Doctor of Philosophy (Medicine) Dissertation

Effect of HIV-1 subtype C Transactivator of transcription (Tat) A21P variant on TAR binding ability, nuclear levels of active positive transcription elongation factor b (P-TEFb) and viral latency.

Zakithi Mkhize
(213571739)

Supervisor: Dr Paradise Madlala

HIV Pathogenesis Programme

Doris Duke Medical Research Institute

Nelson R Mandela School of Medicine

DECLARATION

I, **Zakithi Mkhize**, declare that;

1. The research reported in this thesis, except where otherwise indicated, is my original research.

2. This thesis has not been submitted for any degree or examination at any other university.

3. This thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.

4. This thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:

(a) Their words have been re-written but the general information attributed to them has been referenced,

(b) Where their exact words have been used, then their writing has been placed in italics and inside quotation marks and referenced.

5. This thesis does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the thesis and in the references sections.

Zakithi Mkhize _____ Date: 22 June 2023

Paradise Madlala _ _____ Date: 22 June 2023

ACKNOWLEDGEMENTS

I wish to express my gratitude to:

- My supervisor, Dr Paradise Madlala, for his invaluable and sincere guidance and research expertise throughout the project. For being patient when things were tough and for always believing in me. For also being my mentor and pushing me to be the best I can be.
- Prof Thumbi Ndung'u for granting me the opportunity to pursue my doctoral degree at HPP, funding conference travel and his academic guidance and support.
- Our collaborators at the Ott lab at Gladstone Institutes University of California San Francisco, particularly Daniela Boehm and Dr Melanie Ott for hosting me.
- Our collaborators at Erasmus Medical centre in Rotterdam, Netherlands. Raquel Crespo for her assistance and expertise with protein modelling and RNA IP experiments. Prof Tokameh Mahmoudi for funding and hosting me.
- HPP and SANTHE for housing me as a PhD fellow.
- The National Research Foundation (NRF) for funding.
- University of KwaZulu-Natal College of Health Sciences (CHS) for research conference travel funding.
- Africa Health Research Institute (AHRI) Connect Africa scholarships for funding conference travel.
- Poliomyelitis Research Foundation (PRF) for funding.
- My Mom, Dad and my sister for their continuous support and encouragement.
- My partner for his endless support and carrying me through the stressful times at the very end.
- My friends for their support and belief in me.

Dedication

To all the black girl scientists, current and future.

List of Figures:

Abbreviations

HIV	Human immunodeficiency virus
HIV-1	HIV type 1
AIDS	Acquired immunodeficiency syndrome
PLWH	People living with HIV
TAT	Transactivator of transcription
RNA	Ribonucleic acid
RNAPII	RNA polymerase II
TAR	Transactivation response element
P-TEFb	Positive transcription elongation factor b
CYCT1	Cyclin T1
CDK9	Cyclin dependent kinase 9
7SK snRNA	7SK small nuclear RNA
HMBA	Hexamethylene bisacetamide
HEXIM1	HMBA protein 1
snRNP	small nuclear ribonuclear protein
T/F	Transmitted/founder
TatA21	Tat Alanine 21
TatP21	Tat Proline 21
cART	Combination antiretroviral therapy
CD4	Cluster of differentiation
CCR5	C-C chemokine receptor 5
ORF	Open reading frame
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
LRA	Latency reversing agent
dCA	Didehydro-cortistatin
NF-κB	Nuclear factor kappa light chain enhancer of activated B cells
CDK7	Cyclin dependent kinase 7
CTD	C-terminal domain
NELF	Negative elongation factor

TSS	Transcription start site
BRD4	Bromodomain 4
CBP	CREB binding protein
HAT	Histone acetyltransferase
CSF	Cerebral spinal fluid
IL	Interleukin
MCP-1	Monocyte chemoattractant protein 1
HAND	HIV-1 associated neurocognitive disorder
CXCR4	C-X-C motif chemokine receptor 4
GALT	Gastrointestinal tract
mRNA	Messenger RNA
SDM	Site directed mutagenesis
DMEM	Dulbecco modified eagle medium
FBS	Fetal bovine serum
PBS	Phosphate buffered saline
PEI	Polyethylenimine
PCR	Polymerase chain reaction
qPCR	Qualitative PCR
RLU	Relative light units
HEK	Human embryonic kidney

Abstract

The HIV-1 Transactivator of transcription (Tat) enhances the ability of the viral promoter 5' long terminal repeat (LTR) to drive viral gene transcription and is important for HIV-1 pathogenesis. Tat binds to the transactivator RNA (TAR) element of the 5'LTR and subsequently recruits the host positive transcription elongation factor b (P-TEFb) for efficient viral gene transcription. Inter- and intra-subtype Tat genetic variation that translates to functional differences has been reported. Specifically, HIV-1 subtype C (HIV-1C) exhibiting Alanine at position 21 of the Tat protein (TatA21) was reported to be associated with reduced LTR transcriptional activity compared to Tat exhibiting Proline at position 21 mutation (TatP21). However, the effect of Tat variation on its ability to recruit P-TEFb is unknown. Therefore, this study seek to determine the effect of HIV-1 subtype C TatA21 mutant on the ability of Tat to recruit P-TEFb to 5' LTR to enhance viral gene transcription. To this effect, site-directed mutagenesis (SDM) was performed on the Plasmid pcDNA3.1(+) HIV-1C BL43/02 TatA21 to introduce TatP21 alone or together with other mutations using designed primers and the Q5 DNA polymerase kit. The effect of Tat mutations was measured using Tat transactivation assay where the luciferase activity was the measured output in TZM-bl cell lines and the impact of TatA21 was further assessed on ability of the LTR to drive GFP and Gag expression in Jurkat and A72 cells respectively. Next, protein modelling was performed using Hdock software, followed by RNA immunoprecipitation (RNA IP) was performed using stably expressing TatA21 and TatP21 in Jurkat cells. Lastly, co-immunoprecipitation of TatA21 and TatP21 was performed with cycT1 and CDK9. Our data show TatA21 mutant alone was associated with significantly reduced LTR transcription activity compared to TatP21 ($p = 0.0004$). TatA21 resulted in had significantly lower GFP expression Jurkat cells ($p = 0.0439$) and lower Gag expression in A72 cells compared to TatP21. Although TatA21 reduced the LTR transcription activity compared to TatP21, protein modelling using Hdock software revealed that TatA21 and TatP21 protein structures were the same. Consistently, molecular docking showed that TatA21 had a lower binding affinity than TatP21. The RNA IP showed that TatA21 had significantly reduced affinity to bind to TAR compared to TatP21 ($p = 0.0151$). Moreover, TatA21 and TatP21 formed a complex with cycT1 and CDK9. Taken together, our data shows that HIV-1C TatA21 significantly reduced its transactivation activity but does not affect its ability to recruit P-TEFb. Interestingly, TatP21 is able to bind TAR more efficiently than TatA21 thus revealing a possible mechanism but which the reduced functionality of SDMs and patient derived TatA21 variants was observed. The effect of TatA21 and TatP21 on the propensity of HIV-1 latency development or reversal. To this effect, a recombinant viral vector exhibiting either TatA21 ($C731C_{TatA21}C$) or TatP21 ($C731C_{TatP21}C$) were generated. The $C731C_{TatA21}C$ or $C731C_{TatP21}C$ were separately co-transfected together with VSV-G and R8.91 into Jurkat cells for virus production. This virus was then used to infect Jurkat cells for 3 days. Followed by cell sorting of GFP- cells, which represented either truly negative or latently infected cells was then performed. We were able to successfully generate $C731C_{TatA21}C$ virus and characterized it to a 1.2% reactivation. However, the generation of $C731C_{TatP21}C$ recombinant viral vector was unsuccessful and thus could not be used for comparison. Future studies should involve the characterization of TatP21 in the propensity of latency development and/ or reactivation.

Table of Contents

Outcomes of research
Conferences, publications and achievements

Conferences:

1. 23rd International AIDS Conference, 6-10 July 2020. Virtual Poster (Poster number 2786) presentation: **Mkhize. Z** and Madlala, P.
   Abstract title: Impact of human immunodeficiency virus type 1 subtype C (HIV-1C) transmitted/founder (T/F) viruses *tat* genetic variation on transactivation activity and disease outcome,

2. AIDS 2022, the 24th International AIDS Conference, Montreal, Canada, 29 July to 2 August 2022 (virtual). **Mkhize, Z**, Boehm, D, Ott, M, Ndung'u T and Madlala, P.
   Abstract title: Effect of human immunodeficiency virus type 1 (HIV-1) subtype C transactivator of transcription (Tat) P21A variant on nuclear levels of active positive transcription elongation factor b (P-TEFb).

3. HIV Drug Therapy, Glasgow, 22-26 October 2022 (virtual). **Mkhize, Z**, Boehm, D, Ott, M, Ndung'u T and Madlala, P.
   Abstract title: Effect of HIV-1 subtype C transactivator of transcription (Tat) P21A variant on nuclear levels of active positive transcription elongation factor b (P-TEFb), TAR binding and disease outcome.

Publications:

1. Madlala P, **Mkhize Z**, Naicker S, Khathi SP, Maikoo S, Gopee K, et al. (2023) Genetic variation of the HIV-1 subtype C transmitted/founder viruses long terminal repeat elements and the impact on transcription activation potential and clinical disease outcomes. PLoS Pathog 19(6):e1011194. https://doi.org/10.1371/journal.ppat.1011194.

2. In preparation: **Mkhize Z**, Crespo-Galvan R, Boehm D, Ott M, Mahmoudi, T, Ndung'u, T and Madlala P. Impact of HIV-1 subtype C Tat Alanine 21 amino acid on disease outcome and transcriptional activity via TAR RNA interaction.

Achievements:

1. Runner-up in the category for best poster: College of Health Sciences research syposium 18 Sept 2020. Poster presentation: Mkhize, Z., Ndung'u T., Boehm, D., Ott, M., and Madlala, P. "Impact of Tat genetic variation on its transcriptional activity and ability to recruit Positive Transcription Elongation Factor b (P-TEFB) and disease outcome."
2. FameLab KZN winner and South African Top 10 2021.
3. 2nd runner up': 3 minute thesis (3-MT) competition 2021.

Chapter One

Chapter One:
1.1 Introduction

Despite the human immunodeficiency virus (HIV) being discovered as the causative agent for acquired immunodeficiency syndrome (AIDS) almost four decades ago (Barre-Sinoussi et al., 1983), there is still no cure for this viral infection. The absence of an HIV cure is largely due to the presence of HIV latent viral reservoirs (Siliciano et al., 2003). HIV latent reservoirs are made up of cells infected with replication competent yet transcriptionally silent provirus (Hermankova et al., 2003). Although there are two types of HIV, HIV type 1 (HIV-1) and HIV type 2 (HIV-2), this study focuses on HIV-1 since it the most common type globally (Hemelaar et al., 2019). The lack of understanding of the molecular mechanisms that govern HIV-1 latency at the transcriptional level is the major impediment to HIV-1 cure development.

The viral promoter, 5' long terminal repeat (LTR) is the driver viral gene transcription and important for HIV-1 replication (Gaynor, 1992). The HIV-1 5' LTR houses the novel transactivation response (TAR) element, a ribonucleic acid (RNA) that is 59 nucleotides long and located immediately downstream of the transcription initiation site (TSS) (Isel and Karn, 1999). The HIV-1 Transactivator of transcription (Tat) binds to the TAR element and promote 5' LTR transcription activity (Jones, 1993). Specifically, Tat enhances transcription at the level of RNA polymerase II (RNAPII) elongation, by recruiting host Positive Elongation Factor b (P-TEFb) and Super Elongation Complex (SEC) that jointly promotes viral messenger (m)RNA elongation. The P-TEFb is the host protein composed of cyclin T1 (cycT1) and cyclin-dependent Kinase 9 (CDK9), a kinase that phosphorylates RNA Polymerase II carboxyl-terminal domain (CTD) to stimulate cellular genes transcriptional elongation. (Zhu et al., 1997). The role of HIV-1 Tat in HIV latency establishment and reactivation has been investigated. An earlier study showed that HIV-1 subtype B Tat can inhibit the establishment of HIV-1 latency (Donahue et al., 2012). These studies reported that if Tat can counteract HIV-1 latency by promoting transcriptional initiation during HIV replication if it is present in sufficient amounts (Donahue et al., 2012, Pearson et al., 2008).

Inter- and intra-subtype Tat sequence variation that translates to differential Tat transactivation activity exists (Kandathil et al., 2009, Rossenkhan et al., 2013, Li et al., 2012). A previous study reported that patient derived HIV-1 subtype B Tat mutants exhibiting either Phenylalanine (Phe) or Arginine (Arg) instead of consensus Serine (Ser) at position 46 or 61

of the Tat amino acid sequence were associated with a strong TAR RNA binding affinity respectively, leading to enhanced transactivation activities (Ronsard et al., 2017a). Neogi et al., showed that while position 21 of Tat amino acid sequence contains Proline (referred to as TatP21 in this study) for most HIV-1 subtypes, Alanine at position 21 of Tat (referred to as TatA21 in this study) is a signature rescue for HIV-1 subtype C (HIV-1C) circulating in Indian (Neogi et al., 2012). A subsequent study showed that HIV-1C TatA21 variants also exist in Botswana, and that TatA21 was associated with significantly reduced production of LTR-driven luciferase (Rossenkhan et al., 2013). Furthermore, they showed a moderate positive correlation between Tat-mediated LTR activity and plasma viral loads in people living with HIV-1 (PLWH) (Rossenkhan et al., 2013). Consistently, the data from my Masters project showed that TatA21 variant is circulating in South Africa, associated with significantly reduced transactivation activity and reduced viral loads in PLWH compared TatP21 variants (Mkhize et al., prepared for publication at the time of writing my PhD thesis).

The aforementioned data suggest that TatA21 probably alone or in combination with other mutations may be a predictor of disease outcome and plausibly play a role in the establishment of viral latency. However, the mechanism by which TatA21 reduces 5' LTR transcription activity and viral load has not been fully investigated. Therefore, the aim of study was to investigate effect of HIV-1C TatA21 on the ability of Tat to bind to the TAR element and recruit P-TEFb to enhance viral transcription. Furthermore, this study undertook to determine the effect of HIV-1C TatA21 on the propensity of latency development or reversal. The data from this study will generate new knowledge on the possible role of HIV-1C TatA21 in disease outcome and latency development or versal, which will guide future studies on pharmaceutical antivirals and HIV-1 cure development.

1.2 Research Hypothesis

The HIV-1C TatA21 mutant alone or in combination with other mutations may influence the ability of Tat to bind to the TAR element, recruit the P-TEFb and mediate viral latency.

1.3 Study aims and objectives

This study was divided into two main goals:

Goal 1: To molecularly characterize the ability of HIV-1C TatA21 alone or in combination with other mutations to recruit P-TEFb for efficient viral gene transcription.

We hypothesized that TatA21 alone or in combination with other mutants affects its ability to bind TAR and recruit P-TEFb for efficient viral gene transcription.

The objectives were as follows:

1.1 To replace Alanine with Proline at position 21 of HIV-1 subtype C consensus Tat amino acid sequence using site directed mutagenesis, thus creating TatP21 mutant.
1.2 To assess the effect of TatA21 mutation alone or together with other mutations on Tat expression compared to TatP21.
1.3 To characterize the binding affinity of TatA21 and TatP21 to P-TEFb by performing co-immunoprecipitation (co-IP).
1.4 To determine and compare the 3D structures of TatA21 and TatP21 using protein homology modelling.
1.5 To assess the binding affinity of TatA21 versus TatP21 to bind the TAR element using protein modelling software by determining their docking score.
1.6 To confirm the binding affinity of TatA21 and TatP21 using the *in vitro* technique, RNA immunoprecipitation (IP) assay.

Goal 2: To define the effect of the HIV-1C TatA21 alone or in combination with other mutations on viral latency establishment or reversal.

Here we hypothesized that TatA21 alone or in combination with other mutations may play a role in the development or reversal of viral latency.

The objectives of this goal were as follows:

2.1. Construction HIV-1C consensus 5' LTR, *tat* and 3' LTR minimal genome expressing GFP reporter gene (C731CC) by adapting pEV731, which is a subtype B minimal genome reporter virus developed by Verdin's group (Jordan et al., 2003).
2.2. Construction of C731CC exhibiting TatP21 ($C731C_{TatP21}C$).
2.3. Production of infectious viral stock by transfecting C731CC or $C731C_{P21}C$, R8.9 backbone viral DNAs together with a VSV-G glycoprotein for virus harvesting.
2.4. Infection of Jurkat cells using viral stocks generated in the previous objective.
2.5. Cell sorting of green fluorescent protein (GFP) negative (GFP-) cell populations, representing either truly negative or latent cells was performed using the FACSDiva 6.1.1 software (BD Bioscience) and this population was reactivated by phorbol-12-myristate-13-acetate (PMA).

Chapter Two

2. Background and Literature review:

This chapter describes the context behind the major aim of this study through a review of the literature. Beginning by giving an overview of the current state of HIV epidemic globally, followed by discussing HIV latency as the major barrier to HIV-1 cure development. Specifically, this chapter describes the HIV-1 "functional" and "sterilizing" cure, as well the strategies to achieve them. A focal point in this chapter is the transactivator of transcription (Tat) as an enhancer of viral gene transcription and how Tat genetic variation impacts its functionality. In the process, this study investigates the role of Tat variation on its ability to bind to the crucial TAR element found in the HIV-1 promoter and recruit the host transcription factor positive transcription factor b (P-TEFb) for viral gene transcription. This chapter aims to highlight the important findings and knowledge gaps in literature, which formed the rationale of the project described in this thesis.

2.1 Barrier to HIV cure development

The UNAIDS '*Fast Track Ending the AIDS epidemic by 2030*' seeks to achieve ambitious goals by 2030: (1) 95% of people living with HIV (PLWH) should know their status; (2) 95% of all PLWH should receive combination antiretroviral treatment (cART) and (3) 95% of PLWH who are on cART should have effective suppression of the virus (UNAIDS, 2015). Yet we are currently falling short of these goals, as in 2022, only 76% of PLWH had access to treatment. Despite its undisputable success in suppressing HIV replication, cART is a lifelong regimen, often associated with long-term side effects, and is very expensive, thus creating a huge economic burden on developing countries (Poudel et al., 2017), like South Africa. Furthermore, Eastern and Southern Africa bear the brunt of the HIV epidemic, accounting for approximately 53% of infections worldwide (UNAIDS 2020). Moreover, treatment interruption results in the virus rebounding (Chun et al., 1999) demonstrating that cART is not curative. These factors highlight the urgency to develop an HIV cure.

The major barrier to eradication of HIV infection is the early development of viral reservoirs, cells infected with replication competent and yet transcriptionally silent provirus (reviewed in (Eisele and Siliciano, 2012)). Moreover, interruption of cART results in viral rebound (Chun et al., 1999). The main cause of viral rebound is reactivation of virus from a pool of these

latently infected cells, principally residing in the pool of resting memory cluster of differentiation ($CD4^+$) T cells (Chun et al., 1997, Finzi et al., 1997).

There are two types of HIV, HIV-1 and HIV-2. While HIV-2 is exclusive to West Africa, HIV-1 is globally distributed (Buonaguro et al., 2007). HIV-1 is also more easily transmitted and its progression to AIDS happens at a more rapid rate than HIV-2 (Popper et al., 1999). Therefore, HIV-1 is the focus of study and will be referred to henceforth. HIV-1 is divided in four groups (M, N, O and P) and further dived into 9 different subtypes (A-D, F-H, J-L) and 105 recombinant circulating forms(CRFs), which are unevenly globally distributed, with subtype C predominantly found in southern Africa (reviewed in (Elangovan et al., 2021)). Moreover, subtype C is responsible for approximately 50% of HIV-1 infections worldwide (Bbosa et al., 2019).

2.1.1 HIV-1 latency

HIV-1 is a retrovirus, which integrates its proviral genome into the host genome of its target cells during its replication cycle (reviewed in (Wu, 2004)) (Fig. 2.1). In the absence of cART, HIV-1 preferentially infects activated CD4+ T cells, majority of these cells quickly die following infection (Finzi et al., 1999). However, a small proportion of these infected cells transition into a long-term resting state in which the integrated replication-competent viral genome persists for the lifetime in PLWH (Crooks et al., 2015, Finzi et al., 1999). This is what is referred to as HIV latency and reversed when these latent cells are reactivated (Finzi et al., 1997)(Fig. 2.1). Therefore, HIV-1 latency is a state of non-productive infection (reviewed in (Karn, 2011, Ruelas and Greene, 2013)), where viral proteins are not produced. This HIV-1 persistence during cART was established in the late 1990s (Finzi et al., 1997, Chun et al., 1997, Wong et al., 1997). The emergence of advanced single-cell methodologies and next-generation sequencing capacities, it is now evident that the latent HIV-1 reservoir is far more dynamic, with multiple factors contributing to its maintenance (Cohn et al., 2020). Latently infected cells may be maintained by mechanisms operating at the posttranscriptional level (i.e., inhibition of nuclear RNA export and inhibition of HIV-1 transcription). However, in most latently infected cells, HIV-1 infection appears to be blocked at the transcriptional level (reviewed in (Van Lint et al., 2013)) (Fig. 2.1).

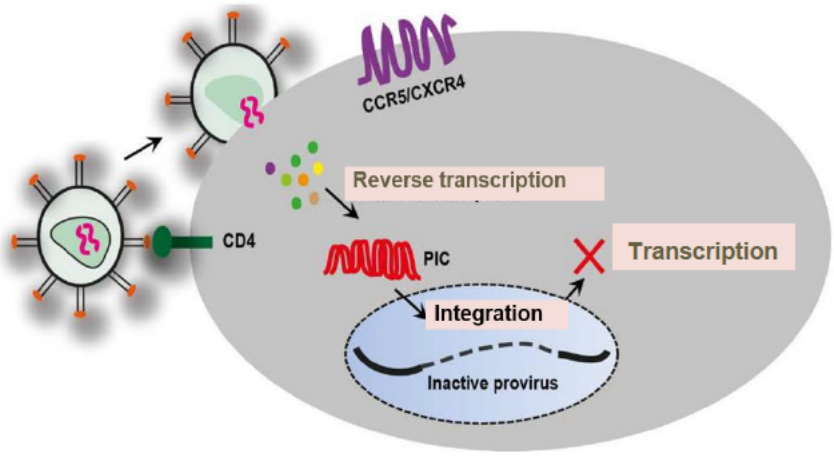

Figure 2.1: Diagram of a latently infected cell CD4 T cell. The virus binds to the CD4 receptor and subsequently engages core-receptors to gain cell entry and release its genome of the host cell, where its RNA genomic material is injected into the cell. Following this virus RNA is reverse transcribed to viral DNA genome, which forms part of a pre-integration complex (PIC) that get imported into the nucleus where viral DNA integrates into the host genome. However, transcription does not occur from integrated viral genome (proviral) DNA in this cell. Thus, there is no virus protein expression and no virions produced thus making the infected cell not to be detectable and unable to be cleared by the immune system while the virus remains hidden from cART.

Latent HIV-1 infection is mostly found in memory $CD4^+$ T cells, which have a long half-life of 44 months *in vivo* (Finzi et al., 1999). Therefore, memory CD4+ T cells allow latent virus to persist within PLWH for decades (Chun et al., 1998a, Finzi et al., 1997). The occurrence of latently infected cells is very low and is estimated to be 0.03–3 infectious units per million resting CD4+ T cells (Siliciano et al., 2003). HIV-1 latently infected cells persist as reservoirs found in sanctuary sites throughout the body of PLWH (Finzi et al., 1997). Moreover, cART is unable to penetrate these reservoirs as it can only target actively replicating virus (Cory et al., 2013). Viral reservoirs are defined as cell types or anatomical sites with replication-competent forms of the virus where it can persist with more stable kinetic properties than in the main pool of actively replicating virus (Blankson et al., 2002, Eisele and Siliciano, 2012).

There are two forms of viral latency which have been observed and can be separated into pre-integration and post-integration latency (Marcello, 2006). The mechanisms of pre-integration latency include poor nuclear transportation of the pre-integration complex (PIC), defective reverse transcriptase and the state of the cell life cycle (Blanco-Rodriguez et al., 2020). The

viral PIC can remain stable for several weeks on centrosomes and possibly integrate into the host genome when the host cell becomes reactivated (Zamborlini et al., 2007). Although macrophages and some tissues found in the brain may retain pre-integrated viruses for a longer period, pre-integration latency appears to be less clinically relevant (Pang et al., 1990, Kelly et al., 2008).

On the other hand, post-integration latency occurs rarely when the integrated provirus fails to effectively express its genome and is reversibly silenced in the host cell genome (reviewed in (Van Lint et al., 2013)). Memory CD4+ T cells transitioning from an active to a quiescent state also exhibit post-integration latency (Siliciano and Greene, 2011). Several factors can contribute to this state including low transcriptional factors, cellular DNA modification, transcriptional suppression, and low cellular proteins such as P-TEFb as well as viral protein Tat expression (Andre et al., 2023). This latent state is exceptionally stable and is only limited by the lifespan of the infected cell and its progeny. However, post-integration latency is a multifactorial phenomenon (Van Lint et al., 2013). Mechanisms that maintain HIV-1 latency *in vivo* are not completely understood. Latently infected cells may be maintained by mechanisms at the post-transcriptional level (i.e., inhibition of nuclear RNA export and inhibition of HIV-1 translation by microRNAs). However, in many latently infected cells, HIV-1 infection appears to be blocked at the transcriptional level (Fig. 2.1). HIV-1 transcriptional repression is crucial to the establishment and maintenance of post-integration latency. Several elements contribute to the transcriptional silencing of integrated HIV-1 proviruses (reviewed in (Colin and Van Lint, 2009, Redel et al., 2010): 1) the site of integration into the host cell genome, the cellular chromatin environment at this site and mechanisms of transcriptional interference; 2) the spatial sub-nuclear positioning of the integrated provirus (reviewed in (Marcello et al., 2010)); 3) the absence of crucial inducible host transcription factors, such as NF-kappaB (Nuclear Factor Kappa-light-chain-enhancer of activated B cells) or NFAT (Nuclear Factor of Activated T-cells), that are excluded from the nuclei of resting cells and transiently activated by various stimuli; 4) the presence of transcriptional repressors, such as CTIP2 (COUP-TF Interacting Protein 2), DSIF (DRB-Sensitivity Inducing Factor), NELF (Negative Elongation Factor) and the family of TRIM proteins (tripartite motif); 5) the chromatin structure of the HIV-1 promoter and the presence of a repressive nucleosome (nuc-1); 6) the epigenetic control of the HIV-1 promoter (histone posttranslational modifications, such as acetylation and methylation, and DNA methylation); 7) The sequestration of the

cellular P-TEFb, composed of cyclin-dependent kinase 9 (CDK9) and human cyclin T1 (cycT1), in an inactive form by the HEXIM-1 (hexamethylene bisacetamide (HMBA)-induced protein 1)/7SK snRNA (7SK small nuclear RNA) regulatory complex; 8) the sub-optimal concentration of the viral transactivator Tat, which promotes transcription by mediating the recruitment to the HIV-1 promoter of the kinase complex P-TEFb, of histone-modifying enzymes and of ATP-dependent chromatin-remodelling complexes required for nucleosomal disruption and transcriptional processivity (reviewed in (Van Lint et al., 2013)).

Previous studies have reported that latently infected memory $CD4^+$ T cells can be reactivated through various stimuli in PLWH on suppressive cART (Archin et al., 2012, Novis et al., 2013, Mann et al., 2020). When these cells come across an antigen or are exposed to specific cytokines or chemokines, proviral transcription is activated and this leads to productive infection (Chun et al., 1998b, Finzi et al., 1997). The suspension or cessation of cART is followed by HIV rebound to high viral loads (Davey et al., 1999). The source of viral rebound after treatment interruption is the latent viral reservoir and thus explains why PLWH must be on life-long treatment (Chavez et al., 2015).

2.1.2 Molecular mechanisms of HIV-1 latency establishment and maintenance

The timing which the latent viral reservoir is seeded is currently debated in the field. An earlier study reported that viral latency is established very early during HIV-1 infection (Chavez et al., 2015), while a subsequent study reported that the reservoir is seeded preceding treatment initiation (Abrahams et al., 2019). Most recently, Eric Hunter's group in 2020 published data showing that the reservoir continues to be seeded throughout infection in antiretroviral therapy naïve HIV-1 infected patients (Brooks et al., 2020). However, the mechanisms that govern viral latency at the transcriptional level are not fully understood. HIV-1 latency is maintained through multiple cellular and molecular mechanisms. Infection of activated CD4+ T cells as they are returning to a resting state to form long-lived memory T cells has been proposed as one possible mechanism that lead to latency establishment (reviewed in (Siliciano and Greene, 2011)). Heterochromatin formation impairs gene expression by impeding transcription factor access to the underlying DNA (reviewed in (Siliciano and Greene, 2011)). These mechanisms

exploit cellular and viral factors as possibly modulators of the viral promoter, 5′ LTR to suppress viral gene transcription thus resulting into viral latency (Elsheikh et al., 2019).

The efficiency of the initial transcription of integrated DNA from 5′ LTR promoter region determines the level of viral RNA in an HIV infected cell. The proviral 5′ LTR promoter contains many *cis*-regulatory elements, which modulate the rate of viral transcription initiation. However, certain cell types and the cell differentiation processes regarding diversity of cell activation signals may contribute to substantial variations in transcriptional activity of LTR (Coffin et al., 1997). It has been proposed that variations in the Tat sequence could modulate transactivation as well as have implications on HIV-1 latency and reactivation (reviewed in (Pluta et al., 2020)). Ronsard and co-workers reported that the Tat variants with an amino acid change of S46F were able to significantly enhance LTR transactivation compared with wild-type Tat (Ronsard et al., 2014). Additionally, the change of S46F caused strong Tat interaction with TAR in *in vitro* and *in silico* models. Contrastingly, a naturally occurring change of the C22S in HIV-1 Oyi strain reduced Tat transactivation activity and was linked with long-term nonprogressive infections (Peloponese et al., 1999). Furthermore, other naturally occurring polymorphisms within Tat identified in HIV-infected patients at acute and/or early infection phase (i.e., P10S, W11R, K19R, A42V, and Y47H) have been shown to significantly impair transactivation activity in the infected CD4+ T lymphocytes (Yukl et al., 2009). These data suggest that certain naturally occurring changes can change Tat transactivation activity (reviewed in (Pluta et al., 2020)).

Several latently infected cell lines have been established after HIV-1 infection, whose proviruses harbour mutations in their Tat–TAR transcriptional axis (Emiliani et al., 1996, Emiliani et al., 1998) or in their NF-κB binding sites in the HIV-1 5’ LTR promoter (Antoni et al., 1994). However, the presence of mutations in Tat–TAR transcriptional axis and NF-κB binding sites has raised questions about the significance of these lines to HIV-1 latency *in vivo* because latent cells are transcriptionally silent. Despite this valid criticism, these mutations have strengthened the concept that transcription inhibition is critical to the establishment and maintenance of HIV latency. An earlier study by Jordan et al in 2001, reported that the site of integration of the HIV-1 provirus into the cell genome impacts its basal HIV gene expression and that HIV Tat activates transcription independent of the chromatin environment (Jordan et al., 2001). Therefore, they predicted that a small fraction of integration sites might lead to such low basal promoter activity resulting in no accumulation of Tat mRNA. This provirus would

be locked in the early phase of transcription, resulting in functional latency. In a follow up study, the same group developed an *in vitro* cell system that is a reflective the state of HIV-1 latency to study how post integration latency is established (Jordan et al., 2003). They used recombinant viruses that express green fluorescent protein (GFP) to selectively enrich for rare cells blocked at the transcriptional level during an acute infection. Using this approach, they isolated many clonal cell lines harbouring HIV-1 in a latent state. The latent provirus can be reactivated at the transcriptional level in these cell lines by a variety of latency reversing agents (LRAs). This study demonstrated that HIV-1 can reproducibly establish latent infection at a low level after acute infection of T cells *in vitro* (Jordan et al., 2003).

2.1.2.1 Clonal expansion of HIV-1 infected cells contributes to HIV-1 reservoir persistence

An early clue providing more information on the nature of the HIV-1 reservoir was the observation that in contrast to the highly diverse population of viral RNA genomes present prior to therapy, the low levels of virus found in blood during cART often included viruses whose genomes had identical sequences (Bailey et al., 2006). Initially, it was not clear if these "predominant plasma clones" arose from isolated pockets of cells infected with single founder viruses (virus clones) or were the result of repeated division of cells derived from a single infected ancestor (clones of infected cells). It is now known that although a large majority of CD4+ T cells that are infected by HIV *in vivo* die within a few days of infection, a small fraction of the infected cells survive and divide, producing large clones, with important clinical consequences through what is referred to as clonal expansion (reviewed in (Coffin and Hughes, 2021))(Fig. 2.2). The long-term persistence of the reservoir depends, to a significant extent, on the clonal expansion of cells infected prior to ART (reviewed in (Coffin and Hughes, 2021)).

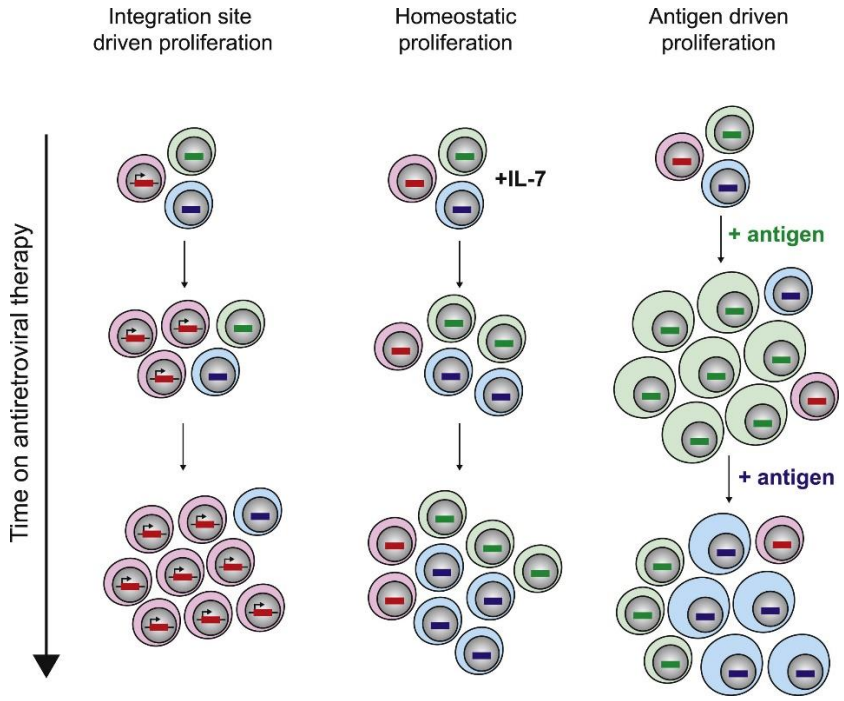

Figure 2.2: HIV-1 persistence through clonal expansion (Cohn et al., 2020). Three independent mechanisms are thought to drive proliferation of latently infected cells. First, the viral integration site may provide a survival advantage allowing preferential proliferation of the infected clone. Second, homeostatic proliferation occurs through homeostatic cytokines, such as IL-7, may signal latently infected cells to divide. Thirdly, latently infected $CD4^+$ T cells with antigen specific T cell receptors may divide in response to recurrent antigen exposure.

As aforementioned, T cell memory resides in long-lived clones of T cells with a half-life of 44 months (Siliciano et al., 2003). This confers an advantage to HIV-1 when integrated into memory T cells (reviewed in (Yeh et al., 2021)). This clonal expansion of memory CD4+ T cells is a normal immune response to antigen stimulation (reviewed in (Yeh et al., 2021)). Of note, memory CD4+ T cells undergo homeostatic proliferation faster than naïve CD4+ T cells (Macallan et al., 2019). When the antigen is removed, memory CD4+ T cells will return to a quiescent memory state which is transcriptionally inactive. Upon cognate antigen stimulation, these resting memory CD4+ T cells become activated and proliferate into a T cell clone again. Although many cells die of viral cytotoxic effects during productive HIV-1 infection, some of HIV-1 infected cells survive and follow the clonal expansion and contraction dynamics of CD4+ T cells (reviewed in (Yeh et al., 2021)).

These transcriptionally inactive cells lack active transcription factors such as NF-κB and NFAT in the nucleus and therefore do not induce effective HIV-1 transcription (Pereira et al., 2000).

By residing in these transcriptionally inactive resting CD4+ T cells, HIV-1 does not make viral antigens and therefore cannot be recognized by the immune system. Furthermore, these HIV-1-infected memory CD4+ T cells undergo homeostatic replenishment through cytokines such as interleukin (IL)-7 (Chomont et al., 2009). These homeostatic cytokines induce T cell homeostatic proliferation without inducing HIV-1 reactivation (Bosque et al., 2011, Wang et al., 2018). Therefore, during homeostatic proliferation, HIV-1-infected cells proliferate but do not express antigens are not recognized by the immune system (reviewed in (Yeh et al., 2021)). Because the HIV-1 LTR promoter has NF-κB and NFAT binding sites, antigen stimulation will also activate HIV-1 and reverse latency. However, antigen activation appears to be stochastic where T cell activation levels follow a gradient. This means that antigen stimulation, which should presumably reactivate all HIV-1 and expose the infected cells to immune clearance, can only activate a subset of HIV-1-infected cells (reviewed in (Yeh et al., 2021)). Therefore, HIV-1-infected cells persist over time by residing in CD4+ T cells, following the normal T cell expansion and replenishment responses and hiding in the quiescent state that cannot be fully reactivated despite T cell activation (reviewed in (Yeh et al., 2021)).

Fairly recent advances in technology show that the persistence of the reservoir occurs through enormous and sustained clonal expansion of cells infected with both defective and intact proviruses (reviewed in (Yeh et al., 2021)). Proliferation of these cells maintains most infected cells during cART as well as shapes the location and disposition of the provirus population (Fig. 2.2). HIV-1 infected cells persist through a dynamic process where they increase by clonal expansion, decrease their clearance by cytotoxic T lymphocyte (CTL) and natural killer (NK) cell killing (reviewed in (Yeh et al., 2021)). Clonal expansion of HIV-1 infected cells contributes to the long-term persistence of the HIV-1 reservoir in cART-suppressed individuals (Cole et al., 2021). More than half of the HIV-1 latent reservoir is maintained through clonal expansion (Bui et al., 2017, Lorenzi et al., 2016, Hosmane et al., 2017), and the number of clonally expanded cells increases over time (Wagner et al., 2014). Latent HIV infection preferentially occurs in memory CD4+ T cells (Chomont et al., 2009). These cells are not transcriptionally active and divide slowly, thus contributing to the maintenance of HIV-1 latency. These clonally expanding HIV-1-infected cells pose a major barrier to an HIV cure (reviewed in (Yeh et al., 2021)).

Despite effective cART, PLWH can still have intermittent low-level of viremia or viral blips (Nettles et al., 2005, Havlir et al., 2001). HIV-1 expressed from clonally expanded HIV-1 infected cells is the possible source of persistent low-level viremia and non-suppressible HIV-1 in plasma (Bailey et al., 2006). Siliciano's group showed that the viruses circulating in the plasma of PLWH on cART during low-level viremia are dominated by a few clones known as the predominant plasma clones (Siliciano et al., 2003).

