Synergy between enhanced transcriptional strength of the viral promoter and the Tat-autoregulatory circuit accelerates latency-establishment in HIV-1C

A thesis submitted in partial fulfilment

of the requirements for the degree of

Doctor of Philosophy

By

Sutanuka Chakraborty

HIV-AIDS Laboratory
Molecular Biology and Genetics Unit
Jawaharlal Nehru Centre for Advanced Scientific Research
BANGALORE- 560064, India

April, 2018

Dedicated to

my family

DECLARATION

I hereby declare that the thesis entitled "**Synergy between enhanced transcriptional strength of the viral promoter and the Tat-autoregulatory circuit accelerates latency-establishment in HIV-1C"** is the result of research work carried out by me under the supervision of Prof. Udaykumar Ranga at the Molecular Biology and Genetics Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore, India. In keeping with the general practice in reporting the scientific observations, due acknowledgments have been made wherever the work described is based on the findings of other investigators. Any oversight due to error of judgment is regretted.

Ms. Sutanuka Chakraborty

Place: Bangalore

Date: 17 April 2018

Jawaharlal Nehru Centre for Advanced Scientific Research
(Autonomous Body under the Department of Science & Technology, Government of India)
Jakkur Campus, Jakkur Post
Bengaluru 560 064, INDIA
Office Tel: +91 (80) 22082830
Office Fax: +91 (80) 2208 2766
+91 (80) 2208 2767
Email: udaykumar@jncasr.ac.in

CERTIFICATE

I hereby certify that the work described in the thesis entitled "**Synergy between enhanced transcriptional strength of the viral promoter and the Tat-autoregulatory circuit accelerates latency-establishment in HIV-1C"** has been carried out by Ms. Sutanuka Chakraborty in the HIV-AIDS Laboratory, Molecular Biology and Genetics Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore, India under my supervision. This work in part or full has not been submitted to any other institute or university for the award of any other degree or diploma.

Udaykumar Ranga PhD
Professor and Head
Molecular Biology and Genetics Unit

Place: Bangalore
Date: 17 April 2018

Acknowledgements

I would like to express my heartfelt gratitude to my supervisor Prof. Udaykumar Ranga for giving me the opportunity to work in his laboratory. He has not only been my PhD mentor but also instilled in me the scientific temperament and curiosity to ask the right kind of academic questions. I sincerely thank him for his invaluable guidance, criticism, counselling, troubleshooting, ingenious ideas and ceaseless enthusiasm.

I would like to thank the former chairmen Prof. M.R.S. Rao and Prof. Anuranjan Anand for the excellent facilities in MBGU to carry out high-quality research. I am indebted to Prof. Tapas K Kundu for investing his valuable time in the proactive discussions related to my research and for his constant support and encouragement. My sincere thanks to Dr Ravi Manjithaya for his suggestions in the confocal data analysis. I would like to acknowledge Prof. Namita Suroliya, Prof. Tapas K Kundu and Dr. Ravi Manjithaya for their course work. I truly appreciate the faculties of MBGU- Prof. Kaustuv Sanyal, Prof. Hemalatha Balaram, Prof. Maneesha Inamdar and Dr. James Chelliah for their insightful suggestions and critical comments during my work presentations.

I especially thank Dr. Meher Prakash from CPMU and his group for the suggestion and analysis of the 'Hill coefficient' data which led to important leads in my project.

I have had the fortune and pleasure of working with a group of very talented and supportive colleagues. I take this opportunity to thank all the past and present members of the lab. My special thanks to Dr. Anand, Dr. Mahesh, Dr. Anjali, Dr. Shilpee, Dr. Prabhu, for their excellent hands-on-training and cultivating the right scientific spirit in the laboratory. My warm thanks to Sreshtha, Ragini, Navneet, Debanjan, Mariyam, Tirath, Malini, Arun, Disha, Deepak, Nityanand, Shambhu, Rishikesh, Rebu, Pavithra, Siddharth, Shubham, Sravan, Chhavi, Haider, Bramha, Sukanya, Parni, Kavita, Neelam, Yuvraj, Swati, Vijeta for making my stay in the lab memorable. My special thanks to Manisha, Surabhi and Neelu for helping me in my work. I cherish the wonderful, fun-filled moments spent with Dr. Anjali, Dr. Shilpee, Shambhu, Debanjan, Mariyam, Sreshtha and Navneet outside the lab. I thank Lakkappa, Sridhar, Lavanya, for their assistance in the lab.

My stay at JNCASR would have been incomplete without the constant support and encouragement from a lovely bunch of friends and hostel mates. I would like to specially thank Pooja, Mariyam, Debanjan, Papri, Srimayee, Niki, Aisha, Abhiroop, Anindita, Sunita, Arnab, Sweta, Dr. Kirthana, Deboshree, Susheela, Manaswini, Dibyashre and, Arkomita who

stood by me through the thick and thin. I thank my other MBGU friends, colleagues and seniors- Shalini, Divyesh, Arindam, Shveta, Rima, Pallabi, Sreedevi, Vijay (JCL), Dr. Amrutha, Dr. Ronak, Dr. Kalpita, Dr. Manpreet, Dr. Vijays (Akhade and J), Dr. Sreyoshi, Dr. Stephanie, Aditya, Shrilaxmi, Sambhavi, Neha, Saloni, Aparna, Arpit, Anindita, Bornika for sharing reagents, equipments and all the good times spent together.

I acknowledge the MBGU central facilities- Dr. Uttara, Swaroopa and Dr. Narendra (flow cytometry), Suma Ma'am and Sunil (Confocal facility), Anitha (Sequencing facility) for extending the technical support. I am thankful to the members and staff of Academic Section, Administrative Section, Library, Complab, Hostel and specially Dhanvantari for their contribution. I wholeheartedly thank JNCASR for the financial support during the entire tenure of my PhD.

I am indebted to my teachers and mentors- my parents, Dr. Swagata Ray, Dr. Aditi Nag Chaudhuri, Dr. Biren Das, Dr. Tapan Datta and Dr. Satyen Saha for their invaluable teaching, well wishes and bountiful inspiration. They have played a pivotal role in the successful completion of my PhD thesis. Additionally, my heartfelt thanks to all the well-wishers who made my stay in Bangalore pleasant, comfortable and enjoyable. I would like to especially thank Srinagesh uncle, Meghna di, Sanchita di, Ravi ji, Ira and Shashank in this regard.

Words fall short to acknowledge my family for their enormous contribution towards my accomplishments. My parents have always been a constant source of inspiration to chase my dreams. I am extremely grateful for their sacrifice, perseverance and commitment to help me succeed in life. I express my heartfelt gratitude to my husband and in-laws for supporting through-out my PhD. Special thanks to my husband, who has been my friend, philosopher and guide in the true sense. I take this opportunity to express my warm regards to all the members of my extended family and other relatives who have knowingly and unknowingly supported me in my academic endeavour.

Last but not the least, I am grateful to the all-merciful Lord, for showering his blessings upon me.

SYNOPSIS OF THE THESIS

Title: Synergy between enhanced transcriptional strength of the viral promoter and the Tat-autoregulatory circuit accelerates latency-establishment in HIV-1C

Submitted by

Sutanuka Chakraborty
HIV-AIDS Laboratory,
Molecular Biology and Genetics Unit
Jawaharlal Nehru Centre for Advanced Scientific Research
Bangalore- 560064, India

Thesis advisor: Prof. Udaykumar Ranga

A world-wide, functional HIV-1 cure appears implausible despite the discovery of the virus over 30 years ago as the replication-competent provirus remains persistent within a small pool of long-lived memory $CD4^+$ T cells. The latent virus is capable of reinitiating fresh rounds of infection, thus, subverting antiretroviral therapy (ART). HIV-1 latency further poses a serious challenge to modelling the event in a laboratory setting. There are few to no markers that can distinguish a latently infected cell from its uninfected counterpart. There exists considerable debate among the HIV-1 researchers as to whether the host factors or the intrinsic viral factors play the pivotal role in regulating viral latency. This complexity has led to two distinct schools of thought - the hypothesis of 'epiphenomenon' or that of 'viral circuitry' to explain HIV-1 latency. While the majority of the latency studies in HIV-1 are supportive of the former hypothesis (Chun TW et al., 1997; Pierson T et al., 2000), compelling pieces of evidence have begun to emerge in recent years in support of the latter (Razooky BS et al., 2015; Weinberger LS et al., 2005).

Given the lack of an appropriate marker to indicate viral latency, HIV-1 vectors encoding fluorescent proteins have emerged as powerful tools to study the phenomenon. Pseudotyped reporter vectors of HIV-1continue to remain as model vectors of choice to elucidate the basic mechanisms of viral latency. The absence of interference from other

viral components simplifies the system of pseudotyped reporter vectors and permits in-depth analyses of the mechanisms regulating latency.

In this backdrop, it must be emphasized that the examination of the mechanisms that regulate latency in HIV-1 subtype C (HIV-1C) is further complicated for two different reasons. First, HIV-1C contains a huge variation in the profile of transcription factor binding sites (TFBS), importantly that of NF-κB. The enhancer of HIV-1C typically contains three elements of NF-κB, and we reported previously the emergence of viral strains containing four of these elements. In contrast, other genetic families of HIV-1 contain only one or two of these motifs (Bachu M et al., 2012; Boullosa J et al., 2014). Second, the additional NF-κB binding elements of HIV-1C (the third and the fourth) are also genetically diverse. Thus, the LTR of HIV-1C (C-LTR) is characterized by the copy-number and genetic variations of the NF-κB motifs that also translates to the enhanced transcriptional strength of the viral promoter, a situation unique for this viral family. The impact of such genetic diversity on the establishment and maintenance of transcriptional silence has not been examined previously. In this backdrop, the present study attempts to examine the influence of the copy-number variation of NF-κB binding elements in the C-LTR. Importantly, the focus of the present study is on the copy-number difference of the NF-κB binding sites, not on the genetic variation of these elements, therefore, on the overall strength of transcription and its influence on viral latency.

Chapter 1 provides the review of literature pertaining to the present thesis. Following a brief introduction to viral latency, the chapter presents an overview and critical evaluation of the two prevailing hypotheses for the establishment of HIV-1 latency: the 'epiphenomenon' hypothesis and the 'viral circuitry' hypothesis. The viral circuitry of HIV-1 is elaborated in greater detail including a comparison of the phenomenon among other prokaryotic and eukaryotic viruses known to establish a latent phase during their life cycles- particularly bacteriophage λ, Cytomegalovirus (CMV), Herpes Simplex Virus (HSV) and Epstein Barr Virus (EBV). Next, a detailed account of the experimental strategies and the models employed for the study of HIV-1 latency are presented, portraying the merits and limitations of each model. The classification of the genetic subtypes of HIV-1 and their global preponderance are reviewed with an emphasis on HIV-1C, the focus of the present thesis. A detailed description of the TFBS profile of the

LTR is presented. The domain architecture and functions of the viral regulatory protein Tat and its possible involvement in regulating viral latency are reviewed (Brigati C et al., 2003; Hetzer C et al., 2005; Karn J et al., 1999). Finally, the current eradication approaches to counter HIV-1 persistence are discussed.

Chapter 2 depicts the technical details regarding the reagents, experimental methodologies, statistical analyses, and software packages used in the present work.

Chapter 3 illustrates the details of the three different Jurkat cell models used here to study HIV-1C latency. The three models differ from one another in the way the LTR and Tat are functionally associated with each other in the context of HIV-1C minimal vector backbone. **The Autonomous Tat-feedback (ATF)** model preserves the natural LTR-Tat feedback axis as in the full-length viral genome (Burnett JC et al., 2009; Weinberger LS et al., 2005). The vector LTR-GFP-IRES-Tat (LGIT) co-expresses EGFP or GFPd2 and Tat from the LTR. Tat expression from this vector is expected to be proportional to the strength of the viral promoter. Thus, in the ATF model, any variation in the LTR-strength would inevitably lead to an associated change in the Tat-feedback strength. Therefore, the ATF model examines the synergistic influence of LTR and Tat-feedback strengths on viral latency. The second model used in the study is **the Disjoint Tat-feedback model (DTF)**. Only EGFP, but not Tat, is expressed under the viral promoter from the vector LTR-GFP (LG). Tat is expressed in the cell under the control of the inducible Tet-ON promoter. That the functional LTR-Tat feedback axis is disrupted in this experimental model, it would be possible to study viral gene expression and latency as a function of varying promoter strength alone under normalized Tat concentration. The **Tunable Tat-feedback (TTF) model** permits the examination of viral transcription at a constant transcriptional strength of the LTR by manipulating only the Tat-feedback strength. This is made possible by employing controlled proteolysis of the Tat protein by attaching a destabilization domain (DD) at the C-terminal end of the viral protein. Tat stability can be reversibly modulated by the use of Shield1, a small molecule regulator (Banaszynski LA et al., 2006). Additionally, an RFP ORF was fused with Tat to create the Tat-RFP fusion protein tagged with the DD domain. The construction of Tat-RFP fusion in LTR-GFPd2-IRES-Tat:RFP:DD (LdGITRD) permits the expression analysis of Tat in real time. We constructed a panel of five LTR variant vectors that differed in the number of functional

NF-κB sites ranging from 4 to 0 in the viral enhancer in each of the above three latency models.

Viral stocks pseudotyped with VSV-G envelope were generated in HEK293T cells using standard experimental procedures (Burnett JC et al., 2009). The relative infectious units (RIU) of the viral stocks were estimated in Jurkat cells using flow cytometry of GFP expression. Following the validation of each experimental cell model, the kinetic profiles of latency were constructed to decipher the influence of promoter and/or Tat-feedback strength on latency.

The transcriptional strength of the LTR was directly proportional to the number of functional NF-κB binding elements in the ATF model. We first used the ATF panel of reporter viral strains to assess the NF-κB copy-number variation regulating gene expression and latency. Jurkat cells were infected independently with each of the five LGIT variant strains (with functional NF-κB sites ranging from 4 to 0) at an RIU of ~0.05-0.06. Three days following the infection, half of the infected cells were activated with a cocktail of T-cell activators (PMA, TNFα, TSA, and HMBA). GFP flow analysis and Tat-RT PCR were performed 24 h following the activation in both the control and activated fractions to evaluate the transcriptional strengths of all the LTR variants. A perfect positive correlation was observed between the number of functional NF-κB motifs in the LTR and GFP or Tat expression from the panel of LTRs. The results, thus, ascertained the additive effect of the NF-κB motifs on the LTR. Of note, the GFP expressing cells segregated into two distinct populations based on the mean fluorescence intensity – the low (GFP^{Low}) and high GFP (GFP^{High}) cells. While the proportion of the GFP^{Low} cells remained comparable across the vectors of the LTR panel, importantly, that of the GFP^{High} cells correlated positively with the copy-number of the NF-κB motifs. The stronger the viral LTR, the higher was the proportion of GFP^{High} cells possibly representing the Tat-transactivated cells harbouring proviruses containing a stronger LTR-Tat feedback circuit.

To understand the influence of NF-κB copy-number variation on viral latency, we generated the kinetic GFP expression profiles of stable Jurkat pools infected with the five viral strains of the ATF model varying in NF-κB copy-numbers from 4 to 0. Jurkat cells

were infected with each variant strain at a low RIU (~0.01-0.02) to ensure a single integration event per cell as confirmed using real-time PCR. The infected cells were expanded for a week under standard culture conditions before activating the cells with a cocktail of global activators (40 ng/ml PMA + 40 ng/ml TNFα + 200 nM TSA + 2.5 mM HMBA) to reactivate the proviruses before the start of the assay. Twenty four hours following the activation, the GFP^{+ve} cells were sorted, and GFP expression was monitored by flow cytometry at multiple time-points. The MFI of GFP expression at the start of the assay (D0 post GFP^{+ve} sort) was directly proportional to the copies of NF-κB motifs among the viral strains despite the percentage of infection remaining constant. Additionally, the initial expression-level of Tat transcripts was also directly proportional to the functional copies of NF-κB motifs in the LTR.

A strong, positive Tat-feedback is probably essential for rapid latency-establishment. Importantly, the GFP^{+ve} cells in the experiment could be further classified into two distinct populations based on the fluorescence intensity – low or basal and high or transactivated GFP phenotypes adopting the phenomenon of phenotypic bifurcation (PheB) as described previously (Weinberger LS et al.; 2005). Following cell infection and activation, all the viral strains of the panel demonstrated three different cell pools – GFP^{-ve} cells, GFP^{Low} cells, and GFP^{High} cells. Of the two GFP^{+ve} cell fractions in the assay, the GFP^{High} fraction represented a stronger Tat feedback circuit, thus, denoting the 'transactivated' population. Of note, although the transcriptional strength of the viral promoter influenced the percent of cells in each of the three populations, the most pronounced impact on the GFP expression as well as latency was manifested on the GFP^{High} fraction.

