGGCGAGGGTGTGGCCCAGCCTTGTGGAGCCAACCAGACCGTGTG
TGAGCCCTGCCTGGACAGCGTGACGTTCTCCGACGTGGTGAGCGCGACCGAGCCGTGCAAGCCGTGCAC
CGAGTGCGTGGGGCTCCAGAGCATGTCGGCGCCGTGCGTGGAGGCCGACGACGCCGTGTGCCGCTGCG

CCTACGGCTACTACCAGGATGAGACGACTGGGCGCTGCGAGGCGTGCCGCGTGTGCGAGGCGGGCTCGG
GCCTCGTGTTCTCCTGCCAGGACAAGCAGAACACCGTGTGCGAGGAGTGCCCCGACGGCACGTATTCCGA
CGAGGCCAACCACGTGGACCCGTGCCTGCCCTGCACCGTGTGCGAGGACACCGAGCGCCAGCTCCGCGA
GTGCACACGCTGGGCCGACGCCGAGTGCGAGGAGATCCCTGGCCGTTGGATTACACGGTCCACACCCCCA
GAGGGCTCGGACAGCACAGCCCCCAGCACCCAGGAGCCTGAGGCACCTCCAGAACAAGACCTCATAGCCA
GCACGGTGGCAGGTGTGGTGACCACAGTGATGGGCAGCTCCCAGCCCGTGGTGACCCGAGGCACCACCG
ACAACCTCATCCCTGTCTATTGCTCCATCCTGGCTGCTGTGGTTGTGGGCCTTGTGGCCTACATAGCCTTCA
AGAGGTGGAACAGCACCGGTTCCGGA***GCCACGAACTTCTCTCTGTTAAAGCAAGCAGGAGACGTGGAAG***
AAAACCCCGGTCCTAGGCTGCAGACGCGT**ATG**GAACAAGCCCCAGAAGACCAAGGGCCACAGAGGGAGC
CATACAATGAATGGACACTAGAGCTTTTAGAGGAACTTAAGAGTGAAGCTGTTAGACATTTTCCTAGGATATG
GCTCCATAACTTAGGACAACATATCTATGAAACTTACGGGGATACTTGGGCAGGAGTGGAAGCCATAATAAG
AATTCTGCAACAACTGCTGTTTATCCATTTCAGAATTGGGTGTCGACATAGCAGAATAGGCGTTACTCGACA
GAGGAGAGCAAGAAATGGAGCCAGTAGATCC**TAG**GGATTCCGGGGGTTGGGGTTTCCCC

Vpr-Rev

TTAAATTTAAGTTTGCTCTCTCTGTCGAGTAACGCCTATTCTGCTATGTCGACACCCAATTCTGAAATGGATA
AACAGCAGTTGTTGCAGAATTCTTATTATGGCTTCCACTCCTGCCCAAGTATCCCCGTAAGTTTCATAGATAT
GTTGTCCTAAGTTATGGAGCCATATCCTAGGAAAATGTCTAACAGCTTCACTCTTAAGTTCCTCTAAAAGCTC
TAGTGTCCATTCATTGTATGGCTCCCTCTGTGGCCCTTGGTCTTCTGGGGCTTGTTC**CAT**ACGCGTCTGCAG
CCT***AGGACCGGGGTTTTCTTCCACGTCTCCTGCTTGCTTTAACAGAGAGAAGTTCGTGGC***TCCGGAACCG
GTGCTGTTCCACCTCTTGAAGGCTATGTAGGCCACAAGGCCCACAACCACAGCAGCCAGGATGGAGCAATA
GACAGGGATGAGGTTGTCGGTGGTGCCTCGGGTCACCACGGGCTGGGAGCTGCCCATCACTGTGGTCACC
ACACCTGCCACCGTGCTGGCTATGAGGTCTTGTTCTGGAGGTGCCTCAGGCTCCTGGGTGCTGGGGGCTG
TGCTGTCCGAGCCCTCTGGGGGTGTGGACCGTGTAATCCAACGGCCAGGGATCTCCTCGCACTCGGCGTC
GGCCCAGCGTGTGCACTCGCGGAGCTGGCGCTCGGTGTCCTCGCACACGGTGCAGGGCAGGCACGGGTC
CACGTGGTTGGCCTCGTCGGAATACGTGCCGTCGGGGCACTCCTCGCACACGGTGTTCTGCTTGTCCTGG
CAGGAGAACACGAGGCCCGAGCCCGCCTCGCACACGCGGCACGCCTCGCAGCGCCCAGTCGTCTCATCC
TGGTAGTAGCCGTAGGCGCAGCGGCACACGGCGTCGTCGGCCTCCACGCACGGCGCCGACATGCTCTGG
AACCCCACGCACTCGGTGCACGGCTTGCACGGCTCGGTCGCGCTCACCACGTCGGAGAACGTCAAACTGA
CAAGACAGGACATACAAAGCGGAACAGGTAGGATCCACAAAGATTGGGGCCAAATCCCTCGCCCAGGTTAA
CAGGATTTTGAAGCCCATCACTGCTGTGTGGTGAACA

The sequence is correct, as the extra G in the forward sequence is not present as an extra C in the reverse sequence (it is a reading mistake). Same for extra A in reverse sequence: no extra T is present in forward sequence.