Clonally expanded HIV-1 infected cells serve as a source of viral rebound after treatment interruptions (reviewed in (Yeh et al., 2021)). The finding that clonally expanded cells are a source for persistent low-level viremia (Simonetti et al., 2016, Halvas et al., 2020) under cART suggests that clonally expanded cells can likely be a source of viral rebound (Yeh et al., 2021). Thus, identifying the source of viral rebound after treatment interruptions is important in HIV-1 cure research.

2.1.2.2 Chromatin organization and epigenetic modifications

The nucleosome is the fundamental functional and structural repeating unit of chromatin (Chereji and Morozov, 2015). A total of 146 nucleotide base pairs (bp) of DNA are wrapped in superhelical turns around an octamer made of two copies of each histone protein H2A, H2B, H3 and H4 (Luger et al., 1997). The linking of each nucleosome core to the next occurs by a segment of linker DNA that varies from 10 to 80 bp in length. This nucleosomal array assemblies further into higher-order condensed structures, which are stabilized by the histone H1 to each nucleosome and to its adjacent linker (Felsenfeld and Groudine, 2003). During the early prophase, homogenously distributed chromatin of interphase begins to form visible thread-like structures of chromatin. This phenomenon is referred to as chromatin condensation (Antonin and Neumann, 2016). The degree of condensation of the chromatin fibre dictates its functional accessibility to protein machineries responsible for processes such as transcription (Zhu and Li, 2016). Euchromatin is an open, less compact and accessible state, whereas heterochromatin is a highly condensed and inaccessible state of DNA. The dynamic switches between heterochromatin and euchromatin, known as chromatin remodelling, thus play a crucial role in gene expression (Trojer and Reinberg, 2007).

Chromatin remodelling complexes regulate gene activity by changing the structure and accessibility of the gene promoter by transcription factors (He et al., 2002). The changes relax the chromatin structure results in the recruitment of transcription factors (TFs) to the site of gene transcription, which removes or repositions the nucleosomes near transcription start sites (TSS) (He et al., 2002). Furthermore, enzymes that add or remove covalent modifications to histones tails can create a histone "code" that affects the structure of the chromatin and creates docking sites for additional regulators (Strahl and Allis, 2000). Changes in chromatin structure and accessibility can then allow greater access to the promoter by core transcriptional regulators including RNA polymerase, thereby facilitating transcription (Jefferys et al., 2021).

Epigenetic processes are key elements in the silencing of HIV-1 gene expression (reviewed in (Verdikt et al., 2021)). During HIV-1 infection and after reverse transcription of the viral RNA genome to form double stranded viral DNA, the vast majority of viral double-stranded DNA remain mainly as linear DNAs (Pang et al., 1990, Sharkey et al., 2000). These linear HIV-1 DNA molecules act as precursors for integration in the host genome and they are also susceptible to multiple other fates (Sloan and Wainberg, 2011). All other unintegrated DNA products are circular in form. After the virus enters the cell, the HIV-1 reverse transcriptase (RT) converts viral RNA into cDNA in the reverse transcription complex (RTC), a nucleoprotein complex derived from the core of the infecting virion (reviewed in (Suzuki and Craigie, 2007)). Thereafter, the newly synthesized viral cDNA is enzymatically processed by integrase (IN) and transitions from the RTC to pre-integration complex (PIC), a large nucleoprotein complex containing both viral and cellular proteins (reviewed in (Suzuki and Craigie, 2007)). After the HIV-1 PIC enters the nucleus, the viral DNA must become integrated into chromosomal DNA of the host for productive infection to occur. However, the viral DNA can also undergo several circularization events that do not encourage subsequent replication and represent dead ends for the virus (Farnet and Haseltine, 1991).

Following dissociation of the PIC proteins, the two ends of the viral DNA can be ligated to each other, to yield 2-long long terminal repeat (LTR) (2-LTR) circles. Inactivation of the host cell nonhomologous DNA end-joining (NHEJ) components Ku70/ 80, ligase IV, and XRCC4 blocks 2-LTR circle formation, thus implicating these factors are involved in the circularization reaction (Li et al., 2001, Jeanson et al., 2002). Circles with one LTR (1-LTR) copy can also be detected. These can be formed either by recombination between the LTRs within the nucleus,

possibly involving action of the cellular MRN complex (Mre11, Rad50, and NBS1) (Kilzer et al., 2003), or as stalled products of reverse transcription that failed to complete the final steps of strand displacement synthesis (reviewed in (Hu and Hughes 2011)). 2-LTR circles are mediated in the nucleus as a protective host response to the presence of double stranded DNA (Bukrinsky et al., 1991).

Integration of a DNA copy of the viral genome into a host cell chromosome is an essential step in the retroviral replication cycle (Coffin et al., 1997). Once integrated, the proviral DNA is replicated along with cellular DNA during cycles of cell division, as with any cellular gene. The provirus serves as the template for transcription of viral RNAs. Some viral RNAs are translated to yield the viral proteins, whereas a portion of the full-length viral RNA is recruited to serve as genomic RNA in progeny virions (reviewed in (Craigie and Bushman, 2012)).

Integration is mediated by the virus-encoded integration (IN) protein, which is introduced into cells during infection along with reverse transcriptase, the viral RNA, and other proteins as a part of the viral core. After the viral DNA is synthesized by reverse transcription in the cytoplasm, it stably associates with IN and other proteins as a high-molecular-weight nucleoprotein complex that is later transported to the nucleus for subsequent integration (reviewed in (Craigie and Bushman, 2012)).

2.1.2.3 Transcriptional interference

Replication competent HIV-1 proviral genomes can be silenced in both activated and resting cells. The most likely cause of this lack of HIV-1 gene expression, in resting cells, is transcriptional interference (reviewed in (Cary et al., 2016)). In this scenario, HIV-1 is its own worst enemy. Retroviruses have identical LTRs on either end of their DNA genome, one for initiation of transcription (5′ LTR), the other for termination of transcription (3′ LTR) (Cullen et al., 1984). In the case of HIV-1, the 5′ LTR initiates transcription 20 times more often than the 3′ LTR, which is occluded by the elongating RNAPII (reviewed in (Cary et al., 2016)).

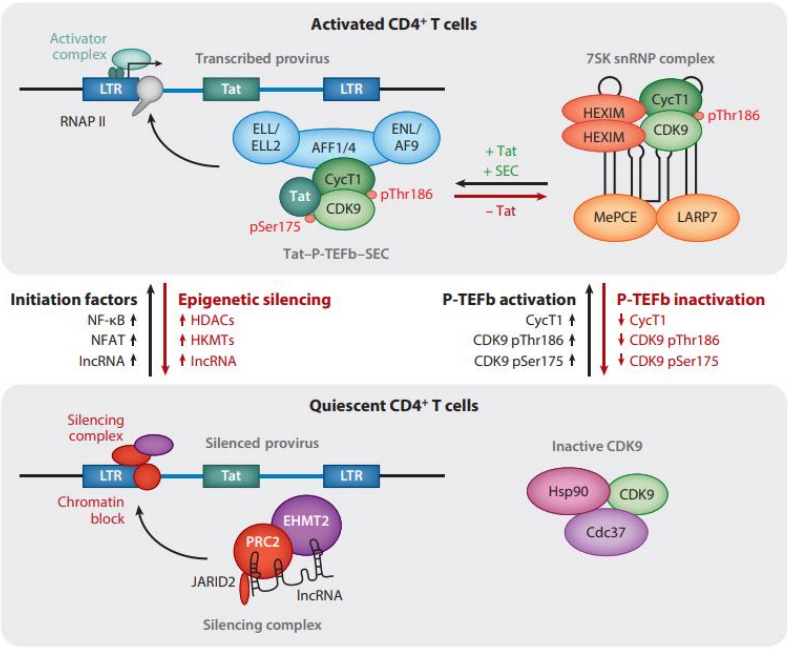

Figure 2.3: Diagram showing the main molecular pathways regulating HIV-1 latency (reviewed in (Mbonye and Karn, 2017)). In CD4+ T cells, most P-TEFb molecules are sequestered into 7SK snRNP, which contains 7SK snRNA, HEXIM1, MePCE, and LARP7. In 7SK snRNP, the CycT1 subunit directly binds to the central loop of 7SK snRNA and HEXIM1, which inhibits the kinase activity of CDK9. Various stimuli, including stress, environmental stimuli, cytokine signaling, PKC activation, and treatment of cells with HDACis, BETis, and other LRAs, release P-TEFb and stimulate CDK9 kinase activities accompanied by phosphprylation of critical T-loop (Thr 186) and central serine (Ser 175) residues. Released P-TEFb, which is now free, can subsequently be recruited to RNAPII early elongation complex paused at promoter proximal regions of many cellular genes by various factors such as transcription factor, Brd4, super elongation complex (SEC), Mediator complex, etc. HIV Tat protein can directly recruit P-TEFb to RNAPII on HIV LTR via binding with viral TAR RNA, since it can complete with HEXIM1/7SK snRNA for P-TEFb.

Transcriptional interference refers to the occlusion or dislodgement of transcription initiation or elongation complexes from the provirus by readthrough transcription of the viral-containing host gene or termination of host-initiated transcripts at the polyadenylation site in the 5′ LTR inactivates the provirus (Han et al., 2008) (Fig. 2.3). HIV-1 transcriptional repression is crucial to the establishment and maintenance of post-integration latency (reviewed in (Van Lint et al., 2013)). Transcriptional interference has been proposed to explain HIV-1 promoter repression when integrated into introns of highly expressed genes (reviewed in (Van Lint et al., 2013)).

Transcriptional interference occurs because of the low-affinity binding between the DNA at the 3′ LTR and transcription factors such as SP1, TFIID, and the initiator element (Lenasi et al., 2008, Greger et al., 1998). The elongating RNAPII from the 5′ LTR displaces these TFs from proviral DNA at the 3′ LTR, then continues to the polyA site in the R region and terminates transcription (Cary et al., 2016). Of note, when the virus integrates in the same orientation as the host gene, the elongating RNAPII from the host gene terminates at the polyA site in the 5′ LTR, curtailing HIV-1 transcription (Lenasi et al., 2008, Greger et al., 1998, Ashe et al., 1995, Han et al., 2008). In this case, the 3′ LTR is not occluded and initiates transcription, producing sterile transcripts that contain TAR. When the virus integrates in the opposite orientation, RNAPII copies HIV-1 antisense transcripts, ignoring both polyA sites in the HIV sense orientation and producing long hybrid mRNA species (reviewed in (Cary et al., 2016)). If the provirus integrates in introns of host genes, the HIV-1 antisense transcripts are spliced out and rapidly degraded. Nevertheless, HIV antisense transcripts have been detected (Landry et al., 2007), and they could provide an estimate of transcriptional interference in latently infected cells (reviewed in (Cary et al., 2016)).

In double helix DNA one strand of DNA is called the sense strand because when you read it in the right direction it provides the code to make a protein. In two-stranded DNA, the sense strand is bonded to an opposite DNA strand that is called the antisense or noncoding strand. During the production of proteins, the anti-sense strand becomes the template for messenger RNA (mRNA). Interestingly, low-level antisense transcription takes place at the 3′ LTR, a mechanism by which latency can be maintained (Kobayashi-Ishihara et al., 2012, Saayman et al., 2014). Sense transcription results in at least 40 coding transcripts due to alternative splicing of the HIV-1 genome (reviewed in (Karn and Stoltzfus, 2012)). Finally, both LTRs also act as a source of negative sense transcription, which could potentially affect the expression of neighbouring genes (Peeters et al., 1996, Bentley et al., 2004). Persistence of transcriptionally silent replication competent HIV-1 is a major barrier to clearance of the virus from patients (reviewed in (Sengupta and Siliciano, 2018)).

2.1.2.3 The potential role of Tat in HIV latency

HIV-1 relies on both cellular and viral factors for efficient transcription of its genome. HIV-1 transcription driven by the viral promoter, 5' LTR is characterized by two phases. The first or early phase, known as basal transcription occurs immediately after integration relying solely on cellular transcription factors and results mainly in short abortive viral mRNA transcripts (Roebuck and Saifuddin, 1999). However, several transcripts elongate throughout the viral genome, and this results in transcription of the *transactivator* of *transcription* (*tat*) gene that gets translated into the HIV-1 Tat protein (referred to as Tat, in this thesis) in the cytoplasm, which is produced early during infection (Das et al., 2011). Newly translated Tat enters the nucleus to activate RNA pol II (reviewed in (Karn and Stoltzfus, 2012)). The second phase or transcriptional elongation occurs when sufficient Tat protein has accumulated in the nucleus to bind and enhance HIV-1 LTR transcription activity (Das et al., 2011). This has been referred to as Tat-mediated transcription, as it enhances viral gene transcription (Das et al., 2011).

The Tat stimulates transcription from the viral 5′ LTR promoter and promotes RNA polymerase II (RNAP II) elongation (Brady and Kashanchi, 2005)(Fig. 2.4). During HIV-1 replication, Tat binds to the TAR RNA element found in the 5' LTR and recruits host transcription factors, namely P-TEFb. The Tat-P-TEFb complex binds the TAR element on the RNA, which leads to increased processivity of RNAPII. Apart from this interaction with P-TEFb, Tat also recruits histone acetyltransferases (HATs) like CNP/p300 complex to the viral promoter to activate the acetylation of nucleosomes to promote transcription of HIV-1 RNA (reviewed in (Nekhai and Jeang, 2006, Vardabasso et al., 2008). The establishment of latency or generation of viral reservoirs is mainly due to several cellular processes that restrict access of the transcription machinery to the HIV-1 promoter region (reviewed in (Karn, 2011)). Latency also can be established by proviruses that carry transactivation-defective *tat* sequence (Tat-C22G) that has been shown to prevent effective expression of viral proteins (Wang et al., 1996), which eventually represses Tat expression and/or function (reviewed in (Ajasin and Eugenin, 2020)). Over the past ten years, studies have demonstrated that latently HIV-infected cells can be reactivated by adding Tat and one of the proposed cure strategies against preventing reactivation of latently infected cells is to repress the provirus and prevent access of Tat to the promoter region of the provirus by using Tat inhibitors (Mbonye and Karn, 2011, Mbonye and

Karn, 2017, Kamori and Ueno, 2017, Siliciano and Greene, 2011, Deeks, 2012, Donahue et al., 2012, Asamitsu et al., 2018, Khoury et al., 2018) (Fig. 2.4).

Therefore, HIV-1 transcription step is an important target for pharmaceutical intervention for HIV-1 cure development purposes. Although HIV-1 transcription step has been well characterised, it remains unclear how the occurrence of latent infection blocks this step of the HIV-1 replication cycle.

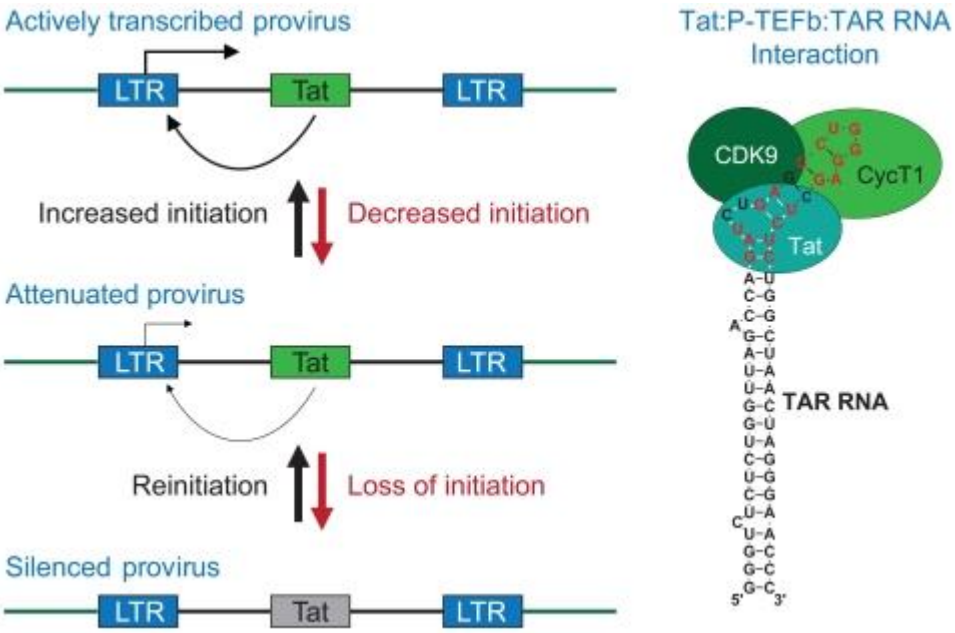

Figure 2.4: Diagram showing the autoregulation of HIV-1 transcription by Tat (reviewed in (Mbonye and Karn, 2011)). Left panels show efficient transcription elongation from the HIV-1 LTR is dependent on Tat. Minor changes in initiation efficiency, due to transcriptional interference or epigenetic silencing, reduce Tat levels in the cell and disproportionately inhibit transcription. This drives the HIV-1 provirus into latency. Re-initiation stimulates Tat production and restores full transcription efficiency. Right panels show recognition of TAR RNA by Tat and P-TEFb. The red bases in TAR are recognized by Tat in the TAR bulge region, and by CycT1 in the TAR loop region.

Suboptimal concentrations or post-translational modifications of Tat hamper Tat-induced transcription elongation. For instance, phosphorylation of Tat by CDK2 results in inhibition of transcription (Ammosova et al., 2006). Moreover, in latent cells P-TEFb is retained in an inactive form by hexamethylene bisacetamide-induced protein (HEXIM-1) and 7SK small nuclear RNA (7SK snRNA) (Yang et al., 2001, Yik et al., 2003)(Fig. 2.3). Tat and

bromodomain containing protein 4 (BRD4) can release P-TEFb by disrupting the inactive complex (Jang et al., 2005). BRD4 represses HIV-1 transcription by competing with Tat for the P-TEFb binding site (Bisgrove et al., 2007).

Previous studies demonstrated that Tat is responsible for directly activating viral transcription in the patient-derived latently infected resting memory $CD4^+$ T cells with no cellular activation (Lin et al., 2003, Lassen et al., 2006). This is also supported by the Jurkat model of latency showing that the introduction of exogenous Tat was sufficient to reactivate most of the latently infected population (Donahue et al., 2012). Similarly, HIV-1 latently infected cells, at least in Jurkat cells, can be reactivated by cellular superinfection in a Tat-dependent manner (Donahue et al., 2013). Moreover, both experimental and computational methods have revealed that Tat is more effective than cellular activation approaches in reactivation of full-length transcription of latent HIV-1 (reviewed in (Kamori and Ueno, 2017)).

Thus, studies which investigate the transcriptional dynamics of HIV-1 are of major importance to eradicating the viral reservoir and in turn edging us closer to an eradicating HIV-1 cure. Understanding the molecular mechanisms of how latency is established and maintained will also be critically important to developing strategies to prevent or eliminate the latent reservoir.

2.2 Different strategies have been explored to eradicate HIV-1 infected cells

Although cART potently inhibits HIV-1 replication, it does not completely eradicate the virus. As aforementioned, HIV-1 persists in cellular and anatomical reservoirs that show minimal decay during cART (Blankson et al., 2002). Many studies conducted during the past 20 years have shown that HIV-1 persists in a small pool of cells harbouring integrated and replication-competent viral genomes (Finzi et al., 1997, Chomont et al., 2009). Most of these cells do not produce viral particles and constitute what is referred to as the latent reservoir of HIV-1 infection (Dufour 2020). HIV-1 latency is maintained through the mechanisms described above in section 2.1.2 such as clonal expansion, epigenetic silencing and transcriptional repression, led to the development of approaches or strategies to target viral latency with the aim of curing HIV-1 infection (Fig. 2.5)

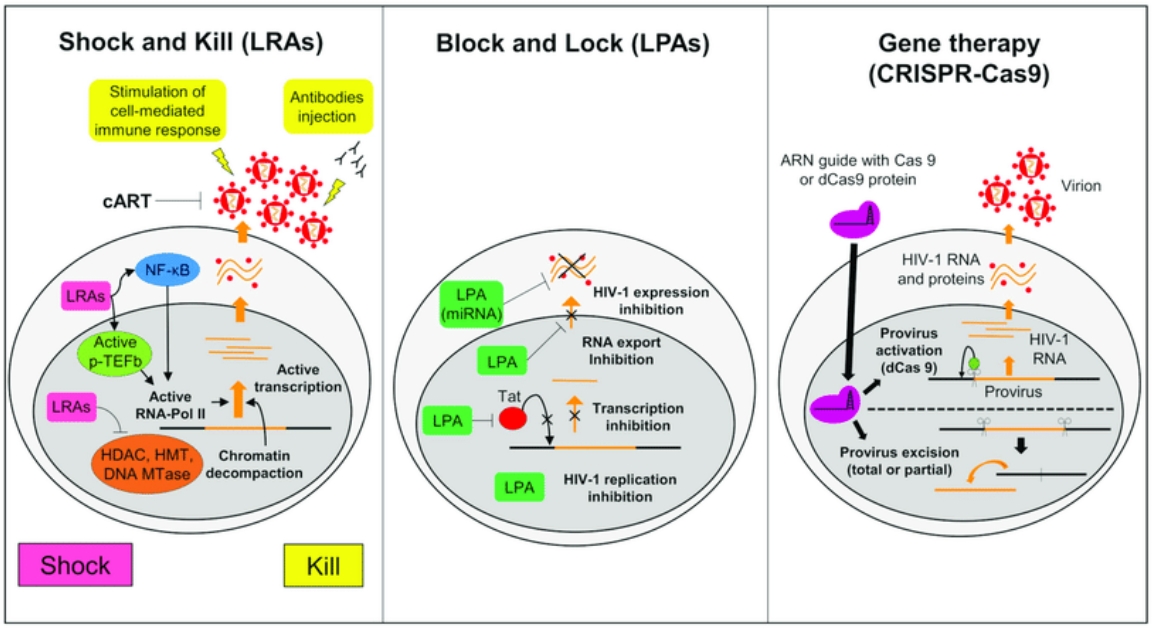

Figure 2.5: Strategies targeting HIV-1 latency (Wallet et al., 2019). The shock and kill strategy relies on latency reversing agents (LRAs) to reactivate HIV-1 transcription in a latently infected cells so that infected cells may be recognised and cleared by host immune response while the produced viruses may be killed through exposure to cART. The block and lock strategy uses latency promoting agents (LPAs) to inhibit Tat and force the cell into a deeper state of latency. Gene therapy using CRISPR-Cas9 gene editing to excise part of the provirus to produce defective virions.

As a result, over the past few years there has been an emergence of strategies to target latency in HIV-1 infected cells. The approaches or strategies that are being explored towards the development of HIV-1 cure fall into two categories, the "sterilizing cure" and the "functional cure." Briefly, the sterilizing cure seeks a total elimination of replication-competent proviruses from a latently infected cell (Martin and Siliciano, 2016). On the other hand, the functional cure, seeks to force the virus to go into deep latency results in cART remission (Kessing et al., 2017). The strategies to achieve these approaches will be discussed below.

2.2.1 Sterilizing versus functional cure

Combined antiretroviral therapy (cART) has enabled the sustained control of viremia to low levels even below limit of detection and prolong the lifespan of PLWH who have access to treatment (Deeks et al., 2013). However, as aforementioned the key obstacle to an HIV-1 cure development is the existence of latent reservoirs comprising of mainly resting memory CD4+ T cells infected with transcriptional dormant viruses (Chun et al., 1998a).

A sterilizing cure refers to the complete elimination of replication-competent HIV-1 proviruses in the body (Xu et al., 2017). The functional cure refers to the long-term control of HIV-1 replication, which involves maintaining a normal CD4+ T cell count and HIV-1 replication below a detectable level (Kmietowicz, 2013), without cART. “Elite controllers” are those patients whose HIV-1 RNA remains below the clinical baseline or limit of detection for a long period without cART. “Elite controllers” studies are expected to provide important clues for the development of therapies or strategies for functional HIV-1 cure, such as therapeutic vaccines and vector-mediated gene transfer therapy (Okulicz and Lambotte, 2011, Cockerham and Hatano, 2015).

This section will describe the “Shock and Kill” strategy which has been proposed to achieve a sterilizing cure, followed by the “Block and Lock” strategy which has been proposed to achieve a functional cure.

2.2.2 The “shock and kill” strategy

One strategy currently under development to target HIV-1 latency to achieve a sterilizing cure is the “Shock and Kill” strategy. The latent HIV-1 is reactivated by latency reversal agents (LRAs) in the ‘shock’ part and then followed by an immune clearance of infected cells in the ‘kill’ part (reviewed in (Kim et al., 2018)). The aim is to eliminate latently infected cells and activate virus transcription, protein expression and virion production using these LRAs, potentially triggering cytolysis or immune-mediated clearance (Deeks, 2012). Latency reversing agents include histone posttranslational modification modulators, toll-like receptor (TLR) agonists, NF-κB stimulators, nonhistone chromatin modulators, extracellular stimulators, and mechanistic target of rapamycin kinase (MTOR) activators (reviewed in (Abner and Jordan, 2019)). Other LRAs which can increase viral gene expression from latency *in vitro*, include: histone deacetylase inhibitors (HDACi), histone methyltransferase (HMT) inhibitors, DNA methyltransferase inhibitors, bromodomain inhibitors, protein kinase C (PKC) agonists (reviewed in (Rasmussen and Lewin, 2016)), as well as PI3K/Akt pathway inhibitors that affect cell survival and agonists for the innate immune receptors TLR7 or TLR9 (Tsai et al., 2017, Offersen et al., 2016). Activation of the integrated provirus results in an increased expression of viral genes and production of virions so that the reactivated cells will be killed either through immune mediated clearance or HIV-1-mediated cytolysis (Campbell and

Spector, 2022). This approach aims to clear the majority of resting CD4+ T cells which contribute to the HIV-1 latent reservoir (Campbell and Spector, 2022).

However, the efficacy of LRAs has not been substantial enough to show a reduction in the size of the HIV-1 reservoir (Yukl et al., 2018). Moreover, the evidence from clinical trials has also confirmed this (Grau-Exposito et al., 2019, Fidler et al., 2020, Gramatica et al., 2021). Only a number of these agents have been taken to clinical trials, however with little success. For example, although the histone deacetylase (HDAC) inhibitors (HDACi) panobinostat and romidepsin both increase total and elongated HIV-1 transcripts, its effect on polyadenylated or multiply spliced transcripts is inadequate (Yukl et al., 2018). Polyadenylation is crucial for HIV-1 RNA nuclear export, stability as well as translation into HIV-1 proteins, which are essential for both productive infection and T cell recognition and killing (Fasken and Corbett, 2009, Wilusz, 2013). It is likely that although these agents increase cell-associated HIV-1 RNA, this will be insufficient to induce productive infection, immune-mediated clearance, and/or HIV-1-mediated cytolysis (Campbell and Spector, 2022). Moreover, both panobinostat and romidepsin, as well as the popularly studied, vorinostat, which are also HDACi, can suppress the CTL response and impair their ability to kill HIV-1-infected CD4+ T cells (Jones et al., 2014). *In vitro* studies also suggest that HDACi suppress the interferon gamma (IFN-γ) response to HIV-1 antigens, impair natural killer cell function, and are potentially cytotoxic over long durations (Lucera et al., 2014, Zhao et al., 2019). These data indicate that reactivation of latently infected cells with LRAs alone is unlikely to lead to elimination of the infected cells and an additional intervention for the "kill" is required (Campbell and Spector, 2022).

Currently and of relevance to this study, a combination strategy to reactivate latent HIV-1 that targets both NF-κB and P-TEFb signalling pathways was shown to be an effective combination strategy (Jiang et al., 2015, Darcis et al., 2015). Importantly, the effect of the ingenol-3-angelate (PEP005) was synergistic with JQ1, a P-TEFb activator. PEP005 is member of the class of anti-cancer ingenol compounds and effectively reactivated HIV from latency in primary CD4+ T cells from HIV infected individuals receiving ART. Importantly, a combination of PEP005 and JQ1, a P-TEFb agonist, reactivated HIV from latency at level on average 7.5-fold higher compared to PEP005 alone. This combination was highly potent in reactivating latent HIV-1 expression both *in vitro* and *ex vivo* (Jiang et al., 2015).

It is crucial to consider the sites of HIV-1 replication persistence in patients whilst adhering to cART to develop a cure (Mediouni et al., 2022). A prominent source of viral rebound is the reservoir of long-lived resting memory CD4+ T cells that harbour replication-competent HIV-1 proviruses and does not raise alarm to host immune surveillance (Xu et al., 2017, Sadowski and Hashemi, 2019). Cell-to-cell HIV-1 spread and the homeostatic proliferation of infected T cells due to chronic inflammation are further drivers of persistent HIV-1 replication during cART (Siliciano et al., 2003). Thus, a reasonable HIV-1 cure strategy would aim to purge the patient of latently infected cells. Although clinical trials seek to minimize latent HIV-1 reservoirs, a significant reduction of reservoir size has not been observed in patients yet, indicating that the reactivation of latent HIV-1 reservoirs alone does not necessarily cause a reservoir clearance (Archin et al., 2012, Elliott et al., 2014, Rasmussen et al., 2014, Sogaard et al., 2015).

However, there is an explanation for this: Firstly, the reactivation process *in vivo*, by current LRAs may not be strong enough due to the low potency to reactivate latent cells Therefore, new potent LRAs need to be developed. In the last few years, numerous LRAs have been developed, including epigenetic compounds such as HDAC inhibitors, HMT inhibitors, bromodomain inhibitors, protein kinase C agonists, and TLR7 agonist (Jiang and Dandekar, 2015, Lim et al., 2018). Due to multiple signalling pathways being involved in HIV-1 latency establishment, a combination therapy was proposed in order to target one or more pathway simultaneously (Margolis and Hazuda, 2013).

The well-known “Berlin patient” represents one successful case of a sterilizing cure. Twelve years ago, the first known patient to be cured of HIV-1 infection was Timothy Brown (referred to as the Berlin Patient) after receiving a bone marrow transplantation from a donor with stem cells harbouring a 32 nucleotides deletion in the HIV-1 core-receptor, CCR5 (termed CCR5 delta (Δ) 32) (Hutter et al., 2009). The CCR5Δ32 mutation was first discovered in 1996 (Dean et al., 1996, Liu et al., 1996), as a genetic mutation which protects cells from HIV infection (Barmania and Pepper, 2013). Genetic analysis of the open reading frame (ORF) of the gene by Liu et al. showed a deletion of 32 base pairs from nucleotides 794 to 825 (Liu et al., 1996). The mutant allele contains 215 amino acids in comparison to the full-length 352 amino acid wild type CCR5. Samson et al. found that the region affected was the second extracellular loop (Samson et al., 1996). The subsequent protein lacked the last three transmembrane domains as

well as regions important in G-protein interaction and signal transduction. Both groups discovered that CD4+ cells with the CCR5Δ32 prevented HIV-1 envelope fusion (Barmania and Pepper, 2013). More recent studies showed that, CCR5Δ32 appears to reduce protection against some other viral infections such as influenza and West Nile virus (Falcon et al., 2015, Lim and Murphy, 2011). Furthermore, it was estimated that an individual who is homozygous for the Δ32 allele has a 21% increase of mortality rate (Wei and Nielsen, 2019).

Following bone marrow transplantation, Mr Brown showed long-term cART-free HIV-1 control indicating the restoration of the immune system's ability to combat infection. Timothy Brown had received a stem cell transplant from a donor with CCR5Δ32. Ultrasensitive assays including the measurement of cell-associated HIV DNA in PBMC and whole transcriptome sequencing which allows the characterization of all types of RNA transcripts, showed that the "Berlin patient" achieved at least a 7500-fold, or nearly four log, reservoir reduction, and he has remained without detectable virus for nine years after transplant (Spragg et al., 2016). This is what is largely referred to as a 'functional cure' in the field. Following this huge achievement, there have been several failed clinical attempts to duplicate this therapy (Elsheikh et al., 2019). These studies demonstrated that, while effective for cure of HIV-1, bone marrow transplantation is a risky procedure and is not tolerated by most patients. Therefore, it is not a scalable treatment (Peterson and Kiem, 2019). In 2019, a second patient (referred to as the London Patient) also showed the same outcome using the same strategy: bone marrow transplantation and very recently the "City of Hope" patient presented at the AIDS 2022 conference.

However, this strategy of HIV cure has not be reproducible in other patients. It will be impossible to use curing 38 million PLWH where a staggering 25% of them are from the sub-Saharan Africa. There has been speculation that some individuals with a natural control of HIV infection with viral-RNA-suppressed <50 c/mL plasma (named "elite" controllers) may be able to achieve functional cure, in which HIV is controlled without the requirement for ART (Jiang et al., 2020, Blazkova et al., 2021). These individuals are considered an ideal example of durable HIV control and have the potential to provide substantial insight into the "natural" mechanisms of functional cure and sustained HIV remission.

2.2.3 The "Block and lock" strategy

The "Block and lock" strategy conclusively seeks to achieve a functional cure, which is a strategy that aims to permanently silence the latent reservoir using latency promoting agents (LPAs) to "block" virus transcription and "lock" the virus promoter in a latent state via repressive epigenetic modifications (reviewed in (Vansant et al., 2020)). With permanent control of the HIV-1 promoter, cART is no longer required (reviewed in (Ahlenstiel et al., 2020)). This approach mimics naturally occurring viral latency by inducing a state of latency, to achieve a "super latency" or "deep latency" (reviewed in (Ahlenstiel et al., 2020)). This is one of the reasons why this approach has been described as more realistic, feasible and has longevity in controlling HIV infection in the absence of cART, However, understanding the process of HIV-1 transcription and the molecular mechanisms involved in regulating HIV-1 latency is important for developing targeted therapies (reviewed in (Ahlenstiel et al., 2020)). A few of the mechanisms proposed in achieving this functional cure are discussed in this section.

Methods used in silencing post-integration of HIV-1 genome, target the trans-regulation mechanisms that suppress viral gene expression by inhibiting both viral and cellular transcription factors such as Tat, P-TEFb, and NF-κB (Barboric et al., 2001, Li et al., 2012, Boehm and Ott, 2017, Neogi et al., 2012). Didehydro-cortistatin A (dCA), the equipotent analog of cortistatin A drug, was developed to inhibit Tat mediated transcription by interacting with the TAR domain of Tat thus preventing Tat from binding to the TAR element and transactivating the 5' LTR (Li et al., 2019, Mediouni et al., 2019). In 2017, the Valente lab used patient-derived cell models and bone marrow/liver/thymus (BLT) mouse latency models to demonstrate that treatment with dCA delayed viral rebound and reduced rebounded viremia (Kessing et al., 2017). The BLT mice were co-treated with dCA and antiretroviral therapy (ART) for four weeks prior to treatment interruption. After ten days, all eight control mice displayed viremia, while the dCA treated mice showed a viral rebound only at day 19. Moreover, dCA induced a high nucleosomal occupancy at the Nuc-1 region of the HIV-1 LTR promoter potentially explaining its long-term effects (Kessing et al., 2017). As a follow up study in 2019, the same group elaborated on this mechanism by showing that dCA promotes tight nucleosome/DNA association by increasing deacetylated histone 3 occupancy at Nuc-1 (Li et al., 2019). Furthermore, dCA enhanced the recruitment of the repressive BAF complex while the activating chromatin remodelling complex PBAF was inhibited. In line with these results, less RNAPII was detected at the transcription start site, even upon stimulation with

LRAs. The lack of effect of dCA on Tat-TAR defective proviruses confirms the specificity of dCA for (Li et al., 2019). Collectively, these results show that dCA inhibits Tat-dependent transcription and induces a repressive epigenetic landscape that hampers HIV-1 reactivation upon treatment interruption (Vansant et al., 2020).

Another inhibitor targeting Tat is the Nullbasic. It was identified by the Harrich laboratory as NullBasic in 2009, which was the first Tat inhibitor to induce a block and lock HIV-1 silencing effect (Meredith et al., 2009). This is Tat mutant with 101 amino acids that has an altered basic domain (amino acids 49–57) where wild-type Tat basic domain (RKKRRQRRR) is replaced with the amino acid sequence GGGGAGGG. Thus, the basic domain, including the TAR binding region, becomes mutated, hence the name NullBasic. HIV-1 transcription inhibition occurs by NullBasic through competition with endogenous Tat (Meredith et al., 2009). In vitro studies have reported that CD4+ T cells transduced with a retroviral vector expressing NullBasic showed suppression of virus transcription and replication (Jin et al., 2016). A more recent in vivo study using retroviral vector delivery of NullBasic to primary human CD4+ T cells and engraftment in a NSG mouse model showed undetectable viral RNA in plasma samples up to day 14 post-infection as well as significantly reduced viral RNA levels in tissue-derived CD4+ T cells (Jin et al., 2019). Although there was no difference in viral mRNA levels at later time points, there were increased levels of CD4+ T cells in NullBasic treated mice, thus suggesting a survival advantage (Jin et al., 2019). NullBasic shows potential as a gene therapy candidate and warrants further investigation to optimize the permanence of silencing (reviewed in (Ahlenstiel et al., 2020)).

Previous studies have shown that viral transcription can be stopped by disrupting P-TEFb formation (Pisell et al., 2001, Rice, 2016). Approaches to manipulate P-TEFb for transcriptional inhibitors' development are ongoing and include the inhibition of CDK9 kinase activity, neutralizing Cyclin T1 or its interaction with Tat, shifting P-TEFb equilibrium, changing P-TEFb protein levels and modulating the interaction between P-TEFb and its recruitment factors (Fujinaga, 2020, Li et al., 2016). The P-TEFb activity will be described in detail in the following sections of this thesis. Thus far, P-TEFb inhibitors lack specificity to HIV transcription, and often result in undesired toxicity (Tahirov et al., 2010). Since P-TEFb functions depend on the kinase activity of CDK9, targeting CDK9 has been extensively studied. The most characterized first-generation CDK9 inhibitor is Flavopiridol, which competes with

ATP for CDK90 s catalytic site at low nanomolar concentrations (Ali et al., 2009), inhibiting HIV Tat-transactivation.

To this effect, the idea of extinguishing HIV-1 replication through the permanent deactivation of provirus transcription through the block and lock strategy, has gained momentum over the last decade due to promising *in vitro* and *in vivo* results from multiple studies (Ramskold et al., 2009, Rands et al., 2014, Kessing et al., 2017, Jiang et al., 2020, Mousseau et al., 2012). This type of functional cure has been observed in individuals who have been termed post-treatment controllers (Saez-Cirion et al., 2013). These individuals interrupted their cART and have not observed viral rebound (Saez-Cirion et al., 2013). In these cases, individuals received cART very soon after infection, which is atypical, given that many individuals are not immediately aware of their HIV-1 status following exposure. Post-treatment controllers are rare, about 5 to 15 percent of PLWH (reviewed in (Vansant et al., 2020)), and it is not fully understood how these individuals maintain viral suppression in the absence of cART (reviewed in (Mediouni et al., 2022)).

However, some controllers have proviruses in deep transcriptional dormancy by encapsulating their proviruses in heterochromatin regions (Ramskold et al., 2009). The "block and lock" approach seeks to mimic this viral suppression, in the first instance using novel small molecules to epigenetically silence HIV-1 replication, followed by treatment interruption thus allowing for a cART-free life. To expand on the "block-and-lock" type of approach by targeting viral or either host transcriptional factors and chromatin regulators, it is important to have a deeper understanding of the mechanisms that regulate HIV-1 latency development and reactivation in cells that comprise latent reservoir (reviewed in (Mediouni et al., 2022)).