In the ATF model, stronger the viral promoter, faster was the establishment of latency. The results of the above analyses from the gene-expression and latency studies in the ATF model unequivocally established that the transcriptional strength of HIV-1C LTR is directly proportional to the number of the functional NF-κB motifs in the enhancer. It is rather paradoxical that a virus that must establish transcriptional silence should possess a strong promoter especially when the other genetic families of HIV-1 do not employ such a strategy. To understand this paradox, we determined the kinetics of transcriptional silence of the panel of ATF reporter viral strains which identified many

interesting issues. The expression of GFP reduced progressively from all the five LTRs up to 16 days. Surprisingly, the rate of GFP switch off was faster from the stronger promoters (3- and 4-κB LTRs) as compared to the three weak promoters (0-, 1- and 2-κB LTRs) in the panel. The kinetics of latency-establishment was the slowest from the 0-κB LTR followed by 1- and 2-κB LTRs suggesting that the presence of the NF-κB binding motifs plays a direct role in establishing HIV-1 latency. Furthermore, the process of latency-establishment was not complete from any of the promoters that even for the 3-κB LTR that demonstrated the fastest GFP switch off kinetics, approximately 56.6% cells remained GFP^{+ve} at day 16 indicating that only 43.4% cells became latent. The inherently activated state of Jurkat cells combined with the nearly 48 h of the half-life of EGFP may underlie the incomplete latency-establishment. To circumvent this technical limitation, we substituted EGFP with GFPd2 that contains a half-life of only 2 h and found that the variant GFP faithfully represented LTR transcriptional activity, unlike EGFP. We generated the NF-κB variant panel with GFPd2 for the subsequent analyses. The mean fluorescence intensities of the LTRs were proportional to the NF-κB number in the panel and ranged from the highest for the 4-κB LTR to the lowest for the 0-κB LTR at the 0 h time point and broadly at the subsequent time points. Collectively, our data of the ATF model are assertive that the transcriptional strength of HIV-1 promoter is an important parameter regulating viral latency. The stronger is the transcriptional strength, the faster is the rate of latency-establishment. GFPd2, a variant form of GFP containing a shorter mean half-life represents the kinetics of the viral transcription more faithfully than the regular EGFP mostly used in the assays.

The gene-expression and kinetic profiles of latency-establishment in the DTF model were conserved across the NF-κB variants. In the ATF model described above, a stronger viral promoter not only contained a difference in the number of NF-κB motifs but also generated more intracellular Tat protein, thus, making it difficult to discern the influence of either factor alone on viral latency. To circumvent this limitation, we removed Tat expression from the LTR and placed it under a Tet-ON promoter. The Tat-transcript levels were consistent across the variants as well as at two different time-points of the assay suggesting accurate functioning of the DTF model. Strikingly, there was no correlation with the copy-number of NF-κB sites and viral gene expression both with and without global activation. Also, under the normalized Tat levels in the DTF model, the

kinetic profiles of latency-establishment did not vary significantly across the LTR variants despite the difference in the copies of NF-κB motifs. Additionally, there was a complete loss of GFP^{High} cells even under the strongest LTR following the abolition of the Tat-feedback circuit suggesting that the basal strength of transcription alone could not produce a significant difference in the latency-establishment when the functional LTR-Tat feedback axis is disrupted.

The TTF model identified two distinct modes of latency-establishment in HIV-1C: Tat-dependent and Tat-independent routes. The TTF model, unlike the ATF model, permitted to keep the transcriptional strength of the viral promoter constant and change only the strength of the Tat-feedback circuit. To this end, Tat was expressed under the LTR and as a fusion protein fused with RFP. Further, Tat-RFP protein was tagged with a degradation domain; therefore, the cellular half-life of Tat-RFP-DD protein could be regulated by using Shield1, a small molecule modulator. We examined latency kinetics of the TTF model using 3- and 1-NF-κB variant LTRs as representatives of strong and weak viral promoters, respectively. Fluorescent microscopy of transfected HEK293T cells at 48 h showed a dose-response of RFP expression with the Shield1 administration. Likewise in Jurkat cells, the analysis of GFP expression by flow cytometry and the quantitation of Tat-transcript expressions by real-time PCR showed proportional enhancement with the Shield1 concentration. The analysis of latency kinetics in the TTF model was restricted only to the GFP^{High} cell population since only these cells represent the Tat-transactivated population. The TTF model also permits the independent evaluation of the LTR-strength and Tat-feedback strength by monitoring the expression of two fluorescent markers-GFPd2 and RFP. The data of the TTF model disclosed the presence of two distinct mechanisms of LTR-silencing, the Tat-dependent and Tat-independent. Tat-dependent latency is characterized by the manifestation of a Tat-RFP^{+ve} GFP^{-ve} phenotype suggestive of the presence of Tat in a cell where the LTR is transcriptionally silent. This phenotype is predominantly represented by the stronger 3-κB LTR. In contrast, the weak 1-κB LTR is only restricted to the Tat-RFP^{-ve} GFP^{-ve} phenotype suggestive of a trajectory of Tat-independent latency. Further, the kinetics of latency-establishment is faster by the Tat-dependent, than Tat-independent route. This observation provided a satisfying explanation as to why the stronger LTR can establish latency at a faster rate. Overall, our experimental strategy using a multi-directional approach collectively ascertained two

important aspects of HIV-1C latency. First, the transcriptional strength of the viral promoter could play a dominant role in viral latency especially when the Tat positive feedback circuit is intact. Second, the data presented here allude to a direct role Tat plays in the repression of LTR transcriptional activity.

In **Chapter 4,** we examined the kinetics of HIV-1C latency using the ATF model in greater detail. We sorted single GFP^{High} cells from all the five LTR variant strains and monitored the profile of GFP expression in the clonal lines during latency. We established at least 16 clonal cell populations representing each of the 5 variant LTRs of the panel to examine GFP expression profile. We found that the isoclonal populations representing each LTR variant demonstrated one of the three phenotypes of latency. All the cells of some clones maintained persistent GFP^{High} expression up to day 21 without down-regulation of GFP. Cells of some other clones demonstrated the opposite that all the cells completely down-regulated GFP expression- the rate of down-regulation being directly proportional to the NF-κB copy-number in the LTR. A few other cell populations exhibited a bimodal distribution of GFP^{High} and GFP^{-ve} cells. Importantly, the cells of the bimodal type were either strongly positive for the GFP expression or completely negative without an intermediate phenotype.

ChIP assay detected the presence of Tat at the latent promoter. To examine the nature of the transcription complexes recruited to the LTRs, we sorted GFP^{High} and GFP^{-ve} fractions of the clonally expanded cells of two independent cell clones representing each of 3- and 4-κB LTRs. The selected clonal cells demonstrated the bimodal gene expression. Thus, in the experimental design, the proviruses of the GFP^{High} and GFP^{-ve} cell fractions were expected to contain the same integration site since they both derived from the same parental cell clone. We performed a ChIP analysis using the chromatin extracted from GFP^{High} and GFP^{-ve} cells and interrogated for the presence of several host factors and chromatin marks including NF-κB and NFAT transcription factors, and HIV-1 Tat on the viral promoters. Reciprocal binding of NF-κB and NFAT members was noted at the active versus latent LTRs. Importantly, we demonstrate the presence of Tat on the latent promoter for the first time although the levels of Tat on the latent LTR were relatively low, approximately, 1.7 - 3.0 fold lower, as compared to those found on the active promoter. The presence of Tat on the latent LTR was highly reproducible, could be

detected up to 7 days from the onset of latency and consistent with our proposition that Tat may be involved in playing an active role in establishing latency. In the subsequent work, we attempted to confirm the presence of Tat in the latent cell by more experiments.

The persistent presence of Tat was noted in the nucleus after LTR silencing. The nuclear levels of Tat are of relevance to HIV-1 latency. We, therefore, determined the nuclear and extra-nuclear levels of Tat in infected Jurkat cells using confocal microscopy as GFP expression from the LTR dropped with the establishment of transcription silence. Jurkat cells infected with the 3-κB LdGIT (ATF) virus were subjected to simultaneous GFP expression analysis by flow cytometry and Tat expression analysis by indirect immunofluorescence staining using a rabbit anti-Tat primary antibody followed by an anti-rabbit secondary antibody conjugated to Alexa-568. The GFP^{+ve} cells were sorted at D0 and monitored up to D16 with readout every 4 days. The percentage of GFP^{+ve} cells reduced to 6.9, 2.2 and 0.8 by days 8, 12 and 16, respectively (Figure 4.5 B). On D16, GFP expression was found below the limit of detection suggesting the establishment of latency. We found a concomitant and progressive reduction in the intracellular Tat levels as measured from 150 cells as a function of time in both nuclear and extra-nuclear compartments (DAPI overlap). However, the concentration of Tat in the extra-nuclear compartment was higher as compared to that of the nuclear compartment. Importantly, the reduction of Tat, particularly in the initial phase (D0 to D4) was significantly much sharper in the extra-nuclear compartment with a slope of value -74.54 ± 16.8 as compared to that in the nuclear compartment which demonstrated a slope of -37.28 ± 3.2. The data thus suggested that the stability of Tat in the nucleus could be higher. The proportion of Tat level in the nucleus to that outside of the nucleus clearly demonstrated that intracellular Tat levels are more stable suggesting a direct role for Tat on HIV latency (Figure 4.5 C). Of note, Tat could still be detected on days 12 and 16 in both nuclear and extra-nuclear compartments above the threshold levels where the expression of GFP could not be detected. The overall intensity of the Tat signal at D12, and D16 was only marginally above the background level; hence an alternate technique with enhanced sensitivity was essential to validate the presence of Tat in the latent cells.

Proximity ligation assay (PLA) confirmed the presence of Tat both in the active and latent cells and indicated an enhanced proximity of Tat to p65 in the active cells

compared to the latent cells. With an intention to increase the sensitivity of Tat detection in the latent cells, we employed the in situ proximity ligation assay (PLA) to estimate Tat qualitatively and quantitatively in the active versus latent cells. Since PLA does not work well in non-adherent cells, and our attempts to adapt the protocol to the Jurkat cells were not successful, we used HEK293 cells in this assay. HEK293 cells were infected with 4-κB virus of the LdGIT panel, (representing the ATF model), the GFP^{+ve} cells were sorted, and the cells were allowed for a week for some of the cells to down-regulate GFP expression. At the time of the assay, approximately 50% of the infected cells expressed GFP. Thus, the cell pool contained both active (GFP^{+ve}) and latent (GFP^{-ve}) cells. The cell pool was stained for Tat using a pair of anti-Tat antibodies and the optimized PLA protocol. Tat-specific staining was evident only in the presence of both the antibodies not only in the GFP^{+ve} cells but also in the GFP^{-ve} cells. Tat staining by PLA was quantitated by determining the number of PLA-dots per cell in a total of 85 GFP^{+ve} cells and 119 GFP^{-ve} cells comprising of three independent experiments. These values were found to be 2.91 ± 2.5 and 2.34 ± 1.9 for GFP^{+ve} and GFP^{-ve} cells, respectively, the difference not being significant. Further, PLA analysis indicated a two-fold higher incidence of proximity of Tat to p65 in the active cells as compared to the latent cells.

In summary, the data from the PLA experiments confirmed the presence of Tat in both active as well as latent cells with physical proximity with p65 preferably in the active cells and possibly with other transcription factors binding to the LTR. ChIP, together with PLA permitted the examination of the molecular complexes differentially recruited to the active versus latent promoter and the interaction of Tat with some of these factors at the single-cell resolution. These results are also consistent with our previous data that suggested a positive correlation between the transcriptional strength of the LTR and the Tat-feedback strength which favoured latency-establishment. Importantly, we identified a reciprocal binding of essential transcription factors such as p50, p65, NFAT1, and NFAT2 to active versus latent LTRs. The presence of Tat at the latent promoter is suggestive of the viral factor playing a critical role in initiating the process of transcriptional silence and its regulation.

In **Chapter 5**, we summarize the data, offer an interpretation of the findings, highlight the novelty and limitations of our work, and make suggestions for future work based on the leads obtained. The results of all the three Jurkat-based ex vivo latency models are consistent in pointing towards a common unifying theme that the transcriptional strength of the viral promoter is a critical determinant of HIV latency. We demonstrated for the first time that the transcriptional strength of the LTR, intricately linked with the strength of the Tat feedback loop, is directly proportional to the rapidity of latency-establishment. The increasing number of NF-κB elements not only enhances the transcriptional strength but also autonomously up-regulates the Tat-feedback strength which appears to play a key role in HIV-1C latency as demonstrated by the present study. The enhanced strengths of transcription and Tat-feedback loop collectively lead to the accumulation of physiological levels of Tat in the system resulting in a possible switch from an active to a repressive form. The presence of Tat on the latent promoter as far as 7 days after the LTR is switched off (ChIP analysis, Figures 4.3 and 4.4) and the disproportionate reduction of Tat levels in the nucleus (immunofluorescence of individual cells, Figure 4.5) collectively allude to the possibility of Tat directly being involved in the repression of transcription from the LTR. In a study examining latency reversal study targeting the NF-κB signalling pathway, a combined approach of time-lapse microscopy and computation using ordinary differential equation (ODE) revealed possible cooperativity of the HIV-1C circuit as opposed to the non-cooperative HIV-1B circuit (Razooky BS et al., 2011). Work is presently in progress in our laboratory to examine how a few selected post-translation modification of Tat may be involved in influencing HIV-1 latency.

Bibliography:

Bachu, M., Yalla, S., Asokan, M., Verma, A., Neogi, U., Sharma, S., Murali, R.V., Mukthey, A.B., Bhatt, R., Chatterjee, S. and Rajan, R.E., 2012. Multiple NF-κB sites in HIV-1 subtype C long terminal repeat confer superior magnitude of transcription and thereby the enhanced viral predominance. *Journal of Biological Chemistry*, *287*(53), pp.44714-44735.

Banaszynski, L.A., Chen, L.C., Maynard-Smith, L.A., Ooi, A.L. and Wandless, T.J., 2006. A rapid, reversible, and tunable method to regulate protein function in living cells using synthetic small molecules. *Cell*, *126*(5), pp.995-1004.

Boullosa, J., Bachu, M., Bila, D., Ranga, U., Süffert, T., Sasazawa, T. and Tanuri, A., 2014. Genetic diversity in HIV-1 subtype C LTR from Brazil and Mozambique generates new transcription factor-binding sites. *Viruses*, *6*(6), pp.2495-2504.

Brigati, C., Giacca, M., Noonan, D.M. and Albini, A., 2003. HIV Tat, its TARgets and the control of viral gene expression. *FEMS microbiology letters*, *220*(1), pp.57-65.

Burnett, J.C., Miller-Jensen, K., Shah, P.S., Arkin, A.P. and Schaffer, D.V., 2009. Control of stochastic gene expression by host factors at the HIV promoter. *PLoS pathogens*, *5*(1), p.e1000260.

Chun, T.W., Carruth, L., Finzi, D., Shen, X., DiGiuseppe, J.A., Taylor, H., Hermankova, M., Chadwick, K., Margolick, J., Quinn, T.C. and Kuo, Y.H., 1997. Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection. *Nature*, *387*(6629), pp.183-188.

Hetzer, C., Dormeyer, W., Schnölzer, M. and Ott, M., 2005. Decoding Tat: the biology of HIV Tat posttranslational modifications. *Microbes and infection*, *7*(13), pp.1364-1369.

Karn, J., 1999. Tackling tat. *Journal of molecular biology*, *293*(2), pp.235-254.

Pierson, T., Hoffman, T.L., Blankson, J., Finzi, D., Chadwick, K., Margolick, J.B., Buck, C., Siliciano, J.D., Doms, R.W. and Siliciano, R.F., 2000. Characterization of chemokine receptor utilization of viruses in the latent reservoir for human immunodeficiency virus type 1. *Journal of virology*, *74*(17), pp.7824-7833.

Razooky, B.S. and Weinberger, L.S., 2011. Mapping the architecture of the HIV-1 Tat circuit: A decision-making circuit that lacks bistability and exploits stochastic noise. *Methods*, *53*(1), pp.68-77.

Razooky, B.S., Pai, A., Aull, K., Rouzine, I.M. and Weinberger, L.S., 2015. A hardwired HIV latency program. *Cell*, *160*(5), pp.990-1001.

Weinberger, L.S., Burnett, J.C., Toettcher, J.E., Arkin, A.P. and Schaffer, D.V., 2005. Stochastic gene expression in a lentiviral positive-feedback loop: HIV-1 Tat fluctuations drive phenotypic diversity. *Cell*, *122*(2), pp.169-182.