pCW57-dNGFR-P2A-VPU colony 4

LNCX-Fwd

GGAAGGAATGGCTAGCATGGGGGGCAGGTGCCACCGGCCGCGCCATGGACGGGCCGCGCCTGCTGCTGT
TGCTGCTTCTGGGGGTGTCCCTTGGAGGTGCCAAGGAGGCATGCCCCACAGGCCTGTACACACACAGCGG
TGAGTGCTGCAAAGCCTGCAACCTGGGCGAGGGTGTGGCCCAGCCTTGTGGAGCCAACCAGACCGTGTGT
GAGCCCTGCCTGGACAGCGTGACGTTCTCCGACGTGGTGAGCGCGACCGAGCCGTGCAAGCCGTGCACC
GAGTGCGTGGGGCTCCAGAGCATGTCGGCGCCGTGCGTGGAGGCCGACGACGCCGTGTGCCGCTGCGC
CTACGGCTACTACCAGGATGAGACGACTGGGCGCTGCGAGGCGTGCCGCGTGTGCGAGGCGGGCTCGGG
CCTCGTGTTCTCCTGCCAGGACAAGCAGAACACCGTGTGCGAGGAGTGCCCCGACGGCACGTATTCCGAC
GAGGCCAACCACGTGGACCCGTGCCTGCCCTGCACCGTGTGCGAGGACACCGAGCGCCAGCTCCGCGAG
TGCACACGCTGGGCCGACGCCGAGTGCGAGGAGATCCCTGGCCGTTGGATTACACGGTCCACACCCCCAG
AGGGCTCGGACAGCACAGCCCCCAGCACCCAGGAGCCTGAGGCACCTCCAGAACAAGACCTCATAGCCAG
CACGGTGGCAGGTGTGGTGACCACAGTGATGGGCAGCTCCCAGCCCGTGGTGACCCGAGGCACCACCGA
CAACCTCATCCCTGTCTATTGCTCCATCCTGGCTGCTGTGGTTGTGGGCCTTGTGGCCTACATAGCCTTCAA
GAGGTGGAACAGCACCGGTTCCGGA***GCCACGAACTTCTCTCTGTTAAAGCAAGCAGGAGACGTGGAAGA***
AAACCCCGGTCCTAGGCTGCAGACGCGT**ATG**CAACCTATAATAGTAGCAATAGTAGCATTAGTAGTAGCAA
TAATAATAGCAATAGTTGTGTGGTCCATAGTAATCATAGAATATAGGAAAATATTAGGACAAGGAAAAATAGA
CAGGTTAATTGATAGACTAATAGAAAGAGCAGAAGACAGTGGCAATGAAGAGTGAAGGAGAAGTATCAGCA
CTTGTGGAGATGGGGGTGGAAATGGGGCACCATGCTCCTTGGGAAATTGATGA

Vpu-Rev

CCCTGGGTGCCCCATTTCCACCCCCATCTCCACAAGTGCTGATACTTCTCCTTCACTCTCATTGCCACTGTC
TTCTGCTCTTTCTATTAGTCTATCAATTAACCTGTCTATTTTTCTTTGTCTTAATATTTTCCTATATTCTATGATT
ACTATGGACCACACAACTATTGCTATTATTATTGCTACTACTAATGCTACTATTGCTACTATTATAGGTTGCAT
ACGCGTCTGCAGCCT***AGGACCGGGGTTTTCTTCCACGTCTCCTGCTTGCTTTAACAGAGAGAAGTTCGTG***
GCTCCGGAACCGGTGCTGTTCCACCTCTTGAAGGCTATGTAGGCCACAAGGCCCACAACCACAGCAGCCA
GGATGGAGCAATAGACAGGGATGAGGTTGTCGGTGGTGCCTCGGGTCACCACGGGCTGGGAGCTGCCCA
TCACTGTGGTCACCACACCTGCCACCGTGCTGGCTATGAGGTCTTGTTCTGGAGGTGCCTCAGGCTCCTGG
GTGCTGGGGGCTGTGCTGTCCGAGCCCTCTGGGGGTGTGGACCGTGTAATCCAACGGCCAGGGATCTCCT
CGCACTCGGCGTCGGCCCAGCGTGTGCACTCGCGGAGCTGGCGCTCGGTGTCCTCGCACACGGTGCAGG
GCAGGCACGGGTCCACGTGGTTGGCCTCGTCGGAATACGTGCCGTCGGGGCACTCCTCGCACACGGTGTT
CTGCTTGTCCTGGCAGGAGAACACGAGGCCCGAGCCCGCCTCGCACACGCGGCACGCCTCGCAGCGCCC
AGTCGTCTCATCCTGGTAGTAGCCGTAGGCGCAGCGGCACACGGCGTCGTCGGCCTCCACGCACGGCGC
CGACATGCTCTGGAGCCCCACGCACTCGGTGCACGGCTTGCACGGCTCGGTCGCGCTCACCACGTCGGA
GAACGTCACGCTGTCCAGGCAGGGCTCACACACGGGTCTGGTTGGCTCCACAAGGCTGGGCCACACCCTC
GCCCAGGTTGCAGGCTTTGCAGCACTCACCGCTGTGTGTGTACAGGCCTGTGGGGCATGCCTCCTTGGCA
CCTCCAAGGGACACCCCCAGAAGCAGCAACAGCAGCAGGCGCGGCCCGTCCATGGCGCGGCCCGGTGGC
ACCTGCCCCCATGCTAACCAAATTTTT

The sequence is correct, as the extra G in the forward sequence is not present as an extra C in the reverse sequence (it is a reading mistake). Same for the extra G in reverse sequence: no extra T is present in forward sequence.