In 2010, structure-based drug design in the Debyser laboratory, identified the first small molecule inhibitors of the interaction between HIV-1 integrase (IN) and the cellular chromatin-thethering factor LEDGF/p75 (Christ et al., 2010). Inhibitors belonging to this class of antivirals, named 'LEDGINs', are unique due to their multimodal mechanism of action affecting both early and late stages of HIV-1 replication. Although initially the idea of using LEDGINs for a functional HIV-1 cure arose when the Debyser lab started investigating their effect on integration sites and latency in 2016 (Vranckx et al., 2016, Debyser et al., 2018). A more recent study Vansant et al. showed that infection of cells with virus produced in the

presence of LEDGINs also resulted in provirus with a more latent phenotype (Vansant et al., 2019).

The suppression of virus replication through the induction of epigenetic silencing or transcriptional gene silencing (TGS) in the HIV-1 promoter is a common feature of all block and lock strategies (reviewed in (Ahlenstiel et al., 2020)). RNA induced epigenetic silencing uses short interfering (si) or short hairpin (sh) RNA to maintain the repressive heterochromatic landscape at the HIV-1 5' LTR promoter resulting in the silencing of transcription (Vansant et al., 2020). Kehellers group designed two siRNAs named 143 and Prom A, which target transcription factor binding sites in the LTR promoter (Suzuki et al., 2008, Ahlenstiel et al., 2015). siRNA 143 binds upstream of Nuc-0 where AP-1 (activator protein 1) and COUP (chicken ovalbumin upstream promoter) transcription binding sites are found. siRNA Prom A targets a unique NF-κB binding site situated between Nuc-0 and Nuc-1. These siRNAs epigenetically silence HIV-1 transcription by recruiting Argonaute 1 (AGO1), histone deacetylase 1 and histone methyl transferases (Mendez et al., 2018). AGO1 is an essential component of the RNA-induced silencing complex (RISC) that binds siRNAs and cleaves the mRNA, through a process called RNA interference (RNAi). Both siRNAs reduced reactivation of the latently infected J-Lat cells by two to three-fold when challenged by different LRAs (Mendez et al., 2018). Therefore, transcriptional gene silencing by siRNAs might be useful in a HIV-1 block-and-lock functional cure via a gene therapy application. The siRNAs could be delivered to cART-treated patients via retroviral vector transduced autologous CD4+ T cells or CD34+ cells in the absence of cART (Ahlenstiel et al., 2015). However, further extensive preclinical evaluation is required.

2.2.4 Gene editing, HIV-1 proviral genome excision by CRISPR/Cas9

Another emerging strategy for HIV-1 cure is to disrupt proviral DNA by gene editing. The disruption of CCR5 by the CRISPR/Cas9 system has been a widely pursued strategy (Spragg et al., 2016). Studies have shown high efficiency of CCR5 gene disruption using both lentiviral delivery to CD4+ T cell lines (Wang et al., 2014) and chimeric adenoviral vector delivery to primary T cells (Li and Lieber, 2019). The bacterial CRISPR/Cas9 system has two elements: the nuclease protein Cas9, which cuts double-stranded DNA, and a single guide RNA (sgRNA)

molecule that guides the Cas9 protein to a specific DNA sequence (Sternberg and Doudna, 2015). After cutting the double-strand DNA open, repair can occur through two basic mechanisms, i.e., non-homologous end joining (NHEJ), a mechanism that allows the cell to randomly insert or delete nucleotides at the break site, and homology directed repair (HDR), a mechanism that enables insertion of a template DNA to correct mutations at the DNA break site (Song, 2017, DeWitt et al., 2017). CRISPR/cas9-induced double strand breaks are mainly repaired by NHEJ mechanisms (Sander and Joung, 2014).

In 2013 Ebina et al. (Ebina et al., 2013) successfully suppressed HIV-1 gene expression in Jurkat cells by targeting HIV-1 LTR with CRISPR/cas9. Subsequently, the CRISPR/cas9 system has been used in the exploration of HIV-1 treatments. For example, a study by Hu et al. (Hu et al., 2014) found that the CRISPR/Cas9 system can be used to identify the specific targets of complete excision and integration of the pre-HIV-1 genome, leading to inactivation of viral gene expression and replication in HIV-1 latently infected cells. This is a potential therapeutic advance in eliminating barriers of all pro-viruses in HIV-1 infected people. In addition to the HIV-1 provirus, other researchers have turned their attention to the HIV-1 receptor. Wang et al. used a lentivirus expressing CCR5-single guide RNA (sgRNA) and Cas9 to knockout the coreceptor CCR5 in $CD4^+$ T cells, making them resistant to HIV-1 (Wang et al., 2014). Two different gRNA combinations targeting both *CXCR4* and *CCR5* were designed by Guo's team. The CRISPR-sgRNA-Cas9 system successfully induced *CXCR4* and *CCR5* gene editing in various cell lines and primary $CD4^+$ T cells, indicating that this CRISPR/Cas9 approach could have applications in the functional cure of HIV/AIDS (Liu et al., 2017). Although this strategy may have advantages over others, it can cause unpredictable damage via the DNA repair mechanism and can result in virus escape (Xun et al., 2021). Gene-editing strategies are limited by the absence of tools or biomarkers to detect the pool of latently infected cells selectively and safely, and the difficulty of using genome editing at a large scale to destroy the provirus (Barouch and Deeks, 2014, Henderson et al., 2020). However, his rapidly developing gene editing technology may make gene editing of the HIV-1 genome achievable.

Based on the literature discussed above, it is evident that transcription is silent in latently infected cells and therefore targeting this step in HIV-1 replication cycle is a major key in achieving either a sterilizing or functional cure. The following section will discuss HIV-1

transcription focusing on the role of Tat in the transcription step, in the context of HIV-1 latency.

2.3 Transactivator of transcription (Tat) enhances HIV-1 viral gene transcription

2.3.3 The functional properties of Tat

Following nuclear entry, HIV-1 DNA genome gets integrated into the host genome to become a provirus which acts as a transcription template, regulated at the transcriptional and posttranscriptional levels (reviewed in (Karn and Stoltzfus, 2012)). Immediately following integration, HIV-1 transcription solely produces short completely spliced mRNAs encoding the viral regulatory proteins Tat and Rev (Kim et al., 1989). As the infection proceeds, transcription rapidly increases, and larger, incompletely spliced mRNAs are produced (reviewed in (Karn and Stoltzfus, 2012)). These encode the remainder of the HIV-1 proteins: Env and the HIV-1 accessory genes Vif, Vpr, and Vpu. Also synthesized late are the full-length unspliced transcripts which act both as the virion genomic RNA and the mRNA for the Gag-Pol polyprotein (Kim et al., 1989, Pomerantz et al., 1990). In the absence of Tat, basal transcription occurs which is characterized by short abortive mRNA transcripts produced during elongation. Transactivator of transcription (Tat) is the HIV-1 protein that enhances viral gene transcription (Ensoli et al., 2021).

HIV-1 transcription is regulated by two different phases based on a sequential back-and-forth between host and viral factors in the HIV promoter (Fig. 2.5) (Li et al., 2021, Perkins et al., 2017). The recruitment of host transcription factors marks the initiation of transcription where host transcription factors such as Nuclear Factor-κB (NF-κB) and SP1, and other general transcription factors, such as transcription factor IID, IIA, IIB, IIE, IIF and IIH (TFIIH) are recruited to their cognate sites on the HIV-1 5’ LTR, forming the pre-initiation complex (PIC). These factors allow the recruitment of the hypo-phosphorylated RNA polymerase II (RNAPII) to the HIV-1 5’ LTR (Richman et al., 2009).

The ATP-dependent DNA helicase, XPB a subunit of TFIIH, then facilitates negative DNA supercoiling that is threaded through the RNAPII active site. The cyclin-dependent kinase 7 (CDK7) subunit of TFIIH next phosphorylates Ser 7 and Ser 5 of the RNAPII C- terminal

domain (CTD), activating the RNAPII (reviewed in (Mori and Valente, 2020)). Transcriptional elongation is not efficient during this phase and RNAPII rapidly aborts viral mRNA transcription due to the scarcity of positive modulators NF-κB and the P-TEFb, sequestrated in an inactive form such as such the NF-κB and the P-TEFb (Sainsbury et al., 2015, Harlen and Churchman, 2017). In addition, abortive viral mRNA transcription is due to the presence of transcriptional repressors such as the negative elongation factor (NELF), DRB sensitivity inducing factor (DSIF), Yin Yang 1 (YY1) and the C-promoter binding factor (CBF) (Dutilleul et al., 2020), the positioning of nucleosome-1 (Nuc-1) downstream of the transcription start site (TSS) (Mori et al., 2020) and negative chromatin remodellers (Barboric et al., 2001, Van Lint et al., 2013). Only short viral transcripts of 60 nucleotides (nascent TAR RNA) downstream of the TSS are synthetized and accumulated (Tripathy et al., 2011). This initial step is designated as the "basal" state of provirus transcription (Agosto et al., 2015), illustrated in Fig 2.6A.

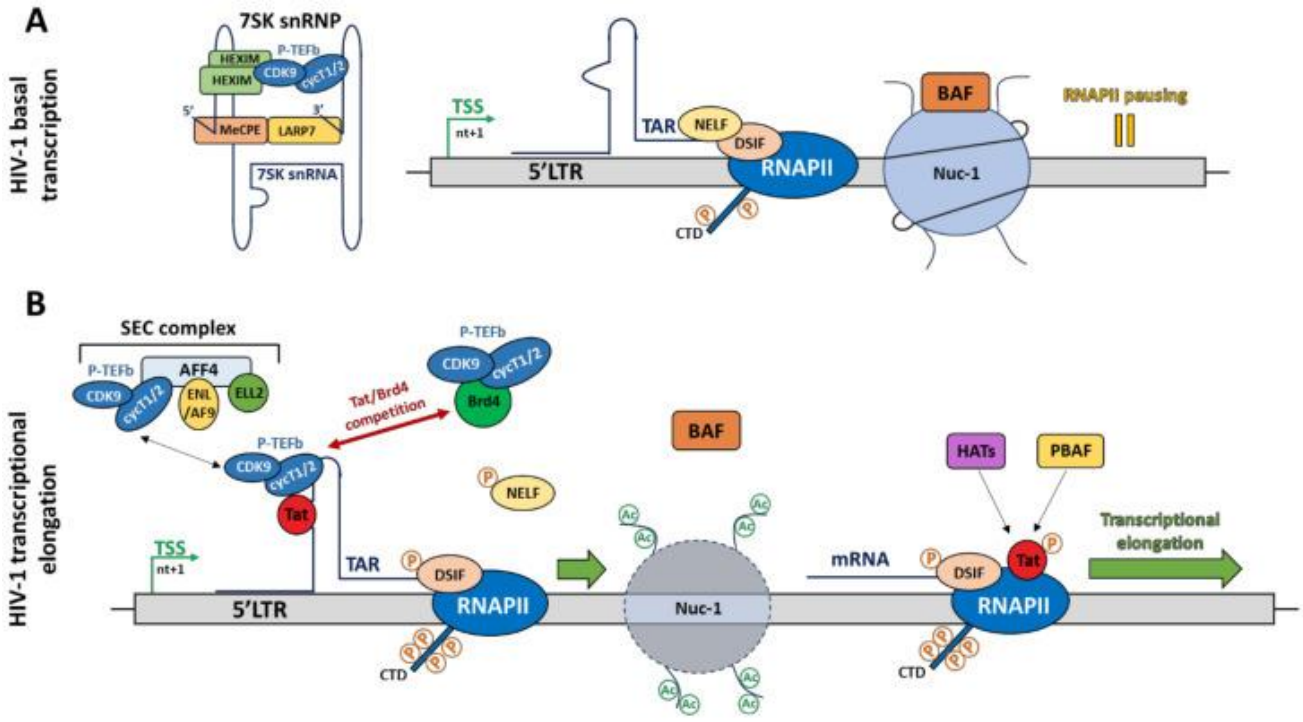

Figure 2.6: Diagram showing the molecular mechanisms involved in the regulation of HIV-1 transcription. HIV-1 transcription is regulated at two phases: the basal transcription and transcriptional elongation. (**A**) shows the initiation of HIV-1 transcription is strongly regulated through cis-regulatory elements located in the 5′ LTR. The subsequent recruitment of cellular transcription factors to their binding sites is sufficient to promote the recruitment of the transcription pre-initiation complex and thus to initiate HIV-1 transcription. However, the

transcriptional elongation is paused due to (1) the recruitment of two pausing factors, called DRB (5,6-dichloro-1-β-d-ribofuranosylbenzimidazole) sensitivity-inducing factor (DSIF) and negative elongation factor (NELF), on the RNA polymerase II (RNAPII) and to (2) the positioning of nuc-1, mediated by BRG1- or HBRM-associated factor (BAF), just downstream of the TSS. HIV-1 transcription is then aborted and results in the formation of the TAR element. During this step, the P-TEFb is sequestered under an inactive form in the 7SK snRNP complex. (**B**) depicts the transcriptional elongation which is mainly regulated by the Tat where it is recruited at the TAR element and then recruits the P-TEFb, as a result inducing the phosphorylation of the RNAPII carboxy terminal domain (CTD) and of the pausing factors which then dissociate. (Dutilleul et al., 2020).

As aforementioned, HIV-1 replication can occur in the basal phase where Tat is absent. As a result of immune activation, transcription factors, such as NF-κB and Bromodomain containing protein 4 (BRD4), translocate into the nucleus (Kaczmarek et al., 2013, Morton et al., 2019) and facilitate inefficient recruitment of P-TEFb (a complex formed by Cyclin T1 and CDK9). P-TEFb will be further described and illustrated in section 2.4. Briefly, once recruited to the HIV-1 5' LTR, P-TEFb phosphorylates the transcriptional repressors NELF and DSIF, along with the RNAPII CTD at the Ser 2 residue which releases NELF from the HIV-1 5'LTR and transforms DSIF into a positive transcription factor (Fujinaga et al., 2004). This is illustrated in figure 2.5B. In addition, BRD4 has been suggested to act as a kinase that could further phosphorylate the Ser 2 residue of RNAPII (Bacher et al., 2021). Ultimately, full-length HIV-1 mRNAs are transcribed, spliced and translated to produce HIV-1 proteins, including the Tat. This second phase of transcription is transcriptional elongation of provirus transcription (Agosto et al., 2015). During the transcriptional elongation process, P-TEFb and Tat undergo multiple association/dissociation cycles allowing a powerful induction of HIV-1 gene expression (Dutilleul et al., 2020).

2.3.1 The role of the transactivator of transcription (Tat)

The preceding section has alluded to the important role which the HIV-1 protein, Tat, plays in transcription. The current section will describe in detail the function of Tat and provide evidence through a survey of the literature for the significance and relevance of the current study. Tat is encoded by two exons of the HIV-1 *tat* gene and its length varies between 99 and 103 amino acids (aa), with the predominant form being 101 residues (Clark et al., 2017). Exon 1 encodes the first 72 amino acids comprising the primary sequence of Tat, which is divided into five domains (Clark et al., 2017). The first domain is the proline-rich domain or acidic N-

terminal region and contains the first 21 amino acids. This domain is responsible for mediating LTR transactivation through interactions with cycT1, in conjunction with the cysteine-rich and core domains (domain 2 and 3). Thus, together making up the domains responsible for transactivation. Most variability occurs in the proline-rich domain, residue 11 is a well conserved tryptophan and is required for efficient secretion of Tat (Rayne et al., 2010). The cysteine-rich domain, residue 22 to 37, has an abundance of highly conserved cysteine residues, which are located at positions 22, 25, 27, 30, 31, 34, and 37.

These closely associated cysteines are responsible for the formation of intra-molecular disulfide bonds (Pierleoni et al., 2010, Koken et al., 1994). Of note, residue 31 may encode a cysteine-to-serine mutation that is prevalent in HIV-1 subtype C and is the subject of much debate because of its potential role in the reduction in neurocognitive impairment in patients infected with subtype C virus (Ranga et al., 2004). The core domain contains residues 38 to 48 and, in conjunction with the proline-rich and cysteine-rich domains, is responsible for interactions with cycT1 (Wei et al., 1998). Together with the cysteine-rich domain, the core domain has also been demonstrated to mediate cofactor binding, specifically with CREB-binding protein (CBP)/p300, histone acetyltransferase (HAT), and the Sp1 transcription factor (Marzio et al., 1998, Jeang et al., 1993). The arginine-rich domain, also referred to as the basic domain, contains a well-conserved sequence, ^{49}RKKRRQRRR^{57}, that is crucial for the interaction with TAR as well as the secretion and uptake of Tat (Rayne et al., 2010, Hauber et al., 1989). The glutamine-rich domain spans from residue 59-72 and contains the remainder of the first exon. In conjunction with the arginine-rich domain, it is referred to as the basic region and is responsible for nuclear localization and mediates binding to CCATT enhancer binding protein (C/EBP) (Hauber et al., 1989, Ruben et al., 1989).

Exon two of Tat is less conserved compared to the first exon and is classically characterized as its own distinct domain (van der Kuyl et al., 2018), but has been demonstrated to be crucial for efficient replication of macrophage-tropic strains of HIV-1 and contributes to mechanisms of viral persistence (Neuveut et al., 2003, Lopez-Huertas et al., 2013). Despite the crucial role of the transactivation domain in enhancing viral gene transcription, genetic variations exists within HIV-1 *tat* gene during the course of infection, which could translate to differential ability of Tat to trans activate the HIV-1 5' LTR-driven transcription (Ronsard et al., 2014). During HIV-1 replication, one of the first proteins produced after viral DNA integration is Tat

(Li et al., 2009). It is required for full-length RNA chain synthesis will not happen without it (Wong et al., 2010). Tat appears to be critical to bring about the rapid increase in genomic length transcription required for the transition from a quiescent to an active viral infection (Roebuck and Saifuddin, 1999).

As aforementioned, the Tat protein, required for production of full-length viral mRNA transcripts, is synthesized early during the replication cycle (Kuciak et al., 2008). P-TEFb is a general RNAPII transcription factor that is required for efficient expression of most protein-coding genes, as well as for production of full-length viral mRNA transcripts from the provirus (Zhu et al., 1997). As briefly mentioned, once recruited to the 5' LTR, P-TEFb phosphorylates the CTD of the RNAPolII thus releasing it from its paused state to allow full-length viral mRNA transcription (Liu et al., 2014). Thus, Tat participates in a positive feedback mechanism that ensures high levels of proviral genes expression and productive infection.

Several other functional properties have been associated with Tat of which many have been observed once the protein resides in extracellular space (Khan et al., 2019). In this regard, Tat has been shown to be secreted through many different mechanisms (Mele et al., 2018), which are dependent on residues at positions 11 and 49-51 and their interactions with phosphatidylinositol-4,5-bisphosphate (PtdIns(4,5)P2), a lipid in the plasma membrane (Rayne et al., 2010). Alongside its vital role in enhancing viral gene transcription, Tat also contributes to the pathogenesis of HIV-1 infection by its capacity to interact with infected or uninfected cells (Ensoli et al., 1993). Moreover, Tat also contributes to the spread of HIV-1 through its effect on the increase of CCR5 and CXCR4 surface expression (Howcroft et al., 1993). It has been found to induce apoptosis in CD4 T-cells (Li et al., 1995). Moreover, several studies suggest a direct effect of Tat in the structural and immunological dysfunctions observed early after infection, in the gastrointestinal tract (GALT) from HIV-1 infected patients (Brenchley et al., 2006). The association of the HIV-1 Tat with several pathologies related to AIDS, makes it an important target for development of therapeutics (Johri et al., 2011).

The functional properties of Tat, however, are dependent on the sequence of amino acid residues. The HIV-1 genome is predisposed to genetic variation, which is caused by factors such as the error-prone viral reverse transcriptase in conjunction with host immune response (Li et al., 2012). Therefore, Tat is also subject to genetic variation during the course of HIV-1

replication resulting in a number of alterations in amino acid residues that have been associated with LTR functional activity (Dampier et al., 2016). Amino acid changes within Tat have resulted in reduced LTR transcription activity (Boven et al., 2007, Ronsard et al., 2017a). Reduced transcription activity has been reported to occur with only a single residue mutation, such as position 11 substitutions to either alanine, phenylalanine, or leucine (Yezid et al., 2009), a glutamine substitution at position 50 (Bres et al., 2002), or an alanine substitution at position 51 (Van Duyne et al., 2008). And an Alanine at position 21 (TatA21), which is a signature residue for HIV-1 subtype C (Neogi et al., 2012) was also reported to reduce the LTR transcription activity and associated with slower rate of disease progress (Rossenkhan et al., 2013). Consistently, the data that is being prepared for publication my current project demonstrate that TatA21 is associated with reduced transcription activity and significantly lower viral loads in patients compared to a Proline at this position (TatP21) (Mkhize et al., manuscript under preparation for submission). However, the mechanism by which these mutations cause altered LTR transcription activity remain not known. Given the primary role of Tat at the HIV-1 transcription step, which is to bind to TAR element and recruits P-TEFb to the 5' LTR for efficient transcription, investigating these pathways may uncover the mechanism by which single amino acid changes in Tat result in reduced LTR transactivation.

2.3.2 The importance of HIV-1 Tat-TAR binding

The result of these post-translational modifications is the synthesis of high levels of full-length viral mRNA transcripts. Tat binds to the TAR sequence with a 1:1 stoichiometry. In the presence of a variety of ligands, including Tat, the bulge region of TAR undergoes a local conformational rearrangement and forms a more stable structure. Unbound TAR is highly dynamic in both the bulge and apical loop. In the bound form, the U23 residue at the 5'-end of the bulge is positioned near G26 and A27 in the major groove, rather than stacked on A22 as in the free TAR. U23 and G26 are brought into close proximity by contacts to the guanidinium group and side chain amide group of a common arginine residue (Aboul-ela et al., 1995) (Fig. 2.7).

Tat binds into the major groove of the TAR stem-loop structure at the level of the UCU bulge. The apical loop of TAR contains the binding site of cyclin T1 (Anand et al., 2008). The importance of the Tat–TAR interaction for viral replication has led to considerable interest in

discovering new inhibitors of viral replication that function by disrupting the formation of the complex. Many different ligands, from small molecules to peptide mimetics, have been developed to bind TAR at the Tat-binding site (Aboul-ela et al., 1995, Aboul-ela et al., 1996, Leeper et al., 2005, Murchie et al., 2004). Early strategies developed for the design of small compounds able to disrupt the Tat–TAR interaction have been summarized in Thomas and Hergenrother in 2008 (reviewed by (Thomas and Hergenrother, 2008)).

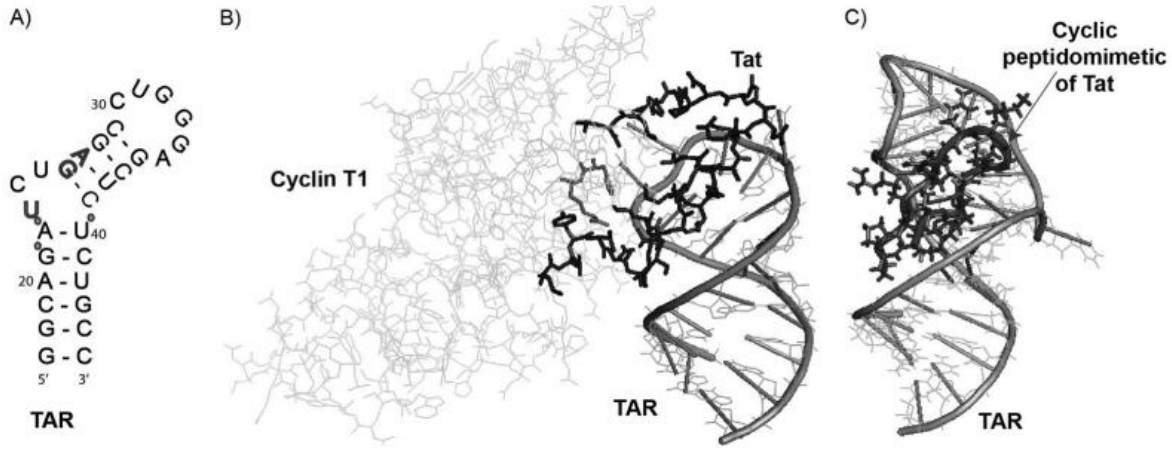

Figure 2.7: TAR RNA structure. A) Secondary structure of HIV-1 TAR sequence; nucleotides in grey are involved in the interaction with Tat and grey dots indicate phosphate groups involved in the interaction with Tat B). Crystal structure of the cyclin T1–Tat–TAR RNA transcription activation complex from EIAV (PDB code: 2W2H). Tat is shown as black sticks and the cyclin T1 in light grey sticks with the two side chains of His79 and Arg80 interacting with TAR in grey sticks. C) NMR structure of the complex of a cyclic peptide (black sticks) bound to BIV TAR RNA (PDB code: 2A9X). The structure reveals a ß-hairpin conformation of the bound peptide in the major groove of the RNA (Blond et al., 2014).

Conformational changes likely appear in the basic region of Tat protein as was shown for equine infectious anemia virus Tat protein (Anand et al., 2008), which shares similarities with the basic domain of HIV-1 Tat (Foucault et al., 2010). Moreover, the basic domain provides structural stability for the Tat protein through electrostatic interactions with its N-terminal part (Pantano et al., 2004). Other Tat domain that can undergo conformational changes is the one that facilitate Tat-P-TEFb complex. X-ray crystallography showed that the first three N-terminal Tat domains (one to 49 residues) sustain extended conformation mostly through interactions with cyclin T1 whereas 50 to 86 residues are not defined (Tahirov et al., 2010).

Nuclear magnetic resonance showcased the structural propensity of Tat protein suggesting that the cysteine-rich domain tends to fold into α-helixes in contrast to the basic domain with extended or β-sheet conformation (To et al., 2016). Comparison analysis of X-ray and nuclear magnetic resonance studies suggests that different fragments of Tat protein can employ different folding mechanisms (To et al., 2016). This flexibility enables Tat to adopt diverse conformations upon interaction with its physiological partners, thus greatly extending its multifunctionality (Kurnaeva et al., 2019).

One of the most important and well-studied domains of Tat protein is the basic domain. While Tat can tolerate up to 40% of sequence mutations without significant changes in its activity, its basic domain is highly conserved among Tat domains (Campbell and Loret, 2009). It is enriched with positively charged Arginine and Lysine residues comprising 49- RKKRRQRRR-57 motif. The basic domain confers many properties to Tat such at regulation of viral transcription and manipulation of cellular processes in favour of HIV-1 replication (Kurnaeva et al., 2019). An earlier study showed that an inhibitor of the Tat/TAR RNA interaction effectively suppressed HIV-1 replication (Hamy et al., 1997). The Tat-TAR interaction as a mechanism which causes reduced LTR transactivation will be investigated in the present study. Data from this study may lead to the identification of HIV-1 subtype C variants that affect viral transcription, disease outcomes and the propensity of latency establishment or reactivation potential. The data from this study may result in identification of Tat-TAR interaction hotspot, which could serve as a new target for drug development.

2.4 Host transcription factor PTEF-b, an essential co-factor for Tat

2.4.1 The P-TEFb

In addition to binding to the TAR element, Tat recruits the P-TEFb to HIV-1 5’ LTR for efficient transcription. Tat increases more stable and active P-TEFb compounds function by inserting grooves at heterodimer interface (Weydert et al., 2016). Normally, P-TEFb is housed within the 7SK small nuclear RNA (snRNA), and its activity is controlled by the 7SK snRNA and hexamethylene bisacetamide (HMBA) inducible protein 1 (HEXIM1), which sequester P-TEFb into a transcriptionally inactive small nuclear ribonuclear protein (snRNP) (Yik et al., 2003). As aforementioned, the Tat binds to the TAR element and recruits P-TEFb (cycT1/CDK9) to promote disassembly of the 7SK/HEXIM/P-TEFb snRNP, activating

transcriptional elongation (McNamara et al., 2013), reviewed in (Ott et al., 2011)). A more recent paper showed that cycT1 is highly upregulated in activated memory $CD4^+$ T cells but may become less associated with T cell activation during HIV-1 replication (Couturier et al., 2019). This is probably due to hijacking of P-TEFb for viral replication which decreases its availability or the transition from productive replication to latency establishment following down regulation of cycT1 (Couturier et al., 2019).

CDK9 also extensively phosphorylates the CTD of RNAPII (Fig. 2.6). This phosphorylation mainly occurs at Ser2 residues of the heptad repeats Y-S-P-T-S-P-S and C-terminal repeats (G-S-Q/R-T-P) of the hSpt5 subunit of DSIF at Thr4 residues (Suh et al., 2016, Schuller et al., 2016). The overall consequence of these phospho-modifications by P-TEFb is to remove elongation being hampered imposed by NELF and DSIF and to stimulate efficient elongation and co-transcriptional processing of proviral mRNA transcripts. An additional CDK9 T-loop phospho-modification occurs at a conserved residue Ser175 (pSer175 CDK9), which was reported in a previous study (Mbonye et al., 2018) due to CDK7 activity in both Jurkat and primary CD4+ T cells. This pSer175 CDK9 is only found on a subset of transcriptionally active P-TEFb (ie. P-TEFb which is free and has dissociated from 7SK snRNP) and is functionally important in being recruited by Tat to enhance Tat's interaction with P-TEFb in order to outcompete BRD4 (Mbonye et al., 2013), a bromodomain-containing protein that is considered to be a major recruiter of P-TEFb to cellular genes (Jang et al., 2005).

This study focused on identifying the pathways that are essential for the generation of transcriptionally active P-TEFb. However, despite that there is a dependent relationship between Tat and P-TEFb, studies which investigate how Tat genetic variation affects its ability to recruit P-TEFb of thus of importance to better understanding this pathway. The P-TEFb is expressed at vanishingly low levels in memory CD4+ T cells and this is one of the key regulatory mechanisms that minimizes HIV-1 transcription in resting cells (Mbonye et al., 2021). It has been shown in a primary $CD4^+$ T cell model of HIV-1 latency that P-TEFb must be up-regulated for reactivation of latent virus (Tyagi et al., 2010). Therefore, the "shock" component of a "Shock and Kill" strategy will require the induction of P-TEFb. Going forward, high-throughput screens for small molecules that up-regulate P-TEFb in resting $CD4^+$ T cells can be used to develop novel LRAs (Rice, 2017). The regulation of P-TEFb in primary T cells is fundamental to control of HIV-1 transcription. Therefore, dissecting the

cellular pathways through which HIV-1 emerges from latency is critical to the development of "Shock and Kill" therapeutic approaches that can be used to effectively eradicate persistent reservoirs of transcriptionally latent but replication-competent HIV proviruses in memory CD4+ T cells of infected individuals. (Mbonye et al., 2021). Tat recruits the P-TEFb, to the transcription start site (TSS) in the LTR promoter (Zhu et al., 1997). Thus, targeting this pathway for the "Lock and Block" strategy can silence HIV transcription resulting in latent infections.

2.4.2 Structural importance of the Tat:P-TEFb complex

HIV-1 has evolved to manipulate both the CDK9:CycT1 form of P-TEFb and the cellular machinery controlling P-TEFb. The structure of the Tat:P-TEFb complex plays an important role in their interaction. It is important to note that Tat folds on the outer surface of the cycT1 cyclin domain (Fig. 2.4). The amino-terminal "activation" domain of Tat binds to the CDK9 T-loop, a region of the molecule that is essential for its enzymatic activity (Tahirov et al. 2010). It is important to note that Tat is a naturally denatured protein, and its 3D structure is not established until it binds to the specific interacting molecule (Shojania and O'Neil, 2006). This makes Tat a versatile protein by having multiple interactions with host proteins to carry out various biological actions which benefit HIV-1 propagation and causing AIDS. In a study by Jean et al., they performed a parallel analysis of *in vitro* translated open reading frames (OPFs) (PLATO) approach to identify Tat binding proteins (Jean et al., 2017). As expected, almost all these proteins were involved in transcription, indicating that the main function of Tat is still in transcription. Similar results have been reported by other studies (Gautier et al., 2009, He et al., 2010, Sobhian et al., 2010). Tat is an intrinsically disordered protein (Debaisieux et al., 2012) and therefore, only nuclear magnetic resonance structures are available for Tat alone (Kurnaeva et al., 2019).

Crystallographic studies have shown that the binding of Tat subtype B to CDK9-CycT1 (P-TEFb) (Tahirov et al., 2010) induces significant conformational changes in P-TEFb, which is essential for the cooperative recognition of the TAR RNA stem-loop structure (Schulze-Gahmen et al., 2013). Although Tat forms many of its interactions with the CycT1 subunit, it

also forms functionally important hydrogen bonding interactions with the activation loop of CDK9, which may potentiate P-TEFb kinase activity (Tahirov et al., 2010).

To develop efficient anti-HIV-1 compounds targeting Tat, an understanding of its structure is crucial (Asamitsu et al., 2018). Since Tat requires P-TEFb to stimulate HIV-1 transcription elongation, Tat/P-TEFb interaction is an attainable target for developing new anti-HIV drugs. In addition to P-TEFb, several reports have shown that Tat interacts with importin α/β and other transcriptional regulators, including protein kinase PKR, Sp1, and the transcriptional coactivators CBP/p300 through the Tat exon 1 region (Jeang et al., 1993, Smith et al., 2017).

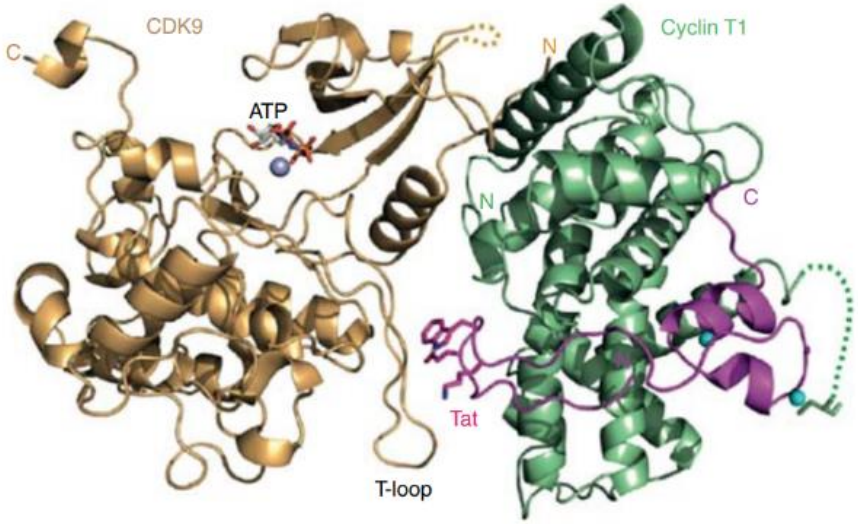

Figure 2.8: Ribbon representation of the Tat•P-TEFb•ATP structure. CDK9 is light orange, Cyclin T1 is pale green and Tat is magenta. The side chains of the CDK9-interacting residues of Tat, Cys261 of Cyclin T1, and ATP analog are drawn as sticks, and the zinc and magnesium atoms are drawn as cyan and light blue spheres, respectively. The dashed lines represent the missing link between Lys88 and Gly of CDK9, and between Leu252 and Cys261 of Cyclin T1 (Tahirov et al., 2010).

In addition to recruiting P-TEFb to proximally stalled polymerases engaged in the HIV-1 5' LTR, Tat stimulates the kinase activity and changes the substrate specificity of P-TEFb (Zhou et al., 2000). Tat adopts a structure complementary to the surface of P-TEFb where it comes into contact, mainly with the Cyclin T1 subunit of P-TEFb, but also with the T-loop of the CDK9 subunit (Fig. 2.8). The structure provides a plausible explanation for ability of Tat to

tolerate up to 40% of sequence variation at certain sites. The structures reveal that Tat interacts with both the cycT1 and CDK9 subunits of P-TEFb and provide evidence for a Tat-induced conformational change in P-TEFb (Tahirov et al., 2010). The Tat:P-TEFb structure provides insight into how Tat functions in recruiting P-TEFb to the HIV-1 5' LTR where Tat acquires an extended conformation mainly through interactions with Cyclin T1. The first step of this process involves a Tat-induced release of P-TEFb from the 7SK snRNP-containing complex (Sedore et al., 2007, Barboric et al., 2007).

The Tat has evolved to bind tightly to P-TEFb as 37% of its folded portion (amino acids 1-49) surface is complementary to the kinase. The amount of surface buried between P-TEFb and Tat is double the average value for stable protein:protein interactions (Janin, 1997). Thus, showing the strong interaction between these two proteins. It is thus imperative to investigate whether the genetic variation in Tat affects its protein structure and thus its functionality in its ability to recruit P-TEFb.

This study is significant as it will inform future therapeutic interventions of combining antiretroviral treatments with latency-inducing strategies to silence the latent reservoirs and lead to a functional cure.

Chapter Three

Chapter Three: Materials and Methods

The materials and methods as well as a brief diagrammatical overview of the experimental procedure (Fig. 3.1) for goal one of this study are described in section 3.1. The materials and methods for goal 2 are described in section 3.2 of this chapter.

A previous study by Rossenkhan et al showed that TatA21 was a signature residue in HIV-1C and resulted in significantly lower transactivation that TatP21. Moreover, data from my own Masters study showed that patients infected with HIV-1C with a TatA21 had a significantly lower viral load than patients with a TatP21. However, in both my study and Roseenkhans study, the mechanism by which TatA21 causes a reduction in functionality and low viral loads in patients is unknown. Therefore, briefly the first goal of the study was to determine the molecular mechanisms by which patient derived HIV-1C TatA21 variant results in reduced transcription activity of the LTR viral promoter in acute infection (data from my Masters degree and Mkhize et al., manuscript under preparing for submission). Therefore, my current PhD project undertook to determine whether a single mutation TatA21 alone or in combination with other mutations results in reduced the Tat function. Firstly, single mutations were introduced into the HIV-1C consensus *tat* contained in the pcTat.BL43.CC (catalog number 11785; NIH AIDS Reagent) (referred to as pCTat) using site directed mutagenesis technique (Fig. 3.1A). The successful introduction of each mutant in the *tat* gene contained in pC-Tat.BL43.CC was confirmed by sequencing. The effect of these single mutants on the function of Tat was assessed using a Tat transactivation assay in TZMbl cells (Fig. 3.1B). Next, the expression levels of the Tat mutants were determined using western blot (Fig. 3.1C). Subsequently, the functionality of the mutated amino acid Proline (P) at position 21 of the Tat protein (TatP21) mutant alone versus consensus (wildtype) amino acid Alanine (A) at position 21 of Tat protein (TatA21) was assessed by nucleofecting them into JLatC, Jurkat cell line harbouring latent HIV-1 subtype C provirus (Maikoo et al., unpublished data from our group) and in A72 cell line, containing the HIV-1 provirus DNA lacking that the *tat* gene. The expression of GFP and Gag mRNA expression under the HIV-1 LTR promoter were assessed by quantitative (q)PCR. Electrophoretic mobility (EMSA) or RNA gel shift assays with P32 labelled Tat was performed to determine whether the TatA21 was able to radioactively bind the TAR RNA. Next, the ability of TatP21 mutant versus of the wild type TatA21 to efficiently bind the TAR RNA was evaluated by molecular docking (Fig. 3.1D). Lastly, ability of TatP21 versus TatA21 to recruit P-TEFb to the TAR element was determined by their co-

immunoprecipitation with cycT1 and CDK9, which together make up the P-TEFb (Fig 3.1E). Each of these methods are described in detail in the following sections.