Abbreviations

AIDS	Acquired immunodeficiency syndrome
ANOVA	Analysis of Variance
AP-1	Activator protein-1
APOBEC3G/F	Apolipoprotein B mRNA editing enzyme, catalytic peptide-like 3G/3F
ATF	Autonomous Tat-feedback
AU	Arbitrary Units
Bcl-2	B-cell lymphoma 2
BLT	Bone marrow-liver-thymus
CCL	Chemokine (C-C motif) ligand
CCR	C-C chemokine receptor type 2
CD	Cluster of differentiation
cDNA	Complementary deoxyribonucleic acid
CDK	Cyclin-dependent kinase
C/EBP	CCAAT-enhancer binding protein
ChIP	Chromatin immunoprecipitation
CNS	Central nervous system
COUP	Chicken ovalbumin upstream promoter
CRFs	Circulating recombinant forms
crs	*cis* repressor sequence
CTD	Carboxy-terminal domain
CTLs	Cytotoxic T lymphocytes
DAPI	4',6-diamidino-2-phenylindole
DEAE	Diethylaminoethyl
DD	Degradation domain
ddPCR	Droplet digital polymerase chain reaction
DMEM	Dulbecco's Modified Eagle's Medium
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphates
DTF	Disjoint Tat-feedback
EBV	Epstein-Barr Virus
EDTA	Ethylenediaminetetraacetic acid
ECFP	Enhanced cyan fluorescent protein
EF1a	Elongation factor-1 alpha
EGFP	Enhanced green fluorescent protein
ELISA	Enzyme-linked immunosorbent assay
ELL2	Elongation Factor For RNA Polymerase II 2
EMSA	Electrophoretic mobility shift assay
env	Envelope
FACS	Fluorescence-activated cell sorting
FAM	Fluorescein amidite
FBS	Fetal bovine serum
FDA	Food and Drug Administration
FKBP	FK506 binding protein
FL_{50}	Fluorescence loss in 50% cells
FSC	Forward scatter
gag	Group specific antigen
GFPd2	Destabilized enhanced green fluorescent protein
gp120	Glycoprotein 120
HAART	Highly active antiretroviral therapy
HAD	HIV-associated dementia
HAT	Histone acetyltransferase
HDAC	Histone deacetylase
HDACi	Histone deacetylase inhibitors
Hex	Hexachlorofluorescein
HEXIM	Hexamethylene bisacetamide-inducible proteins
HIV-1	Human immunodeficiency virus-1
HLA-DR	Human Leukocyte Antigen – antigen D Related

HMBA	Hexamethylene bisacetamide
HMT	Histone methyl transferases
HPCs	Hematopoietic progenitor cells
HSV-1	Herpes simplex virus-1
ICP0	Infected cell protein 0
IE2	Immediate-early 2
I κB-α	Inhibitory kappaB alpha
IL	Interleukin
IRES	Internal ribosome entry site
LARP	La-related proteins
LdGIT	LTR-GFPd2-IRES-Tat
LdGITRD	LTR-GFPd2-IRES-Tat:RFP:DD
LEF-1	Lymphoid enhancer-binding factor-1
LG	LTR-GFP
LGIT	LTR-GFP-IRES-Tat
LRAs	Latency reversal agents
LSM	Laser Scanning Microscopy
LTR	Long terminal repeat
MDDCs	Monocyte-derived Dendritic Cells
MePCE	Methylphosphate capping enzyme
MFI	Mean fluorescent intensity
miRNA	Micro ribonucleic acid
MIEP	Major immediate-early promoter
Nef	Negative regulatory factor
NELF	Negative elongation factor
NFAT	Nuclear factor of activated T-cells
NF-κB	Nuclear factor-KappaB
NHP	Non-human primate model
NOD	Non-obese diabetic
NRE	Negative regulatory element
NSG	NOD SCID gamma
ODE	Ordinary differential equation
ORF	Open reading frame
PBMC	Peripheral blood mononuclear cells
PBS	Phosphate buffer saline
PCR	Polymerase chain reaction
PHA	Phytohaemagglutinin
PheB	Phenotypic bifurcation
PIC	Protease inhibitor cocktail
PKC	Protein kinase C
PLA	Proximity ligation assay
PMA	Phorbol-12-myristate-13-acetate
PMSF	Phenylmethylsulfonyl fluoride
PP1	Protein phosphatase 1
P-TEFb	Positive Transcription Elongation Factor b
PTM	Post translational modification
qPCR	Quantitative polymerase chain reaction
QVOA	Quantitative Viral Outgrowth Assay
RFP	Red fluorescent protein
RFU	Relative fluorescence unit
RIU	Relative infectious unit
RNA	Ribonucleic acid
RPMI	Roswell Park Memorial Institute
RT-PCR	Reverse transcriptase polymerase chain reaction
SCID	Severe combined
SD	Standard deviation
SDS	Sodium dodecyl sulphate
SEM	Standard error of mean
SHIV	Simian/human immunodeficiency virus
SIV	Simian immunodeficiency virus

snRNP	Small nuclear ribonucleo proteins
Sp1	Specificity protein 1
Sp4	Specificity protein 4
SSC	Side scatter
TE	Tris-EDTA
TAR	Transactivation response
Tat	Transactivator of transcription
T_{CM}	Central memory $CD4^{+ve}$ T-cells
T_{EM}	Effector memory $CD4^{+ve}$ T-cells
TFBS	Transcription factor binding sites
Th	T helper cells
TILDA	Tat/rev Induced Limiting Dilution Assay
TNFα	Tumor necrosis factor alpha
TSA	Trichostain A
T_{SCM}	Stem-cell memory $CD4^{+ve}$ T-cells
TTF	Tunable Tat-feedback
T_{TM}	Transitional memory $CD4^{+ve}$ T-cells
U3	Unique to 3’
U5	Unique to 5’
UNAIDS	United Nations Programme on Aquired Immune Deficiency Syndrome
USF	Upstream stimulating factor
VSV-G	Vesicular stomatitis virus glycoprotein

Table of Contents

Chapter 1

The Review of Literature

1.1 Background

The year 1981 marked the discovery of the human immunodeficiency virus (HIV) as the etiological agent of acquired immunodeficiency syndrome (AIDS). Till date, HIV has infected >75 million individuals globally and an estimated 37 million people are presently carrying the infection (UN Joint Programme on HIV/AIDS. MDG 6: 15 years, 15 lessons of hope from the AIDS response- Fact sheet). The first successful highly active anti-retroviral therapy (HAART) was implemented back in 1997 when an initial regimen of three anti-HIV drugs reduced the plasma viral load below the clinical detection limit of 50 copies of HIV-1 RNA/ml (Perelson AS et al., 1997). HAART could prolong the lifespan of infected individuals significantly and was thus, considered an important break-through in the field of medicine. However, a functional HIV-1 cure could not be promised by HAART which continues to be a major challenge till date. Two principal reasons account for the rapid rebound of viraemia on HAART interruption. The first is an ongoing, low-level of viral replication persistent in the immune-privileged, drug-inaccessible anatomic niche such as the CNS and lymphatic system and preferably in the macrophages which are less effective to HIV-1 cytopathy (Ho DD et al., 1986). The second and the most vital parameter for viral re-emergence on HAART termination is the existence of latent HIV-1 reservoirs in the resting $CD4^{+ve}$ T lymphocytes, a phenomenon first demonstrated in vivo by Dr. Siliciano's group (Chun TW et al., 1995).

1.2 HIV-1 latency

HIV-1 latency is defined as the reversibly nonproductive state of the integrated provirus. Following reverse transcription, the viral life cycle could be partially or completely blocked prior to the integration of the viral cDNA into the host chromatin leading to little or no viral production- a phenomenon termed pre-integration latency. A reduced pool of dNTPs in the metabolically inactive, quiescent T-cells (Gao WY et al., 1993) or the deficiency of ATP for the successful import of the pre-integration complex (PIC) (Bukrinsky MI et al., 1991) result in pre-integration latency. Several forms of unintegrated viral DNA exist that could contribute to the pre-integration latency. These include the linear cDNA, which is the primary product of the viral reverse transcription, the 1-LTR circles which arise either by homologous recombination of linear DNAs at the

LTR or by ligation of incomplete reverse transcribed cDNA and the 2-LTR circles that are produced through the process of non-homologous end joining (Farnet CM et al., 1991; Miller MD et al., 1995). However, the pre-integration latency does not account for any clinical consequence since the unintegrated viral DNA persists only for a short while (one to two days) within the host cytoplasm with minimal contribution to the long-term $CD4^{+ve}$ T-cell reservoirs (Pierson TC et al., 2002).The post-integration latency on the other hand, is characterized by the transcriptionally silent but replication-competent provirus that ultimately hinders a sterile HIV-1 eradication. Ho et al. elegantly demonstrated that <1% of the intact, replication-competent proviruses are maximally induced by a diverse combination of latency-reversal agents (LRAs) at any given point (Ho YC et al., 2013). A combined approach of genome-wide sequence analyses and statistical modeling in the study further predicted that the proportion of the provirus purged by the traditional reactivation assays underestimated the true reservoir size by almost 60 folds.

1.2.1 Reservoirs of HIV-1 latency

Resting memory $CD4^{+ve}$ T-cells are the primary reservoir for latent HIV-1 infection and the decay rate of the latent reservoir in these cells is extremely slow, the predicted half-life being 43.9 months. The average estimated time to eradicate the latent reservoir with persistent antiretroviral therapy is between 25 to 60 years (Finzi D et al., 1999; Siliciano JD et al., 2003; Murray JM et al., 2014). These cells are further classified into the central memory (T_{CM}), transitional memory (T_{TM}), effector memory (T_{EM}) and the stem-cell memory T-cells (T_{SCM}). Each subset contributes differently to the latent reservoir pool. T_{CM} and T_{TM} form the bulk of the latent reservoir. In drug-naïve patients with low $CD4^{+ve}$ counts, the T_{TM} with a slow decay rate forms the major reservoir whereas T_{CM} cells constitute the primary latent pool in HAART-treated patients with normal $CD4^{+ve}$ level. Viral persistence in these cells is assisted by homeostatic proliferation with continuous immune activation (Chomont N et al., 2009). Recently, the T_{SCM} subset was suggested of harboring HIV-1 DNA and suppressing the viral load for a median of 7 years since infection (Buzon MJ et al., 2014). Several groups have demonstrated that hematopoietic progenitor cells (HPCs; $CD34^{+ve}$) in the bone marrow might serve as excellent latent reservoirs for HIV-1 subtypes B, C and D (Carter CC et al., 2011; McNamara LA et al., 2013). Further, HIV-1 could latently infect all the HPC subsets and selectively reactivated

by TNFα and HDAC inhibitors but not by HMBA (McNamara LA et al., 2012). Cells of the monocyte/macrophage lineage display $CD4^{+ve}$ as well as CCR5 or CXCR4 receptors on the cell surface and hence susceptible to early HIV-1 infection. The resistance of the monocytes and macrophages to HIV-1 mediated apoptosis, the non-dividing nature of the macrophages and the low turnover rates for CNS resident macrophages support the ability of this lineage to serve as stable HIV-1 reservoirs although, there are only few experimental evidences towards the hypothesis (Kumar A et al., 2014; Narasipura SD et al., 2014).

1.3 The route to HIV-1 latency: The deterministic vs stochastic hypothesis

Soon after the discovery of HIV-1 latency, the next obvious question was which factor(s) absolutely controlled the viral entry to and exit from quiescence. According to the conventional belief, latency-establishment is an 'epiphenomenon' which manifested during transition of the infected, activated $CD4^{+ve}$ T-cells to the resting, memory phenotype. Latency-establishment was regarded as a rare side-effect of the viral tropism for activated $CD4^{+ve}$ T-cells and the phenomenon by itself had no evolutionary consequence (Siliciano RF and Greene WC., 2011; Eisele E and Siliciano RF., 2012). The 'epiphenomenon' hypothesis for HIV-1 latency is deterministic in nature and resulted from several clinical and experimental observations. Activated $CD4^{+ve}$ T cells and not their resting counterparts were found to be highly susceptible to viral infection and catered favorable environment for viral transcription (Siliciano RF et al., 2011; Coffin J et al., 2013). According to the deterministic hypothesis of viral latency-establishment, extrinsic factors such as the activation status of the host-cells, integration in restricted chromatin environment, interference with the host transcription and subsequent epigenetic modifications resulted in the establishment and maintenance of viral latency (reviewed in Van Lint C et al., 2013; Archin NM et al., 2014) (Figure1.1).

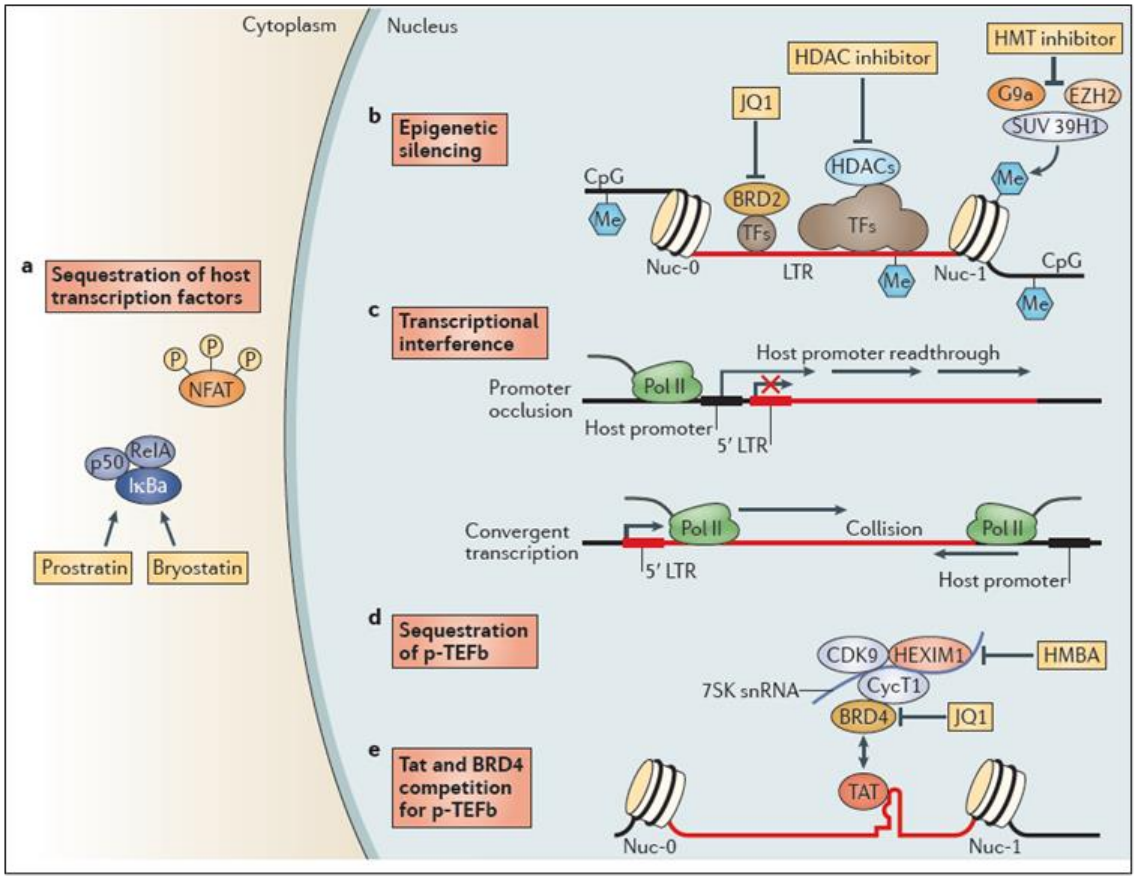


Figure 1.1: The deterministic mechanisms of HIV-1 latency. Various host-cellular factors decide viral entry to and exit from the latent state. Image is reproduced from doi: 10.1038/nrmicro3352 with permission from Springer Nature.

Several recent findings were contradictory to the deterministic model of latency-establishment. First, viral latency in Rhesus monkeys initiated during the 'eclipse phase' within 3 days post infection, long before any detectable viraemia (Whitney JB et al., 2014). Second, uniform activation stimuli to patient-derived resting $CD4^{+ve}$ lymphocytes did not necessarily reactivate all the latent, provirus (Ho YC et al., 2013; Weinberger AD and Weinberger LS., 2013). Third, in a cultured cell model, about half of the infected cells attained post-integration latency without any cellular relaxation (Dahabieh MS et al., 2013). Fourth, the intrinsic viral circuitry in the form of LTR-Tat-positive feedback is sufficient to perturb latency without an associated change in the cellular activation status (Weinberger LS et al., 2005; Burnett JC et al., 2009; Razooky BS et al., 2015). Given the assumption that latency has neither an evolutionary significance nor a selective advantage, the phenotype would have been lost by the natural selection over time and the Tat-feedback circuit would not evolve. Recently, a second group of scientists accounted for the above anomalies by proposing that viral latency is an intrinsic, stochastic property

of the viral life cycle regulated by the LTR-Tat positive feedback circuit independent of the host or environmental influence. A comparison between the salient features of the deterministic and stochastic model is depicted in Figure 1.2.