3.1. Goal One Methodology overview

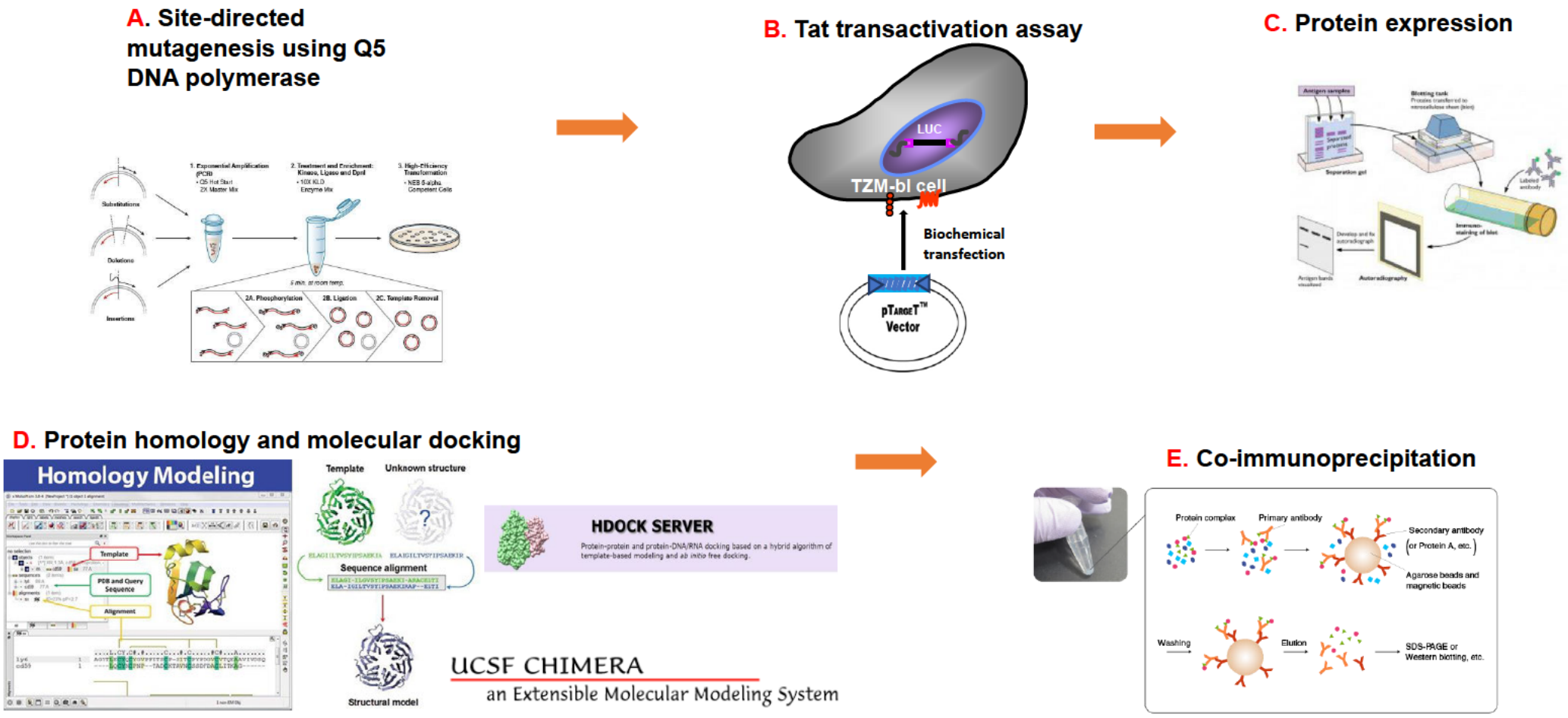

Figure 3.1: Diagrammatical overview of the methods used in achieving goal one of this project. **A**. Site directed mutagenesis using Q5 DNA polymerase. **B**. Tat transactivation assay. **C**. Protein expression using Western blot. **D**. Programs and software used for protein homology modelling and molecular docking. **E**. Co-immunprecipitation of Tat variants with cycT1 and CDK9.

3.1.1 Generation of mutant Tat clones

Introducing the relevant nucleotide substitutions into exon 1 of the subtype C consensus *tat* gene contained in the HIV-1 subtype C BL43/02 Tat Expression Vector pcTat.BL43.CC (catalogue number 11785; NIH AIDS Reagent) (referred to as pCTat) was performed using site directed mutagenesis using the Q5 DNA polymerase kit (New England BioLabs, Ipswich, MA). A total of twelve *tat* sequence primers (Table 3.1) were designed to introduce single point mutations into the pCTat recombinant clone in this thesis. These primers were used to introduce a total of five single HIV-1C Tat mutations (V4I, A21P, Q39L, S46Y and R53K) and one double mutation L35Q/Q39L mutation previously reported to affect Tat transactivation (Rossenkhan et al., 2013, Ronsard et al., 2017b). Both the forward and reverse primers of each mutation annealed at specific positions on the DNA of the pCTat consensus (wildtype) *tat* gene sequence to introduce specific mutations during polymerase chain reaction (PCR) that was performed using the Q5 DNA polymerase enzyme (New England BioLabs, MA, USA). Briefly, a PCR reaction mixture containing 2 nanograms (ng) of pCTat plasmid DNA containing the wildtype *tat* gene sequence together with 1X Q5 reaction buffer, 10nM deoxynucleotides (dNTPs), 0.2µM of the respective primers and PCR grade water to make up the final reaction volume to 25µL was prepared. The PCR was performed using the following conditions: initial denaturation at 98°C for 30s; amplification: 30 cycles of 98°C for 10s, 55°C for 30s and 72°C for 30s; and a final extension of 72°C for 5 mins in with a final hold at 10°C. Following the PCR, the parental pCTat plasmid DNA strand was digested by adding 1 µl endonuclease Dpn1 to digest all dam-methylated and hemi-methylated parental DNA at 37°C for 1 hr. This is to ensure that the parental strand is completely digested and the newly formed strand with the mutation of interest remains. The site-directed mutagenesis (SDM) products were either used immediately or stored at -20°C until use.

3.1.2 Transformation

The SDM products were transformed into JM109 competent cells (Promega) according to the manufacturer's instructions. Briefly, 50 microlitres (µL) JM109 competent cells were thawed on ice for 5-10 mins. Then 2µL of the SDM products were transformed into their respective 50 µL of JM109 competent cells and placed on ice for 20 minutes. The cold of the cells being on ice increases the effectiveness of the heat shock step by increasing the sudden change in

temperature. As this change in temperature affects the structure of the cell wall, allowing the plasmid uptake into the cells. The JM109 competent cells and SDM products (DNA) mixture were heat shocked at 42 ºC for 45 seconds (s) to disrupt the membranes of the cells and allow uptake of the plasmid and placed on ice for 2 minutes (mins) to allow the cells to recover and retain the plasmid inside the cells. Then 950µL of SOC medium was added to the cells to provide nutrients and a favourable environment for the cells to grow and the mixture was incubated in shaking incubator at 150 rpm for 1.5 hours to allow cell proliferation. Thereafter, a volume of 100 µL was then plated on Ampilicin Luria Broth (LB) agar plates as the plasmid contained an Ampilicin resistant gene, and incubated overnight at 37°C. Following overnight incubation, colonies were randomly picked to inoculate 100ml of Luria-Bertani (LB) with Ampilicin for maxi prep, described in the following section. Glycerol stocks of the plasmid were prepared by aliquoting 100 µL of culture into 1ml of LB Ampilicin broth containing 20% glycerol. Glycerol stocks were stored at -80 ºC until use.

3.1.2.1 Confirmation of the recombinant plasmid

Confirmation of the recombinant plasmid was done by sequencing. The contents of the BigDye™ Terminator v3.1 Cycle Sequencing Kit were thawed and placed on ice together with T7 primers. The reagents were vortexed briefly for 2 to 3 seconds and centrifuged briefly with a benchtop microcentrifuge to collect contents to the bottom of the tubes. A sequencing reaction was prepared using 1.5 ml tubes and each reaction transferred into a well of a clear 96 well plate by adding the following: 0.4 µl of BigDye™ Terminator 3.1 Ready Reaction Mix, 2µM forward primer or reverse primer, 20 ng of maxiprep, deionized water was used to bring the reaction to a final volume of 10 µL. The plate was sealed with MicroAmp™ Clear Adhesive Film, vortexed for 2 to 3 seconds, then centrifuged briefly in a centrifuge 5810 R swinging bucket centrifuge at 1,000 x g for 1 minute to collect contents to the bottom of the wells. The sequencing reaction was placed in a thermocycler. Cycle sequencing was then performed using the following conditions: initial incubation at 96°C for 1 min, followed by 25 cycles of denaturation at 96°C for 1 min, annealing at 50°C for 5 secs and extension at 60°C for 4 mins, final hold at 4°C. Sequences were assembled and edited using the Sequencher Program v5.0 (Gene Codes Corporation, Ann Arbor, MI, USA).

3.1.2.2 Plasmid isolation

Plasmid isolation was performed using the Qiagen Maxi Prep kit (Qiagen, California, USA). The centrifuge was first pre-chilled to 4°C. Briefly, 100ml LB broth culture was aliquoted into two 50 ml falcon tubes and the bacteria culture was centrifuged at 5850 rpm for 20 minutes at 4^{o}C. The Qiarack was prepared to hold the plasmid maxi prep filters during this 20 mins spin. After centrifugation to pellet the bacterial cell, the supernatant was then carefully removed using a pipette. The bacterial pellet was then resuspended in 10 mL of pre-chilled (at 4 ^{o}C) Buffer P1. This pre-chilled Buffer P1 was added incrementally, firstly, 1 mL of the Buffer P1 was used to resuspend the pellet in the first 50 mL falcon tube using a P1000 pipette and the same was done for the second tube. The resuspended pellets of the same sample from the two tubes were then combined into one tube and the remaining 8 mL of the Buffer P1 was added to resuspend the combined pellets in the final volume of 10 mL of Buffer P1. The 10 mL of Buffer P2 was then added to the resuspended pellet and this tube was closed tightly and inverted 4-6 times to obtain a homogenous solution in the lysis step. Cell lysis breaks down the cell membrane and allows release of the DNA. The homogenous solution was then incubated at room temperature for 5 mins. A total of 10 mL of pre-chilled (at 4^{o}C) P3 Buffer was then added to the tubes and tubes were inverted 4-6 times for neutralization of the solution. This buffer is used to neutralize the lysate and digest any RNA which may be present in the solution. The lysate was then poured into the qiafilter cartridge with screwed caps and incubated for 10 minutes. The qiafilter cartridge was then held over the qiafilter tip and the cap opened. The plunger was then inserted and pushed to filter the lysate into qiafilter tip where the filtrate was allowed to flow through. The qiafilter was then washed twice with 30 mL of Buffer QC and the QC Buffer allowed to pass through by gravity. Plasmid DNA was then eluted by gravity with 15 mL buffer QF. The 10 mL of room temperature isopropanol was added, and the contents mixed. This was then centrifuged at 5850 rpm for 1 hour at 4^{o}C. The supernatant was then removed, and the pellet washed with 5 mL of room temperature 70% ethanol. The tube was flicked so that white flakes begin to rise and centrifuged at 5850 rpm for 1 hour at 4^{o}C. The supernatant was decanted, and pellet allowed to air dry. The pellet was then dissolved in 200 µL of distilled water and quantified using the Nanodrop2000 (Thermofischer). Plasmids were stored at -80^{o}C.

Table 3.1: Primers used to introduce mutations of interest by site-directed mutagenesis. Each mutation introduced is reflected in the name of the primer.

V4I FORWARD	5' - ATGGAGCCAATCGATCCTAACCTAGAGCCCTGG - 3'
V4I REVERSE	5' - CTAGGATTGGTAGCTTAAGCGGCGGCGGTCCTCGG - 3'
A21P FORWARD	5' - GCCCGAAACTCCTTGCAATAACTGTTTTTGTAAAAAATG - 3'
A21P REVERSE	5' - ACGTTATTGAGGAAGGTCCTTCAGTCGGGCTTTGA - 3'
K29R FORWARD	5' - TTTTTGTAAACGATGTTGCTACCATTGTCTAGTTTG - 3'
K29R REVERSE	5' - ACAACGATGGGCTGAACGTTATTGACAAAAAC -3'
L35Q FORWARD	5' - CTACCATTGTCAAGTTTGCTTTCAGAAAAAAGG - 3'
L35Q REVERSE	5' - CAAACGAAAGGTTCATTTTTTACAACGATGGTAAC - 3'
Q39L FORWARD	5' - AGTTTGCTTTCTTAAAAAAGGCTTAGGCATTTCCC - 3'
Q39L REVERSE	5' - TTTTTTTCCGAGAAGATGGTAACAGATCAAACG - 3'
R53K FORWARD	5' - GAAGAAGCGGAAACAGCGACGAAGCGCTCCTCCGAG - 3'
R53K REVERSE	5' - GTCGCTGCTTTTTAAGGGTACCGTCCTTCTTCGCC - 3'

For each point mutation introduced, two primers were designed, a forward and reverse primer. Primers were designed manually and checked using a primer design software (Thermofischer free online primer designer), to confirm the sequences obtained. Each primer was designed following conventional primer standards such that the primer length is no longer than 27 bp and does not have a GC content of greater than 50%.

3.1.3 Cells and cell culture

The TZM-bl cells were obtained from the NIH AIDS Reagents Program, which are HeLa cell derivatives that stably express-galactosidase and firefly luciferase under the control of the HIV-1 LTR (Platt et al., 2009, Derdeyn et al., 2000). This cell line was used for the Tat transactivation assay due the presence of the *luciferase* gene under the control of the 5' LTR. Briefly, the TZM-bl cells were cultured in Dulbecco modified Eagle medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 100U of penicillin/mL, 100 mg of streptomycin/mL, 2 mM L-glutamine, and 0.1 mM nonessential amino acids for 48 hours at 37°C and 5% CO_2.

3.1.3.1 Jurkat cells

Jurkat cells infected with an HIV-based retroviral vector containing the GFP and HIV-1 subtype C Tat open reading frames both under the control of the HIV-1 subtype C promoter,

in the 5’ long terminal repeat (LTR) referred to as JLATC cells (Maikoo et al, unpublished) were made available for this study. In addition, other Latent cell lines which includes A72 which habour the full-length HIV-1 genome that lacks the *tat* gene and the C66 cell is a clone of JLat C cell line) were nucleofected with the consensus (TatA21) and mutant that (TatP21). The A72 cell line was used to investigate the effect of Tat TatAP21 versus TatA2) on the ability of HIV-1 subtype C 5’ LTR to drive the mRNA expression of GFP and HIV-1 Gag. The C66 cell line was used as a control of determining the effect of TatA21 and TatP21 since it harbours endogenous full-length Tat.

3.1.4 Western Blot

To run the experiment; 1,5 mm glass were cleaned with water and ethanol, dried and clasped together. This was then first filled with separating gel and while solidifying, about 500µL isopropanol was added to ensure that the gel dries evenly. The excess isopropanol was then cleaned with distilled water and dried with Whatman paper. Following this the remainder was filled with the stacking gel. While the gels solidified, approximately 2 million cells were collected for the transfections of TatA21 and TatP21 using the methodology mentioned above. Jurkat cells were nucleofected with TatA21 and TatP21 on day 1 and collected after two weeks for the western blot. The western blot assay was used to determine the expression level of Tat protein mutants. Cells were washed with Phosphate buffered saline (PBS) and at least 500,000 Jurkat cells were lysed per condition in ice-cold lysis buffer (100 mM NaCl, 10 mM Tris, pH 7.5, 1 mM EDTA, 0.5% Non-idet P-40, protease and phosphatase inhibitor cocktail [Roche]). Cell lysates were denatured in prepared Laemmli sample buffer (Bio-Rad) and incubated for 5 min at 95 °C. The proteins were separated by SDS-PAGE and transferred onto nitrocellulose membranes (Bio-Rad, California, USA). Membranes were blocked with 5% non-fat milk in Tris-buffered saline pH 7.4 is used as a washing buffer for electrodes, alkaline phosphatase and peroxidase conjugates in dot blot assay. Furthermore, it acts as a pH stabilizer, which enables washing without disruption of antibody-antigen binding interactions. 0.5% Tween 20 and tris-buffered saline (TBST) was also added and then incubated with primary anti-Flag antibodies. After three washes with TBST, the membranes were incubated with horseradish peroxidase-conjugated secondary antibodies (Rockland Immunochemicals; 1:10,000). Flag-IP buffer (Flag (Sigma-Aldrich #F7425) was used for lysis in all experiments.

The Tat proteins previously isolated from the nucleofections (in loading buffer) were spun briefly and boiled 5 min at 95°C. Using the western blotting apparatus; the stack gel, plastic gel and loading station were placed together in preparation to run the blot. The loading station (which is beige) was filled with running buffer and held together forcefully using my fingertips to make sure there's no leakage. Following this the gel; was loaded with marker/ladder: PageRuler + Prestained Protein Ladder with capillary pipette tips followed by the samples. The first 30 min were run at 100V, then 1h at 120V. Following this the sandwich was prepared by putting the plastic sandwich into transfer buffer tray, black part into buffer, white part sticking out. Plastic Sandwich, 1x sponge, 2 x Whatman paper (size according to membrane), gel (cause proteins run from lower part (black) to upper part (white)), membrane, 2 x Whatman paper, 1x sponge. The membrane was activated with methanol for 1 min. This was then run in transfer buffer at 400 mA, for 1,5hr. Following this the membrane was blocked using 5% Milk and placed in 5-10 ml in 50mL Falcon at room temperature for 30min or overnight at 4°C. The following day the milk was discarded and add primary antibody (Flag) was added and left overnight at 4°C rolling for 2h at RT. The next day the blot was blocked for 1 hr at RT. The primary antibody was placed back into freezer -20°C (even with rest milk). Following this the blot was washed 3x 5min with PBS-T (0,1%), 10ml and secondary antibody (fluorescent) added diluted in milk. This was incubated for 1 hr and wrapped in aluminium foil due to light sensitivity. The blot was washed 3x 10min with PBS-T, 1 x 10 min with PBS, 10 ml and topped up with PBS to visualize. The images were developed using the gel imager.

3.1.5 Tat transactivation and Luciferase assay

The TZM-bl cell lines have *luciferase* (*luc*) gene which is under the control of LTR promoter. The pCTat plasmid containing the mutated Tat (TatP21) was transfected into TZM-bl cells separately from pCTat plasmid containing the consensus Tat (TatA21). Briefly, 40 000 TZM-bl cells were plated per well of 24 well plate in triplicate on day 0. Approximately 70-85% confluency was obtained on day 2. Then the transfection complex was prepared for each reaction by adding 1.5 µg of plasmid DNA (TatA21 or TatP21) to 240 µL of DMEM in an Eppendorf tube. In a separate tube 10 µl of polyethylenimine (PEI) (ThermoScientific, Waltham, MA, USA) was added to 250 µL of DMEM. Therefore, each reaction had two tubes. The two Eppendorf tubes were incubated separately at room temperature for 5 mins. The contents of the two tubes were then combined and incubated at room temperature for an additional 20 mins. Media was then aspirated off the cells and replenished with 1.5 mL of pre-

warmed DMEM (10% FBS, 1 % Pen/Strep). A total of 500µL PEI-DNA mixture was then added to the wells containing cells in 1.5mL media. The plate was gently rocked back and forth to evenly distribute the mixture. The plate was incubated at 37 ºC for 4-6 hours after which the media was replaced with 2 mL of complete media. The following day the media was aspirated off each well and then the luciferase assay was performed.

BrightGlo was thawed in a water bath at an ambient temperature, away from light. Then 1.85 µL of the culture medium was aspirated off the cells to leave 150 µL of culture medium in each well. Then 100 µL of brightGlo was added to each well and incubated for 2 mins in the dark. Cells in each well were mixed by two pipette strokes and 150 µL of lysed cells was transferred into the wells on a black round bottom 96 well plate. The plate was then read immediately using the Victor Nivo Multimode plate reader (PerkinElmer, Massachuettes, USA) and the results obtained using Microsoft Excel.

3.1.6 Functionality of Tat variants in Jurkat cells

To characterize the functionality of TatA21 and TatP21 in Jurkat cells we performed Amaxa nucleofections. Briefly, on the day 1 Jurkat cells were split to obtain a cell number of $3x10^5$ cells/mL in 24 hours. Pre-warmed 500 µL of RPMI medium was aliquoted in 1,5mL labelled tubes with perforated lab using a needle. And 5 mL of pre-warmed RPMI medium with serum and antibiotic (Pen/Strep) was aliquoted into a 6-well plate approximately 2 hours before commencing experiments. A total of 5-8 x 10^6 cells (per nucleofection) were centrifuged at 900 rpm for 10 min (RT). Then 2µg of each plasmid (TatA21 and TatP21) and empty vector (as a negative control) was added into labelled cuvette on the plastic side. After centrifugation the supernatant was discarded and cells resuspended in appropriate volume of solution R (100µL/nucleofection), this was gently mixed. Next, 100µL of resuspended cells was transferred into labelled cuvette – avoid making bubbles (use p1000).

The cells were nucleofected with Amaxa nucleofector (Amaxa II nucleofecter, Biosystems) using program O28 of this equipment. Following this the cells were immediately resuspended in 500µL of pre-warmed, serum-free RPMI lacking antibiotics (which was prepared in advance) using the provided sterile Pasteur pipette which comes with the kit. Avoiding lysed cells by not disturbing the mass of white cells which resemble lysed cells, nucleofected cells

were transferred to 1,5mL tube and allow them to recover at 37 °C in a 5% CO_2 incubator for 15 min. 500µL of the recovered cells were then transferred to 2-4mL of pre-warmed supplemented RPMI and incubated at 37 °C, 5% CO_2. Cells were collected at 12-, 24-, 36- and 48-hours. Fluorescence-activated cell sorting (FACS), RNA isolation and protein isolation were performed to determine nucleofection result. At each time point three tubes of 500µL of cells were collected, one kept for flow cytometry analysis and the other two centrifuged at 900 rpm for 10 mins for RNA protein isolation respectively. For RNA isolation the pellet was resuspended in 1ml of TriZol reagent (Thermofischer), and RNA isolated (protocol described in the following section) immediately or kept at -20°C and for protein isolation the pellet was resuspended in 1ml of cold lysis buffer (50 mM sodium acetate, 150 mM sodium chloride, 10% glycerol (v/v), and dH2O) with protease inhibitors (1:100) and Dithiothreitol (DTT) (1:1000). The tube was then spun at maximum speed (14 000rpm) for 10 mins. The supernatant was then removed and transferred into a new tube. Protein loading buffer was added to the supernatant and boiled for 5 mins if WB to be run immediately otherwise stored at -20 °C.

3.1.7 mRNA expression of tat, GFP and Gag

Total RNA was isolated using Total RNA Zol Out (A&A Biotechnology, Gdansk, Poland) according to the manufacturers protocol. Briefly, 1 mL Trizol was added to 200 µL of cells collected from the nucleofection described above. The samples from nucleofection were mixed by inverting so that they become homogenized and look a paler pink. This was incubated at room temperature for 5 mins and then centrifuged at 10 000 RPM, 15 minutes, 4°C. The supernatant was transferred to a new 15 mL tube and 2 mL of isopropanol was added to 1 mL supernatant. This was thoroughly mixed and 700 µL of the mixture was applied onto the mini columns. This was centrifuged for 1 min at 10 000 RPM. In case of more mixture than 700 µL, this was repeated. Mini columns were transferred into new 2 ml tubes and 700 µL of A1 wash solution was added. This was then centrifuged for 1 min at 10 000 RPM and the flow-through was discarded from the 2mL tubes and the mini columns was placed in the same 2 ml tubes. Another 700 µL of A1 wash solution was added and centrifuged for 1 min at 10 000 RPM. The flow-through was discarded from the 2mL tubes and mini columns were placed again in the same 2 ml tubes. 200 µL of A1 wash solution was added and centrifuged for 2 min at 10 000 RPM. The mini columns were transferred into a new 1.5 mL tubes and 40 µL of sterile water was added directly onto the centre of column membrane. This was then incubated for 3 min at room temperature and centrifuged for 1 min at 10 000 RPM. The

concentration of the RNA was quantified using a NanoDrop and used immediately or stored at -80 °C until use.

The extracted RNA was used for complimentary DNA (cDNA) synthesis, which performed using the Superscript II Reverse transcriptase kit (Thermofischer, California, USA) according to the manufacturers protocol. Reverse transcription reaction comprised of 1 µL Tat specific primers (TatR1) to 5 µl total RNA, 1 µL 10 mM each dNTP Mix 1 and distilled water added to 12 µL final reaction volume. The mixture was heated to 65°C for 5 min to denature any secondary structures of the primers and quick chilled on ice to prevent these structures from re-forming. The contents of the tube were collected by brief centrifugation and the following added: 1 X First-Strand Buffer, 0.1 M DTT and 1 µL RNaseOUT™. The contents of the tube were mixed and incubated at 25°C for 2 min. Following this, 1 µL of SuperScript™ II RT was added and mixed by pipetting gently up and down and sterile, distilled water added to a 20 µL final volume. The tubes were incubated at 25°C for 10 min, 42°C for 50 min and inactivated at 70°C for 15 min. The synthesized cDNA was used immediately as a template measure Tat and GFP expression or stored at -20°C until use.

Qualitative PCR (qPCR) reactions were performed using SBYR Green Master mix (BioRad) and 2 µl of cDNA. Each PCR reaction consisted of 3 mmol/µL MgCl2, the respective primers at 0.5 pmol/µL, 1 µL Fast Start SYBR Green I (Roche), 1 µg cDNA and water to make up the total reaction volume to 10 µL. Glyceraldehyde 3-phosphate dehydrogenase (GAPDH), (GenBank accession number NM_002046), determined to be the most suitable reference gene. The SYBR green qPCR was performed using the following cycling conditions were used: 15s 95 °C and then 30 s at 55 °C and 60 s at 60 °C for 35 cycles.

3.1.8 Electrophoretic mobility shift assays (EMSA)

As aforementioned, our preliminary data demonstrated that TatA21 mutation resulted in significantly lower ($p = 0.0181$) viral loads at one year post infection (Mkhize et al., unpublished). Therefore, we wanted to investigate the effect of Tat P21A mutation on the ability of Tat to bind TAR element on the 5' LTR, the vial promoter. To investigate whether Tat mutant P21A had an effect of the ability of Tat to bind TAR, EMSA was performed. TAR RNAs (WT, Δbulge and Δloop) were synthesized in in vitro transcription reactions with the

Riboprobe system (Promega) as previously described (Kaehlcke et al., 2003). Transcripts were treated with 2 U of DNase I (Promega), extracted with a phenol:chloroform mixture and purified over an illustra MicroSpin G50 column (GE Healthcare). TAR expressing vectors pEV101, pEV102, pEV103 (TAR WT, dBulge, dLoop) were linearized with HindIII. P32-labeled TAR RNAs (WT, Δbulge and Δloop) were synthesized in *in vitro* transcription reaction with the Riboprobe system (Promega) as previously described (Kaehlcke et al., 2003). Transcripts were treated with 2 U of DNase I (Promega), extracted with a phenol:chloroform mixture and purified over an illustra MicroSpin G50 column (GE Healthcare). Gel-mobility reactions (16 µl final volume) were carried out in binding buffer (50 mM Tris, pH 7.4, 0.5 mM EGTA, 150 mM NaCl, 2% glycerol, 0.2% Tween 20, 0.5 mM DTT, 90 mM ZnSO4, 0.005% BSA and 100 µM ATP) and contained 2×10^4 cpm TAR probes/reaction and the indicated concentrations of Tat 2.5ng, 10ng and 40ng. Supershift experiments were conducted in the presence of 7 µg of α-Set7/9-KMT7 antibodies or corresponding amounts of pre-immune rabbit serum. Reactions were incubated for 30 min at 30°C and separated on a pre-run 4% Tris-glycine gel (Pagans et al., 2010).

3.1.9 Protein modelling and molecular docking

HIV-1 Tat subtype C protein structures exhibiting the consensus (TatA21) and mutated (TatP21) amino acid at position 21 were generated by homology modelling using Modeller v9 software and Chimera UCSF software (Pettersen et al., 2004) based on a solved crystal structure of Tat subtype B (PDB: 3MIA, chain C) (Tahirov et al., 2010). These models were used as the models for molecular docking experiments. To investigate the binding of TatA21 vs TatP21 to TAR *in silico in silico* docking experiments were performed using HDOCK online website server (Yan et al., 2017) using a solved structure of TAR RNA (PDB: 6MCE, chain A) (Pham et al., 2018) as a receptor and the generated homology models for TatA21 and TatP21 structures. For comparison, molecular docking was also performed for Tat subtype B (PDB: 3MIA, chain C) (Tahirov et al., 2010) using TAR RNA (PDB: 6MCE, chain A) (Pham et al., 2018) as a receptor. No ligand/receptor binding residues were pre-specified.

Protein modelling was used to model the structures of different Tat variants. This is because there are no solved protein structures of Tat subtype C, Tat subtype B structures were used as a reference in UCSF Chimera to model what the Tat subtype C structures would look like.

Next, 3D models of TatA21 and TatP21 were generated using UCSF Chimera (Pettersen et al., 2004). Lastly, molecular docking analysis was used to determine the binding affinity of TatA21 vs TatP21 to the TAR RNA using the Hdock server. The resulting solutions were ranked based on the highest binding affinity (or lowest binding energy). Figures were created using Biovia discovery studios.

3.1.10 Sequential co-immunoprecipitation of chromatin (co-IP) assays

To perform co-immunoprecipitation (co-IP) experiments, TatA21 and TatP21 plasmids were transfected into TZMbl cells as previously described. The cell lysate was collected at 24 hours following transfection. At least 500,000 cells were lysed per condition. Cells were washed with PBS and lysed in ice-cold lysis buffer (100 mM NaCl, 10 mM Tris, pH 7.5, 1 mM EDTA, 0.5% Nonidet P-40, protease and phosphatase inhibitor cocktail [Roche]). Cell lysates were denatured in prepared Laemmli sample buffer (Thermofischer) and incubated for 5 min at 95 °C. The proteins were separated by SDS-PAGE and transferred onto nitrocellulose membranes (Bio-Rad). Membranes were blocked with 5% non-fat milk in Tris-buffered saline pH 7.4 and 0.5% Tween 20 (TBST) and then incubated with primary antibodies CDK9 (Cell Signaling #2316S), Cyclin T1 (Santa Cruz #sc-8127 and #sc-8128 at 4°C overnight. The following day, the membrane was washed three washes with TBST, the membranes were incubated with horseradish peroxidase-conjugated secondary antibodies (Rockland Immunochemicals; 1:10,000) HA (clone 3F10 – old lab stock, probably from Roche), GAPDH (Santa Cruz #sc-365062). The blot was developed using previously described western blot method in section 3.1.4.

3.2 Goal Two Methodology overview:

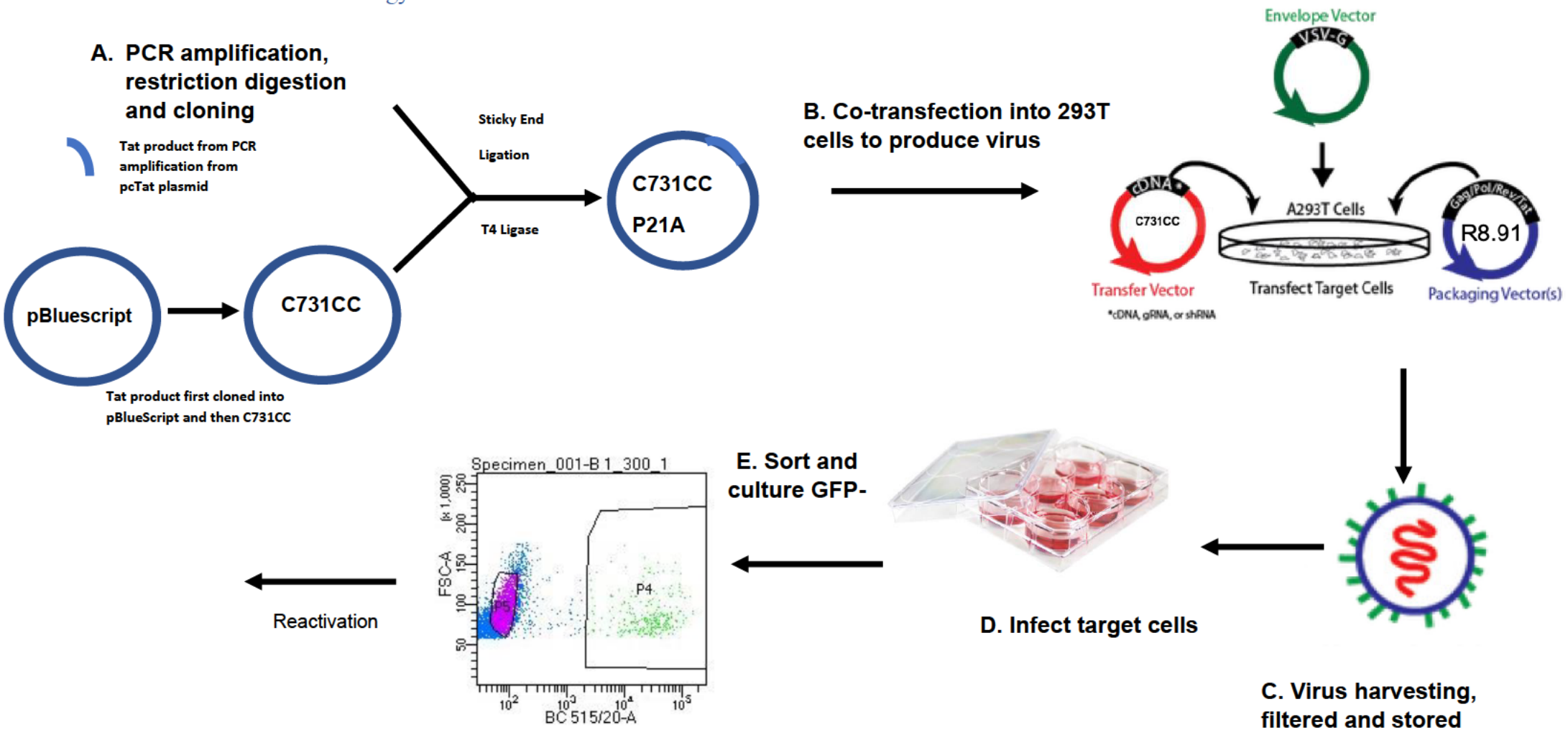

Figure 3.2: Goal two methodology overview. Generation of recombinant C731C$_{TatP21}$C. **B**. Co-transfection together with VSV-G and R8.91 into Jurkat cells. **C**. Virus collection. **D.** Virus used to infect target cells. **E**. Cell sorting of GFP- cells.

3.2.1 TatP21 plasmid construction

The construction of the viral vector exhibiting a mutant HIV-1C TatP21 was achieved by introduction the mutant TatP21 in place of the HIV-1C consensus Tat (TatA21) into a viral vector comprising of HIV-1C 5'LTR-TatA21-IRES-GFP-3'LTR (referred to as C731CC). The C731CC is available in the Madlala lab (Maikoo et al., data unpublished), which was adapted pEV731 minimal genome reporter virus published by Eric Verdin's group (Jordan et al., 2001). Briefly, TatP21 was PCR amplified from pCTatP21 plasmid as described in section 2.1.1 above, this time using primers TatC_Cla1F 5' - ATC GATACTCGACAGAGGAGAGCA AG – 3' and TatC_BsiW1R 5' - GCA TTACATGTACTACTTACTGCGCAT GC – 3' to insert restriction sites (bold and underlined), ClaI and BsiWI for cloning into the pBluescript plasmid. The TatP21 PCR product was analyzed on 1% agarose Gel electrophoresis to confirm the correct size band.

Next, the TatP21 PCR product was then digested using BsiWI (New England BioLabs, Ipswich, MA) and ClaI (New England BioLabs, Ipswich, MA). Briefly, restriction digestion of TaP21 PCR product was performed in a double digestion reaction consisting of 1X CutSmart Cut Buffer®, 1U of BsiWI and ClaI each, 100ng of the TatP21 PCR product and sterile water to make it up to a total of 10µL reaction volume was performed at 37 °C for 1 hour. The same restriction enzymes, BsiWI and ClaI were used to perform a double restriction digestion of 10 µg of pBlueScript plasmid to linearize and create the same sticky ends as the TatP21 PCR product. Briefly, a 10 µL digestion reaction consisted of 1X CutSmart® Buffer, 10U of BsiWI and ClaI each, 10µg of the pBluescript plasmid DNA and sterile water to make it up to a final volume of 10µL at 37 °C for 1 hour.

The TatP21 product was cloned into the pBlueScript and transformed as previously described. Following this the pBlueScript TatP21 was cloned into the C731CC to create $C731C_{TatP21}C$ using T4 DNA ligase as previously described. However, this was unsuccessful.

Alternative approach was, 100ng of TatP21 gene block (WhiteScientific (Pty) Ltd) (gBlock sequence of Tat exon 1 and 2 shown in supplementary material) was digested using BsiWI and ClaI as described for TatP21 PCR product. The C731CC was also digested with the same restriction enzymes, BsiWI and ClaI to cut out TatA21 and linearize the vector to have $C731C_{\Delta}C$ as described for pBluescript above, that is the C731CC contract that is now lacking the *tat* gene. The digested TatP21 gene block was then cloned into a viral into C731CC to

generated C731C$_{TatP21}$C. Ligation was performed using T4 DNA ligase where 0.5µL of T4 DNA ligase was combined with and 1 µL of T4 DNA Ligase Buffer, 50 ng of digested C731CC, and 80 ng of digested TatP21 gene block. Restriction Digest of Tat geneblock-C731CC clone to confirm insert- using ClaI and BamHI-HF was performed using the same digestion conditions described above. (Rossenkhan et al., 2013).

Following this the confirmed C731C$_{TatP21}$C vector construct was transformed into DH5α cells. Briefly, the DH5α cells were thawed on ice and split into 1.5 ml tubes containing 50 µL of competent cells. Following this 2µL of the plasmid was inserted into the 50µL cells and the tube gently flicked to mix the solution. The tubes were placed on ice for 20 mins. Following this the cells were heat shocked for 45 secs and placed back on ice for 2 mins. The tubes then had 900 µL of LB broth added to them and incubated at 37 °C for 1 hour. Following this 100 µL of the cells was plated onto Ampilicin LB agar plates and incubated at 37 °C overnight. The following day colonies were inoculated for plasmid maxiprep as described above using the Qiagen MaxiPrep kit. The plasmids were stored at -80°C and used for virus production.

3.2.2 Production of HIV-1C TatP21 minimal GFP reporter

Virus particle production and cell infections was performed according to (Jordan et al., 2003). Briefly, 293T 2 million cells were plated in a 10 centimeter (cm) petri dish a day before the experiment in supplemented DMEM (10% FBS, pen/strep, HEPES). However, just before we add our transfection mix to the cells serum free DMEM was used to replace the supplemented DMEM. Then following plasmids were required for transfection: On day 3 of cell culture, a 10:1 PEI:DNA ratio was prepared by adding 12,5 µg (4,5 µg packaging [R8.91], 2 µg Envelope [VSV-g], and 6 µg Lentivirus [C731CC]) plasmids into a 1.5 mL tube and this was topped up with serum free DMEM media to 500 µL. In 125 µL PEI was added in a separate 1.5mL tube and this was topped up with serum free DMEM media a total volume of 500 µL media. The tube containing PEI was incubated for 5 min at RT to allow the solution to homogenous. The content of the 1.5 mL tube containing 12.5 µg plasmid DNA was mixed with the content of 1.5 mL containing PEI tube and the mixture was incubated at RT for 30 min. The 1 mL mix drop wise slowly to the cells while mixing and incubated for 6 hours. After 6 hours, the media was changed and then incubated at 37 °C overnight. For Virus collection: collect virus: 24, 36 and 48 hours after media change.

Jurkat cells were then plated in a 6-well plate at a decent density 400 000 cells/mL and infect with different dilutions of the C731CC virus. At day 4 post-infection, BD LSRFortessa™ flow cytometer (Becton Dickinson, San Jose, CA) was used to determine the percentage of GFP positive cells representing infection. The C731CC virus dilution that yielded around 5% GFP positive cells, which represent about 95% either truly negative (C731CC uninfected) or latently infected cells was selected.

3.2.3 Cell sorting and reactivation

A two-parameter analysis to distinguish GFP-derived fluorescence from background fluorescence was used: GFP was measured in FL1 and cellular autofluorescence was monitored in FL2. Electronic compensation was applied during analysis. Analysis was gated on live cells according to forward and side scatter. A gate (R2) containing GFP-positive cells was drawn compared to an uninfected control, and the data shown refer to the percentage of cells in R2 or mean fluorescence intensity (MFI) of those cells. Cell sorting was carried out with a FACSVantage (Becton Dickinson).

The Jurkat cells were grown to ~95% viability and were infected at low multiplicity of infection (MOI), such that approximately 5% of cells were infected. Briefly, 4 x 105 cells/mL were seeded in a 6 well plate followed by the addition of 1/1 dilution (untitrated) of each of the virus harvested at the different time points. After determining which time point gave the most infectious virus (determined by the percentage GFP positive cells by flow cytometry detailed below), the virus at that time point was then diluted to determine which dilution gives ~5% GFP positive cells. This would suggest that the remaining ~95% GFP negative cells are either truly negative (uninfected) or latently infected. The following virus dilutions were used: 1/1; 1/5; 1/25; 1/125; 1/625. Following 96 hours of incubation at 37°C in a humidified 95% air-5% CO2 atmosphere, the virus dilution that gave ~5% GFP positive cells was then sorted for the GFP negative cell population using a BD FACSAria™ Fusion Flow Cytometer (Becton Dickinson) and treated with latency reversing agents (LRAs).

GFP expression was analyzed by flow cytometry (gating shown in supplementary material). The live population was defined by forward versus side scatter profiles. Gating for SSC-H vs SSC-W and FSC-H vs FSC-W was used to exclude doublets. Cells were further gated by using forward scatter versus GFP intensity to differentiate between GFP-positive and -negative cells.

Immediately following cell sorting of the GFP negative cells, the cells were centrifuged at 1500 rpm for 10 minutes and resuspended in RPMI 1640 medium containing L-glutamine (Thermofisher Scientific, United States) supplemented with 10% Fetal bovine serum (FBS), 100 µg/ml penicillin-streptomycin, and 1% HEPES. The following LRAs were added, alone or in combinations as follows: 2mM phorbol 12-myristate 13-acetate (PMA); 2 µg/mL prostratin; 1500nM suberoylanilide hydroxamic acid (SAHA); and 2 µg/mL tumor necrosis factor alpha (TNF-α). This was followed by 24 hours incubation at 37°C in a humidified 95% air-5% CO2 atmosphere. The percentage reactivation was then measured as the percentage GFP positive cells by flow cytometry as described above.

3.2.4 Statistical analysis

All statistical analyses were performed using GraphPad Prism software. The data obtained from this study was used to perform paired T-tests. A * $p > 0.05$, **, $p > 0.01$ and ***, $p > 0.001$.

Chapter Four

Chapter Four: Results

4.1 Single Tat mutants resulted in reduced LTR transactivation

Genetic variation within the patient derived HIV-1C Tat has been reported and of up to 38% sequence variation in Tat can tolerated without any change in its transactivation potential (Opi et al., 2002). In 2012, Neogi *et al* showed that inter and intra subtype variation existed with signature residue Alanine 21 in Tat subtype C (Neogi et al., 2012), once again which is referred to as TatA21 in this thesis. In a study by Rossenkhan *et al*, they showed that patient derived Tat variants resulted in differential functionality. Moreover, they showed by SDM that the presence of selected single mutants (V41, P21, Y29R, L35Q/Q39L, R53K) resulted in the increase or decrease of Tat transactivation activity (Rossenkhan et al., 2013). The data from my Masters study showed that patient derived HIV-1C Tat exhibited differential transactivation activity. Some patients with mutations previously reported mutation that reduces Tat activity, namely TatA21, showed almost abolished Tat activity. Therefore, in this study we hypothesized that TatA21 alone or in combination with other mutations may be responsible for reduction in Tat activity. To this effect, SDM was used to introduce Tat mutations which have been previously reported to increase Tat activity (A21P, Y29R, L35Q/Q39L) and reduce Tat activity (V4I, L35Q, Q39L, S46Y and R53K) individually or in combination (Fig. 4.1). The introduction of A21P (TatP21) was to compare the effect of this mutation to the consensus TatA21 which has been reported to reduce Tat activity (Rossenkhan et al., 2013). Furthermore, patient derived Tat variants exhibiting TatA21 together with other mutations were associated with lower viral loads in people living with HIV (PLWH) as shown by data from my Masters project (Mkhize et al., manuscript under preparation for publication at the time of submitting this thesis). The sequences of the Tat variants are shown below (Fig. 4.1).

Figure 4.1: Sequences of Tat mutants introduced in HIV-1 subtype C consensus Tat (using site directed mutagenesis). The sequences were aligned to a HIV-1C consensus. The dots represent amino acid residues which are the same as the consensus whereas a letter represents an amino acid residue that is different from the consensus.

Each of the mutations which were introduced are shown on the right with the consensus sequence as the first sequence shown in full. Each letter represents an amino acid in the consensus sequence, the presence of the dot shows that the amino acid in sequenced sample is the same as the consensus sequence and the presence of a letter indicates that the amino acid at that position differs from the consensus. Each single mutant introduced is shown by the presence of a letter to represent the amino acid which was introduced to the consensus sequence to create a single mutant.

To investigate the functionality of the mutants above, a Tat transactivation assay was performed. Next, consistent with a previous report (Rossenkhan et al., 2013) we hypothesized that mutations which have been previously reported to reduce Tat activity will show a lower relative change in LTR and Tat mutants previously reported to increase Tat activity will show a higher relative change in LTR activity than the consensus. For the purposes of this study TatP21 is depicted as the reference TatP21 in the figure below (Fig. 4.2). The measure of the relative change in LTR activity is an indication of Tat transactivation activity.

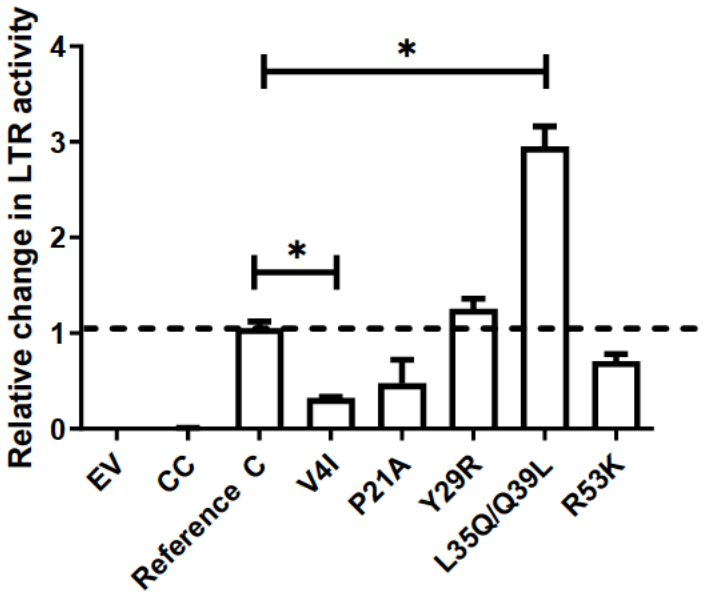

Figure 4.2: Relative change in LTR activity for single mutants reported to increase (Y29R/ L35Q/Q39L) or decrease (V4I, P21A, R53K) Tat transactivation activity and co-immunoprecipitation obtained from site-directed mutagenesis. EV – empty vector, CC – cell control (cultured cells only). The reference C is TatP21 subtype C. A * indicates statistically significant difference from the reference.

As expected, the cell control (CC) and empty vector (EV) showed no LTR transcription activity. Consistent with a previous report V4I, P21A and R53K were associated with reduced transcription activity compared to the reference TatP21 in this study (Rossenkhan et al., 2013). Specifically, V4I was associated with significantly lower relative change in LTR activity ($p = 0.0162$) than the reference TatP21. Consistent with previous study (Rossenkhan et al., 2013), to the data from this study show thatL35Q/Q39L double mutation was significantly associated with significantly higher LTR transcription activity compared to reference TatP21 ($p = 0.0349$). Although the Y29R mutation was reported to significantly enhance the LTR transcription activity (Rossenkhan et al., 2013), it exhibited marginally higher LTR transactivation ability than the consensus Tat.

4.1.2 Protein expression of TatA21 and TatP21 was similar

Consistent with a previous study (Rossenkhan et al., 2013)the data from this study clearly demonstrate that TatA21 significantly reduced the LTR transcription activity compared to TatP21. However, the mechanisms by which TatA21 results in reduced transactivation activity remained to be determined. Therefore, we hypothesized that the TatA21 was associated with

reduced Tat protein expression compared to TatP21. To determine whether the protein expression of these two Tat protein variants, TatA21 and TatP21 differed, Western blot analysis was performed. The band sizes between the two variants were identical even when the film was exposed for three different time periods, each resulted in the same size band (Fig. 4.3).

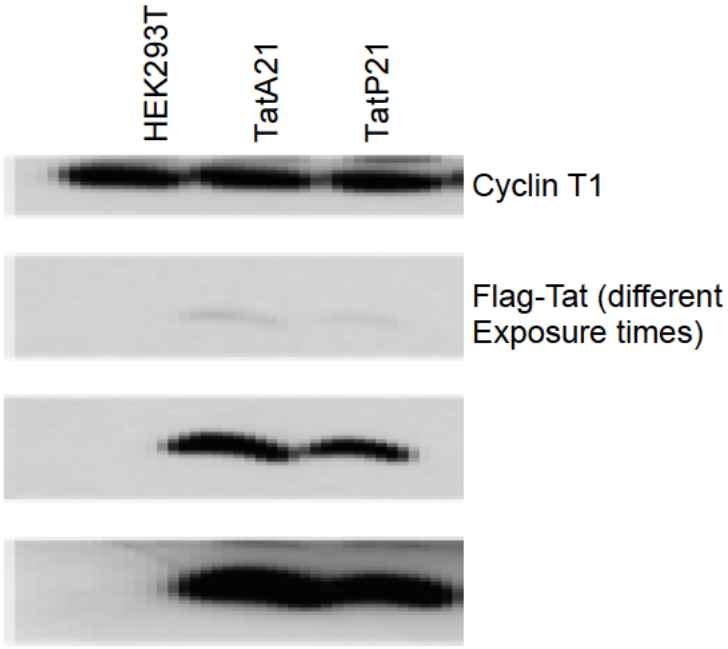

Figure 4.3: Overexpression of Tat mutants in HEK293T cells. The lane with HEK293T is cells only. Cyclin T1 was expression in HEK293T, TatA21 and TatP21. The TatA21 lane shows expression of a band in the flag-Tat as well as the TatP21. The first row showing flag-Tat exposure the film was exposed for approximately 10 secs, the following row shows film exposure for 1 min and the third row shows expression after film exposure for 5 mins.

Although TatA21 showed significantly reduced transactivation activity compared TatP21 in TZM-bl cells, TZM-bl cells harbours a subtype B LTR (Pagans et al., 2010). A previous study has reported that HIV-1 subtype B exhibit TatP21 (Neogi et al., 2012). Therefore, we reasoned that the differential transactivation activity observed between TatA21 and TatP21 could be due to the fact that the LTR activity of a subtype B was being assessed using subtype C Tat variants (TatA21 and TatP21). Therefore, we hypothesized that TatA21 and TatP21 may transactivate subtype C LTR to similar levels. To this effect, the ability of TatA21 and TatP21 to transactivate subtype C LTR contained in subtype C LTR-Tat-GFP-LTR (C731CC) latently infected Jurket cell lines, J-LatC cells (referred to as JL-C66 in this study) which has the consensus TatA21 and A72 cell lines latently infected with LTR±GFP construct (Jordan et al.,

2001) which does not have endogenous Tat. To this effect, TatA21 and TatP21 were nucleofected into JLat-C66 cell line, and Clone A72 cell lines. Following this we performed quantitative PCR (qPCR) to determine the mRNA expression of different targets (Gag, GFP and CypA as a house keeping gene) and the expression of Gag was further assessed by FACS. The results show that CypA was ubiquitously expressed in both conditions (TatA21 and TatP21) (Fig. 4.4A). The percentage GFP positive cells as measured by FACS was higher in TatP21 than TatA21 at 24 hours (Fig. 4.4B). We investigated whether there would be differences in GFP expression when we transfected the Tat variants into the A72 cell line. Results showed that GFP expression was significantly higher ($p = 0.0439$) in TatP21 than TatA21. Moreover, although not significant, Gag expression was slightly higher in TatP21.

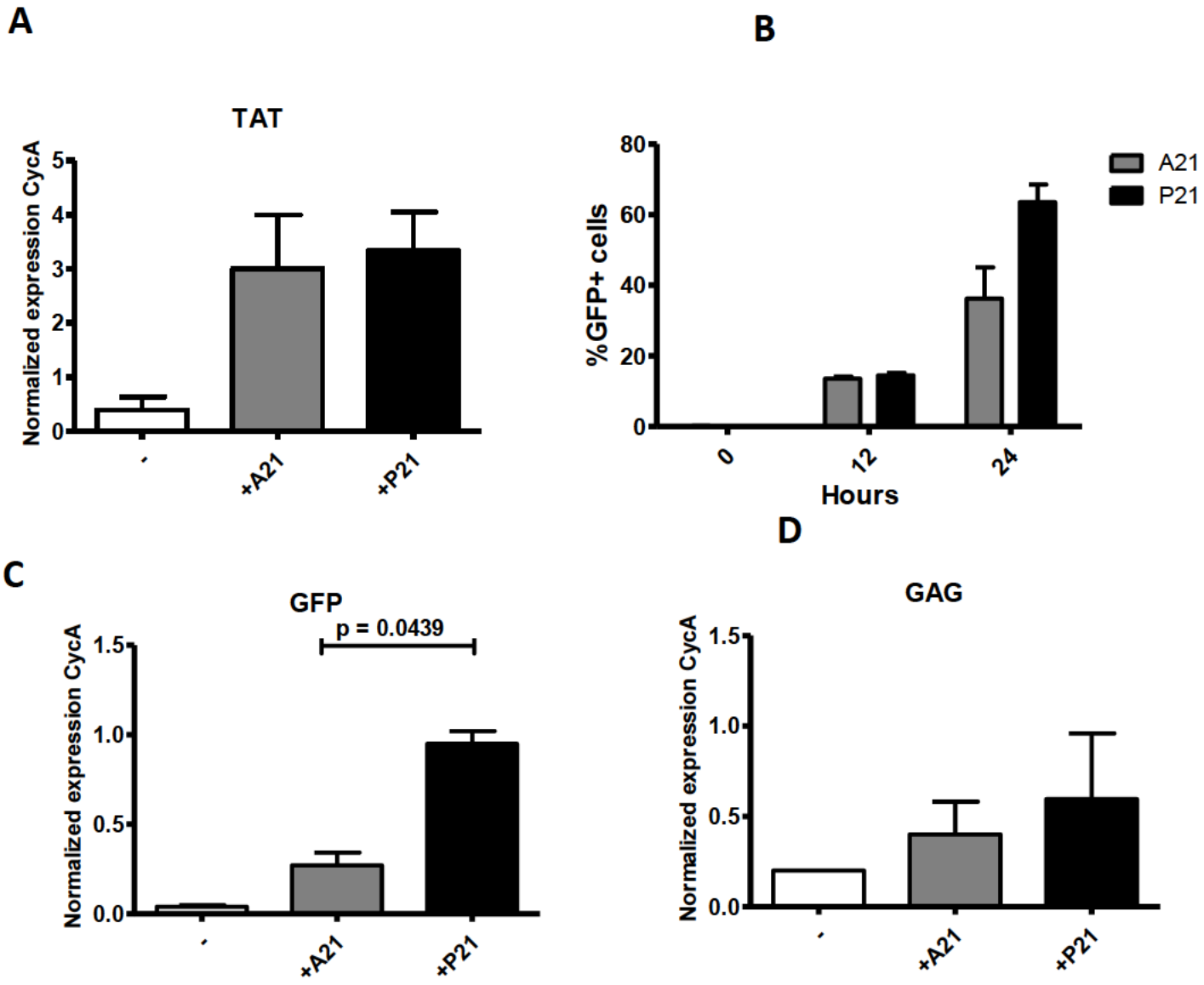

Figure 4.4: Tat mRNA expression. **A.** mRNA expression of Tat in C66 cell line. **B.** Percentage GFP positive cells in TatA21 and TatP21 as measured by Flow cytometry after 12 and 24 hours in C66 cell line. **C.** mRNA expression of GFP as measured by qPCR in A72 cell line. **D.** mRNA expression of Gag as measured by qPCR in A72 cell line.

Taken together our data show that the house keeping gene CypA was equally expressed in both conditions, TatA21 and TatP21.Interestingly, the data from this study confirms that TatA21 is associated with lower LTR driven gene expression in different cell lines but does not affect the expression of Tat. Therefore, the mechanisms by which TatA21 reduce LTR activity remains to be determined.

4.1.3 Electrophoretic mobility (RNA shift) assay

As aforementioned, our preliminary data demonstrated that Tat P21A mutation resulted in significantly lower ($p = 0.0181$) viral loads at one year post infection (Mkhize et al., unpublished). Therefore, we wanted to investigate the effect of Tat P21A mutation on the ability of Tat to bind TAR element on the 5' LTR, the vial promoter. To investigate whether Tat mutant P21A had an effect of the ability of Tat to bind TAR, EMSA was performed. Our results revealed a Tat-TAR complex as shown by the presence of a band where this complex forms (Fig. 4.5).

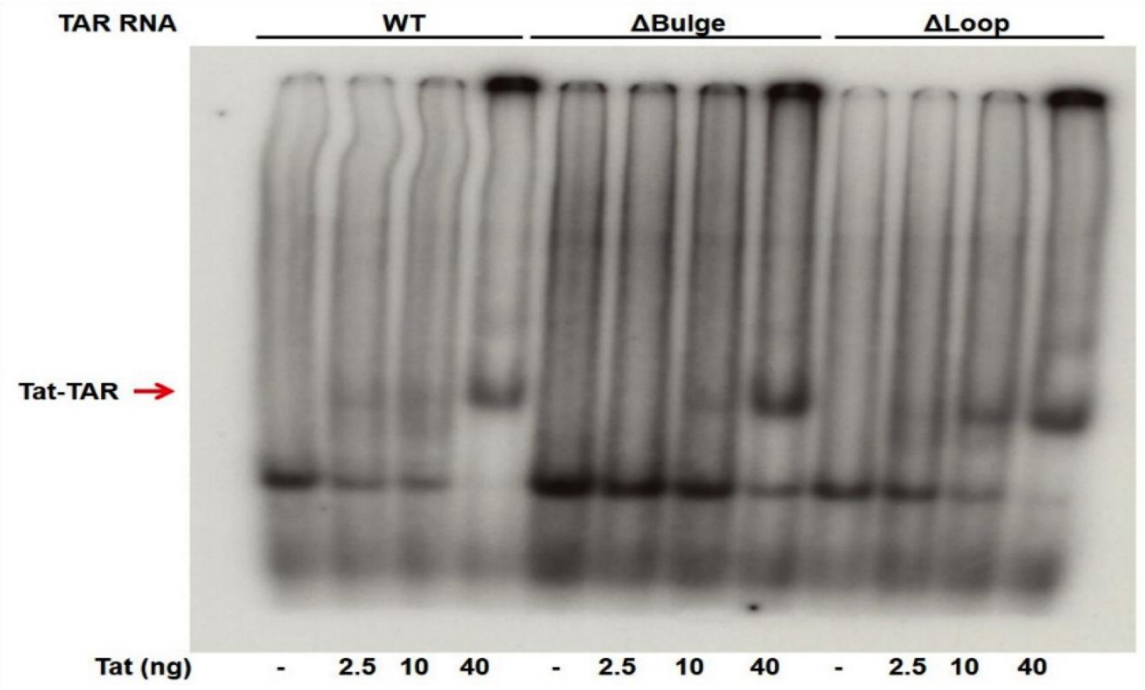

Figure 4.5: RNA Gel Shift experiments of TatA21. P32-labeled TAR RNAs (WT, Δbulge and Δloop) were synthesized in in vitro transcription reactions with the Riboprobe system (Promega). The presence of the Tat-TAR complex is shown by the red arrow. Different concentrations of Tat in nanograms (ng) were assessed.

The presence of the Tat-TAR complex suggests that the TatA21 mutant is able to bind TAR in the wildtype (WT). Interestingly, there is a Tat-TAR complex present in the TAR without the bulge (Δbulge) and without the loop (Δloop) in the TAR RNA loop structure. Thus, suggesting that TatA21 is able to bind anywhere along the TAR RNA and thus this mechanism is responsible for reduced transactivation activity occurs. Taken together, our data demonstrates that although Tat P21A mutant results in reduced transactivation, it is still able to bind P-TEFb

and TAR during HIV replication. Thus, further investigation is required to tease out the mechanism responsible for observed reduced transactivation.

4.1.4 Tat protein and TAR RNA homology modelling

The structure of Tat subtype B has been well studied and well elucidated in literature using NMR and X-ray crystallography (Tahirov et al., 2010, Schulze-Gahmen et al., 2013). However, there is paucity of Tat subtype C structure elucidation which is studied here. To elucidate the structures if TatA21 vs TatP21 and the basis for molecular docking in the following section, homology modelling was performed. The three-dimensional (3D) structures showed no differences nor conformational changes. The tail end of the protein has been reported to be flexible and can take any direction. Here the bulk of the proteins represented by the beta sheets flat representation is the same between the two variants. Given that most of the structures which have been elucidated for Tat are for subtype B, this was used a model to produce the structures below. The sequences showed an approximately 70% similarity between the two subtypes and thus these figures represent an accurate representation of each variant.

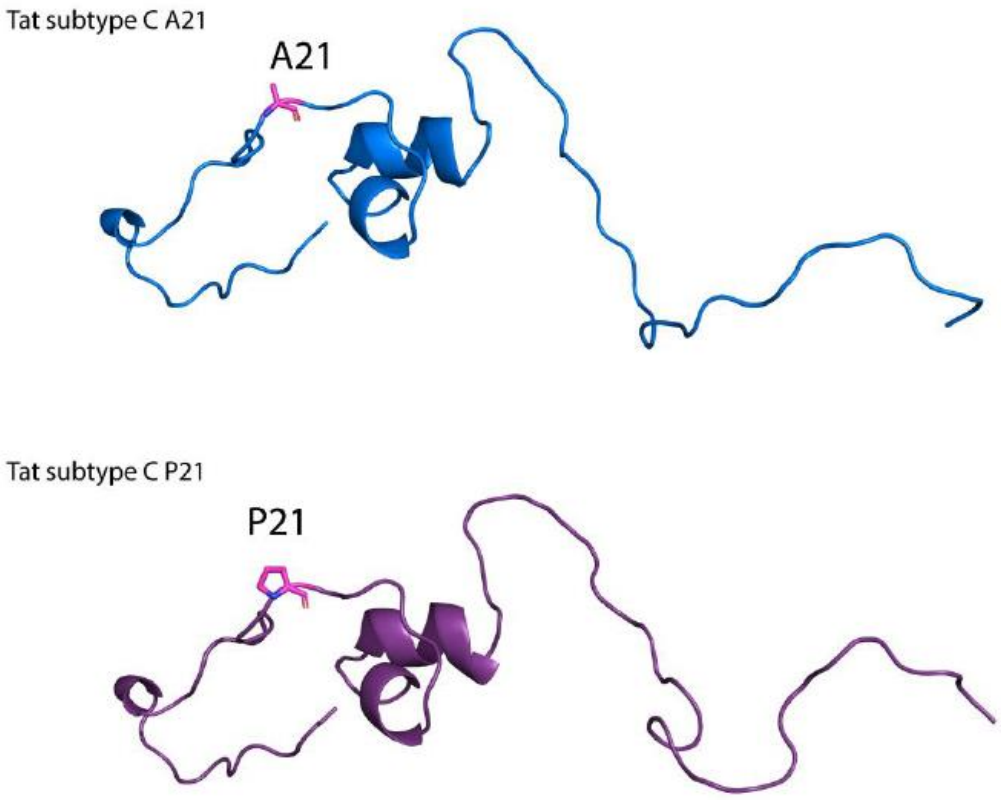

Figure 4.6: Homology models of Tat proteins. **A**. Homology model of TatA21 **B**. Homology model of TatP21. Each amino acid structure is shown in pink with the proline being part of the three-carbon R-group and has a rotationally constrained rigid-ring structure whilst alanine contains an amine group and a carboxylic acid group which are both attached the central carbon atom which has a methyl group side chain.

Protein structures can be verified in several ways. One such is through Ramachandran plots. Here we show that each of the structures produced do not fall in the forbidden areas on the graph. These are classified as the two right panels in each graph. Where the structures fall on each graph is highlighted in yellow.

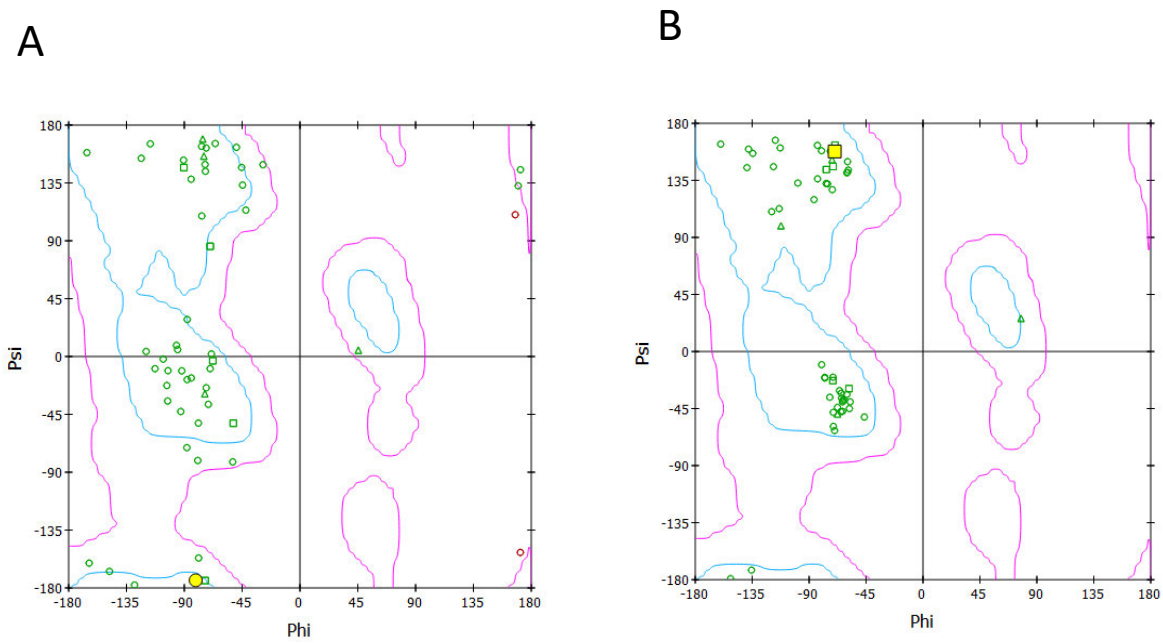

Figure 4.7: Ramachandran plots of protein docking models. **A.** Wildtype Tat transactivation domain docked to subtype C TAR sequence. **B.** Wildtype Tat transactivation domain docked to subtype C TAR sequence. Highlighted yellow residue is where the amino acid at position 21 falls within the allowed regions of the Ramachandran plot.

It is well known that Tat binds to the TAR element in the 5' LTR during HIV replication (Dingwall et al., 1989). We wanted to elucidate the structure of the TAR RNA in subtype B and C given that the Tat transactivation assay was performed with TZM-bl cells which have a subtype B LTR. The sequences of the two TAR elements are very similar with only two nucleotide differences. The 3D structures here also reveal that they share the same structure and thus have the same functionality when bound to Tat.

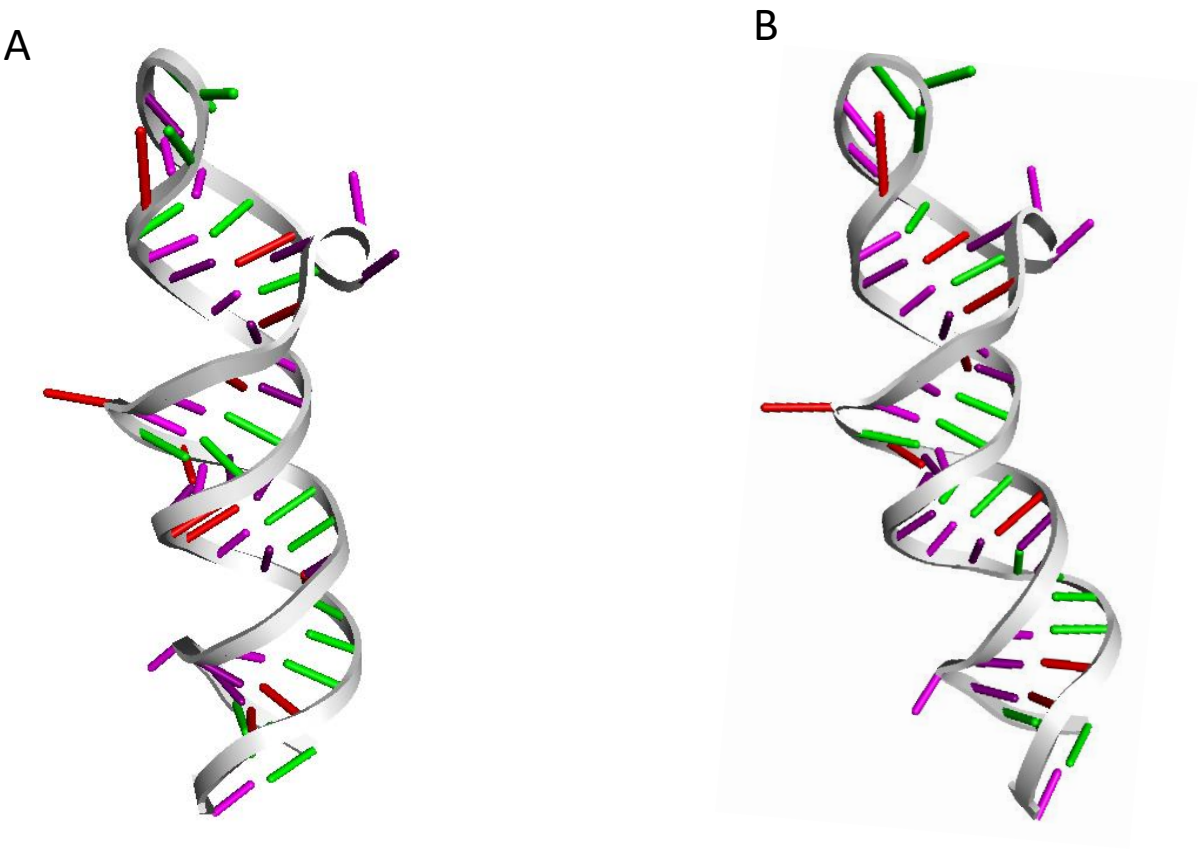

Figure 4.8: TAR RNA 3D structures. A. Subtype B TAR RNA. B. Subtype C RNA.

The TAR RNA 3D structures are shown visually where the grey represents the helix backbone of the structure with the different coloured cylindrical shapes representing different nucleotides. There is similarity in the TAR RNA structures of subtype B and C.

4.1.5 Molecular docking of Tat to TAR RNA revealed different docking scores

To determine the binding affinity of TatA21 vs TatP21 to the TAR RNA *in silico*, we performed molecular docking using the Hdock server. Each Tat variant modelled in this study was used a template target in the server. The TAR RNA sequences were the ligand input. The results showed that TatP21 is able to bind more efficiently to subtype C and B TAR RNA sequences. This was showed by the molecular docking score. The lower score represents high binding affinity.

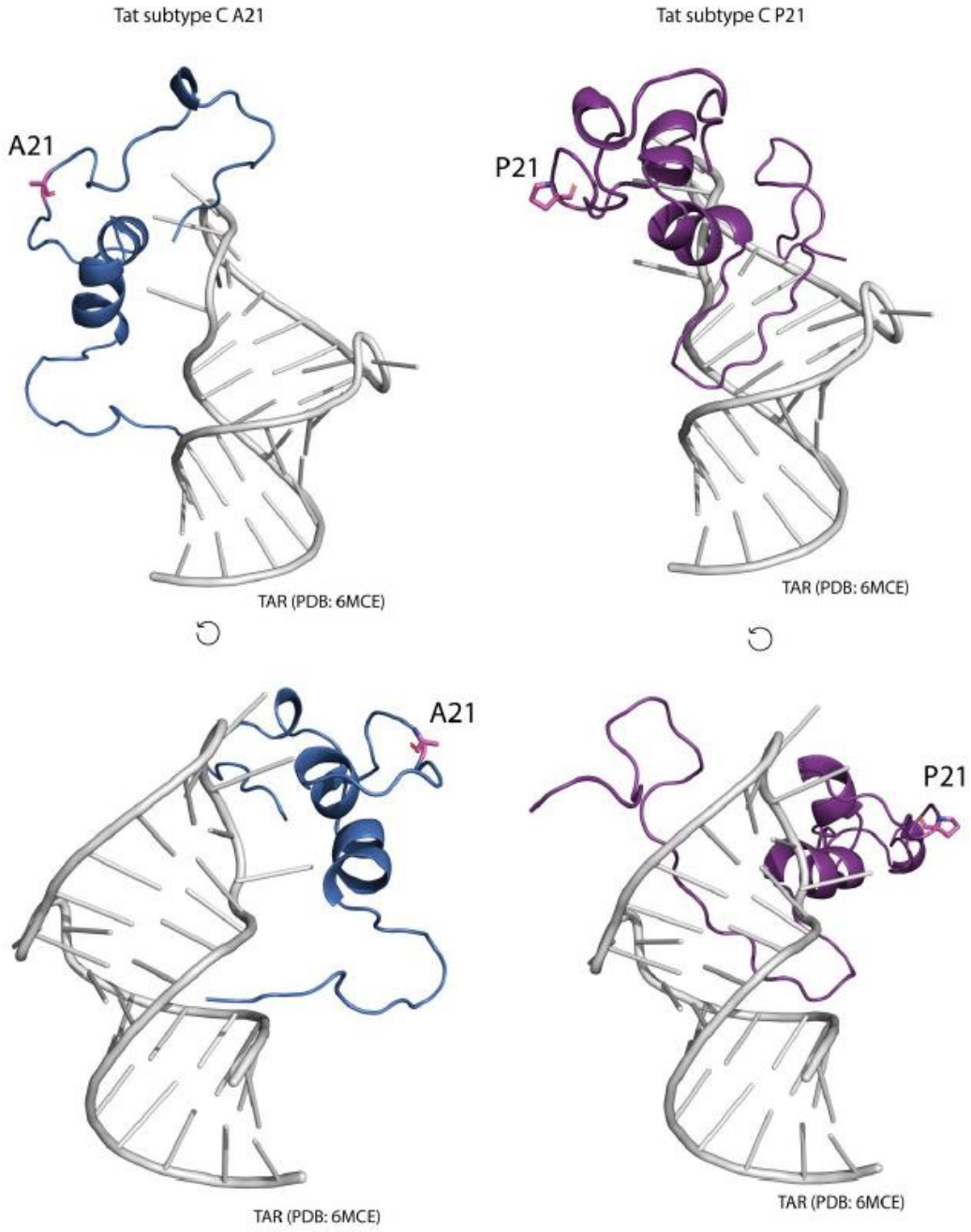

Figure 4.9: Molecular docking of TatA21 and TatP21. **A.** Wildtype Tat transactivation domain docked to subtype C TAR sequence using Hdock server. Wildtype Tat transactivation domain docked to subtype C TAR sequence using Hdock server. Tat subtype C wildtype to TAR subtype C -213,16. **B**. Tat subtype C P21 to TAR RNA subtype C -260,43. Tat models were made using UCSF Chimera and Modeller v5 plug in as a template for docking. Structure visualization was by Biovia Discovery studio software. RNA immunoprecipitation. Stably expressing TatA21 and TatP21 in A72 cell line were used to perform RNA immunoprecipitation. mRNA expression of TAR RNA mRNA expression was significantly higher in TatP21 than TatA21.

Table 2: Docking scores of Tat bound to TAR element using Hdock server

	TatA21 for subtype B – TAR (PDB: 6MCE)	TatA21 for subtype C– TAR (PDB: 6MCE)	TatP21for subtype C – TAR (PDB: 6MCE)
Docking score (Hdock)	-234.46	-241.30	-273.38
Confidence score	0.8441	0.8613	0.9218

The use of docking scores reveals the binding affinity of proteins, in this case Tat, to nucleic acid, which in this case is TAR RNA. A more negative score suggests a higher binding affinity. The molecular docking score of TatP21 to TAR reveals that this variant is able to bind more efficiently to TAR than TatA21 with a confidence score of 0.9218.

4.1.6 RNA immunoprecipitation showed that TatP21 has a higher binding affinity to TAR

To confirm our results from the molecular docking, we wanted to perform molecular docking *in vitro* using RNA immunoprecipitation of TatA21 and TatP21 with TAR RNA.

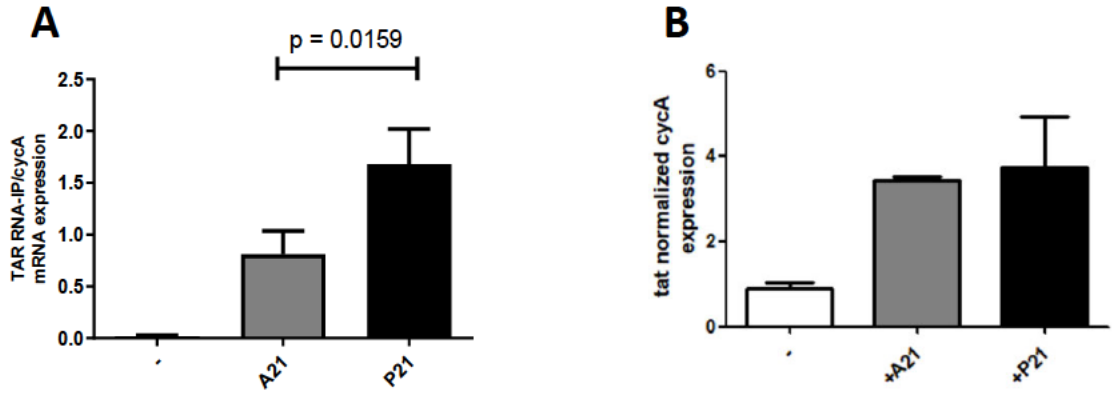

Figure 4.10: RNA immunoprecipitation. **A**. Stably expressing TatA21 and TatP21 in A72 cell line were used to perform RNA immunoprecipitation. **B**. mRNA expression of TAR RNA mRNA expression was significantly higher in TatP21 than TatA21.

Our results showed that although Tat expression was similar in TatA21 and TatP21 (Fig. 4.10B), there was more TAR RNA expressed in TatP21 in mRNA expression experiments (Fig. 4.10A). Thus, suggesting that TatP21 is able to bind more strongly to TAR than TatA21.