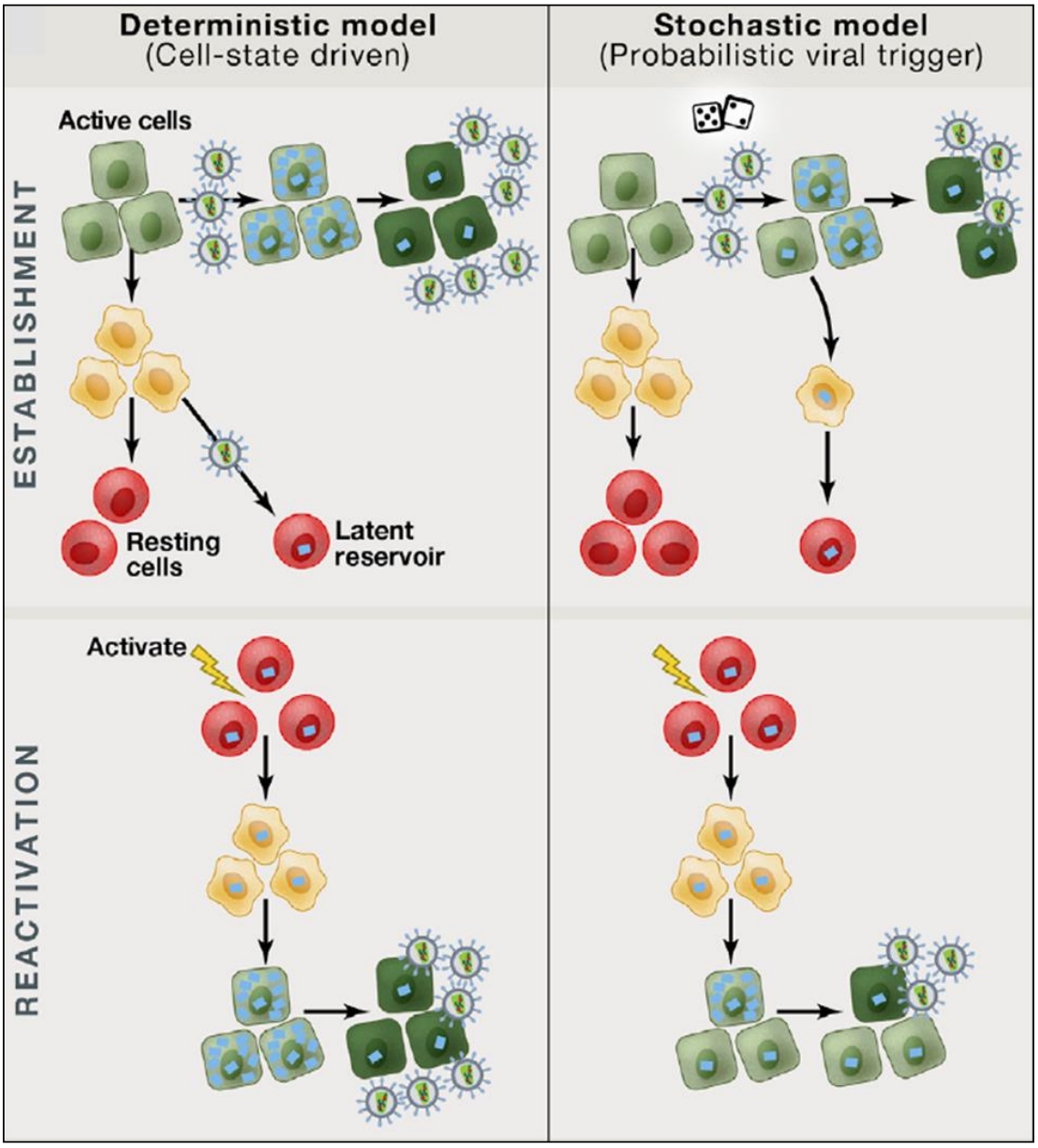


Figure 1.2: The deterministic vs stochastic model for HIV-1 latency. In the deterministic model (left), the activation status of the host-cell plays the key role in viral latency establishment as well as reactivation. In the stochastic model, the path to latency or productive infection is a chance event for every infected cell (right). Image is reproduced from doi: 10.1016/j.cell.2013.09.039 with permission from Elsevier Inc.

1.3.1 The LTR-Tat-feedback is the master regulatory-circuit for stochastic fate-selection in HIV-1

The 'stochastic hypothesis' for viral latency claims that the virus integrates two apparently paradoxical scenarios. On the one hand, viral gene expression is resistant to the global, environmental effects on the infected host-cell while on the other hand, the LTR-Tat positive feedback circuit is highly robust and sensitive to minute fluctuations in the Tat molecules to switch between the active or latent states (Pai A et al., 2017). The model elegantly explains the above paradox by providing a detailed account of the architecture and mode of action of the LTR-Tat feedback axis. The LTR-Tat positive feedback is considered the master regulatory circuit that decides between the active and latent viral phenotypes. Two distinct components are central to the functioning of the circuit- a weak, bursty promoter (5' LTR) and a strong Tat-positive feedback (Pai A et al., 2017). The LTR responds only weakly to the extracellular stimuli. For instance, TNFα, a strong inducer of the NF-κB pathway can enhance the promoter activity only by two folds while the positive Tat-feedback transactivates the LTR by >50 folds (Weinberger LS et al., 2005; Weinberger LS et al., 2008; Razooky BS et al., 2011; Karn J et al., 2012). The strong positive feedback supposedly masks the relatively weak effects of the external stimuli. The only factors that can significantly alter the overall output from the LTR are those that modulate or disrupt the Tat-feedback strength, the working principle for which has been used to design small molecules to eradicate latent reservoir (Mousseau G et al., 2015; Dar RD et al., 2015). All positive-feedback circuits are characterized by their ability to not only amplify the mean expression levels but also enhance the fluctuation levels (genetic noise). HIV-1 is no exception. The magnitude of genetic noise, combined with the cellular factors (such as activation status) at a given time has resulted in a totally probabilistic decision for HIV-1 active or latent phenotype (Singh A et al., 2010).

Tat is subject to mechanisms of internal, molecular switching in addition to the simple, positive feedback architecture. A recent report suggests that in order to maintain a predominantly latent state in adverse situations, Tat undergoes specific post-translational modifications. Enzymatic conversion of the acetylated Tat (Tat_A) to a deacetylated form (Tat_D) with a longer stability of the latter constitutes the feedback-resistor module that accounts for the robust functioning of the circuit (Weinberger LS and Shenk T., 2006).

1.3.2 Transcriptional master circuits in other viral families vs HIV-1

Transcriptional master circuits controlling the alternate replication fates are present in several families of both prokaryotic and eukaryotic viruses. In this context, the lysis-lysogeny decision in bacteriophage λ is the most researched phenomenon. The λ phage infects the bacterium *E coli* and may exist either as an integrated prophage in a lysogenic state with no viral gene expression or rapidly replicate to lyse the host cell. Fate selection in λ phage is governed by two key proteins- the CI or the λ repressor and the Cro or lytic activator, synthesized from the P_{RM} and P_R promoters respectively, in opposite orientations. Two essential properties of the λ circuit regulate the alternate phenotypes. (1) The Cro and CI proteins form a mutual repression circuit; over-expression of one inhibits the other but activates its own synthesis. Hence, both positive and negative feedback circuits are involved (Arkin A et al., 1998). (2) The ability of λ repressor to multimerize sets a functional threshold for its action. λ repressor proteins should octamerize to bind to the promoter-operator region and this cooperativity is central to the establishment of a 'bistable' circuit manifesting either lysis or lysogeny (Dodd IB et al., 2001).

A unique feature of some of the eukaryotic viral circuits is the rate-versus-level trade-off. Here, a rapid up-regulation of a viral protein is essential for an efficient viral replication but the same molecule is cytotoxic at saturating levels. This is the typical case of the CMV transcriptional circuit. The 86 kDa protein-immediate-early 2 (IE2) is a promiscuous transactivator of CMV which is essential for viral replication but at high levels might prematurely damage the cell even before an optimal production of viral particles (Dwarakanath RS et al., 2001; Sanders RL et al., 2008; Stinski MF et al., 2008). Hence, the CMV master regulatory circuit adopts means to quickly express the IE2 protein and at the same time restrict the physiological levels below the steady-state threshold (Stinski MF et al., 2008). This is achieved through a transcription accelerator circuit composed of an IE2 negative-feedback loop. Teng MW et al. decoded the transcription accelerator circuit using an integrated approach of computation and time-lapse experiments (Teng MW et al., 2012). The circuit was demonstrated to be highly cooperative with an estimated hill value (H) of ~7 for the negative feedback. Further, the exceptionally strong major immediate-early promoter (MIEP) is shown to be auto-

repressed by the IE2 through the 14 nucleotide *cis* repressor sequence (crs) (Macias MP et al., 1993).

Epstein Barr virus (EBV) and Herpes Simplex Virus-1 (HSV-1) exploit cooperativity of transactivator proteins as a regulatory parameter to control alternate replication fates. EBV encodes two transactivator proteins- Rta and Zta. Rta initiates an excitatory-feedback loop (Ragoczy T et al., 2001) and augments the expression of several other viral genes including the second viral transactivator protein Zta (Sarisky RT et al., 1996). Expression of Rta above a cooperative threshold disrupts EBV latency (Ragoczy T et al., 2001). Therefore, a latent state is maintained by a relatively stable Rta off state. In HSV-1, the ICP0 protein functions as the transactivator to reactivate the virus from latency and induce the lytic phase (Cai W et al., 1993). Multiple mechanisms permit ICP0 to transactivate its own promoter and activate viral gene expression (Roizman B et al., 2005). Evidences show that the LAT RNAs (anti-sense RNAs expressed in the latently infected cells and have the potential to bind ICP0 mRNA) might inhibit ICP0 post-transcriptionally (Kent JR et al., 2003).

Interestingly, the HIV-1 circuit appears to be significantly different from that of the λ phage or the above-mentioned eukaryotic viral circuits. First, there is no evidence of a repressor molecule or a negative feedback circuit involved in the case of HIV-1. Second, a single transactivator protein Tat participates in a positive feedback circuit. Further, the HIV-1 circuit lacks cooperativity and Tat functions as a monomer (Razooky BS et al., 2011) creating a void in understanding the mechanism of HIV-1 latency.

1.4 Experimental models of HIV-1 latency

HIV-1 latency research is curbed to a large extent due to the deficiency of a latency marker. Besides, the minute fraction of latently infected cells, approximately 1 in 10^6 circulating $CD4^{+ve}$ lymphocytes (Chun TW et al., 1997) and the high background of defective proviruses are additional obstacles in the study of HIV-1 latency. Nevertheless, sincere efforts have made possible the generation of multiple in-vitro and in-vivo latency models that recapitulate the phenomenon in the laboratory settings. Latency mechanisms have been investigated in both transformed as well as primary T-cell models; the latter arising from either (1) infected and activated $CD4^{+ve}$ T-cells during their transition to a

resting memory phenotype or (2) direct infection of the resting memory $CD4^{+ve}$ cells with the virus (Figure 1.3).

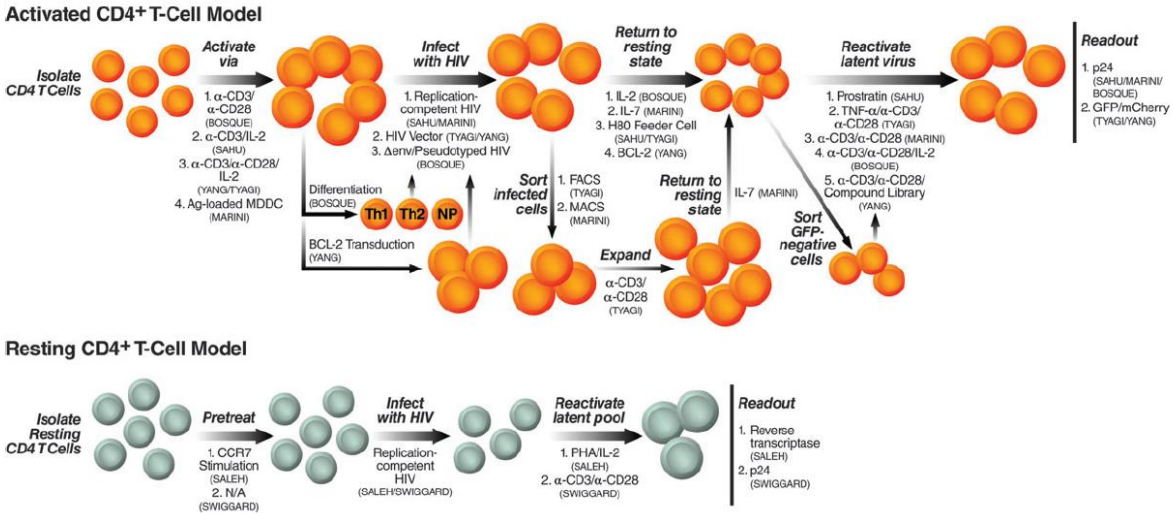


Figure 1.3: The primary cell-culture models for HIV-1 latency. A summary of the activated (above) and resting (below) $CD4^{+ve}$ T-cell models is presented. Image is reproduced from doi: 10.1111/j.1574-6976.2012.00335.x with permission from Oxford University Press.

1.4.1 Transformed T-cell line models

The ACH2 T-cell line (Folks TM et al., 1989) and the promonocytic U1 cell line (Folks TM et al., 1988) were among the earliest transformed T-cell line models for HIV-1 latency. Post-transcriptional regulation of latency was first confirmed in the above models by detecting singly as well as multiply-spliced HIV-1 RNA species in unstimulated cells with little or no trace of full-length genomic RNA (Pomerantz RJ et al., 1990). However, physiological relevance of these models are seriously questioned due to mutations in Tat or in its RNA target TAR in U1 (Emiliani S et al., 1998) and ACH2 (Emiliani S et al., 1996), respectively. Jordan et al recently developed a Jurkat cell-based in vitro cell line model termed J-Lat (Jordan A et al., 2003) (Figure 1.4) where the cells contain a stably integrated GFP-expressing viral construct (either a minimal virus or full-length molecular clones). GFP enabled sensitive, flow and microscopy-based detection of the promoter activity in real time; GFP^{+ve} and GFP^{-ve} phenotypes being indicative of the active and silent LTRs, respectively. Additionally, the model permitted for the first time fluorescence-based sorting of single cells to examine latency. Although the activated phenotype of the transformed cell-line models ensured easy infectivity and prolonged duration of the experimentation, the major drawback of the models lies in the inability of the cells to represent the resting phenotypes of the natural reservoirs.

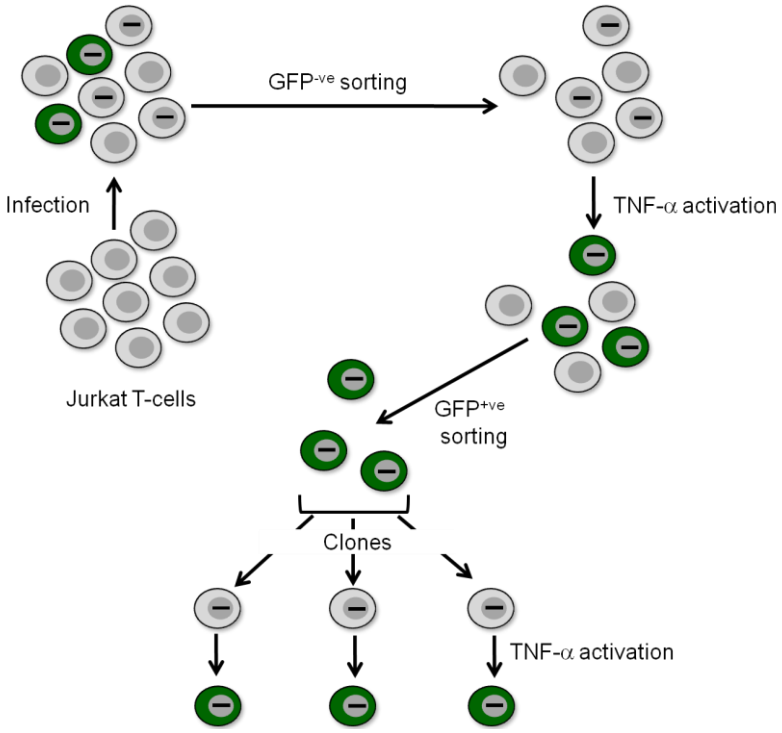


Figure 1.4: The J-Lat in vitro model for HIV-1 latency. Jurkat T-cells were infected with GFP-labelled HIV-1 particles co-expressing Tat. Following infection, a fraction of the infected cells were GFP^{+ve} while the others harboured silent provirus with no GFP expression. The GFP^{-ve} pool was sorted and treated with TNFα to induce the NF-κB-mediated gene expression from the reversibly silenced provirus. Single GFP^{+ve} cells were sorted to eventually establish clonal lines. Individual J-Lat clones expressed low GFP but could be stimulated to express high levels of GFP upon TNFα stimulation. A detailed account of the J-Lat model is available in Jordan A et al., 2003

1.4.2 Activated primary T-cell models

The first activated primary latency model was developed by Sahu GK et al. by isolating $CD4^{+ve}$ T-cells from healthy donors, activating them using immobilized αCD3 antibodies in the continued presence of IL-2 followed by infection with a replication-competent virus. Latency was established during the transition to a quiescent, memory phenotype and promoted by a brain tumor derived feeder cell line, H80 through an unknown mechanism. Although 5% infected cells attained latency this way, a significant fraction still retained the active phenotype ($CD69^{+ve}$ and low p24 signals) (Sahu GK et al., 2006). Tyagi et al used αCD3/αCD28 antibodies to activate the primary T cells in the presence of IL-2 before infecting with a Δgag HIV-1 vector that was also mutated for Tat and encoded a GFP-ORF in the place of Nef. Infected cells were sorted, expanded in αCD3/αCD28/IL-2 containing medium and co-cultured with the H80 feeder cells. The latently infected $CD4^{+ve}$ T-cells in this model were mostly of the central memory

phenotype with pronounced GFP expression upon CD3/CD28 stimulation. The model identified important aspects of epigenetic silencing and the role of the transcription elongation factor, P-TEFb in HIV latency (Tyagi M et al., 2010). A model closely representing the natural context was developed from activated healthy $CD4^{+ve}$ T-cells by co-culturing these cells with antigen-loaded monocyte-derived dendritic cells (Ag-MDDCs) before infecting the cells with a replication-competent HIV-1. The infected and activated cells were maintained for 4 weeks in IL-2 to obtain a large fraction of central memory T-cells ($CD25^{-ve}$) but with incomplete quiescence (Marini A et al., 2008).