4.1.7 Co-immunoprecipitation of TatA21 and TatP21 with cycT1 and CDK9 was the same

It has been well studied that during HIV replication Tat will recruit P-TEFb which will hyperphosphorylate the CTD of the RNAPII and phosphorylate negative transcription factors which will dissociate upon phosphorylation, thus releasing the RNAPII from its paused state and allowing mRNA elongation to proceed (Zhou et al., 1998, Phatnani and Greenleaf, 2006). To determine whether TatA21 can efficiently recruit P-TEFb, co-immunoprecipitation studies were carried out of Tat variants, TatA21 and TatP21, with cycT1 and CDK9, components of the P-TEFb. Interestingly, our data demonstrate that Tat P21A co-immunoprecipitated with CDK9, endogenous cyclin T1 and HA-cyclin T1 at the same level as the consensus as shown by the presence of bands for both the 10% input and the Flag-IP Tat mutant. This suggests that this mutant, although when present results is almost abolished transactivation and significantly lower viral loads in patients, it is still able to recruit and bind P-TEFb during HIV replication.

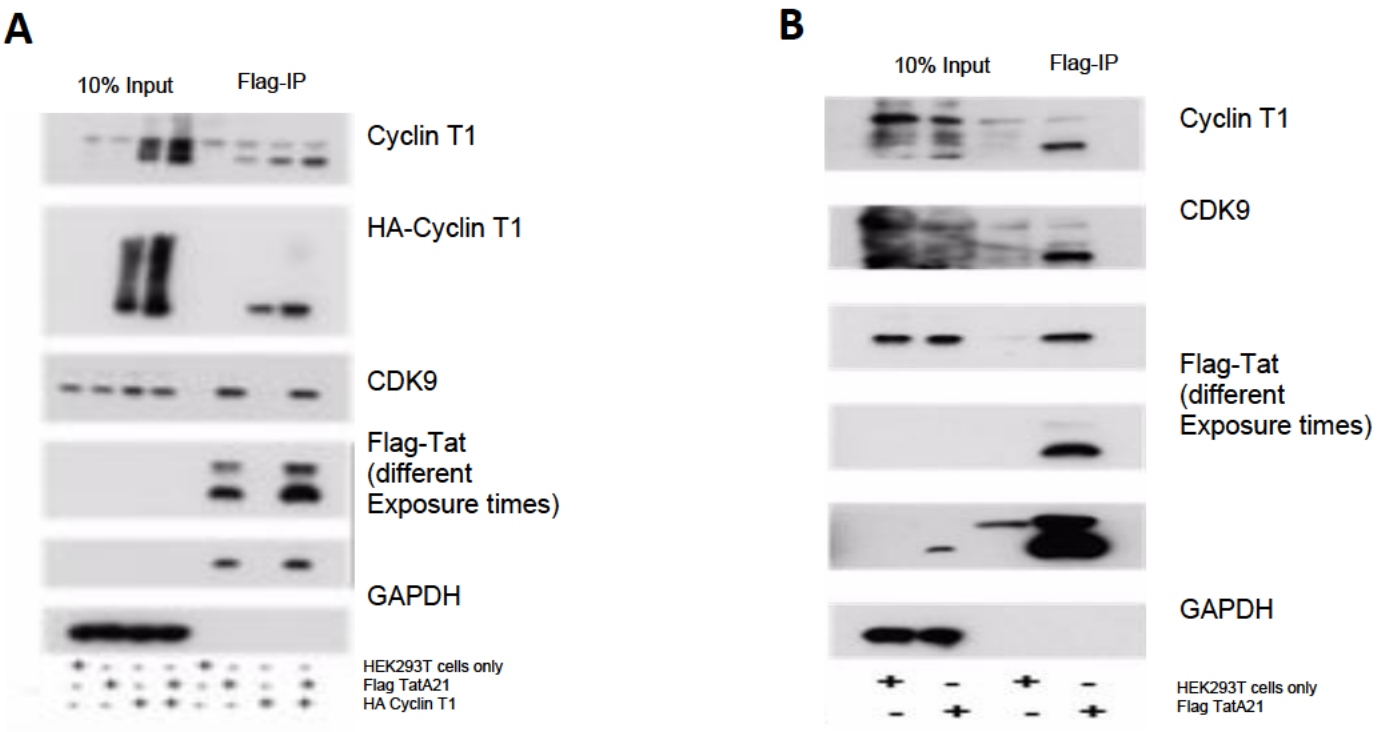

Figure 4.11: Co-immunoprecipitation of TatA21 and TatP21 with cycT1 and CDK9. pcTat is consensus subtype C Tat. A: Co-immunoprecipitation of reference Tat P21 with CycT1 and CDK9. B. Flag-IP with pcTat A21 and HA-Cyclin T1.

4.2. Virus production of the constructs created

In the investigation of TatA21 and TatP21 in the propensity of latency establishment and/or reactivation, we infected a culture of the lymphocytic cell line Jurkat with viral particles containing this vector and used differential fluorescence-activated cell sorting (FACS) based on GFP expression. First, we infected Jurkat cells with the LTR±Tat±IRES±GFP virus at a low m.o.i. and isolated GFP-negative cells by FACS 4 days after infection. This population presumably harboured both uninfected cells and cells with transcriptionally silenced proviruses.

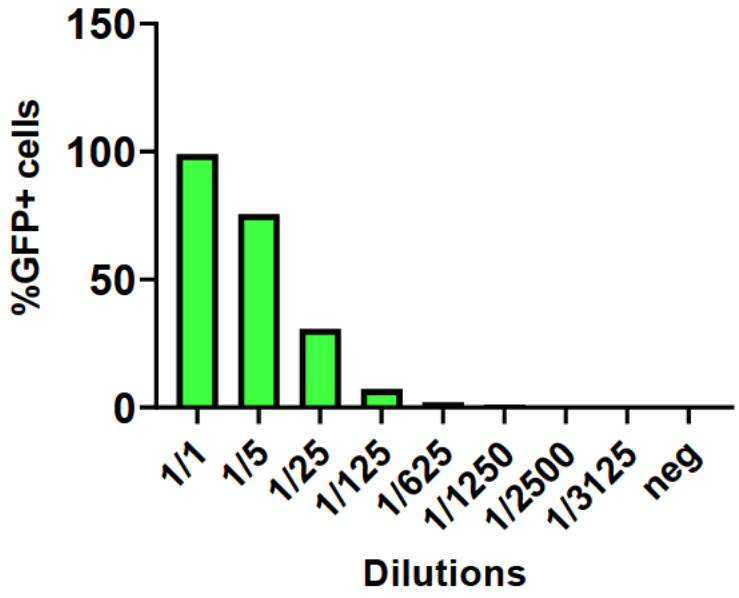

Figure 4.12: Graph showing percentage GFP+ cells at each C731CC virus dilution.

Undiluted (1/1) virus gave almost 100% (99.1%) percentage GFP positive (GFP+) cells. The FACS plots are also shown in the supplementary material (Fig. S1) for unstimulated cells, 1/1 and 1/200. When the virus was diluted 1/5 it gave percentage GFP+ cells of 75.6%, followed by approximately 30% GFP+ cells when diluted 1/25. Percentage GFP+ cells was 7.2% when the virus was diluted 1/125 and 2% when diluted 1/625. Only 1.1% of cells were GFP+ when the virus was diluted 1/1250. No cells were GFP+ when the virus was diluted 1/2500, 1/3150 and for the negative control.

4.2.2 Reactivation of latently infected cells

To activate HIV-1C expression, 95% of GFP negative cells, which represent either truly uninfected or latently infected cells, were treated with PMA.

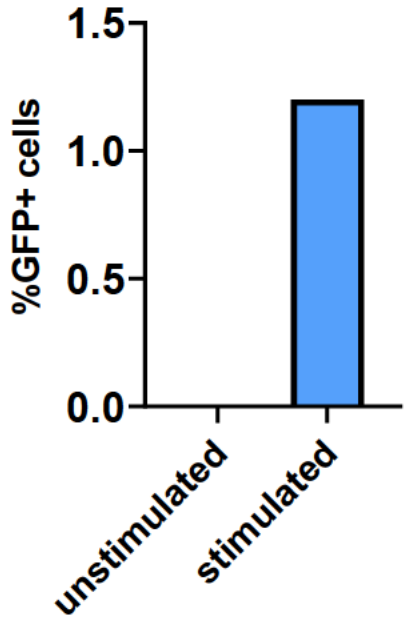

Figure 4.13: Percent GFP positive cells of GFP- cells isolated using cell sorting, unstimulated and stimulated with PMA.

GFP negative cells after reactivation with PMA showed approximately 1.2% GFP positive cells, which suggest that only about 1.2% of the cells were latently infected.

4.3 Introduction of TatP21 into C731CC

Consistent with a previous study (Rossenkhan et al., 2013) the preliminary data from my Masters project (Mkhize et al., manuscript under preparation for publication) show PLWH derived tat variants exhibiting TatA21 were associated with reduced LTR transcription activity and viral load. Specifically, the data from the current study demonstrate TatA21 alone was responsible to significantly reduced LTR activity and binding to the TAR RNA element compared to TatP21. These data could explain low viral load or viral replication associated with TatA21 compared to TatP21 variants in PLWH. However, it remained to be determined

whether TatA21 mediate the propensity of latency development of reversal when compared to TatP21. To study the effect of TatA21 compared to TatP21 on latency establishment and reversal, we had to introduce this TatP21 into the C731CC to create C731C$_{TatP21}$C. To this effect, we cloned TatP21 PCR product amplified from pCTatP21 plasmid in section into the pBlueScript vector and to subclone the TatP21 into the C731CC vector that was deleted of TatA21 fragment. To determine whether the cloning into C731CC was successful we performed a diagnostic restriction digestion of the C731CC to excise the Tat fragment. However, we were unable to failure of cloning was due to high recombination rate of the C731CC viral vector.

Chapter Five

Chapter 5: Discussion and Conclusion

5.1 Discussion

A great challenge to overcoming HIV-1 infection is due to the high genetic diversity in the virus which often leads to failure of clinical trials (Santoro and Perno, 2013). The HIV-1 genome sequence varies from one geographical region to the other, with each region having HIV-1 subtypes, recombinants and mutations which are unique to them (Rodriguez et al., 2009). Moreover, these genetic variations differ in their potential to cause pathogenesis and/or disease outcome (Kiguoya et al., 2017). However, the functional and clinical consequences of the substantial genetic variations of Tat have not been fully elucidated. There are several viral factors which necessitate efficient HIV-1 replication; however, Tat-mediated viral transcription is one of the most desirable targets for therapeutic drugs against HIV-1 replication (Shin et al., 2020).

Subtype-specific Tat variations have been reported (Neogi et al., 2012), which exhibit widely differing viral activities including their ability to activate the HIV-1 LTR promoter (Li et al., 2012). In a study by Rossenkhan et al, they showed that patients showed differential transactivation attributed to inter-patient sequence variation. Furthermore, they showed that they were single point mutations which were responsible for increasing or reducing Tat transactivation of the LTR promoter. They showed that TatA21 resulted in significantly lower transactivation than TatP21 (Rossenkhan et al., 2013). The unpublished data from my Masters project showed that 50% of patients infected with HIV-1C had TatA21 and this variant was associated with significantly lower LTR transcription activity and viral loads in during acute infection in PLWH from cohorts established in Durban, South Africa. However, the mechanism by which TatA21 reduces the transactivation activity of Tat as well as the viral loads of PLWH is unknown. Therefore, in the current study we hypothesized that the TatA21 variant affects its ability to recruit the P-TEFb, which is a cellular protein required for maximum LTR transcription potential and may play a role in the development and/or reversal of viral latency. Our data showed that TatA21 did not affect its ability to recruit P-TEFb. Instead, the mechanism which may be responsible for its reduced transactivation is through TAR binding.

Consistent with previous reports, our data showed that the presence of single point mutations (V4I, A21, L35Q/Q39L, Q39L, R53K) affected the ability of Tat to transactivate the LTR as

shown by the TZM-bl Tat transactivation assay. Particularly, the TatA21 had lower activity than TatP21, mutant Tat. However, the Tat expressed in these two variants was the same suggesting that TatA21 does not affect protein expression via Western blot (Fig. 4.3).

We further characterized the effect of TatA21 vs TatP21 on the expression of GFP using different cell lines to measure the functionality. After 24 hours of incubation after nucleofection, the percentage GFP positive cells was significantly lower ($p = 0.0439$) in the cell lines transfected with TatA21 compared to Tat P21, suggesting that TatA21 has a lower transactivation activity than TatP21. Furthermore, our data show that there was a slightly lower Gag expression in TatA21. Taken together, these data suggest that TatA21 is associated lower viral gene transcription or replication since the Gag protein of HIV-1 are central players in virus particle assembly, release and maturation, as well as function in productive infection establishment (Waheed and Freed, 2012).

Although we saw a reduction in Tat transactivation of TatA21 compared to TatP21, the proteins were expressed at the same level, thus protein expression did not account for the differences in transactivation. During HIV replication, Tat recruits P-TEFb and therefore we hypothesized that the ability of TatA21 to recruit P-TEFb may be affected. However, we found this hypothesis not true as we showed that the ability of TatA21 to recruit P-TEFb was not affected.

The structure of proteins represents an important component of understanding their function. An earlier study described that a mutation in Tat is associated with an effect on one or more roles of the residue concerned (reviewed in (Spector et al., 2019)). The roles which may be affected are protein stability or folding, ligand binding, catalysis, regulation by allosteric and other mechanisms and post-translational modification (Wang and Moult, 2001). Early studies showed that under reducing conditions, which is pre-incubation of Tat with as little as 10 mM DTT and β-mercaptoethanol, Tat loses its transactivation capability due to reduction of cysteine residues (Koken et al., 1994). Reducing conditions inhibit Zn binding of these residues and make the cysteine rich region flexible. Thus, this domain was shown not to be able to perform its function efficiently. In this study we showed through homology modelling of Tat subtype C using Tat subtype B as a template (Tahirov et al., 2010) that TatA21 and TatP21 have the same backbone protein structure. Proline has a bulky protein structure in comparison to

Alanine. Understanding the structure-function relationship in respect to the exogenous roles of Tat may have important clinical implications, both for the development of new vaccines against AIDS targeting Tat (reviewed in (Campbell and Loret, 2009).

The HIV-1 Tat interacts with various host factors that ensure its binding affinity to TAR (reviewed in (Gotora et al., 2023)). Tat protein mutations influence the pathogenesis HIV-1. Tat mutations may also influence Tat-TAR binding and subsequent viral transcription. Of the important functional domains of HIV-1 Tat, mutational analysis has shown that the Arginine-rich basic region (amino acids 47-59) is required for binding to TAR RNA (reviewed in (Romani et al., 2010)). The basic domain of the Tat forms an alpha helix during Tat-TAR binding (Campbell et al., 2004). Modifications in amino acid sequence on the functional groups of Tat proteins have also been shown to affect hydrogen bonding to TAR RNA, which can lower the binding affinity by up to 20-fold (Hamy et al., 1993). The amino acid substitution S46F in the Tat core region could lead to a conformational change to Tat resulting in more hydrogen bond interactions than in the wild-type making it a highly potent transactivator (Ronsard et al., 2017b). In addition, the K51R mutation was shown to make Tat more flexible in this location, giving it a direct hydrogen interaction, which is more non-rigid than in the wild-type (Pantano et al., 2002). Thus, from this knowledge, Tat mutations may affect Tat-TAR interaction and the rate of transcription and ultimately the rate of viral replication (reviewed in (Gotora et al., 2023)). In this study we focused on the major interacting partner TAR to understand how TatA21 results in reduced transactivation.

We performed electrophoretic mobility shift assays (EMSAs) using radioactively labelled TatA21 to investigate its' binding affinity to TAR. In the setting of excess TAR RNA probe, the intensity of Tat–TAR RNA complex correlated with Tat addition in a dose-dependent manner as there was more Tat present. The presence of the complex showed that TatA21 was able to bind TAR, however, we did not have TatP21 for comparison purposes. This experiment could not be repeated due to high cost. Alternatively, we used *in silico* experiments to determine the binding affinity of TatA21 and TatP21 to the TAR RNA element. In silico analyses have been previously used in Tat-TAR interactions (Ronsard et al., 2017b). In the current study, molecular docking revealed that TatP21 has a higher binding affinity to bind TAR RNA then TatA21. Thus, suggesting that TatA21 binds less efficiently to the TAR element than TatP21, resulting in reduced transactivation and subsequently reduced viral loads

in patient with this Tat variant. Interestingly, our data demonstrated that the P-TEFb is still able to be recruited by P-TEFb through co-immunoprecipitation with cycT1 and CDK9 which both comprise the P-TEFb. This pathway thus not affected by genetic variants in Tat studied here during Tat transactivation.

Thus, suggesting that the mechanism by which Tat uses to reduce its transactivation when there is an Alanine at position 21 rather than a Proline is through the Tat-TAR interaction. This study confirms and further highlights the Tat-TAR interaction as a crucial target for latency inducing and achieving the block and lock strategy.

In the second goal of the study, we hypothesized that TatA21 may play a role in the development or reversal of viral latency. Jordan et al 2001 reported that to study HIV-1 transcription in its natural context (i.e., integrated in the host genome), an HIV-1-derived retroviral vector should be used to generate a library of Jurkat clones containing single integration of an HIV-1 mini genome. We were able to successfully construct an HIV-1C genome minimal GFP reporter virus using the C731CC and establish latent T cell lines harbouring minimal HIV-1C (consensus LTR and tat) reporter virus.

Following this, we were able to produce virus of the construct described above. To activate HIV expression, we treated this population with PMA and purified GFP positive cells by FACS. The C731CC cells reactivated to 1.2%. To study the effect of the Tat mutations characterized in this study, TatA21 and TatP21, we attempted to introduce TatP21 into the C731CC to produce virus and characterize its propensity for latency. However, TatP21 did not result in successful virus production as shown by the FACS results. This was mainly due to recombination of the plasmid. We were, however, able to characterize TatA21 as the C731CC contained a TatA21 although we could not compare it to TatP21 in the context of propensity of HIV-1 latency. We hypothesize that TatP21 will show a higher reactivation after stimulation of the latently produced virus. TatP21 did not result in successful virus production as shown by the FACS results. This was mainly due to recombination which occurred at the time of writing this thesis. The lab group members are currently working on optimizing this protocol in the lab to obtain successful virus production.

Future studies should characterize the propensity of TatP21 for comparison to TatA21. As well as explore the production of Tat subtype C crystal structures for protein studies. This will give us more insight into the protein structure and potential interaction with P-TEFb.

5.2 Conclusions

This study showed that Tat transactivation is affected by a single point mutation. Namely TatA21 results in significantly reduced transactivation compared to TatP21. Furthermore, we showed that although TatA21 results in reduced transactivation than TatP21, the protein is expressed at similar levels. We also showed that TatA21 showed reduced transactivation activity and GFP expression in J-LatC cells. In this study we also modelled the Tat subtype C proteins TatA21 and TatP21. We showed that although TatA21 results in reduced transactivation activity, it is still able to efficiently bind P-TEFb. However, the binding affinity of TatA21 to the TAR element is lower than that of TatP21. This study is the first study to explicitly report on the mechanism behind reduced transactivation of TatA21 being caused by its interaction with the HIV TAR RNA using molecular docking. Understanding the mechanism by which Tat uses to reduce its transactivation ability is important to the development of therapies which target Tat. Moreover, the Tat-TAR is an important target in targeting latency by forcing the cells to go into a deep latency. This study is significant as it will inform future therapeutic interventions of combining antiretroviral treatments with latency-purging strategies to accelerate the depletion of latent reservoirs and lead to a cure.

Reference list:

ABNER, E. & JORDAN, A. 2019. HIV "shock and kill" therapy: In need of revision. *Antiviral Res,* 166**,** 19-34.

ABOUL-ELA, F., KARN, J. & VARANI, G. 1995. The structure of the human immunodeficiency virus type-1 TAR RNA reveals principles of RNA recognition by Tat protein. *J Mol Biol,* 253**,** 313-32.

ABOUL-ELA, F., KARN, J. & VARANI, G. 1996. Structure of HIV-1 TAR RNA in the absence of ligands reveals a novel conformation of the trinucleotide bulge. *Nucleic Acids Res,* 24**,** 3974-81.

ABRAHAMS, M. R., JOSEPH, S. B., GARRETT, N., TYERS, L., MOESER, M., ARCHIN, N., COUNCIL, O. D., MATTEN, D., ZHOU, S., DOOLABH, D., ANTHONY, C., GOONETILLEKE, N., KARIM, S. A., MARGOLIS, D. M., POND, S. K., WILLIAMSON, C. & SWANSTROM, R. 2019. The replication-competent HIV-1 latent reservoir is primarily established near the time of therapy initiation. *Sci Transl Med,* 11.

AGOSTO, L. M., GAGNE, M. & HENDERSON, A. J. 2015. Impact of Chromatin on HIV Replication. *Genes (Basel),* 6**,** 957-76.

AHLENSTIEL, C., MENDEZ, C., LIM, S. T., MARKS, K., TURVILLE, S., COOPER, D. A., KELLEHER, A. D. & SUZUKI, K. 2015. Novel RNA Duplex Locks HIV-1 in a Latent State via Chromatin-mediated Transcriptional Silencing. *Mol Ther Nucleic Acids,* 4**,** e261.

AHLENSTIEL, C. L., SYMONDS, G., KENT, S. J. & KELLEHER, A. D. 2020. Block and Lock HIV Cure Strategies to Control the Latent Reservoir. *Front Cell Infect Microbiol,* 10**,** 424.

AJASIN, D. & EUGENIN, E. A. 2020. HIV-1 Tat: Role in Bystander Toxicity. *Front Cell Infect Microbiol,* 10**,** 61.

AMMOSOVA, T., BERRO, R., JEREBTSOVA, M., JACKSON, A., CHARLES, S., KLASE, Z., SOUTHERLAND, W., GORDEUK, V. R., KASHANCHI, F. & NEKHAI, S. 2006. Phosphorylation of HIV-1 Tat by CDK2 in HIV-1 transcription. *Retrovirology,* 3**,** 78.

ANAND, K., SCHULTE, A., VOGEL-BACHMAYR, K., SCHEFFZEK, K. & GEYER, M. 2008. Structural insights into the cyclin T1-Tat-TAR RNA transcription activation complex from EIAV. *Nat Struct Mol Biol,* 15**,** 1287-92.

ANDRE, M., NAIR, M. & RAYMOND, A. D. 2023. HIV Latency and Nanomedicine Strategies for Anti-HIV Treatment and Eradication. *Biomedicines,* 11.

ANTONI, B. A., RABSON, A. B., KINTER, A., BODKIN, M. & POLI, G. 1994. NF-kappa B-dependent and -independent pathways of HIV activation in a chronically infected T cell line. *Virology,* 202**,** 684-94.

ANTONIN, W. & NEUMANN, H. 2016. Chromosome condensation and decondensation during mitosis. *Curr Opin Cell Biol,* 40**,** 15-22.

ARCHIN, N. M., LIBERTY, A. L., KASHUBA, A. D., CHOUDHARY, S. K., KURUC, J. D., CROOKS, A. M., PARKER, D. C., ANDERSON, E. M., KEARNEY, M. F., STRAIN, M. C., RICHMAN, D. D., HUDGENS, M. G., BOSCH, R. J., COFFIN, J. M., ERON, J. J., HAZUDA, D. J. & MARGOLIS, D. M. 2012. Administration of vorinostat disrupts HIV-1 latency in patients on antiretroviral therapy. *Nature,* 487**,** 482-5.

ASAMITSU, K., FUJINAGA, K. & OKAMOTO, T. 2018. HIV Tat/P-TEFb Interaction: A Potential Target for Novel Anti-HIV Therapies. *Molecules,* 23.

ASHE, M. P., GRIFFIN, P., JAMES, W. & PROUDFOOT, N. J. 1995. Poly(A) site selection in the HIV-1 provirus: inhibition of promoter-proximal polyadenylation by the downstream major splice donor site. *Genes Dev,* 9**,** 3008-25.

BACHER, S., MEIER-SOELCH, J., KRACHT, M. & SCHMITZ, M. L. 2021. Regulation of Transcription Factor NF-kappaB in Its Natural Habitat: The Nucleus. *Cells,* 10.

BAILEY, J. R., SEDAGHAT, A. R., KIEFFER, T., BRENNAN, T., LEE, P. K., WIND-ROTOLO, M., HAGGERTY, C. M., KAMIREDDI, A. R., LIU, Y., LEE, J., PERSAUD, D., GALLANT, J. E., COFRANCESCO, J., JR., QUINN, T. C., WILKE, C. O., RAY, S. C., SILICIANO, J. D., NETTLES, R. E. & SILICIANO, R. F. 2006. Residual human immunodeficiency virus type 1

viremia in some patients on antiretroviral therapy is dominated by a small number of invariant clones rarely found in circulating CD4+ T cells. *J Virol,* 80**,** 6441-57.

BARBORIC, M., NISSEN, R. M., KANAZAWA, S., JABRANE-FERRAT, N. & PETERLIN, B. M. 2001. NF-kappaB binds P-TEFb to stimulate transcriptional elongation by RNA polymerase II. *Mol Cell,* 8**,** 327-37.

BARBORIC, M., YIK, J. H., CZUDNOCHOWSKI, N., YANG, Z., CHEN, R., CONTRERAS, X., GEYER, M., MATIJA PETERLIN, B. & ZHOU, Q. 2007. Tat competes with HEXIM1 to increase the active pool of P-TEFb for HIV-1 transcription. *Nucleic Acids Res,* 35**,** 2003-12.

BAROUCH, D. H. & DEEKS, S. G. 2014. Immunologic strategies for HIV-1 remission and eradication. *Science,* 345**,** 169-74.

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

BBOSA, N., KALEEBU, P. & SSEMWANGA, D. 2019. HIV subtype diversity worldwide. *Curr Opin HIV AIDS,* 14**,** 153-160.

BENTLEY, K., DEACON, N., SONZA, S., ZEICHNER, S. & CHURCHILL, M. 2004. Mutational analysis of the HIV-1 LTR as a promoter of negative sense transcription. *Arch Virol,* 149**,** 2277-94.

BISGROVE, D. A., MAHMOUDI, T., HENKLEIN, P. & VERDIN, E. 2007. Conserved P-TEFb-interacting domain of BRD4 inhibits HIV transcription. *Proc Natl Acad Sci U S A,* 104**,** 13690-5.

BLANCO-RODRIGUEZ, G., GAZI, A., MONEL, B., FRABETTI, S., SCOCA, V., MUELLER, F., SCHWARTZ, O., KRIJNSE-LOCKER, J., CHARNEAU, P. & DI NUNZIO, F. 2020. Remodeling of the Core Leads HIV-1 Preintegration Complex into the Nucleus of Human Lymphocytes. *J Virol,* 94.

BLANKSON, J. N., PERSAUD, D. & SILICIANO, R. F. 2002. The challenge of viral reservoirs in HIV-1 infection. *Annu Rev Med,* 53**,** 557-93.

BLAZKOVA, J., GAO, F., MARICHANNEGOWDA, M. H., JUSTEMENT, J. S., SHI, V., WHITEHEAD, E. J., SCHNECK, R. F., HUITING, E. D., GITTENS, K., COTTRELL, M., BENKO, E., KOVACS, C., LACK, J., SNELLER, M. C., MOIR, S., FAUCI, A. S. & CHUN, T. W. 2021. Distinct mechanisms of long-term virologic control in two HIV-infected individuals after treatment interruption of anti-retroviral therapy. *Nat Med,* 27**,** 1893-1898.

BLOND, A., ENNIFAR, E., TISNE, C. & MICOUIN, L. 2014. The design of RNA binders: targeting the HIV replication cycle as a case study. *ChemMedChem,* 9**,** 1982-96.

BOEHM, D. & OTT, M. 2017. Host Methyltransferases and Demethylases: Potential New Epigenetic Targets for HIV Cure Strategies and Beyond. *AIDS Res Hum Retroviruses,* 33**,** S8-S22.

BOSQUE, A., FAMIGLIETTI, M., WEYRICH, A. S., GOULSTON, C. & PLANELLES, V. 2011. Homeostatic proliferation fails to efficiently reactivate HIV-1 latently infected central memory CD4+ T cells. *PLoS Pathog,* 7**,** e1002288.

BOVEN, L. A., NOORBAKHSH, F., BOUMA, G., VAN DER ZEE, R., VARGAS, D. L., PARDO, C., MCARTHUR, J. C., NOTTET, H. S. & POWER, C. 2007. Brain-derived human immunodeficiency virus-1 Tat exerts differential effects on LTR transactivation and neuroimmune activation. *J Neurovirol,* 13**,** 173-84.

BRADY, J. & KASHANCHI, F. 2005. Tat gets the "green" light on transcription initiation. *Retrovirology,* 2**,** 69.

BRENCHLEY, J. M., PRICE, D. A., SCHACKER, T. W., ASHER, T. E., SILVESTRI, G., RAO, S., KAZZAZ, Z., BORNSTEIN, E., LAMBOTTE, O., ALTMANN, D., BLAZAR, B. R., RODRIGUEZ, B., TEIXEIRA-JOHNSON, L., LANDAY, A., MARTIN, J. N., HECHT, F. M., PICKER, L. J., LEDERMAN, M. M., DEEKS, S. G. & DOUEK, D. C. 2006. Microbial translocation is a cause of systemic immune activation in chronic HIV infection. *Nat Med,* 12**,** 1365-71.

BRES, V., TAGAMI, H., PELOPONESE, J. M., LORET, E., JEANG, K. T., NAKATANI, Y., EMILIANI, S., BENKIRANE, M. & KIERNAN, R. E. 2002. Differential acetylation of Tat coordinates its interaction with the co-activators cyclin T1 and PCAF. *EMBO J,* 21**,** 6811-9.

BROOKS, K., JONES, B. R., DILERNIA, D. A., WILKINS, D. J., CLAIBORNE, D. T., MCINALLY, S., GILMOUR, J., KILEMBE, W., JOY, J. B., ALLEN, S. A., BRUMME, Z. L. & HUNTER, E. 2020. HIV-1 variants are archived throughout infection and persist in the reservoir. *PLoS Pathog,* 16**,** e1008378.

BUI, J. K., SOBOLEWSKI, M. D., KEELE, B. F., SPINDLER, J., MUSICK, A., WIEGAND, A., LUKE, B. T., SHAO, W., HUGHES, S. H., COFFIN, J. M., KEARNEY, M. F. & MELLORS, J. W. 2017. Proviruses with identical sequences comprise a large fraction of the replication-competent HIV reservoir. *PLoS Pathog,* 13**,** e1006283.

BUKRINSKY, M. I., STANWICK, T. L., DEMPSEY, M. P. & STEVENSON, M. 1991. Quiescent T lymphocytes as an inducible virus reservoir in HIV-1 infection. *Science,* 254**,** 423-7.

BUONAGURO, L., TORNESELLO, M. L. & BUONAGURO, F. M. 2007. Human immunodeficiency virus type 1 subtype distribution in the worldwide epidemic: pathogenetic and therapeutic implications. *J Virol,* 81**,** 10209-19.

CAMPBELL, G. R. & LORET, E. P. 2009. What does the structure-function relationship of the HIV-1 Tat protein teach us about developing an AIDS vaccine? *Retrovirology,* 6**,** 50.

CAMPBELL, G. R., PASQUIER, E., WATKINS, J., BOURGAREL-REY, V., PEYROT, V., ESQUIEU, D., BARBIER, P., DE MAREUIL, J., BRAGUER, D., KALEEBU, P., YIRRELL, D. L. & LORET, E. P. 2004. The glutamine-rich region of the HIV-1 Tat protein is involved in T-cell apoptosis. *J Biol Chem,* 279**,** 48197-204.

CAMPBELL, G. R. & SPECTOR, S. A. 2022. Current strategies to induce selective killing of HIV-1-infected cells. *J Leukoc Biol,* 112**,** 1273-1284.

CARY, D. C., FUJINAGA, K. & PETERLIN, B. M. 2016. Molecular mechanisms of HIV latency. *J Clin Invest,* 126**,** 448-54.

CHAVEZ, L., CALVANESE, V. & VERDIN, E. 2015. HIV Latency Is Established Directly and Early in Both Resting and Activated Primary CD4 T Cells. *PLoS Pathog,* 11**,** e1004955.

CHEREJI, R. V. & MOROZOV, A. V. 2015. Functional roles of nucleosome stability and dynamics. *Brief Funct Genomics,* 14**,** 50-60.

CHOMONT, N., EL-FAR, M., ANCUTA, P., TRAUTMANN, L., PROCOPIO, F. A., YASSINE-DIAB, B., BOUCHER, G., BOULASSEL, M. R., GHATTAS, G., BRENCHLEY, J. M., SCHACKER, T. W., HILL, B. J., DOUEK, D. C., ROUTY, J. P., HADDAD, E. K. & SEKALY, R. P. 2009. HIV reservoir size and persistence are driven by T cell survival and homeostatic proliferation. *Nat Med,* 15**,** 893-900.

CHRIST, F., VOET, A., MARCHAND, A., NICOLET, S., DESIMMIE, B. A., MARCHAND, D., BARDIOT, D., VAN DER VEKEN, N. J., VAN REMOORTEL, B., STRELKOV, S. V., DE MAEYER, M., CHALTIN, P. & DEBYSER, Z. 2010. Rational design of small-molecule inhibitors of the LEDGF/p75-integrase interaction and HIV replication. *Nat Chem Biol,* 6**,** 442-8.

CHUN, T. W., DAVEY, R. T., JR., ENGEL, D., LANE, H. C. & FAUCI, A. S. 1999. Re-emergence of HIV after stopping therapy. *Nature,* 401**,** 874-5.

CHUN, T. W., ENGEL, D., BERREY, M. M., SHEA, T., COREY, L. & FAUCI, A. S. 1998a. Early establishment of a pool of latently infected, resting CD4(+) T cells during primary HIV-1 infection. *Proc Natl Acad Sci U S A,* 95**,** 8869-73.

CHUN, T. W., ENGEL, D., MIZELL, S. B., EHLER, L. A. & FAUCI, A. S. 1998b. Induction of HIV-1 replication in latently infected CD4+ T cells using a combination of cytokines. *J Exp Med,* 188**,** 83-91.

CHUN, T. W., STUYVER, L., MIZELL, S. B., EHLER, L. A., MICAN, J. A., BASELER, M., LLOYD, A. L., NOWAK, M. A. & FAUCI, A. S. 1997. Presence of an inducible HIV-1 latent reservoir during highly active antiretroviral therapy. *Proc Natl Acad Sci U S A,* 94**,** 13193-7.

CLARK, E., NAVA, B. & CAPUTI, M. 2017. Tat is a multifunctional viral protein that modulates cellular gene expression and functions. *Oncotarget*.

COCKERHAM, L. R. & HATANO, H. 2015. Elite control of HIV: is this the right model for a functional cure? *Trends Microbiol,* 23**,** 71-5.

COFFIN, J. M. & HUGHES, S. H. 2021. Clonal Expansion of Infected CD4+ T Cells in People Living with HIV. *Viruses,* 13.

COFFIN, J. M., HUGHES, S. H. & VARMUS, H. E. 1997. The Interactions of Retroviruses and their Hosts. *In:* COFFIN, J. M., HUGHES, S. H. & VARMUS, H. E. (eds.) *Retroviruses*. Cold Spring Harbor (NY).

COHN, L. B., CHOMONT, N. & DEEKS, S. G. 2020. The Biology of the HIV-1 Latent Reservoir and Implications for Cure Strategies. *Cell Host Microbe,* 27**,** 519-530.

COLE, B., LAMBRECHTS, L., GANTNER, P., NOPPE, Y., BONINE, N., WITKOWSKI, W., CHEN, L., PALMER, S., MULLINS, J. I., CHOMONT, N., PARDONS, M. & VANDEKERCKHOVE, L. 2021. In-depth single-cell analysis of translation-competent HIV-1 reservoirs identifies cellular sources of plasma viremia. *Nat Commun,* 12**,** 3727.

COLIN, L. & VAN LINT, C. 2009. Molecular control of HIV-1 postintegration latency: implications for the development of new therapeutic strategies. *Retrovirology,* 6**,** 111.

CORY, T. J., SCHACKER, T. W., STEVENSON, M. & FLETCHER, C. V. 2013. Overcoming pharmacologic sanctuaries. *Curr Opin HIV AIDS,* 8**,** 190-5.

COUTURIER, J., OROZCO, A. F., LIU, H., BUDHIRAJA, S., SIWAK, E. B., NEHETE, P. N., SASTRY, K. J., RICE, A. P. & LEWIS, D. E. 2019. Regulation of cyclin T1 during HIV replication and latency establishment in human memory CD4 T cells. *Virol J,* 16**,** 22.

CRAIGIE, R. & BUSHMAN, F. D. 2012. HIV DNA integration. *Cold Spring Harb Perspect Med,* 2**,** a006890.

CROOKS, A. M., BATESON, R., COPE, A. B., DAHL, N. P., GRIGGS, M. K., KURUC, J. D., GAY, C. L., ERON, J. J., MARGOLIS, D. M., BOSCH, R. J. & ARCHIN, N. M. 2015. Precise Quantitation of the Latent HIV-1 Reservoir: Implications for Eradication Strategies. *J Infect Dis,* 212**,** 1361-5.

CULLEN, B. R., LOMEDICO, P. T. & JU, G. 1984. Transcriptional interference in avian retroviruses--implications for the promoter insertion model of leukaemogenesis. *Nature,* 307**,** 241-5.

DARCIS, G., KULA, A., BOUCHAT, S., FUJINAGA, K., CORAZZA, F., AIT-AMMAR, A., DELACOURT, N., MELARD, A., KABEYA, K., VANHULLE, C., VAN DRIESSCHE, B., GATOT, J. S., CHERRIER, T., PIANOWSKI, L. F., GAMA, L., SCHWARTZ, C., VILA, J., BURNY, A., CLUMECK, N., MOUTSCHEN, M., DE WIT, S., PETERLIN, B. M., ROUZIOUX, C., ROHR, O. & VAN LINT, C. 2015. An In-Depth Comparison of Latency-Reversing Agent Combinations in Various In Vitro and Ex Vivo HIV-1 Latency Models Identified Bryostatin-1+JQ1 and Ingenol-B+JQ1 to Potently Reactivate Viral Gene Expression. *PLoS Pathog,* 11**,** e1005063.

DAS, A. T., HARWIG, A. & BERKHOUT, B. 2011. The HIV-1 Tat protein has a versatile role in activating viral transcription. *J Virol,* 85**,** 9506-16.

DAVEY, R. T., JR., BHAT, N., YODER, C., CHUN, T. W., METCALF, J. A., DEWAR, R., NATARAJAN, V., LEMPICKI, R. A., ADELSBERGER, J. W., MILLER, K. D., KOVACS, J. A., POLIS, M. A., WALKER, R. E., FALLOON, J., MASUR, H., GEE, D., BASELER, M., DIMITROV, D. S., FAUCI, A. S. & LANE, H. C. 1999. HIV-1 and T cell dynamics after interruption of highly active antiretroviral therapy (HAART) in patients with a history of sustained viral suppression. *Proc Natl Acad Sci U S A,* 96**,** 15109-14.

DEBAISIEUX, S., RAYNE, F., YEZID, H. & BEAUMELLE, B. 2012. The ins and outs of HIV-1 Tat. *Traffic,* 13**,** 355-63.

DEBYSER, Z., VANSANT, G., BRUGGEMANS, A., JANSSENS, J. & CHRIST, F. 2018. Insight in HIV Integration Site Selection Provides a Block-and-Lock Strategy for a Functional Cure of HIV Infection. *Viruses,* 11.

DEEKS, S. G. 2012. HIV: Shock and kill. *Nature,* 487**,** 439-40.

DEEKS, S. G., LEWIN, S. R. & HAVLIR, D. V. 2013. The end of AIDS: HIV infection as a chronic disease. *Lancet,* 382**,** 1525-33.

DERDEYN, C. A., DECKER, J. M., SFAKIANOS, J. N., WU, X., O'BRIEN, W. A., RATNER, L., KAPPES, J. C., SHAW, G. M. & HUNTER, E. 2000. Sensitivity of human immunodeficiency virus type 1 to the fusion inhibitor T-20 is modulated by coreceptor specificity defined by the V3 loop of gp120. *J Virol,* 74**,** 8358-67.

DEWITT, M. A., CORN, J. E. & CARROLL, D. 2017. Genome editing via delivery of Cas9 ribonucleoprotein. *Methods,* 121-122**,** 9-15.

DINGWALL, C., ERNBERG, I., GAIT, M. J., GREEN, S. M., HEAPHY, S., KARN, J., LOWE, A. D., SINGH, M., SKINNER, M. A. & VALERIO, R. 1989. Human immunodeficiency virus 1 tat protein binds trans-activation-responsive region (TAR) RNA in vitro. *Proc Natl Acad Sci U S A,* 86**,** 6925-9.

DONAHUE, D. A., BASTARACHE, S. M., SLOAN, R. D. & WAINBERG, M. A. 2013. Latent HIV-1 can be reactivated by cellular superinfection in a Tat-dependent manner, which can lead to the emergence of multidrug-resistant recombinant viruses. *J Virol,* 87**,** 9620-32.

DONAHUE, D. A., KUHL, B. D., SLOAN, R. D. & WAINBERG, M. A. 2012. The viral protein Tat can inhibit the establishment of HIV-1 latency. *J Virol,* 86**,** 3253-63.

DUTILLEUL, A., RODARI, A. & VAN LINT, C. 2020. Depicting HIV-1 Transcriptional Mechanisms: A Summary of What We Know. *Viruses,* 12.

EBINA, H., MISAWA, N., KANEMURA, Y. & KOYANAGI, Y. 2013. Harnessing the CRISPR/Cas9 system to disrupt latent HIV-1 provirus. *Sci Rep,* 3**,** 2510.

EISELE, E. & SILICIANO, R. F. 2012. Redefining the viral reservoirs that prevent HIV-1 eradication. *Immunity,* 37**,** 377-88.

ELANGOVAN, R., JENKS, M., YUN, J., DICKSON-TETTEH, L., KIRTLEY, S., HEMELAAR, J., ISOLATION, W.-U. N. F. H. & CHARACTERISATION 2021. Global and Regional Estimates for Subtype-Specific Therapeutic and Prophylactic HIV-1 Vaccines: A Modeling Study. *Front Microbiol,* 12**,** 690647.

ELLIOTT, J. H., WIGHTMAN, F., SOLOMON, A., GHNEIM, K., AHLERS, J., CAMERON, M. J., SMITH, M. Z., SPELMAN, T., MCMAHON, J., VELAYUDHAM, P., BROWN, G., RONEY, J., WATSON, J., PRINCE, M. H., HOY, J. F., CHOMONT, N., FROMENTIN, R., PROCOPIO, F. A., ZEIDAN, J., PALMER, S., ODEVALL, L., JOHNSTONE, R. W., MARTIN, B. P., SINCLAIR, E., DEEKS, S. G., HAZUDA, D. J., CAMERON, P. U., SEKALY, R. P. & LEWIN, S. R. 2014. Activation of HIV transcription with short-course vorinostat in HIV-infected patients on suppressive antiretroviral therapy. *PLoS Pathog,* 10**,** e1004473.

EMILIANI, S., FISCHLE, W., OTT, M., VAN LINT, C., AMELLA, C. A. & VERDIN, E. 1998. Mutations in the tat gene are responsible for human immunodeficiency virus type 1 postintegration latency in the U1 cell line. *J Virol,* 72**,** 1666-70.

EMILIANI, S., VAN LINT, C., FISCHLE, W., PARAS, P., JR., OTT, M., BRADY, J. & VERDIN, E. 1996. A point mutation in the HIV-1 Tat responsive element is associated with postintegration latency. *Proc Natl Acad Sci U S A,* 93**,** 6377-81.

ENSOLI, B., BUONAGURO, L., BARILLARI, G., FIORELLI, V., GENDELMAN, R., MORGAN, R. A., WINGFIELD, P. & GALLO, R. C. 1993. Release, uptake, and effects of extracellular human immunodeficiency virus type 1 Tat protein on cell growth and viral transactivation. *J Virol,* 67**,** 277-87.

ENSOLI, B., MORETTI, S., BORSETTI, A., MAGGIORELLA, M. T., BUTTO, S., PICCONI, O., TRIPICIANO, A., SGADARI, C., MONINI, P. & CAFARO, A. 2021. New insights into pathogenesis point to HIV-1 Tat as a key vaccine target. *Arch Virol,* 166**,** 2955-2974.

FARNET, C. M. & HASELTINE, W. A. 1991. Circularization of human immunodeficiency virus type 1 DNA in vitro. *J Virol,* 65**,** 6942-52.

FASKEN, M. B. & CORBETT, A. H. 2009. Mechanisms of nuclear mRNA quality control. *RNA Biol,* 6**,** 237-41.

FELSENFELD, G. & GROUDINE, M. 2003. Controlling the double helix. *Nature,* 421**,** 448-53.

FIDLER, S., STOHR, W., PACE, M., DORRELL, L., LEVER, A., PETT, S., KINLOCH-DE LOES, S., FOX, J., CLARKE, A., NELSON, M., THORNHILL, J., KHAN, M., FUN, A., BANDARA, M., KELLY, D., KOPYCINSKI, J., HANKE, T., YANG, H., BENNETT, R., JOHNSON, M., HOWELL, B., BARNARD, R., WU, G., KAYE, S., WILLS, M., BABIKER, A., FRATER, J. & GROUP, R. T. S. 2020. Antiretroviral therapy alone versus antiretroviral therapy with a kick and kill approach, on measures of the HIV reservoir in participants with recent HIV infection (the RIVER trial): a phase 2, randomised trial. *Lancet,* 395**,** 888-898.

FINZI, D., BLANKSON, J., SILICIANO, J. D., MARGOLICK, J. B., CHADWICK, K., PIERSON, T., SMITH, K., LISZIEWICZ, J., LORI, F., FLEXNER, C., QUINN, T. C., CHAISSON, R. E., ROSENBERG, E., WALKER, B., GANGE, S., GALLANT, J. & SILICIANO, R. F. 1999.

Latent infection of CD4+ T cells provides a mechanism for lifelong persistence of HIV-1, even in patients on effective combination therapy. *Nat Med,* 5**,** 512-7.

FINZI, D., HERMANKOVA, M., PIERSON, T., CARRUTH, L. M., BUCK, C., CHAISSON, R. E., QUINN, T. C., CHADWICK, K., MARGOLICK, J., BROOKMEYER, R., GALLANT, J., MARKOWITZ, M., HO, D. D., RICHMAN, D. D. & SILICIANO, R. F. 1997. Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy. *Science,* 278**,** 1295-300.

FOUCAULT, M., MAYOL, K., RECEVEUR-BRECHOT, V., BUSSAT, M. C., KLINGUER-HAMOUR, C., VERRIER, B., BECK, A., HASER, R., GOUET, P. & GUILLON, C. 2010. UV and X-ray structural studies of a 101-residue long Tat protein from a HIV-1 primary isolate and of its mutated, detoxified, vaccine candidate. *Proteins,* 78**,** 1441-56.

FUJINAGA, K., IRWIN, D., HUANG, Y., TAUBE, R., KUROSU, T. & PETERLIN, B. M. 2004. Dynamics of human immunodeficiency virus transcription: P-TEFb phosphorylates RD and dissociates negative effectors from the transactivation response element. *Mol Cell Biol,* 24**,** 787-95.

GAUTIER, V. W., GU, L., O'DONOGHUE, N., PENNINGTON, S., SHEEHY, N. & HALL, W. W. 2009. In vitro nuclear interactome of the HIV-1 Tat protein. *Retrovirology,* 6**,** 47.

GAYNOR, R. 1992. Cellular transcription factors involved in the regulation of HIV-1 gene expression. *AIDS,* 6**,** 347-63.

GOTORA, P. T., VAN DER SLUIS, R. & WILLIAMS, M. E. 2023. HIV-1 Tat amino acid residues that influence Tat-TAR binding affinity: a scoping review. *BMC Infect Dis,* 23**,** 164.

GRAMATICA, A., SCHWARZER, R., BRANTLEY, W., VARCO-MERTH, B., SPERBER, H. S., HULL, P. A., MONTANO, M., MIGUELES, S. A., ROSENTHAL, D., HOGAN, L. E., JOHNSON, J. R., PACKARD, T. A., GRIMMETT, Z. W., HERZIG, E., BESNARD, E., NEKORCHUK, M., HSIAO, F., DEEKS, S. G., SNAPE, M., KIERNAN, B., ROAN, N. R., LIFSON, J. D., ESTES, J. D., PICKER, L. J., VERDIN, E., KROGAN, N. J., HENRICH, T. J., CONNORS, M., OTT, M., PILLAI, S. K., OKOYE, A. A. & GREENE, W. C. 2021. Evaluating a New Class of AKT/mTOR Activators for HIV Latency Reversing Activity Ex Vivo and In Vivo. *J Virol,* 95.

GRAU-EXPOSITO, J., LUQUE-BALLESTEROS, L., NAVARRO, J., CURRAN, A., BURGOS, J., RIBERA, E., TORRELLA, A., PLANAS, B., BADIA, R., MARTIN-CASTILLO, M., FERNANDEZ-SOJO, J., GENESCA, M., FALCO, V. & BUZON, M. J. 2019. Latency reversal agents affect differently the latent reservoir present in distinct CD4+ T subpopulations. *PLoS Pathog,* 15**,** e1007991.

GREGER, I. H., DEMARCHI, F., GIACCA, M. & PROUDFOOT, N. J. 1998. Transcriptional interference perturbs the binding of Sp1 to the HIV-1 promoter. *Nucleic Acids Res,* 26**,** 1294-301.

HALVAS, E. K., JOSEPH, K. W., BRANDT, L. D., GUO, S., SOBOLEWSKI, M. D., JACOBS, J. L., TUMIOTTO, C., BUI, J. K., CYKTOR, J. C., KEELE, B. F., MORSE, G. D., BALE, M. J., SHAO, W., KEARNEY, M. F., COFFIN, J. M., RAUSCH, J. W., WU, X., HUGHES, S. H. & MELLORS, J. W. 2020. HIV-1 viremia not suppressible by antiretroviral therapy can originate from large T cell clones producing infectious virus. *J Clin Invest,* 130**,** 5847-5857.

HAMY, F., ASSELINE, U., GRASBY, J., IWAI, S., PRITCHARD, C., SLIM, G., BUTLER, P. J., KARN, J. & GAIT, M. J. 1993. Hydrogen-bonding contacts in the major groove are required for human immunodeficiency virus type-1 tat protein recognition of TAR RNA. *J Mol Biol,* 230**,** 111-23.

HAMY, F., FELDER, E. R., HEIZMANN, G., LAZDINS, J., ABOUL-ELA, F., VARANI, G., KARN, J. & KLIMKAIT, T. 1997. An inhibitor of the Tat/TAR RNA interaction that effectively suppresses HIV-1 replication. *Proc Natl Acad Sci U S A,* 94**,** 3548-53.

HAN, Y., LIN, Y. B., AN, W., XU, J., YANG, H. C., O'CONNELL, K., DORDAI, D., BOEKE, J. D., SILICIANO, J. D. & SILICIANO, R. F. 2008. Orientation-dependent regulation of integrated HIV-1 expression by host gene transcriptional readthrough. *Cell Host Microbe,* 4**,** 134-46.

HARLEN, K. M. & CHURCHMAN, L. S. 2017. The code and beyond: transcription regulation by the RNA polymerase II carboxy-terminal domain. *Nat Rev Mol Cell Biol,* 18**,** 263-273.

HAUBER, J., MALIM, M. H. & CULLEN, B. R. 1989. Mutational analysis of the conserved basic domain of human immunodeficiency virus tat protein. *J Virol,* 63**,** 1181-7.
HAVLIR, D. V., BASSETT, R., LEVITAN, D., GILBERT, P., TEBAS, P., COLLIER, A. C., HIRSCH, M. S., IGNACIO, C., CONDRA, J., GUNTHARD, H. F., RICHMAN, D. D. & WONG, J. K. 2001. Prevalence and predictive value of intermittent viremia with combination hiv therapy. *JAMA,* 286**,** 171-9.
HE, G., YLISASTIGUI, L. & MARGOLIS, D. M. 2002. The regulation of HIV-1 gene expression: the emerging role of chromatin. *DNA Cell Biol,* 21**,** 697-705.
HE, N., LIU, M., HSU, J., XUE, Y., CHOU, S., BURLINGAME, A., KROGAN, N. J., ALBER, T. & ZHOU, Q. 2010. HIV-1 Tat and host AFF4 recruit two transcription elongation factors into a bifunctional complex for coordinated activation of HIV-1 transcription. *Mol Cell,* 38**,** 428-38.
HEMELAAR, J., ELANGOVAN, R., YUN, J., DICKSON-TETTEH, L., FLEMINGER, I., KIRTLEY, S., WILLIAMS, B., GOUWS-WILLIAMS, E., GHYS, P. D. & CHARACTERISATION, W.-U. N. F. H. I. 2019. Global and regional molecular epidemiology of HIV-1, 1990-2015: a systematic review, global survey, and trend analysis. *Lancet Infect Dis,* 19**,** 143-155.
HENDERSON, L. J., REOMA, L. B., KOVACS, J. A. & NATH, A. 2020. Advances toward Curing HIV-1 Infection in Tissue Reservoirs. *J Virol,* 94.
HERMANKOVA, M., SILICIANO, J. D., ZHOU, Y., MONIE, D., CHADWICK, K., MARGOLICK, J. B., QUINN, T. C. & SILICIANO, R. F. 2003. Analysis of human immunodeficiency virus type 1 gene expression in latently infected resting CD4+ T lymphocytes in vivo. *J Virol,* 77**,** 7383-92.
HOSMANE, N. N., KWON, K. J., BRUNER, K. M., CAPOFERRI, A. A., BEG, S., ROSENBLOOM, D. I., KEELE, B. F., HO, Y. C., SILICIANO, J. D. & SILICIANO, R. F. 2017. Proliferation of latently infected CD4(+) T cells carrying replication-competent HIV-1: Potential role in latent reservoir dynamics. *J Exp Med,* 214**,** 959-972.
HOWCROFT, T. K., STREBEL, K., MARTIN, M. A. & SINGER, D. S. 1993. Repression of MHC class I gene promoter activity by two-exon Tat of HIV. *Science,* 260**,** 1320-2.
HU, W., KAMINSKI, R., YANG, F., ZHANG, Y., COSENTINO, L., LI, F., LUO, B., ALVAREZ-CARBONELL, D., GARCIA-MESA, Y., KARN, J., MO, X. & KHALILI, K. 2014. RNA-directed gene editing specifically eradicates latent and prevents new HIV-1 infection. *Proc Natl Acad Sci U S A,* 111**,** 11461-6.
ISEL, C. & KARN, J. 1999. Direct evidence that HIV-1 Tat stimulates RNA polymerase II carboxyl-terminal domain hyperphosphorylation during transcriptional elongation. *J Mol Biol,* 290**,** 929-41.
JANG, M. K., MOCHIZUKI, K., ZHOU, M., JEONG, H. S., BRADY, J. N. & OZATO, K. 2005. The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription. *Mol Cell,* 19**,** 523-34.
JANIN, J. 1997. The kinetics of protein-protein recognition. *Proteins,* 28**,** 153-61.
JEAN, M. J., POWER, D., KONG, W., HUANG, H., SANTOSO, N. & ZHU, J. 2017. Identification of HIV-1 Tat-Associated Proteins Contributing to HIV-1 Transcription and Latency. *Viruses,* 9.
JEANG, K. T., CHUN, R., LIN, N. H., GATIGNOL, A., GLABE, C. G. & FAN, H. 1993. In vitro and in vivo binding of human immunodeficiency virus type 1 Tat protein and Sp1 transcription factor. *J Virol,* 67**,** 6224-33.
JEANSON, L., SUBRA, F., VAGANAY, S., HERVY, M., MARANGONI, E., BOURHIS, J. & MOUSCADET, J. F. 2002. Effect of Ku80 depletion on the preintegrative steps of HIV-1 replication in human cells. *Virology,* 300**,** 100-8.
JEFFERYS, S. R., BURGOS, S. D., PETERSON, J. J., SELITSKY, S. R., TURNER, A. W., JAMES, L. I., TSAI, Y. H., COFFEY, A. R., MARGOLIS, D. M., PARKER, J. & BROWNE, E. P. 2021. Epigenomic characterization of latent HIV infection identifies latency regulating transcription factors. *PLoS Pathog,* 17**,** e1009346.
JIANG, C., LIAN, X., GAO, C., SUN, X., EINKAUF, K. B., CHEVALIER, J. M., CHEN, S. M. Y., HUA, S., RHEE, B., CHANG, K., BLACKMER, J. E., OSBORN, M., PELUSO, M. J., HOH, R., SOMSOUK, M., MILUSH, J., BERTAGNOLLI, L. N., SWEET, S. E., VARRIALE, J. A., BURBELO, P. D., CHUN, T. W., LAIRD, G. M., SERRAO, E., ENGELMAN, A. N., CARRINGTON, M., SILICIANO, R. F., SILICIANO, J. M., DEEKS, S. G., WALKER, B. D.,

LICHTERFELD, M. & YU, X. G. 2020. Distinct viral reservoirs in individuals with spontaneous control of HIV-1. *Nature,* 585**,** 261-267.

JIANG, G. & DANDEKAR, S. 2015. Targeting NF-kappaB signaling with protein kinase C agonists as an emerging strategy for combating HIV latency. *AIDS Res Hum Retroviruses,* 31**,** 4-12.

JIANG, G., MENDES, E. A., KAISER, P., WONG, D. P., TANG, Y., CAI, I., FENTON, A., MELCHER, G. P., HILDRETH, J. E., THOMPSON, G. R., WONG, J. K. & DANDEKAR, S. 2015. Synergistic Reactivation of Latent HIV Expression by Ingenol-3-Angelate, PEP005, Targeted NF-kB Signaling in Combination with JQ1 Induced p-TEFb Activation. *PLoS Pathog,* 11**,** e1005066.

JOHRI, M. K., MISHRA, R., CHHATBAR, C., UNNI, S. K. & SINGH, S. K. 2011. Tits and bits of HIV Tat protein. *Expert Opin Biol Ther,* 11**,** 269-83.

JONES, K. A. 1993. Tat and the HIV-1 promoter. *Curr Opin Cell Biol,* 5**,** 461-8.

JONES, R. B., O'CONNOR, R., MUELLER, S., FOLEY, M., SZETO, G. L., KAREL, D., LICHTERFELD, M., KOVACS, C., OSTROWSKI, M. A., TROCHA, A., IRVINE, D. J. & WALKER, B. D. 2014. Histone deacetylase inhibitors impair the elimination of HIV-infected cells by cytotoxic T-lymphocytes. *PLoS Pathog,* 10**,** e1004287.

JORDAN, A., BISGROVE, D. & VERDIN, E. 2003. HIV reproducibly establishes a latent infection after acute infection of T cells in vitro. *EMBO J,* 22**,** 1868-77.

JORDAN, A., DEFECHEREUX, P. & VERDIN, E. 2001. The site of HIV-1 integration in the human genome determines basal transcriptional activity and response to Tat transactivation. *EMBO J,* 20**,** 1726-38.

KACZMAREK, K., MORALES, A. & HENDERSON, A. J. 2013. T Cell Transcription Factors and Their Impact on HIV Expression. *Virology (Auckl),* 2013**,** 41-47.

KAEHLCKE, K., DORR, A., HETZER-EGGER, C., KIERMER, V., HENKLEIN, P., SCHNOELZER, M., LORET, E., COLE, P. A., VERDIN, E. & OTT, M. 2003. Acetylation of Tat defines a cyclinT1-independent step in HIV transactivation. *Mol Cell,* 12**,** 167-76.

KAMORI, D. & UENO, T. 2017. HIV-1 Tat and Viral Latency: What We Can Learn from Naturally Occurring Sequence Variations. *Front Microbiol,* 8**,** 80.

KANDATHIL, A. J., KANNANGAI, R., ABRAHAM, O. C., PULIMOOD, S. A. & SRIDHARAN, G. 2009. Amino acid sequence divergence of Tat protein (exon1)of subtype B and C HIV-1 strains: Does it have implications for vaccine development? *Bioinformation,* 4**,** 237-41.

KARN, J. 2011. The molecular biology of HIV latency: breaking and restoring the Tat-dependent transcriptional circuit. *Curr Opin HIV AIDS,* 6**,** 4-11.

KARN, J. & STOLTZFUS, C. M. 2012. Transcriptional and posttranscriptional regulation of HIV-1 gene expression. *Cold Spring Harb Perspect Med,* 2**,** a006916.

KELLY, J., BEDDALL, M. H., YU, D., IYER, S. R., MARSH, J. W. & WU, Y. 2008. Human macrophages support persistent transcription from unintegrated HIV-1 DNA. *Virology,* 372**,** 300-12.

KESSING, C. F., NIXON, C. C., LI, C., TSAI, P., TAKATA, H., MOUSSEAU, G., HO, P. T., HONEYCUTT, J. B., FALLAHI, M., TRAUTMANN, L., GARCIA, J. V. & VALENTE, S. T. 2017. In Vivo Suppression of HIV Rebound by Didehydro-Cortistatin A, a "Block-and-Lock" Strategy for HIV-1 Treatment. *Cell Rep,* 21**,** 600-611.

KHAN, N., LAKPA, K. L., HALCROW, P. W., AFGHAH, Z., MILLER, N. M., GEIGER, J. D. & CHEN, X. 2019. BK channels regulate extracellular Tat-mediated HIV-1 LTR transactivation. *Sci Rep,* 9**,** 12285.

KHOURY, G., DARCIS, G., LEE, M. Y., BOUCHAT, S., VAN DRIESSCHE, B., PURCELL, D. F. J. & VAN LINT, C. 2018. The Molecular Biology of HIV Latency. *Adv Exp Med Biol,* 1075**,** 187-212.

KIGUOYA, M. W., MANN, J. K., CHOPERA, D., GOUNDER, K., LEE, G. Q., HUNT, P. W., MARTIN, J. N., BALL, T. B., KIMANI, J., BRUMME, Z. L., BROCKMAN, M. A. & NDUNG'U, T. 2017. Subtype-Specific Differences in Gag-Protease-Driven Replication Capacity Are Consistent with Intersubtype Differences in HIV-1 Disease Progression. *J Virol,* 91.

KILZER, J. M., STRACKER, T., BEITZEL, B., MEEK, K., WEITZMAN, M. & BUSHMAN, F. D. 2003. Roles of host cell factors in circularization of retroviral dna. *Virology,* 314**,** 460-7.

KIM, S. Y., BYRN, R., GROOPMAN, J. & BALTIMORE, D. 1989. Temporal aspects of DNA and RNA synthesis during human immunodeficiency virus infection: evidence for differential gene expression. *J Virol,* 63**,** 3708-13.

KIM, Y., ANDERSON, J. L. & LEWIN, S. R. 2018. Getting the "Kill" into "Shock and Kill": Strategies to Eliminate Latent HIV. *Cell Host Microbe,* 23**,** 14-26.

KMIETOWICZ, Z. 2013. Early HIV treatment led to "functional cure" in 14 patients, report researchers. *BMJ,* 346**,** f1759.

KOBAYASHI-ISHIHARA, M., YAMAGISHI, M., HARA, T., MATSUDA, Y., TAKAHASHI, R., MIYAKE, A., NAKANO, K., YAMOCHI, T., ISHIDA, T. & WATANABE, T. 2012. HIV-1-encoded antisense RNA suppresses viral replication for a prolonged period. *Retrovirology,* 9**,** 38.

KOKEN, S. E., GREIJER, A. E., VERHOEF, K., VAN WAMEL, J., BUKRINSKAYA, A. G. & BERKHOUT, B. 1994. Intracellular analysis of in vitro modified HIV Tat protein. *J Biol Chem,* 269**,** 8366-75.

KUCIAK, M., GABUS, C., IVANYI-NAGY, R., SEMRAD, K., STORCHAK, R., CHALOIN, O., MULLER, S., MELY, Y. & DARLIX, J. L. 2008. The HIV-1 transcriptional activator Tat has potent nucleic acid chaperoning activities in vitro. *Nucleic Acids Res,* 36**,** 3389-400.

KURNAEVA, M. A., SHEVAL, E. V., MUSINOVA, Y. R. & VASSETZKY, Y. S. 2019. Tat basic domain: A "Swiss army knife" of HIV-1 Tat? *Rev Med Virol,* 29**,** e2031.

LANDRY, S., HALIN, M., LEFORT, S., AUDET, B., VAQUERO, C., MESNARD, J. M. & BARBEAU, B. 2007. Detection, characterization and regulation of antisense transcripts in HIV-1. *Retrovirology,* 4**,** 71.

LASSEN, K. G., RAMYAR, K. X., BAILEY, J. R., ZHOU, Y. & SILICIANO, R. F. 2006. Nuclear retention of multiply spliced HIV-1 RNA in resting CD4+ T cells. *PLoS Pathog,* 2**,** e68.

LEEPER, T. C., ATHANASSIOU, Z., DIAS, R. L., ROBINSON, J. A. & VARANI, G. 2005. TAR RNA recognition by a cyclic peptidomimetic of Tat protein. *Biochemistry,* 44**,** 12362-72.

LENASI, T., CONTRERAS, X. & PETERLIN, B. M. 2008. Transcriptional interference antagonizes proviral gene expression to promote HIV latency. *Cell Host Microbe,* 4**,** 123-33.

LI, C. & LIEBER, A. 2019. Adenovirus vectors in hematopoietic stem cell genome editing. *FEBS Lett,* 593**,** 3623-3648.

LI, C., MORI, L. & VALENTE, S. T. 2021. The Block-and-Lock Strategy for Human Immunodeficiency Virus Cure: Lessons Learned from Didehydro-Cortistatin A. *J Infect Dis,* 223**,** 46-53.

LI, C., MOUSSEAU, G. & VALENTE, S. T. 2019. Tat inhibition by didehydro-Cortistatin A promotes heterochromatin formation at the HIV-1 long terminal repeat. *Epigenetics Chromatin,* 12**,** 23.

LI, C. J., FRIEDMAN, D. J., WANG, C., METELEV, V. & PARDEE, A. B. 1995. Induction of apoptosis in uninfected lymphocytes by HIV-1 Tat protein. *Science,* 268**,** 429-31.

LI, L., DAHIYA, S., KORTAGERE, S., AIAMKITSUMRIT, B., CUNNINGHAM, D., PIRRONE, V., NONNEMACHER, M. R. & WIGDAHL, B. 2012. Impact of Tat Genetic Variation on HIV-1 Disease. *Adv Virol,* 2012**,** 123605.

LI, L., OLVERA, J. M., YODER, K. E., MITCHELL, R. S., BUTLER, S. L., LIEBER, M., MARTIN, S. L. & BUSHMAN, F. D. 2001. Role of the non-homologous DNA end joining pathway in the early steps of retroviral infection. *EMBO J,* 20**,** 3272-81.

LI, W., LI, G., STEINER, J. & NATH, A. 2009. Role of Tat protein in HIV neuropathogenesis. *Neurotox Res,* 16**,** 205-20.

LIM, S. Y., OSUNA, C. E., HRABER, P. T., HESSELGESSER, J., GEROLD, J. M., BARNES, T. L., SANISETTY, S., SEAMAN, M. S., LEWIS, M. G., GELEZIUNAS, R., MILLER, M. D., CIHLAR, T., LEE, W. A., HILL, A. L. & WHITNEY, J. B. 2018. TLR7 agonists induce transient viremia and reduce the viral reservoir in SIV-infected rhesus macaques on antiretroviral therapy. *Sci Transl Med,* 10.

LIN, X., IRWIN, D., KANAZAWA, S., HUANG, L., ROMEO, J., YEN, T. S. & PETERLIN, B. M. 2003. Transcriptional profiles of latent human immunodeficiency virus in infected individuals: effects of Tat on the host and reservoir. *J Virol,* 77**,** 8227-36.

LIU, R. D., WU, J., SHAO, R. & XUE, Y. H. 2014. Mechanism and factors that control HIV-1 transcription and latency activation. *J Zhejiang Univ Sci B,* 15**,** 455-65.

LIU, Z., CHEN, S., JIN, X., WANG, Q., YANG, K., LI, C., XIAO, Q., HOU, P., LIU, S., WU, S., HOU, W., XIONG, Y., KONG, C., ZHAO, X., WU, L., LI, C., SUN, G. & GUO, D. 2017. Genome editing of the HIV co-receptors CCR5 and CXCR4 by CRISPR-Cas9 protects CD4(+) T cells from HIV-1 infection. *Cell Biosci,* 7**,** 47.

LOPEZ-HUERTAS, M. R., MATEOS, E., SANCHEZ DEL COJO, M., GOMEZ-ESQUER, F., DIAZ-GIL, G., RODRIGUEZ-MORA, S., LOPEZ, J. A., CALVO, E., LOPEZ-CAMPOS, G., ALCAMI, J. & COIRAS, M. 2013. The presence of HIV-1 Tat protein second exon delays fas protein-mediated apoptosis in CD4+ T lymphocytes: a potential mechanism for persistent viral production. *J Biol Chem,* 288**,** 7626-7644.

LORENZI, J. C., COHEN, Y. Z., COHN, L. B., KREIDER, E. F., BARTON, J. P., LEARN, G. H., OLIVEIRA, T., LAVINE, C. L., HORWITZ, J. A., SETTLER, A., JANKOVIC, M., SEAMAN, M. S., CHAKRABORTY, A. K., HAHN, B. H., CASKEY, M. & NUSSENZWEIG, M. C. 2016. Paired quantitative and qualitative assessment of the replication-competent HIV-1 reservoir and comparison with integrated proviral DNA. *Proc Natl Acad Sci U S A,* 113**,** E7908-E7916.

LUCERA, M. B., TILTON, C. A., MAO, H., DOBROWOLSKI, C., TABLER, C. O., HAQQANI, A. A., KARN, J. & TILTON, J. C. 2014. The histone deacetylase inhibitor vorinostat (SAHA) increases the susceptibility of uninfected CD4+ T cells to HIV by increasing the kinetics and efficiency of postentry viral events. *J Virol,* 88**,** 10803-12.

LUGER, K., MADER, A. W., RICHMOND, R. K., SARGENT, D. F. & RICHMOND, T. J. 1997. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature,* 389**,** 251-60.

MACALLAN, D. C., BUSCH, R. & ASQUITH, B. 2019. Current estimates of T cell kinetics in humans. *Curr Opin Syst Biol,* 18**,** 77-86.

MANN, J. F. S., PANKRAC, J., KLEIN, K., MCKAY, P. F., KING, D. F. L., GIBSON, R., WIJEWARDHANA, C. N., PAWA, R., MEYEROWITZ, J., GAO, Y., CANADAY, D. H., AVINO, M., POON, A. F. Y., FOSTER, C., FIDLER, S., SHATTOCK, R. J. & ARTS, E. J. 2020. A targeted reactivation of latent HIV-1 using an activator vector in patient samples from acute infection. *EBioMedicine,* 59**,** 102853.

MARCELLO, A. 2006. Latency: the hidden HIV-1 challenge. *Retrovirology,* 3**,** 7.

MARCELLO, A., DHIR, S. & DIEUDONNE, M. 2010. Nuclear positional control of HIV transcription in 4D. *Nucleus,* 1**,** 8-11.

MARGOLIS, D. M. & HAZUDA, D. J. 2013. Combined approaches for HIV cure. *Curr Opin HIV AIDS,* 8**,** 230-5.

MARTIN, A. R. & SILICIANO, R. F. 2016. Progress Toward HIV Eradication: Case Reports, Current Efforts, and the Challenges Associated with Cure. *Annu Rev Med,* 67**,** 215-28.

MARZIO, G., TYAGI, M., GUTIERREZ, M. I. & GIACCA, M. 1998. HIV-1 tat transactivator recruits p300 and CREB-binding protein histone acetyltransferases to the viral promoter. *Proc Natl Acad Sci U S A,* 95**,** 13519-24.

MBONYE, U. & KARN, J. 2011. Control of HIV latency by epigenetic and non-epigenetic mechanisms. *Curr HIV Res,* 9**,** 554-67.

MBONYE, U. & KARN, J. 2017. The Molecular Basis for Human Immunodeficiency Virus Latency. *Annu Rev Virol,* 4**,** 261-285.

MBONYE, U., LESKOV, K., SHUKLA, M., VALADKHAN, S. & KARN, J. 2021. Biogenesis of P-TEFb in CD4+ T cells to reverse HIV latency is mediated by protein kinase C (PKC)-independent signaling pathways. *PLoS Pathog,* 17**,** e1009581.

MBONYE, U., WANG, B., GOKULRANGAN, G., SHI, W., YANG, S. & KARN, J. 2018. Cyclin-dependent kinase 7 (CDK7)-mediated phosphorylation of the CDK9 activation loop promotes P-TEFb assembly with Tat and proviral HIV reactivation. *J Biol Chem,* 293**,** 10009-10025.

MBONYE, U. R., GOKULRANGAN, G., DATT, M., DOBROWOLSKI, C., COOPER, M., CHANCE, M. R. & KARN, J. 2013. Phosphorylation of CDK9 at Ser175 enhances HIV transcription and is a marker of activated P-TEFb in CD4(+) T lymphocytes. *PLoS Pathog,* 9**,** e1003338.

MCNAMARA, R. P., MCCANN, J. L., GUDIPATY, S. A. & D'ORSO, I. 2013. Transcription factors mediate the enzymatic disassembly of promoter-bound 7SK snRNP to locally recruit P-TEFb for transcription elongation. *Cell Rep,* 5**,** 1256-68.

MEDIOUNI, S., KESSING, C. F., JABLONSKI, J. A., THENIN-HOUSSIER, S., CLEMENTZ, M., KOVACH, M. D., MOUSSEAU, G., DE VERA, I. M. S., LI, C., KOJETIN, D. J., EVANS, D. T. & VALENTE, S. T. 2019. The Tat inhibitor didehydro-cortistatin A suppresses SIV replication and reactivation. *FASEB J,* 33**,** 8280-8293.

MEDIOUNI, S., LYU, S., SCHADER, S. M. & VALENTE, S. T. 2022. Forging a Functional Cure for HIV: Transcription Regulators and Inhibitors. *Viruses,* 14.

MELE, A. R., MARINO, J., CHEN, K., PIRRONE, V., JANETOPOULOS, C., WIGDAHL, B., KLASE, Z. & NONNEMACHER, M. R. 2018. Defining the molecular mechanisms of HIV-1 Tat secretion: PtdIns(4,5)P2 at the epicenter. *Traffic*.

MENDEZ, C., LEDGER, S., PETOUMENOS, K., AHLENSTIEL, C. & KELLEHER, A. D. 2018. RNA-induced epigenetic silencing inhibits HIV-1 reactivation from latency. *Retrovirology,* 15**,** 67.

MORI, L., JENIKE, K., YEH, Y. J., LACOMBE, B., LI, C., GETZLER, A., MEDIOUNI, S., CAMERON, M., PIPKIN, M., HO, Y. C., RAMIREZ, B. C. & VALENTE, S. 2020. The XPB Subunit of the TFIIH Complex Plays a Critical Role in HIV-1 Transcription and XPB Inhibition by Spironolactone Prevents HIV-1 Reactivation from Latency. *J Virol*.

MORI, L. & VALENTE, S. T. 2020. Key Players in HIV-1 Transcriptional Regulation: Targets for a Functional Cure. *Viruses,* 12.

MORTON, E. L., FORST, C. V., ZHENG, Y., DEPAULA-SILVA, A. B., RAMIREZ, N. P., PLANELLES, V. & D'ORSO, I. 2019. Transcriptional Circuit Fragility Influences HIV Proviral Fate. *Cell Rep,* 27**,** 154-171 e9.

MOUSSEAU, G., CLEMENTZ, M. A., BAKEMAN, W. N., NAGARSHETH, N., CAMERON, M., SHI, J., BARAN, P., FROMENTIN, R., CHOMONT, N. & VALENTE, S. T. 2012. An analog of the natural steroidal alkaloid cortistatin A potently suppresses Tat-dependent HIV transcription. *Cell Host Microbe,* 12**,** 97-108.

MURCHIE, A. I., DAVIS, B., ISEL, C., AFSHAR, M., DRYSDALE, M. J., BOWER, J., POTTER, A. J., STARKEY, I. D., SWARBRICK, T. M., MIRZA, S., PRESCOTT, C. D., VAGLIO, P., ABOUL-ELA, F. & KARN, J. 2004. Structure-based drug design targeting an inactive RNA conformation: exploiting the flexibility of HIV-1 TAR RNA. *J Mol Biol,* 336**,** 625-38.

NEKHAI, S. & JEANG, K. T. 2006. Transcriptional and post-transcriptional regulation of HIV-1 gene expression: role of cellular factors for Tat and Rev. *Future Microbiol,* 1**,** 417-26.

NEOGI, U., GUPTA, S., SAHOO, P. N., SHET, A., RAO, S. D., RANGA, U. & PRASAD, V. R. 2012. Genetic characterization of HIV type 1 Tat exon 1 from a southern Indian clinical cohort: identification of unique epidemiological signature residues. *AIDS Res Hum Retroviruses,* 28**,** 1152-6.

NETTLES, R. E., KIEFFER, T. L., KWON, P., MONIE, D., HAN, Y., PARSONS, T., COFRANCESCO, J., JR., GALLANT, J. E., QUINN, T. C., JACKSON, B., FLEXNER, C., CARSON, K., RAY, S., PERSAUD, D. & SILICIANO, R. F. 2005. Intermittent HIV-1 viremia (Blips) and drug resistance in patients receiving HAART. *JAMA,* 293**,** 817-29.

NEUVEUT, C., SCOGGINS, R. M., CAMERINI, D., MARKHAM, R. B. & JEANG, K. T. 2003. Requirement for the second coding exon of Tat in the optimal replication of macrophage-tropic HIV-1. *J Biomed Sci,* 10**,** 651-60.

NOVIS, C. L., ARCHIN, N. M., BUZON, M. J., VERDIN, E., ROUND, J. L., LICHTERFELD, M., MARGOLIS, D. M., PLANELLES, V. & BOSQUE, A. 2013. Reactivation of latent HIV-1 in central memory CD4(+) T cells through TLR-1/2 stimulation. *Retrovirology,* 10**,** 119.

OFFERSEN, R., NISSEN, S. K., RASMUSSEN, T. A., OSTERGAARD, L., DENTON, P. W., SOGAARD, O. S. & TOLSTRUP, M. 2016. A Novel Toll-Like Receptor 9 Agonist, MGN1703, Enhances HIV-1 Transcription and NK Cell-Mediated Inhibition of HIV-1-Infected Autologous CD4+ T Cells. *J Virol,* 90**,** 4441-4453.

OKULICZ, J. F. & LAMBOTTE, O. 2011. Epidemiology and clinical characteristics of elite controllers. *Curr Opin HIV AIDS,* 6**,** 163-8.