An important latency model in the activated T-cell category was generated by Bosque et al. (Bosque A et al., 2009) which produced both central as well as effector memory cells. The activated $CD4^{+ve}$ T-cells were cultured for several days in three different conditions to produce Th1, Th2 and non-polarized $CD4^{+ve}$ T-cells (Messi M et al., 2003). While the Th1 and Th2 populations closely resembled both T_{EM} and T_{CM} in vivo, the nonpolarized population resembled the T_{CM} alone. An env-deficient, pseudotyped-HIV-1 was used to infect all the three cell types. That Lck and NFAT but not NF-κB was critical for latency reactivation in memory T-cells (Bosque A et al., 2009) was revealed in this model. Yang et al. developed a latency model in primary $CD4^{+ve}$ cells which were previously exposed to Bcl-2, an anti-apoptotic protein that is expected to prolong the lives of the infected cells without a requirement for feeder cells during their transition to the resting phenotype. The model helped evaluate the potential of a combination of drugs (~ 4000) in purging the latent virus (Yang Z et al., 2009).

1.4.3 Resting primary T-cell models

Direct infection of resting $CD4^{+ve}$ T-cells without prior stimulation was first demonstrated by Swiggard et al. (Swiggard WJ et al., 2005). A pool of naïve, freshly isolated T_{CM} and T_{EM} were infected with a replication-competent HIV-1 by a novel centrifugation technique termed spinoculation. Integration in these cells occurred at a much lower frequency than activated $CD4^{+ve}$ T-cells. After 3 days of infection, 4.5% cells expressed HIV-1 gag upon αCD3/αCD28 induction showing that a fraction of cells were latently infected in this model by direct spinoculation with the viral supernatant. A second resting T-cell model was generated by Saleh et al. based on a previous observation that majority

of the HIV-1 infected resting $CD4^{+ve}$ T-cells express CCR7, a lymphoid organ homing receptor. Freshly isolated resting $CD4^{+ve}$ lymphocytes were first stimulated with the CCR7 ligands- CCL19 and CCL21. These chemokines increased the susceptibility of resting $CD4^{+ve}$ T-cells to infection by a replication-competent HIV-1 virus without inducing either CD69 or CD25 expression. Both reverse transcripts and integrated provirus were detected in these cells showing the prevalence of latent provirus (Saleh S et al., 2007).

1.4.4 Animal models

Zack and colleagues generated a humanized mouse model (Brooks DG et al., 2001) by co-transplanting human fetal thymus and liver tissue into congenitally immunodeficient C.B-17 *scid/scid* mice deficient in functional T and B cells (McCune JM et al., 1988). HIV-1 was directly injected into the Thy/Liv implants. This led to a high frequency of naïve $CD4^{+ve}$ T-cells being latently infected during thymopoiesis. The reactivation properties of prostatin, IL-7 and a gp120-targeted immunotoxin with minimal influence on the cellular activation or proliferation status were first investigated in the model (Korin YD et al., 2002; Scripture-Adams DD et al., 2002; Brooks et al., 2003). Engraftment of the human immune system was later significantly improved in the humanized NSG (NOD SCID gamma) mice and a highly robust humanized BLT (bone marrow–liver–thymus) mouse model was generated from the NSG mice (Lan P et al., 2006; Melkus MW et al., 2006; Tonomura N et al., 2008) and extensively used in pre-clinical trials. Recently, the model demonstrated the longitudinal, distribution pattern of residual viraemia during HAART in a comprehensive panel of anatomic sites (Denton PW et al., 2014).

Non-human primate (NHP) models generated by infecting macaques with SIV or its recombinant chimera strains such as RT-SHIV and SHIV offer excellent tools to evaluate several high-risk experimental strategies that may not be permitted in human trials for ethical reasons. The kinetics of disease progression in these models closely resembles that of humans. Using two RT inhibitors, the plasma viraemia in the SIV-macaque model of Shen et al. could be reduced to undetectable levels and the condition was then used to measure latent virus in blood, spleen, thymus and lymph nodes by different methods (Shen A et al., 2003). Animal models for HIV-1 latency closely resemble viral dissemination pattern in ART treated individuals. The NHP models have significantly

helped our understanding of the human immune responses against HIV-1 infection (Schmitz JE et al., 1999). They hold future promise to develop strategies that would eliminate latent viral reservoir.

1.5 Strategies to detect viral latency

Precise detection of virus in latently infected cells involves three different strategies: (1) purification of $CD4^{+ve}$ cells, (2) estimation of integrated HIV-1 DNA in the isolated cells and (3) evaluation of the replication-competence of the latent provirus using latency-reversing agents (LRA). A practical limitation of the three-step approach is the mutually exclusive nature of the proviral load and replication-competence. Once the cell is lysed to measure the integrated provirus distinct from the unintegrated DNA, the possibility of viral reactivation is lost. Again, replication-competence measurement by LRA-mediated viral reactivation eliminates the feasibility of obtaining the integration frequency. In a recent study, $CD4^{+ve}$ cells were purified from patient blood, treated with a cocktail of three RT inhibitors and an integrase inhibitor to block any further reverse transcription or integration of the circular HIV-1 DNA and then, exposed to drugs for viral reactivation. This was followed by the quantification of RNA from the cell-supernatant using RT-PCR. The method identified only the proviral sequences that were integrated, intact and replication-competent, thereby overcoming the above practical limitations (Monie D et al., 2005). Ho et al. precisely quantified the sizes of defective, intact inducible and intact non-inducible proviral genomes in individual patients by reconstructing near full-length clones from each fraction and measuring their growth kinetics in vitro (Ho YC et al., 2013). Nevertheless, the traditional three-step approach has significantly contributed to the reservoir-size determination, delineation of the persistence mechanisms, and development of novel eradication strategies. Assays pertaining to each step are briefly discussed as follows:-

1.5.1 Purification of resting $CD4^{+ve}$ T-cells

For the examination of HIV-1 latency, resting $CD4^{+ve}$ T-cells harboring latent proviral genomes should be enriched from a complex mixture of peripheral blood cells or

lymphoid tissues of the patients. In general, flow-cytometry-based separation of cells using cellular activation markers is a robust and widely accepted method that allows simultaneous gating on multiple subsets. In this context, CD69, CD25 and human leukocyte antigen (HLA-DR) are used as markers of early, intermediate and late stages of T-cell activation, respectively (Hermankova M et al., 2003).

1.5.2 Estimation of proviral integration in resting $CD4^{+ve}$ T-cells

Majority of the HIV-1 DNA in the natural infections exist in an unintegrated form either as the linear viral cDNAs or as the unintegrated 1-LTR or 2-LTR circles. The challenge is to differentiate and specifically quantitate the chromatinized and replication-competent HIV-1 DNA from the unintegrated forms. Regular PCRs fall short in the above precision (Bukrinsky MI et al., 1991). Two advanced PCR-based assays are routinely used to capture the integrated DNA. The first is an inverse-PCR assay- the original method to identify the presence of latent infection in the resting $CD4^{+ve}$ T-cells in vivo (Chun TW et al., 1997). Genomic DNA from the patient $CD4^{+ve}$ cells is digested with a 4-bp overhang enzyme and permitted to self-ligation under the conditions of limiting dilution followed by PCR using outward-directed primers against HIV-1 sequence. The amplicon identifies the chromatin-provirus junctions. This technique was subsequently employed to determine the integration site of the latent provirus and to demonstrate the introns of transcriptionally active genes as the sites of preferential integration (Han Y et al., 2004). The second approach, the Alu PCR, amplifies genomic DNA with one primer against HIV-1 DNA and the other against the retro-element Alu present pervasively in the host genome in multiple copies (Butler SL et al., 2001; O'doherty U et al., 2002; Brussel A et al., 2005; Yamamoto N et al., 2006). Both the techniques underestimate the actual integration frequencies since the amplicon size in either case varies with the distance between the integration site and the nearest restriction site (inverse PCR) or the nearest Alu element (Alu PCR). Droplet digital PCR (ddPCR) enhances precision and sensitivity by partitioning the template into thousands of droplets such that each droplet bears a single target DNA. A ddPCR that identifies latent HIV-1 integration was reported (Strain MC et al., 2013).

1.5.3 Replication-competence determination of the latent virus

Integrated HIV-1 DNA measured by the above methods is ~2 logs higher than the equivalent viral particles released by reactivation. Thus, only a minor fraction of the proviral sequences are intact, and represent the replication-competent viral genomes. Till date the **enhanced culture assay** or the **quantitative viral outgrowth assay** (QVOA) remains the gold standard for measuring replication-competent virus. To this end, infected resting $CD4^{+ve}$ T-cells are isolated, limit diluted, induced with suitable mitogens such as PHA and irradiated allogenic PBMCs or αCD3/αCD28 and measured for viral release by p24 ELISA (Chun TW et al., 1997) (Figure 1.5).

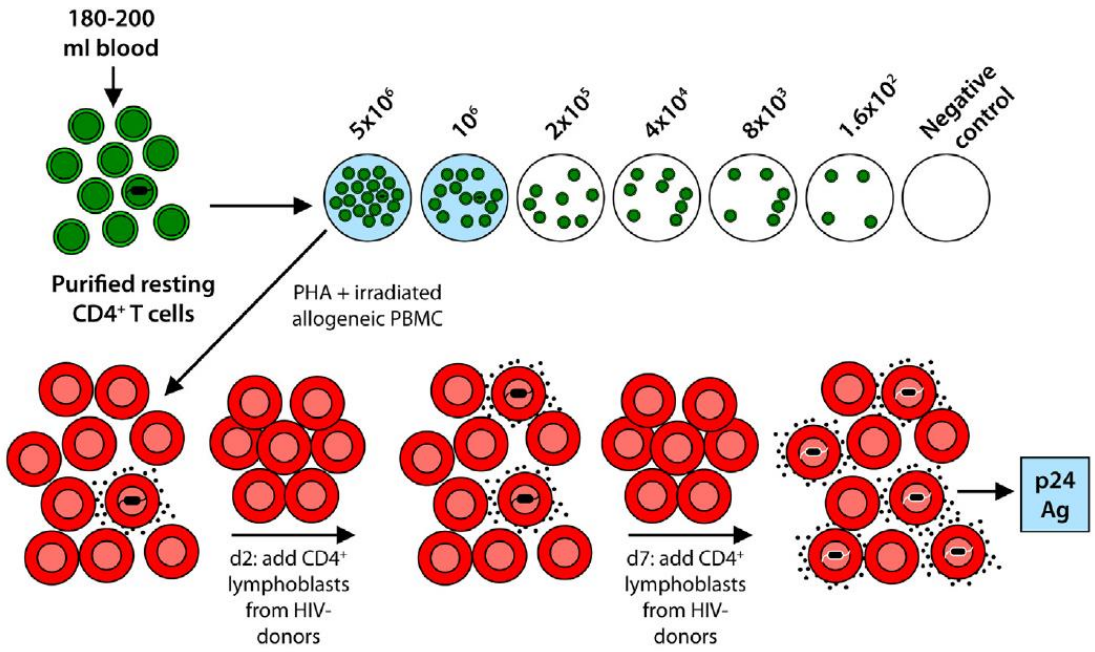


Figure 1.5: Schematic representation of the viral outgrowth assay (VOA). Image is reproduced from doi: 10.1016/j.jaci.2014.05.026 with permission from Elsevier Inc.

A few significant limitations of the QVOA are noted. Firstly, to determine the tiny fraction of latent cells, a large volume of blood (~ 200 ml) is required from each subject. Secondly, the p24 released by a single latently infected cell is often below the detection level hence, the viral particles should be amplified by co-culturing with permissive cells such as CD8-depleted PBMCs from healthy donors. Thirdly, donor PBMCs widely differ in their infection potential. To overcome the above limitations, several innovations of QVOA are in progress towards making it highly robust, reproducible, less time- and labour-intensive and more practically reliable (reviewed in Norton NJ et al., 2017).A

novel method termed Tat/rev Induced Limiting Dilution Assay (TILDA) measures the frequency of latently infected cells with a precision intermediate between the DNA-PCRs and the QVOA. The assay targets the cells with inducible, multiply-spliced HIV-1 RNAs since, these species are absent in latently infected cells and synthesised only upon reactivation. TILDA reported the median frequency of latent cells in ART-treated subjects to be 24 cells/million which was 48 times more and 6-7 times less than the frequencies estimated by the QVOA and PCR-based assays, respectively (Procopio FA et al., 2015).

1.6 The HIV-1 promoter: the 5' Long Terminal Repeat (5' LTR)

The ~ 9 kb HIV-1 genome contains two identical long terminal repeats (LTR) at either ends. While the 5' LTR acts as the viral promoter, the 3' LTR participates in the cleavage and polyadenylation of the full-length viral mRNA essentially functioning as the transcription terminator. Interestingly, both the 5' and 3' LTRs are Tat-inducible and can initiate viral transcription and also transcribe the TAR regions. The 3' LTR might even transcribe in the absence of 5' LTR (Klaver B and Berkhout B., 1994). The exact mechanism(s) that account for the preferential usage of 5' LTR in HIV-1 for transcription still remains to be deciphered. A diverse group of transcription factors simultaneously occupies the multiple transcription factor binding sites (TFBS) at the LTR to ensure efficient transcription initiation. The LTR is divided into the U3, R and U5 regions (Figure 1.6); U3 and U5 elements being exclusive at the 3' and 5' ends of the viral RNA genome, respectively. During reverse transcription, they are reconstituted into the complete U3-R-U5 configuration generating two identical LTRs at the ends of the proviral DNA. The U3 element which is the most populated region with TFBS is further segmented into the modulatory (nt -454 to -104), the enhancer (nt -105 to -79) and the core-promoter regions (nt -78 to -1) (Pereira LA et al., 2000). Each U3 sub-segment bears landmark TFBS which are broadly conserved across the HIV-1 subtypes.

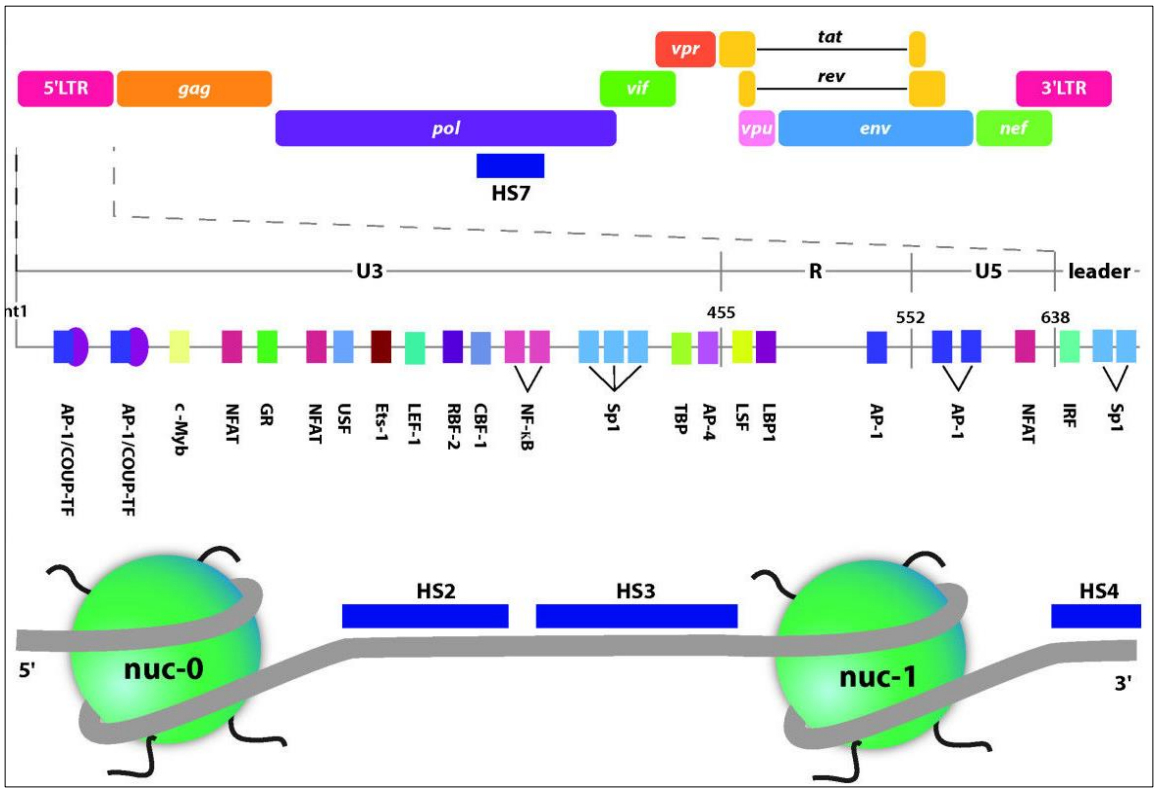


Figure 1.6: The organization of the HIV-1 LTR. The viral promoter or the 5' LTR is divided into the U3, R and U5 segments. A wide array of transcription factor binding sites (TFBS) spans the LTR, particularly the U3 region. Specific positions of nucleosomes and the hypersensitive sites (HS) are indicated. Image is reproduced from doi: 10.1186/1742-4690-6-111 under appropriate licensing terms of BioMed Central Ltd.

1.6.1 The modulatory region

The modulatory region houses binding sites for CCAAT/enhancer-binding protein (C/EBP), activating transcription factor/cyclic AMP response element-binding protein (ATF/CREB), lymphocyte enhancer factor (LEF-1), and nuclear factor of activated T-cells (NFAT) (Shaw JP et al., 1988; Tesmer VM et al., 1993; Henderson AJ et al., 1995; Krebs FB et al., 1998). C/EBP sites are particularly important in inducing HIV-1 transcription in the monocyte-macrophage lineage and less in the $CD4^{+ve}$ T-cells (Henderson AJ et al., 1997). A negative regulatory element (NRE; nt -340 to -184) has been identified within the modulatory region; deletions within the element enhanced viral transcription and replication (Rosen CA et al., 1985; Siekevitz M et al., 1987). Further, multiple host factors such as cMyb, NFAT, COUP and USF have been proposed to function at the modulatory region (Shaw JP et al., 1988; Dasgupta P et al., 1990; Cooney AJ et al., 1991; Giacca M et al., 1992).

1.6.2 The enhancer

The enhancer element primarily consists of the binding sites for NF-κB and related proteins. The tandemly repeated, 10 nucleotides long NF-κB motifs are central to the canonical and non-canonical NF-κB signaling and often converge with the protein kinase C (PKC) pathway during T-cell activation. NF-κB proteins not only initiate transcription, but also promote elongation of viral transcription in a Tat-independent manner. Elimination of the NF-κB motifs significantly reduced transcription initiation as well as Tat mediated transactivation (Nabel G and Baltimore D et al., 1987; Zeichner SL et al., 1991; West MJ et al., 2001).

1.6.3 The core-promoter

The core-promoter marks the presence of the TATAA box, located 29–24 nucleotides upstream of the transcriptional start site and the specificity protein (Sp) binding sites, which are three tandem GC-rich binding sites interacting with transcription factors- Sp1 through Sp4. The TATAA box binds TATAA-binding protein and a number of other factors comprising the RNA polymerase II (pol II) transcription complex. Studies have shown that while the Sp1 and Sp4 are potential activators of HIV-1 transcription, a truncated version of Sp3 might repress viral gene expression (Jones KA et al., 1886; Jones KA et al., 1994; Majello B et al., 1994).

1.6.4 Nucleosome occupancy at the LTR

DNase-1 hypersensitivity and micrococcal nuclease assays have revealed the maps of nucleosome assemblies at the 5'LTR. Two DNase-1 hypersensitive sites within the 5'LTR- nt +223 to +325 and nt +390 to +449 have been identified in chronically infected $CD4^{+ve}$ T lymphocytes as well as monocyte/macrophage cell lines (Verdin E et al., 1991). Further, studies indicate that two nucleosomes nuc-0 and nuc-1 are precisely organized on the integrated viral promoter and their positions are spatially and temporally altered depending on the transcriptional status of the LTR. In a transcriptionally silent state, nuc-0 is situated from nt +40 to +200 and nuc-1 is positioned from nt +465 to +610 (Verdin E et al., 1993). It is proposed that displacement of nuc-1 by chromatin remodelers or viral

protein mediated mechanisms is a prerequisite for HIV-1 transcription as observed in response to T-cell activation stimuli (Rafati H et al., 2011).The nucleosome organization at the HIV-1 LTR is presented (Figure 1.6).

1.6.5 Subtype-specific LTR polymorphism and variations within the HIV-1C enhancer

About 90% of the HIV-1 infections belong to the major group (Group M) of HIV-1 which is further classified into at least ten distinct subtypes or clades termed A through J, and several circulating recombinant forms (CRFs) based on their genomic sequences. Recent studies identified that the most prevalent HIV-1 subtypes are A, B, and C, with subtype C accounting for almost 50% of all HIV-1 infections worldwide. Further, the geographical distribution of the subtypes is not uniform which could be the result of accidental trafficking leading to a 'founder effect' or a prevalent route of transmission. The prevalence of subtype A infection is higher in the areas of central and eastern Africa (Kenya, Uganda, Tanzania, and Rwanda), and in Eastern Europe. Subtype B is predominant in western and central Europe, the Americas, the Australia, countries of Southeast Asia, northern Africa and the Middle East. Subtype C viruses are wide-spread in the countries with more than 80% of all global HIV-1 infections, such as South Africa and India (Myers G et al., 1994; Buonaguro L et al., 2007; Hemelaar J et al., 2012).

Extensive genetic and functional characterization of the TFBS and other regulatory elements in the LTR have identified numerous subtype-specific differences within the binding sites of NF-κB, NF-AT, USF, the TATA box, and the TAR region (Montano MA et al., 1997; Naghavi MH et al., 1999; Jeening RE et al., 2000). However, a major hotspot for polymorphism within the C-LTR itself is the enhancer region comprising of the NF-κB motifs (Figure 1.7). In mammals, the NF-κB family of transcription factors comprises of five members namely p65 (RelA), RelB, c-Rel, p50/p105 (NF-κB1) and p52/p100 (NF-κB2) that assemble as homo or hetero-dimers to either activate or repress the transcription of cellular genes (Baldwin Jr AS 1996; Ghosh S et al., 1998). All the five members share a conserved, approximately 300 amino acid long, Rel-homology domain (RHD) at the amino terminal necessary for dimerization of the proteins, DNA binding, and nuclear transport. The NF-κB dimmers recognize and bind a 10-bp DNA consensus

sequence- 5′-GGGPuNNPyPyCC-3′ (Ghosh G et al 1995; Müller CW et al 1995; Chen FE et al 1998).

All the HIV-1 subfamilies possess two genetically identical NF-κB motifs (5'-GGGACTTTCC-3'; referred to here as the H-κB motif) with the exception of subtypes E and C. While a single H-κB motif is present in the E-LTR (Ranjbar S et al., 2006; Roof P et al., 2002), a large proportion of C-LTRs contains three or four NF-κB sites (Bachu M et al., 2012a; Bachu M et al. 2012b; Bjorndal A et al., 1999; Munkanta M et al., 2005; Novitsky V et al., 2002). The additional κB motifs in the C-LTR are also genetically divergent and unique for subtype C. We refer to the third NF-κB site as the C-κB motif (5'-GGGGCGTTCC-3') characterized by the variations at the 4th and 6th nucletotide residues and the fourth NF-κB motif as the F-κB motif (5'-GGGACTTTCT-3'), which differs from the canonical H-κB motif by only the last residue at position 10 (C to T).

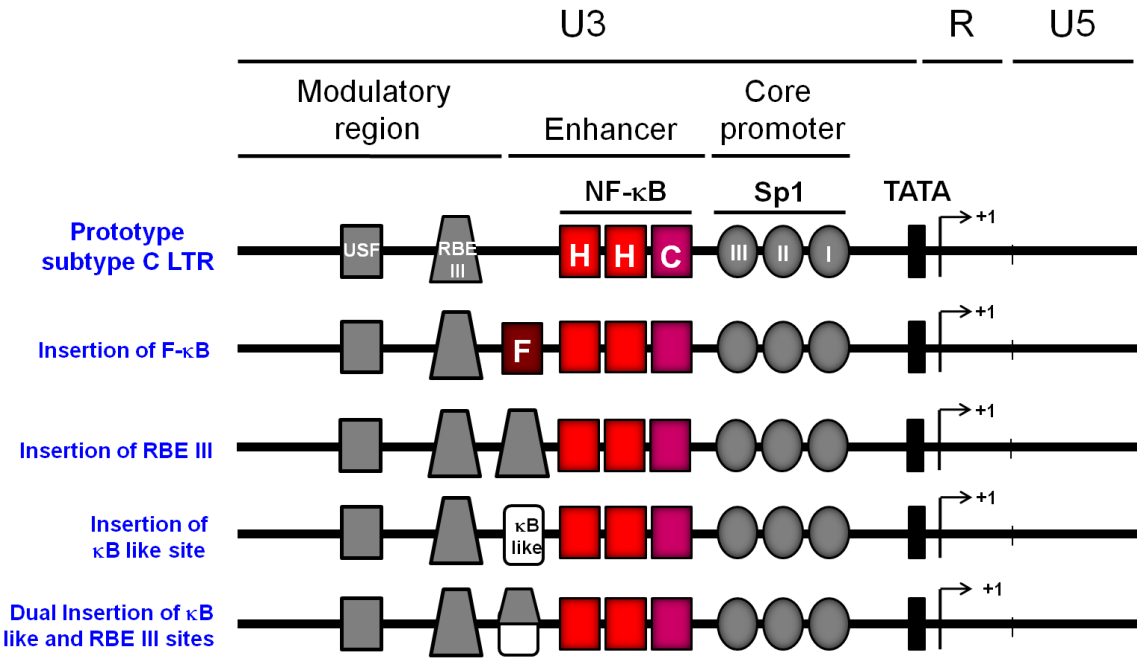


Figure 1.7: The enhancer is the hotspot of TFBS polymorphism in HIV-1C. The top panel represents a typical C-LTR highlighting the important TFBS including the RBEIII, NF-κB, Sp1, TATA box and the USF-1 site. A canonical C-enhancer houses three NF-κB motifs, two identical, HIV-specific H-κB sites and an additional C-κB site, unique to subtype C. The arrow marks the direction and the start site of viral transcription. The sequences obtained from the enhancer region of 25 Indian and 4 African HIV-1C isolates showed four distinct variations with respect to TFBS insertions (Bachu et al., 2012a) comprising of a fourth NF-κB (F-kB), a κB-like, a RBEIII or both kB-like and RBEIII motifs.

1.7 HIV-1 Tat and its influence on latency

HIV-1 Tat (transactivator of transcription) and its role in viral latency have been extensively researched. Tat is a small regulatory protein comprising of 86 to 101 amino acids and varying in molecular weight from 14 to 16 kDa (Ruben S et al., 1989). The ORF of Tat is made up of two exons and structurally divided into six functional domains (Figure 1.8). Tat is a versatile protein with multiple functions. Its primary role is to enhance viral transcription through interaction with the transactivation response (TAR) element and the P-TEFb complex followed by chromatin remodeling and phosphorylation of the stalled RNA Pol II (Yedavalli VS et al., 2003; Wong K et al., 2005; Ammosova T et al., 2006). Other activities of Tat include facilitation of reverse transcription (Liang C and Wainberg MA, 2002), immune suppression (Gupta S et al., 2008), apoptosis induction (Kim TA et al., 2003; Miura Y et al., 2003; Giacca M et al., 2005; Poggi A and Zocchi MR, 2006) and contribution in HIV-associated dementia (HAD) (Cheng J et al., 1997; Pocernich CB et al., 2005; Haughey NJ and Mattson MP., 2012).

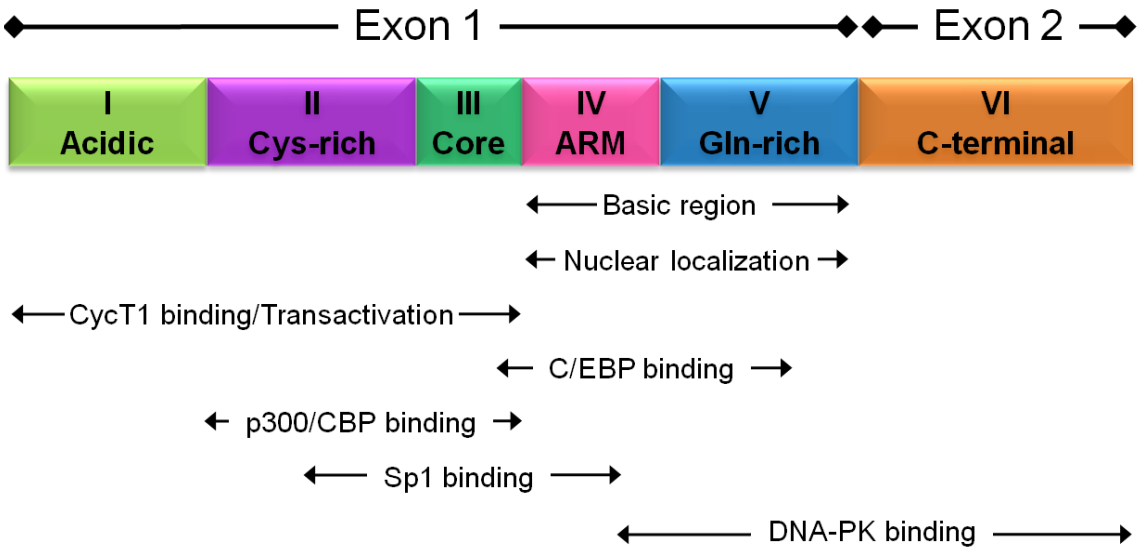


Figure 1.8: Diagrammatic representation of the structure of HIV-1 Tat. Structural domains including the two exons and the six major functional domains of a typical HIV-1 Tat protein are indicated. The binding sites for important cellular co-factors and their positions are marked (Li L et al., 2012).

HIV-1 Tat is known to regulate several host genes. Recently, a comprehensive, genome-wide analysis in T-cells revealed that HIV-1 Tat occupies ~3,000 locations on the host chromatin and almost 2,000 host genes are differentially regulated by Tat, a quarter of

which are direct targets comprised of both activated and repressed genes in the presence of Tat (Reeder JE et al., 2015). Interestingly, the authors identified that unlike the Tat-TAR binding in the case of viral transcription, Tat uses unique mechanisms to regulate host gene transcription, both at the levels of initiation and elongation. To up-regulate initiation of host genes, Tat recruits RNA Pol II at the site of transcription by DNA looping while in the case of elongation, Tat either enhances or restricts P-TEFb recruitment through interaction with master transcriptional regulators bound at specific promoters and enhancers (Reeder JE et al., 2015).

1.7.1 The role of Tat in the establishment of latency

Tat is predominantly a negative inducer of HIV-1 latency as hypothesized by the deterministic model of viral latency (Section 1.3). Limited physiological Tat or its inactivation owing to the host-cell relaxation result in short, non-polyadenylated transcripts typically <100 bp thereby restricting full-length transcription beyond the TAR hairpin loop (Feng S et al., 1988; Roy S et al., 1990; Yedavalli VS et al., 2003). As a general notion, when present in sufficient quantities, Tat abrogates latency establishment (Pearson R et al., 2008; Donahue DA et al., 2012). A near permanent latent state could be induced in primary, latently infected $CD4^{+ve}$ lymphocytes from ART-treated individuals as well as in cultured cell-line models such as J-Lat, promyelocytic OM-10.1 and HeLa-CD4 using the Tat-inhibitor didehydro-cortistatin A (dCA). Further, it was demonstrated that dCA could inhibit viral reactivation from latency using several prominent LRAs such as HDAC inhibitors, PKC activators, cytokines as well as upon antigenic stimulation of latently infected primary cells long after the withdrawal of dCA (Mousseau G et al., 2015). Using in silico as well as experimental approaches, the stochastic latency model proposes that a random fluctuation in the Tat levels is the sole driving force towards latency-establishment or latency-reversal in a probabilistic manner (Weinberger LS et al., 2005). Further experiments in this direction revealed that the virus might play additional tricks to modulate the noisy Tat expression by hijacking the host transcription factors to recruit them at its own LTR (Burnett JC et al., 2009).

The role of Tat in actively silencing the viral promoter has been significantly overlooked. A single group demonstrated that viral latency could also initiate in the continued

presence of Tat. The study showed an initial, rapid transactivation of viral genes followed by gradual LTR-repression while Tat was still present in the system (Drysdale CM et al., 1991). Using a novel cell-fusion assay, the authors analyzed both the Tat-induced viral gene expression levels and the associated cellular responses. A temporal kinetic analysis showed a rapid mRNA accumulation from the HIV-1 LTR (within 6 h) with subsequent decline in the transcript levels post-Tat induction. Control genes such as *neo* or actin were not subjected to such a sequential enhancement and down-regulation by Tat. The authors attribute several molecular mechanisms to explain the coupled phenomena of initial transactivation and subsequent repression of LTR in the continued presence of Tat. First, a Tat-induced, inhibitory protein may be synthesized by the de novo pathway rather than the salvage route for the late repression of LTR. Second, a post-transcriptional Tat inhibition may be central to the phenomenon as indicated by reduced mRNA levels after 9 h of Tat-induction while the translational efficiency continued to increase even after 24 h. Third, the authors propose a model where Tat as a positive factor continually shifts the equilibrium towards enhanced viral transcription. At saturating Tat levels, a slight disturbance in the equilibrium favoring Tat over negative-cellular factors may rapidly reverse the phenomenon resulting in gene-repression (Drysdale CM et al., 1991). However, most of the assays were performed by transient transfection of plasmids containing LTR and a downstream reporter gene. Validation of the above experiments with integrated proviral DNA and in suitable target cells would provide better insights into the phenomenon.

1.7.2 The role of Tat in the reversal of latency

Evidence shows that Tat is sufficient to initiate viral transcription in patient-derived latent cells (Lin X et al., 2003; Lassen KG et al., 2006) and in Jurkat cell line-based models (Donahue DA et al., 2012) without the need of cellular activation. A combined experimental and computational approach by Razooky BS et al. demonstrated that Tat is more efficient than T-cell activating agents in reactivating full length, latent provirus (Razooky BS et al., 2015). In the same study, the authors show that deprivation of activating stimuli resulted in marked reduction of activation markers in primary infected T-cells but did not affect the molecular levels of Tat or the viral transcriptional activity (measured by GFP expression). Of the six functional domains, domains II and III

spanning the amino acid residues 22 to 48 are reported to play active roles in transactivation (Kuppuswamy M et al., 1989; Ruben S et al., 1989). In a recent mutational study, the authors generated an attenuated Tat protein Tat-R5M4 comprising of V36A, Q66A, V67A, S66A, and S77A mutations and demonstrated its enhanced reactivation-potential compared to wild-type Tat, albeit with limited toxicity and immunogenicity. The variant Tat also triggered efficient transcription in combination with HDACi from a broad range of replication-competent viral isolates holding promise as a potential LRA (Geng G et al., 2016).