OPI, S., PELOPONESE, J. M., JR., ESQUIEU, D., CAMPBELL, G., DE MAREUIL, J., WALBURGER, A., SOLOMIAC, M., GREGOIRE, C., BOUVERET, E., YIRRELL, D. L. & LORET, E. P. 2002. Tat HIV-1 primary and tertiary structures critical to immune response against non-homologous variants. *J Biol Chem,* 277**,** 35915-9.

OTT, M., GEYER, M. & ZHOU, Q. 2011. The control of HIV transcription: keeping RNA polymerase II on track. *Cell Host Microbe,* 10**,** 426-35.

PAGANS, S., KAUDER, S. E., KAEHLCKE, K., SAKANE, N., SCHROEDER, S., DORMEYER, W., TRIEVEL, R. C., VERDIN, E., SCHNOLZER, M. & OTT, M. 2010. The Cellular lysine methyltransferase Set7/9-KMT7 binds HIV-1 TAR RNA, monomethylates the viral transactivator Tat, and enhances HIV transcription. *Cell Host Microbe,* 7**,** 234-44.

PANG, S., KOYANAGI, Y., MILES, S., WILEY, C., VINTERS, H. V. & CHEN, I. S. 1990. High levels of unintegrated HIV-1 DNA in brain tissue of AIDS dementia patients. *Nature,* 343**,** 85-9.

PANTANO, S., TYAGI, M., GIACCA, M. & CARLONI, P. 2002. Amino acid modification in the HIV-1 Tat basic domain: insights from molecular dynamics and in vivo functional studies. *J Mol Biol,* 318**,** 1331-9.

PANTANO, S., TYAGI, M., GIACCA, M. & CARLONI, P. 2004. Molecular dynamics simulations on HIV-1 Tat. *Eur Biophys J,* 33**,** 344-51.

PEARSON, R., KIM, Y. K., HOKELLO, J., LASSEN, K., FRIEDMAN, J., TYAGI, M. & KARN, J. 2008. Epigenetic silencing of human immunodeficiency virus (HIV) transcription by formation of restrictive chromatin structures at the viral long terminal repeat drives the progressive entry of HIV into latency. *J Virol,* 82**,** 12291-303.

PEETERS, A., LAMBERT, P. F. & DEACON, N. J. 1996. A fourth Sp1 site in the human immunodeficiency virus type 1 long terminal repeat is essential for negative-sense transcription. *J Virol,* 70**,** 6665-72.

PELOPONESE, J. M., JR., COLLETTE, Y., GREGOIRE, C., BAILLY, C., CAMPESE, D., MEURS, E. F., OLIVE, D. & LORET, E. P. 1999. Full peptide synthesis, purification, and characterization of six Tat variants. Differences observed between HIV-1 isolates from Africa and other continents. *J Biol Chem,* 274**,** 11473-8.

PEREIRA, L. A., BENTLEY, K., PEETERS, A., CHURCHILL, M. J. & DEACON, N. J. 2000. A compilation of cellular transcription factor interactions with the HIV-1 LTR promoter. *Nucleic Acids Res,* 28**,** 663-8.

PERKINS, M. J., BRADLEY, W. P., LALANI, T., AGAN, B. K., WHITMAN, T. J., FERGUSON, T. M., OKULICZ, J. F. & GANESAN, A. 2017. Brief Report: Prevalence of Posttreatment Controller Phenotype Is Rare in HIV-Infected Persons After Stopping Antiretroviral Therapy. *J Acquir Immune Defic Syndr,* 75**,** 364-369.

PETTERSEN, E. F., GODDARD, T. D., HUANG, C. C., COUCH, G. S., GREENBLATT, D. M., MENG, E. C. & FERRIN, T. E. 2004. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem,* 25**,** 1605-12.

PHAM, V. V., SALGUERO, C., KHAN, S. N., MEAGHER, J. L., BROWN, W. C., HUMBERT, N., DE ROCQUIGNY, H., SMITH, J. L. & D'SOUZA, V. M. 2018. HIV-1 Tat interactions with cellular 7SK and viral TAR RNAs identifies dual structural mimicry. *Nat Commun,* 9**,** 4266.

PHATNANI, H. P. & GREENLEAF, A. L. 2006. Phosphorylation and functions of the RNA polymerase II CTD. *Genes Dev,* 20**,** 2922-36.

PIERLEONI, R., MENOTTA, M., ANTONELLI, A., SFARA, C., SERAFINI, G., DOMINICI, S., LAGUARDIA, M. E., SALIS, A., DAMONTE, G., BANCI, L., PORCU, M., MONINI, P., ENSOLI, B. & MAGNANI, M. 2010. Effect of the redox state on HIV-1 tat protein multimerization and cell internalization and trafficking. *Mol Cell Biochem,* 345**,** 105-18.

PLATT, E. J., BILSKA, M., KOZAK, S. L., KABAT, D. & MONTEFIORI, D. C. 2009. Evidence that ecotropic murine leukemia virus contamination in TZM-bl cells does not affect the outcome of neutralizing antibody assays with human immunodeficiency virus type 1. *J Virol,* 83**,** 8289-92.

PLUTA, A., JAWORSKI, J. P. & CORTES-RUBIO, C. N. 2020. Balance between Retroviral Latency and Transcription: Based on HIV Model. *Pathogens,* 10.

POMERANTZ, R. J., TRONO, D., FEINBERG, M. B. & BALTIMORE, D. 1990. Cells nonproductively infected with HIV-1 exhibit an aberrant pattern of viral RNA expression: a molecular model for latency. *Cell,* 61**,** 1271-6.

POPPER, S. J., SARR, A. D., TRAVERS, K. U., GUEYE-NDIAYE, A., MBOUP, S., ESSEX, M. E. & KANKI, P. J. 1999. Lower human immunodeficiency virus (HIV) type 2 viral load reflects the difference in pathogenicity of HIV-1 and HIV-2. *J Infect Dis,* 180**,** 1116-21.

POUDEL, A. N., NEWLANDS, D. & SIMKHADA, P. 2017. The economic burden of HIV/AIDS on individuals and households in Nepal: a quantitative study. *BMC Health Serv Res,* 17**,** 76.

RAMSKOLD, D., WANG, E. T., BURGE, C. B. & SANDBERG, R. 2009. An abundance of ubiquitously expressed genes revealed by tissue transcriptome sequence data. *PLoS Comput Biol,* 5**,** e1000598.

RANDS, C. M., MEADER, S., PONTING, C. P. & LUNTER, G. 2014. 8.2% of the Human genome is constrained: variation in rates of turnover across functional element classes in the human lineage. *PLoS Genet,* 10**,** e1004525.

RANGA, U., SHANKARAPPA, R., SIDDAPPA, N. B., RAMAKRISHNA, L., NAGENDRAN, R., MAHALINGAM, M., MAHADEVAN, A., JAYASURYAN, N., SATISHCHANDRA, P., SHANKAR, S. K. & PRASAD, V. R. 2004. Tat protein of human immunodeficiency virus type 1 subtype C strains is a defective chemokine. *J Virol,* 78**,** 2586-90.

RASMUSSEN, T. A. & LEWIN, S. R. 2016. Shocking HIV out of hiding: where are we with clinical trials of latency reversing agents? *Curr Opin HIV AIDS,* 11**,** 394-401.

RASMUSSEN, T. A., TOLSTRUP, M., BRINKMANN, C. R., OLESEN, R., ERIKSTRUP, C., SOLOMON, A., WINCKELMANN, A., PALMER, S., DINARELLO, C., BUZON, M., LICHTERFELD, M., LEWIN, S. R., OSTERGAARD, L. & SOGAARD, O. S. 2014. Panobinostat, a histone deacetylase inhibitor, for latent-virus reactivation in HIV-infected patients on suppressive antiretroviral therapy: a phase 1/2, single group, clinical trial. *Lancet HIV,* 1**,** e13-21.

RAYNE, F., DEBAISIEUX, S., YEZID, H., LIN, Y. L., METTLING, C., KONATE, K., CHAZAL, N., AROLD, S. T., PUGNIERE, M., SANCHEZ, F., BONHOURE, A., BRIANT, L., LORET, E., ROY, C. & BEAUMELLE, B. 2010. Phosphatidylinositol-(4,5)-bisphosphate enables efficient secretion of HIV-1 Tat by infected T-cells. *EMBO J,* 29**,** 1348-62.

REDEL, L., LE DOUCE, V., CHERRIER, T., MARBAN, C., JANOSSY, A., AUNIS, D., VAN LINT, C., ROHR, O. & SCHWARTZ, C. 2010. HIV-1 regulation of latency in the monocyte-macrophage lineage and in CD4+ T lymphocytes. *J Leukoc Biol,* 87**,** 575-88.

RICE, A. P. 2017. The HIV-1 Tat Protein: Mechanism of Action and Target for HIV-1 Cure Strategies. *Curr Pharm Des,* 23**,** 4098-4102.

RICHMAN, D. D., MARGOLIS, D. M., DELANEY, M., GREENE, W. C., HAZUDA, D. & POMERANTZ, R. J. 2009. The challenge of finding a cure for HIV infection. *Science,* 323**,** 1304-7.

RODRIGUEZ, M. A., DING, M., RATNER, D., CHEN, Y., TRIPATHY, S. P., KULKARNI, S. S., CHATTERJEE, R., TARWATER, P. M. & GUPTA, P. 2009. High replication fitness and transmission efficiency of HIV-1 subtype C from India: Implications for subtype C predominance. *Virology,* 385**,** 416-24.

ROEBUCK, K. A. & SAIFUDDIN, M. 1999. Regulation of HIV-1 transcription. *Gene Expr,* 8**,** 67-84.

ROMANI, B., ENGELBRECHT, S. & GLASHOFF, R. H. 2010. Functions of Tat: the versatile protein of human immunodeficiency virus type 1. *J Gen Virol,* 91**,** 1-12.

RONSARD, L., GANGULI, N., SINGH, V. K., MOHANKUMAR, K., RAI, T., SRIDHARAN, S., PAJANIRADJE, S., KUMAR, B., RAI, D., CHAUDHURI, S., COUMAR, M. S., RAMACHANDRAN, V. G. & BANERJEA, A. C. 2017a. Impact of Genetic Variations in HIV-1 Tat on LTR-Mediated Transcription via TAR RNA Interaction. *Front Microbiol,* 8**,** 706.

RONSARD, L., LATA, S., SINGH, J., RAMACHANDRAN, V. G., DAS, S. & BANERJEA, A. C. 2014. Molecular and genetic characterization of natural HIV-1 Tat Exon-1 variants from North India and their functional implications. *PLoS One,* 9**,** e85452.

RONSARD, L., RAI, T., RAI, D., RAMACHANDRAN, V. G. & BANERJEA, A. C. 2017b. In silico Analyses of Subtype Specific HIV-1 Tat-TAR RNA Interaction Reveals the Structural Determinants for Viral Activity. *Front Microbiol,* 8**,** 1467.

ROSSENKHAN, R., MACLEOD, I. J., SEBUNYA, T. K., CASTRO-NALLAR, E., MCLANE, M. F., MUSONDA, R., GASHE, B. A., NOVITSKY, V. & ESSEX, M. 2013. tat Exon 1 exhibits functional diversity during HIV-1 subtype C primary infection. *J Virol,* 87**,** 5732-45.

RUBEN, S., PERKINS, A., PURCELL, R., JOUNG, K., SIA, R., BURGHOFF, R., HASELTINE, W. A. & ROSEN, C. A. 1989. Structural and functional characterization of human immunodeficiency virus tat protein. *J Virol,* 63**,** 1-8.

RUELAS, D. S. & GREENE, W. C. 2013. An integrated overview of HIV-1 latency. *Cell,* 155**,** 519-29.

SAAYMAN, S., ACKLEY, A., TURNER, A. W., FAMIGLIETTI, M., BOSQUE, A., CLEMSON, M., PLANELLES, V. & MORRIS, K. V. 2014. An HIV-encoded antisense long noncoding RNA epigenetically regulates viral transcription. *Mol Ther,* 22**,** 1164-1175.

SADOWSKI, I. & HASHEMI, F. B. 2019. Strategies to eradicate HIV from infected patients: elimination of latent provirus reservoirs. *Cell Mol Life Sci,* 76**,** 3583-3600.

SAEZ-CIRION, A., BACCHUS, C., HOCQUELOUX, L., AVETTAND-FENOEL, V., GIRAULT, I., LECUROUX, C., POTARD, V., VERSMISSE, P., MELARD, A., PRAZUCK, T., DESCOURS, B., GUERGNON, J., VIARD, J. P., BOUFASSA, F., LAMBOTTE, O., GOUJARD, C., MEYER, L., COSTAGLIOLA, D., VENET, A., PANCINO, G., AUTRAN, B., ROUZIOUX, C. & GROUP, A. V. S. 2013. Post-treatment HIV-1 controllers with a long-term virological remission after the interruption of early initiated antiretroviral therapy ANRS VISCONTI Study. *PLoS Pathog,* 9**,** e1003211.

SAINSBURY, S., BERNECKY, C. & CRAMER, P. 2015. Structural basis of transcription initiation by RNA polymerase II. *Nat Rev Mol Cell Biol,* 16**,** 129-43.

SANDER, J. D. & JOUNG, J. K. 2014. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat Biotechnol,* 32**,** 347-55.

SANTORO, M. M. & PERNO, C. F. 2013. HIV-1 Genetic Variability and Clinical Implications. *ISRN Microbiol,* 2013**,** 481314.

SCHULLER, R., FORNE, I., STRAUB, T., SCHREIECK, A., TEXIER, Y., SHAH, N., DECKER, T. M., CRAMER, P., IMHOF, A. & EICK, D. 2016. Heptad-Specific Phosphorylation of RNA Polymerase II CTD. *Mol Cell,* 61**,** 305-14.

SCHULZE-GAHMEN, U., UPTON, H., BIRNBERG, A., BAO, K., CHOU, S., KROGAN, N. J., ZHOU, Q. & ALBER, T. 2013. The AFF4 scaffold binds human P-TEFb adjacent to HIV Tat. *Elife,* 2**,** e00327.

SEDORE, S. C., BYERS, S. A., BIGLIONE, S., PRICE, J. P., MAURY, W. J. & PRICE, D. H. 2007. Manipulation of P-TEFb control machinery by HIV: recruitment of P-TEFb from the large form by Tat and binding of HEXIM1 to TAR. *Nucleic Acids Res,* 35**,** 4347-58.

SENGUPTA, S. & SILICIANO, R. F. 2018. Targeting the Latent Reservoir for HIV-1. *Immunity,* 48**,** 872-895.

SHARKEY, M. E., TEO, I., GREENOUGH, T., SHAROVA, N., LUZURIAGA, K., SULLIVAN, J. L., BUCY, R. P., KOSTRIKIS, L. G., HAASE, A., VERYARD, C., DAVARO, R. E., CHEESEMAN, S. H., DALY, J. S., BOVA, C., ELLISON, R. T., 3RD, MADY, B., LAI, K. K., MOYLE, G., NELSON, M., GAZZARD, B., SHAUNAK, S. & STEVENSON, M. 2000. Persistence of episomal HIV-1 infection intermediates in patients on highly active anti-retroviral therapy. *Nat Med,* 6**,** 76-81.

SHIN, Y., KIM, H. G., PARK, C. M., CHOI, M. S., KIM, D. E., CHOI, B. S., KIM, K. & YOON, C. H. 2020. Identification of novel compounds against Tat-mediated human immunodeficiency virus-1 transcription by high-throughput functional screening assay. *Biochem Biophys Res Commun,* 523**,** 368-374.

SHOJANIA, S. & O'NEIL, J. D. 2006. HIV-1 Tat is a natively unfolded protein: the solution conformation and dynamics of reduced HIV-1 Tat-(1-72) by NMR spectroscopy. *J Biol Chem,* 281**,** 8347-56.

SILICIANO, J. D., KAJDAS, J., FINZI, D., QUINN, T. C., CHADWICK, K., MARGOLICK, J. B., KOVACS, C., GANGE, S. J. & SILICIANO, R. F. 2003. Long-term follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting CD4+ T cells. *Nat Med,* 9**,** 727-8.

SILICIANO, R. F. & GREENE, W. C. 2011. HIV latency. *Cold Spring Harb Perspect Med,* 1**,** a007096.

SIMONETTI, F. R., SOBOLEWSKI, M. D., FYNE, E., SHAO, W., SPINDLER, J., HATTORI, J., ANDERSON, E. M., WATTERS, S. A., HILL, S., WU, X., WELLS, D., SU, L., LUKE, B. T., HALVAS, E. K., BESSON, G., PENROSE, K. J., YANG, Z., KWAN, R. W., VAN WAES, C., ULDRICK, T., CITRIN, D. E., KOVACS, J., POLIS, M. A., REHM, C. A., GORELICK, R., PIATAK, M., KEELE, B. F., KEARNEY, M. F., COFFIN, J. M., HUGHES, S. H., MELLORS, J. W. & MALDARELLI, F. 2016. Clonally expanded CD4+ T cells can produce infectious HIV-1 in vivo. *Proc Natl Acad Sci U S A,* 113**,** 1883-8.

SLOAN, R. D. & WAINBERG, M. A. 2011. The role of unintegrated DNA in HIV infection. *Retrovirology,* 8**,** 52.

SMITH, K. M., HIMIARI, Z., TSIMBALYUK, S. & FORWOOD, J. K. 2017. Structural Basis for Importin-alpha Binding of the Human Immunodeficiency Virus Tat. *Sci Rep,* 7**,** 1650.

SOBHIAN, B., LAGUETTE, N., YATIM, A., NAKAMURA, M., LEVY, Y., KIERNAN, R. & BENKIRANE, M. 2010. HIV-1 Tat assembles a multifunctional transcription elongation complex and stably associates with the 7SK snRNP. *Mol Cell,* 38**,** 439-51.

SOGAARD, O. S., GRAVERSEN, M. E., LETH, S., OLESEN, R., BRINKMANN, C. R., NISSEN, S. K., KJAER, A. S., SCHLEIMANN, M. H., DENTON, P. W., HEY-CUNNINGHAM, W. J., KOELSCH, K. K., PANTALEO, G., KROGSGAARD, K., SOMMERFELT, M., FROMENTIN, R., CHOMONT, N., RASMUSSEN, T. A., OSTERGAARD, L. & TOLSTRUP, M. 2015. The Depsipeptide Romidepsin Reverses HIV-1 Latency In Vivo. *PLoS Pathog,* 11**,** e1005142.

SONG, M. 2017. The CRISPR/Cas9 system: Their delivery, in vivo and ex vivo applications and clinical development by startups. *Biotechnol Prog,* 33**,** 1035-1045.

SPECTOR, C., MELE, A. R., WIGDAHL, B. & NONNEMACHER, M. R. 2019. Genetic variation and function of the HIV-1 Tat protein. *Med Microbiol Immunol,* 208**,** 131-169.

SPRAGG, C., DE SILVA FEELIXGE, H. & JEROME, K. R. 2016. Cell and gene therapy strategies to eradicate HIV reservoirs. *Curr Opin HIV AIDS,* 11**,** 442-9.

STERNBERG, S. H. & DOUDNA, J. A. 2015. Expanding the Biologist's Toolkit with CRISPR-Cas9. *Mol Cell,* 58**,** 568-74.

STRAHL, B. D. & ALLIS, C. D. 2000. The language of covalent histone modifications. *Nature,* 403**,** 41-5.

SUH, H., FICARRO, S. B., KANG, U. B., CHUN, Y., MARTO, J. A. & BURATOWSKI, S. 2016. Direct Analysis of Phosphorylation Sites on the Rpb1 C-Terminal Domain of RNA Polymerase II. *Mol Cell,* 61**,** 297-304.

SUZUKI, K., JUELICH, T., LIM, H., ISHIDA, T., WATANEBE, T., COOPER, D. A., RAO, S. & KELLEHER, A. D. 2008. Closed chromatin architecture is induced by an RNA duplex targeting the HIV-1 promoter region. *J Biol Chem,* 283**,** 23353-63.

SUZUKI, Y. & CRAIGIE, R. 2007. The road to chromatin - nuclear entry of retroviruses. *Nat Rev Microbiol,* 5**,** 187-96.

TAHIROV, T. H., BABAYEVA, N. D., VARZAVAND, K., COOPER, J. J., SEDORE, S. C. & PRICE, D. H. 2010. Crystal structure of HIV-1 Tat complexed with human P-TEFb. *Nature,* 465**,** 747-51.

THOMAS, J. R. & HERGENROTHER, P. J. 2008. Targeting RNA with small molecules. *Chem Rev,* 108**,** 1171-224.

TO, V., DZANANOVIC, E., MCKENNA, S. A. & O'NEIL, J. 2016. The Dynamic Landscape of the Full-Length HIV-1 Transactivator of Transcription. *Biochemistry,* 55**,** 1314-25.

TRIPATHY, M. K., ABBAS, W. & HERBEIN, G. 2011. Epigenetic regulation of HIV-1 transcription. *Epigenomics,* 3**,** 487-502.

TROJER, P. & REINBERG, D. 2007. Facultative heterochromatin: is there a distinctive molecular signature? *Mol Cell,* 28**,** 1-13.

TSAI, A., IRRINKI, A., KAUR, J., CIHLAR, T., KUKOLJ, G., SLOAN, D. D. & MURRY, J. P. 2017. Toll-Like Receptor 7 Agonist GS-9620 Induces HIV Expression and HIV-Specific Immunity in Cells from HIV-Infected Individuals on Suppressive Antiretroviral Therapy. *J Virol,* 91.

TYAGI, M., PEARSON, R. J. & KARN, J. 2010. Establishment of HIV latency in primary CD4+ cells is due to epigenetic transcriptional silencing and P-TEFb restriction. *J Virol,* 84**,** 6425-37.

UNAIDS. 2015. *Understanding Fast-Track: Accelerating action to end the AIDS epidemic by 2030* [Online]. [Accessed].

VAN DER KUYL, A. C., VINK, M., ZORGDRAGER, F., BAKKER, M., WYMANT, C., HALL, M., GALL, A., BLANQUART, F., BERKHOUT, B., FRASER, C., CORNELISSEN, M. & COLLABORATION, B. 2018. The evolution of subtype B HIV-1 tat in the Netherlands during 1985-2012. *Virus Res,* 250**,** 51-64.

VAN DUYNE, R., EASLEY, R., WU, W., BERRO, R., PEDATI, C., KLASE, Z., KEHN-HALL, K., FLYNN, E. K., SYMER, D. E. & KASHANCHI, F. 2008. Lysine methylation of HIV-1 Tat regulates transcriptional activity of the viral LTR. *Retrovirology,* 5**,** 40.

VAN LINT, C., BOUCHAT, S. & MARCELLO, A. 2013. HIV-1 transcription and latency: an update. *Retrovirology,* 10**,** 67.

VANSANT, G., BRUGGEMANS, A., JANSSENS, J. & DEBYSER, Z. 2020. Block-And-Lock Strategies to Cure HIV Infection. *Viruses,* 12.

VANSANT, G., VRANCKX, L. S., ZURNIC, I., VAN LOOVEREN, D., VAN DE VELDE, P., NOBLES, C., GIJSBERS, R., CHRIST, F. & DEBYSER, Z. 2019. Impact of LEDGIN treatment during virus production on residual HIV-1 transcription. *Retrovirology,* 16**,** 8.

VARDABASSO, C., MANGANARO, L., LUSIC, M., MARCELLO, A. & GIACCA, M. 2008. The histone chaperone protein Nucleosome Assembly Protein-1 (hNAP-1) binds HIV-1 Tat and promotes viral transcription. *Retrovirology,* 5**,** 8.

VERDIKT, R., HERNALSTEENS, O. & VAN LINT, C. 2021. Epigenetic Mechanisms of HIV-1 Persistence. *Vaccines (Basel),* 9.

VRANCKX, L. S., DEMEULEMEESTER, J., SALEH, S., BOLL, A., VANSANT, G., SCHRIJVERS, R., WEYDERT, C., BATTIVELLI, E., VERDIN, E., CERESETO, A., CHRIST, F., GIJSBERS, R. & DEBYSER, Z. 2016. LEDGIN-mediated Inhibition of Integrase-LEDGF/p75 Interaction Reduces Reactivation of Residual Latent HIV. *EBioMedicine,* 8**,** 248-264.

WAGNER, T. A., MCLAUGHLIN, S., GARG, K., CHEUNG, C. Y., LARSEN, B. B., STYRCHAK, S., HUANG, H. C., EDLEFSEN, P. T., MULLINS, J. I. & FRENKEL, L. M. 2014. HIV latency. Proliferation of cells with HIV integrated into cancer genes contributes to persistent infection. *Science,* 345**,** 570-3.

WAHEED, A. A. & FREED, E. O. 2012. HIV type 1 Gag as a target for antiviral therapy. *AIDS Res Hum Retroviruses,* 28**,** 54-75.

WALLET, C., DE ROVERE, M., VAN ASSCHE, J., DAOUAD, F., DE WIT, S., GAUTIER, V., MALLON, P. W. G., MARCELLO, A., VAN LINT, C., ROHR, O. & SCHWARTZ, C. 2019. Microglial Cells: The Main HIV-1 Reservoir in the Brain. *Front Cell Infect Microbiol,* 9**,** 362.

WANG, W., YE, C., LIU, J., ZHANG, D., KIMATA, J. T. & ZHOU, P. 2014. CCR5 gene disruption via lentiviral vectors expressing Cas9 and single guided RNA renders cells resistant to HIV-1 infection. *PLoS One,* 9**,** e115987.

WANG, Z., GURULE, E. E., BRENNAN, T. P., GEROLD, J. M., KWON, K. J., HOSMANE, N. N., KUMAR, M. R., BEG, S. A., CAPOFERRI, A. A., RAY, S. C., HO, Y. C., HILL, A. L., SILICIANO, J. D. & SILICIANO, R. F. 2018. Expanded cellular clones carrying replication-competent HIV-1 persist, wax, and wane. *Proc Natl Acad Sci U S A,* 115**,** E2575-E2584.

WANG, Z., MORRIS, G. F., RICE, A. P., XIONG, W. & MORRIS, C. B. 1996. Wild-type and transactivation-defective mutants of human immunodeficiency virus type 1 Tat protein bind human TATA-binding protein in vitro. *J Acquir Immune Defic Syndr Hum Retrovirol,* 12**,** 128-38.

WANG, Z. & MOULT, J. 2001. SNPs, protein structure, and disease. *Hum Mutat,* 17**,** 263-70.

WEI, P., GARBER, M. E., FANG, S. M., FISCHER, W. H. & JONES, K. A. 1998. A novel CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its high-affinity, loop-specific binding to TAR RNA. *Cell,* 92**,** 451-62.

WEYDERT, C., DE RIJCK, J., CHRIST, F. & DEBYSER, Z. 2016. Targeting Virus-host Interactions of HIV Replication. *Curr Top Med Chem,* 16**,** 1167-90.

WILUSZ, J. 2013. Putting an 'End' to HIV mRNAs: capping and polyadenylation as potential therapeutic targets. *AIDS Res Ther,* 10**,** 31.

WONG, J. K., CAMPBELL, G. R. & SPECTOR, S. A. 2010. Differential induction of interleukin-10 in monocytes by HIV-1 clade B and clade C Tat proteins. *J Biol Chem,* 285**,** 18319-25.

WONG, J. K., HEZAREH, M., GUNTHARD, H. F., HAVLIR, D. V., IGNACIO, C. C., SPINA, C. A. & RICHMAN, D. D. 1997. Recovery of replication-competent HIV despite prolonged suppression of plasma viremia. *Science,* 278**,** 1291-5.

WU, Y. 2004. HIV-1 gene expression: lessons from provirus and non-integrated DNA. *Retrovirology,* 1**,** 13.

XU, W., LI, H., WANG, Q., HUA, C., ZHANG, H., LI, W., JIANG, S. & LU, L. 2017. Advancements in Developing Strategies for Sterilizing and Functional HIV Cures. *Biomed Res Int,* 2017, 6096134.

XUN, J., ZHANG, X., GUO, S., LU, H. & CHEN, J. 2021. Editing out HIV: application of gene editing technology to achieve functional cure. *Retrovirology,* 18, 39.

YAN, Y., ZHANG, D., ZHOU, P., LI, B. & HUANG, S. Y. 2017. HDOCK: a web server for protein-protein and protein-DNA/RNA docking based on a hybrid strategy. *Nucleic Acids Res,* 45, W365-W373.

YANG, Z., ZHU, Q., LUO, K. & ZHOU, Q. 2001. The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription. *Nature,* 414, 317-22.

YEH, Y. J., YANG, K., RAZMI, A. & HO, Y. C. 2021. The Clonal Expansion Dynamics of the HIV-1 Reservoir: Mechanisms of Integration Site-Dependent Proliferation and HIV-1 Persistence. *Viruses,* 13.

YEZID, H., KONATE, K., DEBAISIEUX, S., BONHOURE, A. & BEAUMELLE, B. 2009. Mechanism for HIV-1 Tat insertion into the endosome membrane. *J Biol Chem,* 284, 22736-46.

YIK, J. H., CHEN, R., NISHIMURA, R., JENNINGS, J. L., LINK, A. J. & ZHOU, Q. 2003. Inhibition of P-TEFb (CDK9/Cyclin T) kinase and RNA polymerase II transcription by the coordinated actions of HEXIM1 and 7SK snRNA. *Mol Cell,* 12, 971-82.

YUKL, S., PILLAI, S., LI, P., CHANG, K., PASUTTI, W., AHLGREN, C., HAVLIR, D., STRAIN, M., GUNTHARD, H., RICHMAN, D., RICE, A. P., DAAR, E., LITTLE, S. & WONG, J. K. 2009. Latently-infected CD4+ T cells are enriched for HIV-1 Tat variants with impaired transactivation activity. *Virology,* 387, 98-108.

YUKL, S. A., KAISER, P., KIM, P., TELWATTE, S., JOSHI, S. K., VU, M., LAMPIRIS, H. & WONG, J. K. 2018. HIV latency in isolated patient CD4(+) T cells may be due to blocks in HIV transcriptional elongation, completion, and splicing. *Sci Transl Med,* 10.

ZAMBORLINI, A., LEHMANN-CHE, J., CLAVE, E., GIRON, M. L., TOBALY-TAPIERO, J., ROINGEARD, P., EMILIANI, S., TOUBERT, A., DE THE, H. & SAIB, A. 2007. Centrosomal pre-integration latency of HIV-1 in quiescent cells. *Retrovirology,* 4, 63.

ZHAO, M., DE CRIGNIS, E., ROKX, C., VERBON, A., VAN GELDER, T., MAHMOUDI, T., KATSIKIS, P. D. & MUELLER, Y. M. 2019. T cell toxicity of HIV latency reversing agents. *Pharmacol Res,* 139, 524-534.

ZHOU, M., HALANSKI, M. A., RADONOVICH, M. F., KASHANCHI, F., PENG, J., PRICE, D. H. & BRADY, J. N. 2000. Tat modifies the activity of CDK9 to phosphorylate serine 5 of the RNA polymerase II carboxyl-terminal domain during human immunodeficiency virus type 1 transcription. *Mol Cell Biol,* 20, 5077-86.

ZHOU, Q., CHEN, D., PIERSTORFF, E. & LUO, K. 1998. Transcription elongation factor P-TEFb mediates Tat activation of HIV-1 transcription at multiple stages. *EMBO J,* 17, 3681-91.

ZHU, P. & LI, G. 2016. Structural insights of nucleosome and the 30-nm chromatin fiber. *Curr Opin Struct Biol,* 36, 106-15.

ZHU, Y., PE'ERY, T., PENG, J., RAMANATHAN, Y., MARSHALL, N., MARSHALL, T., AMENDT, B., MATHEWS, M. B. & PRICE, D. H. 1997. Transcription elongation factor P-TEFb is required for HIV-1 tat transactivation in vitro. *Genes Dev,* 11, 2622-32.

Supplementary material:

Geneblock (gBlock) sequence: Tat exon 1 and 2

GACGGTATCGATGCCACCATGGAGCCA GTAGATCCTAACCTAGAGCCCTGGAAC
CATCCAGGAAGTCAGCCTAAAACTCCT TGTAATAAGTGTTATTGTAAACACTGTA
GCTATCATTGTCTAGTTTGCTTTCAGAC AAAAGGCTTAGGCATTTCCTATGGCAG
GAAGAAGCGGAGACAGCGACGAAGC ACTCCTCCAAGCAGTGAGGATCATCAA
AATCTTATATCAAAGCACGTACGACCCA CCTCCCAATCCCGGGGGGACCCGACA
GGCCCGAAGGAATCGAAGAAGAAGG TGGAGAGAGAGACAGAGACAGATCC
ATTCGATGGAGGCGACTACAAGGACG ACGATGACAAGTAGAAGCTTGGTACC
GAGCTGGGATCCGCCCCT

Gating strategy for flow cytometry:

The live population was defined by forward versus side scatter profiles. Gating for SSC-H vs SSC-W and FSC-H vs FSC-W was used to exclude doublets. Cells were further gated by using forward scatter versus GFP intensity to differentiate between GFP-positive and -negative cells.

S1: Establishment of latent T cell lines harbouring minimal HIV-1C (consensus LTR and tat) reporter virus.

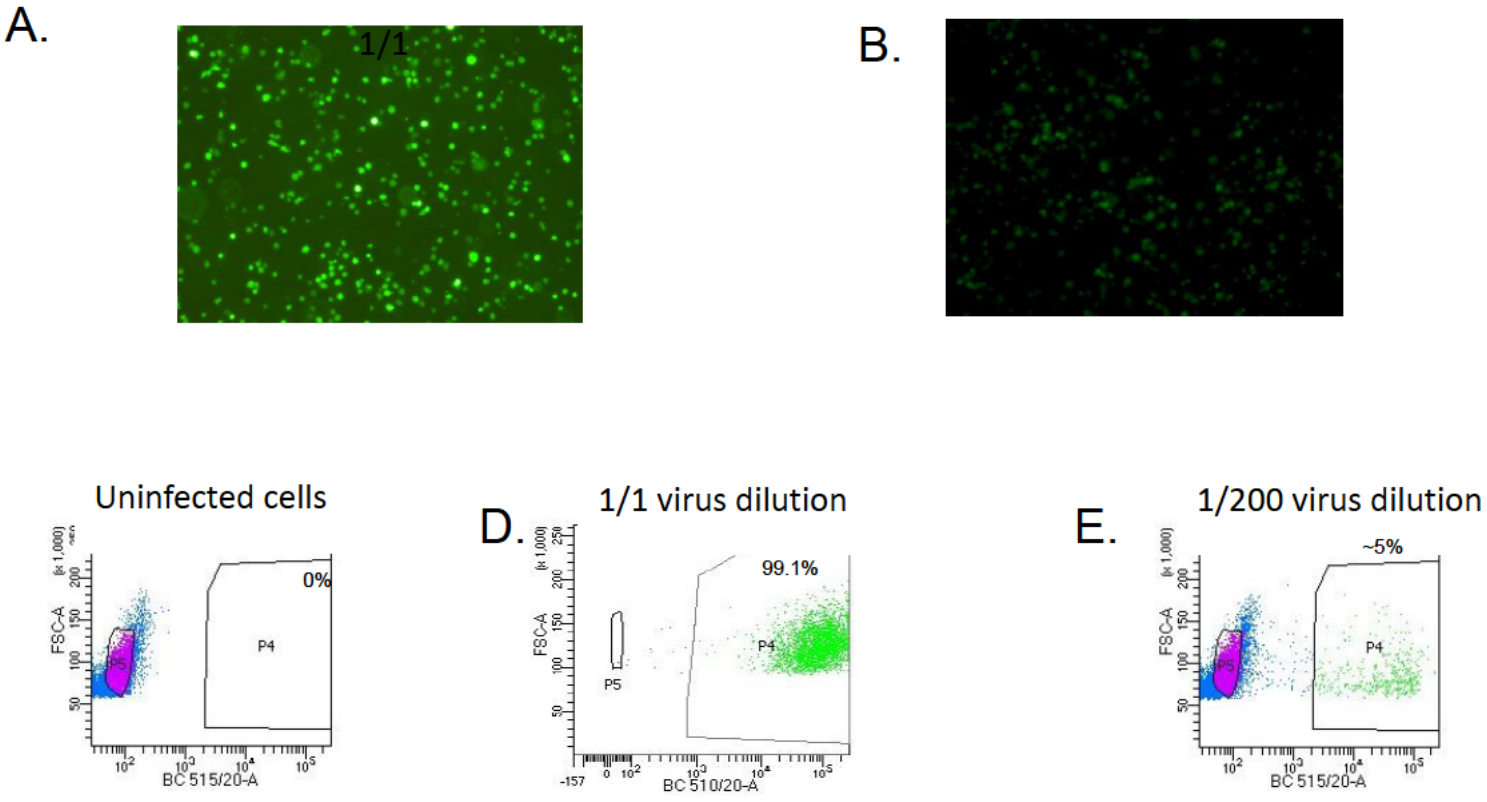

Figure S1: **A**. GFP positive cells of C731CC virus diluted 1/1 viewed under the microscope. **B**. GFP positive cells of C731CC virus diluted 1/300 viewed under the microscope. **C**. FACS plot of uninfected cells. **D**. FACS plot of 1/1 C731CC virus gated on GFP positive cells. **E**. FACS plot of 1/200 C731CC virus gated on GFP positive cells.

S2 Construction of HIV-1C genome minimal GFP reporter virus

Jordan et al 2001 reported that to study HIV-1 transcription in its natural context (i.e., integrated in the host genome), an HIV-1-derived retroviral vector should be used to generate a library of Jurkat clones containing single integration of an HIV-1 mini genome. To study the effect of HIV-1C transcription, subtype B LTR and Tat was replaced with subtype C LTR and Tat. We used an HIV-based retroviral vector containing the Tat and GFP open reading frames both under the control of the HIV promoter in the 5' LTR. We confirmed successful cloning into C731CC using restriction digestion and gel electrophoresis.

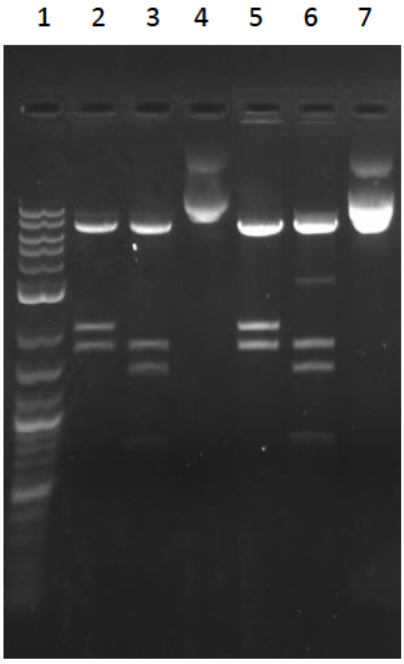

Figure 4.1: Gel electrophoresis of restriction digestion of C731CC showing successful LTR and Tat cloning. Lane: 1: 2 Log Ladder Marker, 2: C731CC cut with XmaI, 3: C731CC cut with XmaI and XhoI, 4: Uncut C731CC plasmid. 5: C731CC cut with XmaI, 6: C731CC cut with XmaI and XhoI, 7: Uncut C731CC plasmid.

Lane 2 shows the C731CC plasmid digested with the XmaI. Lane 3 shows C731CC digested with XmaI and XhoI to produce 7183 bp, 1741 bp, 1444 bp and a 672 bp sized fragment. The same banding pattern is seen in lane 6. Lane 4 shows the uncut C731CC plasmid which is

11 090 bp. It is also shown in lane 7. This gel shows that HIV-1C 5’LTR, Tat and 3’LTR into the pEV vector was successful. The gel shows successful cloning of Tat into C731CC.