Mechanism of Tat-mediated latency reactivation and transactivation. During viral latency, the proviral LTR is sequestered at the Nuc-1 and characterized by the presence of HDACs, HMTs and their associated modifications. Further repression of the LTR occurs as a result of the histone methylation caused by the polycomb repressing complex 2 (PRC2) (Friedman J et al., 2011). At this stage, there is a low level of transcription initiation but severely compromised elongation. The negative elongation factor (NELF) complex recruited at the stalled RNA Pol II forces the generation of abortive transcripts (Yamaguchi Y et al., 2002; Zhang Z et al., 2007) (Figure 1.9). Upon suitable stimulation, NF-κB members, primarily the p50-p65 heterodimer, initiate viral transcription by displacing the repressive HMTs with activating histone acetyltransferases (HATs) and other chromatin remodelling complexes (Sheppard KA et al., 1999). This pre-initiation complex reverses chromatin restrictions at the LTR. Following transcription through the TAR element, the Tat/P-TEFb complex (comprising of CDK9 and CycT1) interacts with the TAR RNA and recruits other accessory elongation factors including ELL2 to the elongation complex. This association activates the CDK9 kinase to hyperphosphorylate the CTD of RNA polymerase II (Isel C and Karn J., 1999; Kim YK et al., 2002), Spt5, and NELF-E. Phosphorylated NELF-E is released from the RNA Pol II (Fujinaga K., et al., 2004) and hyperphosphorylated RNA Pol II (Ramanathan Y et al., 2001) and Spt5 (Ivanov D et al., 2000; Bourgeois CF et al., 2002) allow enhanced transcriptional elongation of the full HIV-1 genome thereby initiating a positive Tat-feedback cascade. During quiescence, high levels of transcriptionally inactive P-TEFb predominate in the cytoplasm in a snRNP complex containing 7SKRNA, HEXIM, and the RNA binding proteins MePCE and LARP7 (Nguyen VT et al., 2001; Yang Z et al., 2001; Yik JH et al.,

2003., He N et al., 2008; Krueger BJ et al., 2010; Sobhian B et al., 2010). Increased Tat concentrations displace HEXIM to form a stable complex with P-TEFb (Figure 1.9).

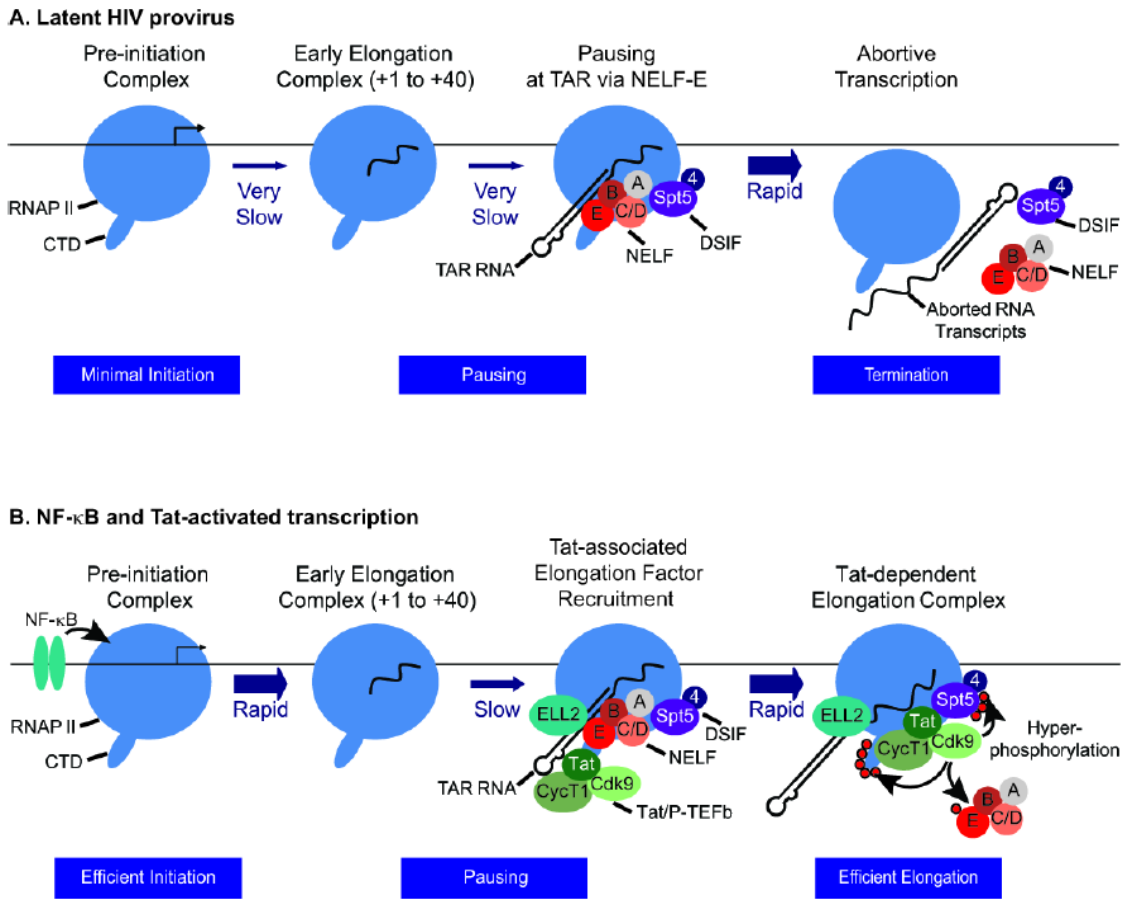


Figure 1.9: Tat-mediated latency reversal. The sequence of events in HIV-1 transcription in the absence (top panel) or the presence (bottom panel) of Tat is depicted. The mechanism of Tat-driven transcription elongation is detailed in the section 1.7.2. Image is reproduced from doi: 10.1097/COH.0b013e328340ffbb with permission from Wolters Kluwer Health, Inc.

1.7.3 The influence of post-translational modifications (PTMs) of Tat on viral transcription and latency

The first exon of the Tat-ORF comprising of the initial 72 amino acid residues is sufficient for its transactivation property (Garcia JA et al., 1988; Kuppuswamy M et al., 1989). Multiple conserved residues in the cysteine- and arginine-rich domains within the Tat-exon1 enable its role as a cofactor for various enzymatic processes in the cell. This results in specific post-translational modifications of these residues with significant consequences to viral gene-expression and latency. A few post-translational modifications

of Tat with important functional consequences to HIV-1 transcription are depicted (Table 1.1).

Table 1.1 Tat-PTMs affecting viral transcription and silencing

Specific Modification	Enzyme	Effect of modification
K50/51 acetylation	• KAT3B (p300/CBP) • KAT2A (GCN5)	• Essential for viral transactivation. • K50-acetylated Tat can no longer form ternary complex with TAR and cyclin T1. • Dissociates Tat from TAR and P-TEFb for subsequent interaction with PCAF bromodomain and elongating RNA Pol II complex.
K28 acetylation	PCAF/KAT2B	• Enhances viral transcription. • Promotes interaction of Tat with cyclin T1 and PCAF
K50 deacetylation	SIRT1	• Enables recycling of Tat for fresh transactivation. • Promotes physical interaction of Tat and PCAF during late transcription.
K71 ubiquitination	Hdm2	• Enhances viral transcription. • Possibly promotes interaction of Tat with cyclin T1 and TAR RNA at the initiation of transcription cycle.
S16 phosphorylation S46 phosphorylation	CDK2/cyclin E complex	• Enhances Tat transcriptional activity.
S62 phosphorylation S68 phosphorylation	PKR	• Promotes Tat-TAR interaction and enhances transactivation.
K50 methylation K51 methylation	Set 7/9 (KMT7)	• Essential for Tat-mediated transactivation, particularly during the early phase of transactivation cycle. • Promotes Tat-TAR interaction.
R52/53 dimethylation	PRMT6	• Interferes with Tat/TAR/P-TEFb complex formation. • Increases Tat stability by preventing proteasomal degradation.
K28 deacetylation	HDAC6	• Reduces Tat transcriptional activity. • Triggers Tat export from the nucleus to cytoplasm.

1.8 Eradication strategies for HIV-1 latency

Two aspects of HIV-1 cure (1) a **sterile cure** which involves complete removal of the virus and (2) a **functional cure** which causes cessation of viral proliferation in the absence of HAART. The latent reservoir should be purged as the ultimate goal for a functional cure. Appropriately, novel strategies to eradicate the latent reservoir are being developed at an exponential rate. The only reported case of a sterile HIV-1 cure is that of the 'Berlin patient' (Hütter G et al., 2009) who simultaneously contracted HIV-1 and

acute myeloid leukemia. The patient was unexpectedly cured of HIV-1 following hematopoietic stem cell transplantation from a mutant donor of CCR5-recetor along with an ongoing myeloablative regimen both treatments being targeted to cure the leukemia.

1.8.1 Strategies to purge the latent reservoir

The most widely accepted strategy to eliminate the latent reservoir is the use of small-molecule, pharmacological modulators to reverse latency in patients undergoing HAART. The early trials involved the stimulation of the resting $CD4^{+ve}$ T-cells with the global T-cell activators OKT3 and IL-2 in HAART-treated patients. However, such an application led to toxic side effects with profound immune activation (Prins JM et al., 1999). This 'shock' (viral reactivation) and 'kill' (HAART-mediated clearance) approach was improved by administering small-molecule HDAC inhibitors which selectively induced the viral transcription without a noticeable change in T-cell activation. Vorinostat and Panobinostat were among the early drugs in this category to be used in the clinical trials (Archin NM et al., 2012; Elliot J et al., 2013; Rasmussen T et al., 2014). Subsequent strategies involved the application of selected combinations of LRAs that targeted specific targets for viral reactivation. Recently, Ho et al. demonstrated that nearly 10% of the intact, replication-competent viruses that did not respond to the maximal mitogens-stimulation in the first round may still get reactivated upon subsequent stimulation with the same inducers (Ho YC et al., 2013) Likewise, synergistic effects of HDAC or HMT inhibitors and activators of the PKC/NF-κB pathway have been particularly effective in primary cells and cell line models although their potential in the in vivo systems remains to be estimated (Reuse S et al., 2009; Burnett JC et al., 2010; Fernandez G et al., 2010). Several FDA approved LRAs that target distinct epigenetic and non-epigenetic targets are currently being considered from human trials (reviewed in Xing S et al., 2013).

Cellular miRNAs have recently offered a great promise towards reactivation of replication-competent virus from the T-cells of HAART-treated patients. A cluster of suppressive cellular micro RNAs (miRNAs) were identified in the infected resting $CD4^{+ve}$ cells and assisted in viral persistence. Recently, a study identified five putative miRNA binding sites in within the 3'UTR, the common element in all the HIV-1 mRNA species (spliced or unspliced). The five miRNAs targets were not only conserved across several

HIV-1 strains but they also restricted the functions of many HIV-1 proteins including those of the key regulatory factors Tat and Rev. Mutated as well as antisense sequences of the five miRNAs neutralized their inhibitory effects resulting in increased viral proliferation. Moreover, in patients on HAART therapy, a combination of the anti-miRNA inhibitors rescued the inhibited transcripts and enhanced viral production by several folds (Huang J et al., 2007).

1.8.2 Targeted elimination of persistently infected cells

Latency-reversal and the concomitant viral cytopathic effects may not be sufficient to totally eradicate the latent reservoir in infected individuals. In such cases, continued expression of viral antigens is likely to lead to chronic immune activation and immune exhaustion. The effector cells targeting the viral antigens are either lost or survive with compromised antiviral or proliferative potentials (Khaitan A et al., 2011). An appropriate scheme to augment virus-specific immune response is, therefore, critical to purge the persistently infected cells. Both T-cell effector functions and B-cell mediated antibody responses are severely compromised in an HIV-1 infection (Moir S et al., 2012). A **therapeutic vaccine** is expected to restore as well as enhance the immune responses by the controlled administration of the viral antigens. Among the several tested therapeutic vaccines, whole inactivated virus, recombinant proteins or virus, DNA vectors or dendritic cell presentation of autologous antigens are worthy of special mention (reviewed in Garcia F et al., 2012). Although significant improvement in the immune responses has been noted, none of the therapeutic vaccines could bring HAART to a permanent halt (Gay C et al., 2014; Casazza JP et al., 2013). **Cell based therapies** such as the adoptive transfer of HIV-1 specific CTLs to target the persistently infected cells (Chapuis AG et al., 2011) are also in progress. Following viral reactivation, the gag-stimulated CTLs prove superior over the freshly injected $CD8^{+ve}$ T-cells in eliminating the infected $CD4^{+ve}$ T-cells (Shan L et al., 2012). Further, a panel of multiple overlapping peptides of various HIV-1 antigens was more effective than individual peptides in clearing the latent $CD4^{+ve}$ reservoir following reactivation (Sung JA et al., 2015). Several other cytotoxic immune cells- the natural killer cells, the lymphokine-activated killer cells (Groscurth P et al., 1989) and the γδ cells (Hudspeth K et al., 2012) have been reported to efficiently eliminate infected T-cells using mechanisms complementary to that of the

CTLs (Vivier E et al., 2011). **Gene therapies** using engineered effector cells are much more effective in clearing the targets. For instance, a recent study used patient PBMCs with genetically engineering TCR to redirect the effector T-cells to the HIV-1 antigens (Varela-Rohena A et al., 2008; Porter DL et al., 2012). However novel tools as these should be carefully implemented owing to their potential for off target effects (Heslop HE., 2013). **Immunotoxins**, defined as bi-functional chimeric proteins comprising of a targeting domain (antibody or ligand) and a toxin effector domain are a recent addition to the existing strategies of clearing persistently infected cells (Pincus SH et al., 1996). In a study involving humanized mouse model, the combined effect of the immunotoxin 3B3-PE38 and HAART reduced the RNA levels in the tissue by several logs than by the use of HAART alone (Denton P W et al., 2014).

1.9 Bibliography

Ammosova, T., Berro, R., Jerebtsova, M., Jackson, A., Charles, S., Klase, Z., Southerland, W., Gordeuk, V.R., Kashanchi, F. and Nekhai, S., 2006. Phosphorylation of HIV-1 Tat by CDK2 in HIV-1 transcription. *Retrovirology*, *3*(1), p.78.

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. and Richman, D.D., 2012. Administration of vorinostat disrupts HIV-1 latency in patients on antiretroviral therapy. *Nature*, *487*(7408), p.482.

Archin, N.M., Sung, J.M., Garrido, C., Soriano-Sarabia, N. and Margolis, D.M., 2014. Eradicating HIV-1 infection: seeking to clear a persistent pathogen. *Nature Reviews Microbiology*, *12*(11), p.750.

Arkin, A., Ross, J. and McAdams, H.H., 1998. Stochastic kinetic analysis of developmental pathway bifurcation in phage λ-infected Escherichia coli cells. *Genetics*, *149*(4), pp.1633-1648.

Bachu, M., Mukthey, A.B., Murali, R.V., Cheedarla, N., Mahadevan, A., Shankar, S.K., Satish, K.S., Kundu, T.K. and Ranga, U., 2012a. Sequence insertions in the HIV type 1 subtype c viral promoter predominantly generate an additional NF-κB binding site. *AIDS research and human retroviruses*, *28*(10), pp.1362-1368.

Bachu, M., Yalla, S., Asokan, M., Verma, A., Neogi, U., Sharma, S., Murali, R.V., Mukthey, A.B., Bhatt, R., Chatterjee, S. and Rajan, R.E., 2012b. Multiple NF-κB sites in HIV-1 subtype C long terminal repeat confer superior magnitude of transcription and thereby the enhanced viral predominance. *Journal of Biological Chemistry*, *287*(53), pp.44714-44735.

Baldwin Jr, A.S., 1996. The NF-κB and IκB proteins: new discoveries and insights. *Annual review of immunology*, *14*(1), pp.649-681.

Bjorndal, A., Sonnerborg, A., Tscherning, C., Albert, J. and Fenyo, E.M., 1999. Phenotypic characteristics of human immunodeficiency virus type 1 subtype C isolates of Ethiopian AIDS patients. *AIDS research and human retroviruses*, *15*(7), pp.647-653.

Bosque, A. and Planelles, V., 2009. Induction of HIV-1 latency and reactivation in primary memory CD4+ T cells. *Blood*, *113*(1), pp.58-65.

Bourgeois, C.F., Kim, Y.K., Churcher, M.J., West, M.J. and Karn, J., 2002. Spt5 cooperates with human immunodeficiency virus type 1 Tat by preventing premature RNA release at terminator sequences. *Molecular and cellular biology*, *22*(4), pp.1079-1093.

Brooks, D.G., Kitchen, S.G., Kitchen, C.M., Scripture-Adams, D.D. and Zack, J.A., 2001. Generation of HIV latency during thymopoiesis. *Nature medicine*, *7*(4), p.459.

Brooks, D.G., Arlen, P.A., Gao, L., Kitchen, C.M. and Zack, J.A., 2003. Identification of T cell-signaling pathways that stimulate latent HIV in primary cells. *Proceedings of the National Academy of Sciences*, *100*(22), pp.12955-12960.

Brussel, A., Delelis, O. and Sonigo, P., 2005. Alu-LTR real-time nested PCR assay for quantifying integrated HIV-1 DNA. In *Human Retrovirus Protocols* (pp. 139-154). Humana Press.

Bukrinsky, M.I., Stanwick, T.L., Dempsey, M.P. and Stevenson, M., 1991. Quiescent T lymphocytes as an inducible virus reservoir in HIV-1 infection. *Science*, *254*(5030), pp.423-427.

Buonaguro, L., Tornesello, M.L. and Buonaguro, F.M., 2007. Human immunodeficiency virus type 1 subtype distribution in the worldwide epidemic: pathogenetic and therapeutic implications. *Journal of virology*, *81*(19), pp.10209-10219.

Burnett, J.C., Miller-Jensen, K., Shah, P.S., Arkin, A.P. and Schaffer, D.V., 2009. Control of stochastic gene expression by host factors at the HIV promoter. *PLoS pathogens*, *5*(1), p.e1000260.

Burnett, J.C., Lim, K.I., Calafi, A., Rossi, J.J., Schaffer, D.V. and Arkin, A.P., 2010. Combinatorial latency reactivation for HIV-1 subtypes and variants. *Journal of virology*, *84*(12), pp.5958-5974.

Butler, S.L., Hansen, M.S. and Bushman, F.D., 2001. A quantitative assay for HIV DNA integration in vivo. *Nature medicine*, *7*(5), p.631.

Buzon, M.J., Sun, H., Li, C., Shaw, A., Seiss, K., Ouyang, Z., Martin-Gayo, E., Leng, J., Henrich, T.J., Li, J.Z. and Pereyra, F., 2014. HIV-1 persistence in CD4+ T cells with stem cell–like properties. *Nature medicine*, *20*(2), p.139.

Cai, W.E.I.Z.H.O.N.G., Astor, T.L., Liptak, L.M., Cho, C., Coen, D.M. and Schaffer, P.A., 1993. The herpes simplex virus type 1 regulatory protein ICP0 enhances virus replication during acute infection and reactivation from latency. *Journal of virology*, *67*(12), pp.7501-7512.

Carter, C.C., McNamara, L.A., Onafuwa-Nuga, A., Shackleton, M., Riddell IV, J., Bixby, D., Savona, M.R., Morrison, S.J. and Collins, K.L., 2011. HIV-1 utilizes the CXCR4 chemokine receptor to infect multipotent hematopoietic stem and progenitor cells. *Cellhost & microbe*, *9*(3), pp.223-234.

Casazza, J.P., Bowman, K.A., Adzaku, S., Smith, E.C., Enama, M.E., Bailer, R.T., Price, D.A., Gostick, E., Gordon, I.J., Ambrozak, D.R. and Nason, M.C., 2013. Therapeutic vaccination

expands and improves the function of the HIV-specific memory T-cell repertoire. *The Journal of infectious diseases*, *207*(12), pp.1829-1840.

Chapuis, A.G., Casper, C., Kuntz, S., Zhu, J., Tjernlund, A., Diem, K., Turtle, C.J., Cigal, M.L., Velez, R., Riddell, S. and Corey, L., 2011. HIV-specific CD8+ T cells from HIV+ individuals receiving HAART can be expanded ex vivo to augment systemic and mucosal immunity in vivo. *Blood*, *117*(20), pp.5391-5402.

Chen, F.E., Huang, D.B., Chen, Y.Q. and Ghosh, G., 1998. Crystal structure of p50/p65 heterodimer of transcription factor NF-κB bound to DNA. *Nature*, *391*(6665), p.410.

Cheng, J., Nath, A., Knudsen, B., Hochman, S., Geiger, J.D., Ma, M. and Magnuson, D.S.K., 1997. Neuronal excitatory properties of human immunodeficiency virus type 1 Tat protein. *Neuroscience*, *82*(1), pp.97-106.

Chun, T.W., Finzi, D., Margolick, J., Chadwick, K., Schwartz, D. and Siliciano, R.F., 1995. In vivo fate of HIV-1-infected T cells: quantitative analysis of the transition to stable latency. *Nature medicine*, *1*(12), p.1284.

Chun, T.W., Carruth, L., Finzi, D., Shen, X., DiGiuseppe, J.A., Taylor, H., Hermankova, M., Chadwick, K., Margolick, J., Quinn, T.C. and Kuo, Y.H., 1997. Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection. *Nature*, *387*(6629), pp.183-188.

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. and Schacker, T.W., 2009. HIV reservoir size and persistence are driven by T cell survival and homeostatic proliferation. *Nature medicine*, *15*(8), p.893.

Coffin, J. and Swanstrom, R., 2013. HIV pathogenesis: dynamics and genetics of viral populations and infected cells. *Cold Spring Harbor perspectives in medicine*, *3*(1), p.a012526.

Cooney, A.J., Tsai, S.Y., O'Malley, B.W. and Tsai, M.J., 1991. Chicken ovalbumin upstream promoter transcription factor binds to a negative regulatory region in the human immunodeficiency virus type 1 long terminal repeat. *Journal of virology*, *65*(6), pp.2853-2860.

Dahabieh, M.S., Ooms, M., Simon, V. and Sadowski, I., 2013. A doubly fluorescent HIV-1 reporter shows that the majority of integrated HIV-1 is latent shortly after infection. *Journal of virology*, *87*(8), pp.4716-4727.

Dar, R.D., Hosmane, N.N., Arkin, M.R., Siliciano, R.F. and Weinberger, L.S., 2014. Screening for noise in gene expression identifies drug synergies. *Science*, *344*(6190), pp.1392-1396.

Dasgupta, P., Saikumar, P., Reddy, C.D. and Reddy, E.P., 1990. Myb protein binds to human immunodeficiency virus 1 long terminal repeat (LTR) sequences and transactivates LTR-mediated transcription. *Proceedings of the National Academy of Sciences*, *87*(20), pp.8090-8094.

Denton, P.W., Long, J.M., Wietgrefe, S.W., Sykes, C., Spagnuolo, R.A., Snyder, O.D., Perkey, K., Archin, N.M., Choudhary, S.K., Yang, K. and Hudgens, M.G., 2014. Targeted cytotoxic therapy kills persisting HIV infected cells during ART. *PLoS pathogens*, *10*(1), p.e1003872.

Dodd, I.B., Perkins, A.J., Tsemitsidis, D. and Egan, J.B., 2001. Octamerization of λ CI repressor is needed for effective repression of P RM and efficient switching from lysogeny. *Genes & development*, *15*(22), pp.3013-3022.

Donahue, D.A., Kuhl, B.D., Sloan, R.D. and Wainberg, M.A., 2012. The viral protein Tat can inhibit the establishment of HIV-1 latency. *Journal of virology*, *86*(6), pp.3253-3263.

Drysdale, C.M. and Pavlakis, G.N., 1991. Rapid activation and subsequent down-regulation of the human immunodeficiency virus type 1 promoter in the presence of Tat: possible mechanisms contributing to latency. *Journal of virology*, *65*(6), pp.3044-3051.

Dwarakanath, R.S., Clark, C.L., McElroy, A.K. and Spector, D.H., 2001. The use of recombinant baculoviruses for sustained expression of human cytomegalovirus immediate early proteins in fibroblasts. *Virology*, *284*(2), pp.297-307.

Eisele, E. and Siliciano, R.F., 2012. Redefining the viral reservoirs that prevent HIV-1 eradication. *Immunity*, *37*(3), pp.377-388.

Elliot, J. *et al.* in Highlights of the 20th conference on retroviruses and opportunistic infections. 2013.*Top. Antivir. Med.* (21)Abstract 50LB

Emiliani, S., Van Lint, C., Fischle, W., Paras, P., Ott, M., Brady, J. and Verdin, E., 1996. A point mutation in the HIV-1 Tat responsive element is associated with postintegration latency. *Proceedings of the National Academy of Sciences*, *93*(13), pp.6377-6381.

Emiliani, S., Fischle, W., Ott, M., Van Lint, C., Amella, C.A. and Verdin, E., 1998. Mutations in the tat gene are responsible for human immunodeficiency virus type 1 postintegration latency in the U1 cell line. *Journal of virology*, *72*(2), pp.1666-1670.

Emerman, M. and Temin, H.M., 1986. Comparison of promoter suppression in avian and murine retrovirus vectors. *Nucleic acids research*, *14*(23), pp.9381-9396.

Farnet, C.M. and Haseltine, W.A., 1991. Circularization of human immunodeficiency virus type 1 DNA in vitro. *Journal of virology*, *65*(12), pp.6942-6952.

Feng, S. and Holland, E.C., 1988. HIV-1 tat trans-activation requires the loop sequence within tar. *Nature*, *334*(6178), p.165.

Fernandez, G. and Zeichner, S.L., 2010. Cell line-dependent variability in HIV activation employing DNMT inhibitors. *Virology journal*, *7*(1), p.266.

Finzi, D., Blankson, J., Siliciano, J.D., Margolick, J.B., Chadwick, K., Pierson, T., Smith, K., Lisziewicz, J., Lori, F., Flexner, C. and Quinn, T.C., 1999. Latent infection of CD4+ T cells provides a mechanism for lifelong persistence of HIV-1, even in patients on effective combination therapy. *Nature medicine*, *5*(5), p.512.

Folks, T.M., Justement, J., Kinter, A., Schnittman, S., Orenstein, J., Poli, G. and Fauci, A.S., 1988. Characterization of a promonocyte clone chronically infected with HIV and inducible by 13-phorbol-12-myristate acetate. *The Journal of Immunology*, *140*(4), pp.1117-1122.

Folks, T.M., Clouse, K.A., Justement, J., Rabson, A., Duh, E., Kehrl, J.H. and Fauci, A.S., 1989. Tumor necrosis factor alpha induces expression of human immunodeficiency virus in a

chronically infected T-cell clone. *Proceedings of the National Academy of Sciences*, *86*(7), pp.2365-2368.

Friedman, J., Cho, W.K., Chu, C.K., Keedy, K.S., Archin, N.M., Margolis, D.M. and Karn, J., 2011. Epigenetic silencing of HIV-1 by the Histone H3 lysine 27 Methyltransferase Enhancer of Zeste 2 (EZH2). *Journal of virology*, pp.JVI-00836.

Fujinaga, K., Irwin, D., Huang, Y., Taube, R., Kurosu, T. and Peterlin, B.M., 2004. Dynamics of human immunodeficiency virus transcription: P-TEFb phosphorylates RD and dissociates negative effectors from the transactivation response element. *Molecular and cellular biology*, *24*(2), pp.787-795.

Gao, W.Y., Cara, A., Gallo, R.C. and Lori, F., 1993. Low levels of deoxynucleotides in peripheral blood lymphocytes: a strategy to inhibit human immunodeficiency virus type 1 replication. *Proceedings of the National Academy of Sciences*, *90*(19), pp.8925-8928.

Garcia, J.A., Harrich, D., Pearson, L., Mitsuyasu, R. and Gaynor, R.B., 1988. Functional domains required for tat-induced transcriptional activation of the HIV-1 long terminal repeat. *The EMBO journal*, *7*(10), pp.3143-3147.

García, F., León, A., Gatell, J.M., Plana, M. and Gallart, T., 2012. Therapeutic vaccines against HIV infection. *Human vaccines & immunotherapeutics*, *8*(5), pp.569-581.

Gay, C. *et al.*,2014 in Special issue: abstracts from the 2014 conference on retroviruses and opportunistic infections. *Top. Antivir. Med.* (22), e-1 Abstract 344

Geng, G., Liu, B., Chen, C., Wu, K., Liu, J., Zhang, Y., Pan, T., Li, J., Yin, Y., Zhang, J. and Huang, F., 2016. Development of an attenuated tat protein as a highly-effective agent to specifically activate HIV-1 latency. *Molecular Therapy*, *24*(9), pp.1528-1537.

Giacca, M., Gutierrez, M.I., Menzo, S., Di Fagagna, F.D.A. and Flaschi, A., 1992. A human binding site for transcription factor USF/MLTF mimics the negative regulatory element of human immunodeficiency virus type 1. *Virology*, *186*(1), pp.133-147.

Giacca, M., 2005. HIV-1 Tat, apoptosis and the mitochondria: a tubulin link? *Retrovirology*, *2*(1), p.7.

Ghosh, G., Van Duyne, G., Ghosh, S. and Sigler, P.B., 1995. Structure of NF-κB p50 homodimer bound to a κB site. *Nature*, *373*(6512), p.303.

Ghosh, S., May, M.J. and Kopp, E.B., 1998. NF-κB and Rel proteins: evolutionarily conserved mediators of immune responses. *Annual review of immunology*, *16*(1), pp.225-260.

Groscurth, P., 1989. Cytotoxic effector cells of the immune system. *Anatomy and embryology*, *180*(2), pp.109-119.

Gupta, S., Boppana, R., Mishra, G.C., Saha, B. and Mitra, D., 2008. HIV-1 Tat suppresses gp120-specific T cell response in IL-10-dependent manner. *The Journal of Immunology*, *180*(1), pp.79-88.

Han, Y., Lassen, K., Monie, D., Sedaghat, A.R., Shimoji, S., Liu, X., Pierson, T.C., Margolick, J.B., Siliciano, R.F. and Siliciano, J.D., 2004. Resting CD4+ T cells from HIV-1-infected

individuals carry integrated HIV-1 genomes within actively transcribed host genes. *J Virol*, *78*(12), pp.6122-6133.

Haughey, N.J. and Mattson, M.P., 2002. Calcium dysregulation and neuronal apoptosis by the HIV-1 proteins Tat and gp120. *Journal of acquired immune deficiency syndromes (1999)*, *31*, pp.S55-61.

He, N., Jahchan, N.S., Hong, E., Li, Q., Bayfield, M.A., Maraia, R.J., Luo, K. and Zhou, Q., 2008. A La-related protein modulates 7SK snRNP integrity to suppress P-TEFb-dependent transcriptional elongation and tumorigenesis. *Molecular cell*, *29*(5), pp.588-599.

Hemelaar, J., 2012. The origin and diversity of the HIV-1 pandemic. *Trends in molecular medicine*, *18*(3), pp.182-192.

Henderson, A.J., Zou, X. and Calame, K.L., 1995. C/EBP proteins activate transcription from the human immunodeficiency virus type 1 long terminal repeat in macrophages/monocytes. *Journal of virology*, *69*(9), pp.5337-5344.

Henderson, A.J. and Calame, K.L., 1997. CCAAT/enhancer binding protein (C/EBP) sites are required for HIV-1 replication in primary macrophages but not CD4+ T cells. *Proceedings of the National Academy of Sciences*, *94*(16), pp.8714-8719.

Hermankova, M., Siliciano, J.D., Zhou, Y., Monie, D., Chadwick, K., Margolick, J.B., Quinn, T.C. and Siliciano, R.F., 2003. Analysis of human immunodeficiency virus type 1 gene expression in latently infected resting CD4+ T lymphocytes in vivo. *Journal of virology*, *77*(13), pp.7383-7392.

Heslop, H.E., 2013. Genetic engineering of T-cell receptors: TCR takes to titin. *Blood*, *122*(6), pp.853-854.

Ho, D.D., Rota, T.R. and Hirsch, M.S., 1986. Infection of monocyte/macrophages by human T lymphotropic virus type III. *The Journal of clinical investigation*, *77*(5), pp.1712-1715.

Ho, Y.C., Shan, L., Hosmane, N.N., Wang, J., Laskey, S.B., Rosenbloom, D.I., Lai, J., Blankson, J.N., Siliciano, J.D. and Siliciano, R.F., 2013. Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure. *Cell*, *155*(3), pp.540-551.

Huang, J., Wang, F., Argyris, E., Chen, K., Liang, Z., Tian, H., Huang, W., Squires, K., Verlinghieri, G. and Zhang, H., 2007. Cellular microRNAs contribute to HIV-1 latency in resting primary CD4+ T lymphocytes. *Nature medicine*, *13*(10), p.1241.

Hudspeth, K., Fogli, M., Correia, D.V., Mikulak, J., Roberto, A., Della Bella, S., Silva-Santos, B. and Mavilio, D., 2012. Engagement of NKp30 on Vδ1 T cells induces the production of CCL3, CCL4, and CCL5 and suppresses HIV-1 replication. *Blood*, *119*(17), pp.4013-4016.

Hütter, G., Nowak, D., Mossner, M., Ganepola, S., Müßig, A., Allers, K., Schneider, T., Hofmann, J., Kücherer, C., Blau, O. and Blau, I.W., 2009. Long-term control of HIV by CCR5 Delta32/Delta32 stem-cell transplantation. *New England Journal of Medicine*, *360*(7), pp.692-698.

Ivanov, D., Kwak, Y.T., Guo, J. and Gaynor, R.B., 2000. Domains in the SPT5 protein that modulate its transcriptional regulatory properties. *Molecular and cellular biology*, *20*(9), pp.2970-2983.

Isel, C. and Karn, J., 1999. Direct evidence that HIV-1 tat stimulates RNA polymerase II carboxyl-terminal domain hyperphosphorylation during transcriptional elongation1. *Journal of molecular biology*, *290*(5), pp.929-941.

Jeeninga, R.E., Hoogenkamp, M., Armand-Ugon, M., de Baar, M., Verhoef, K. and Berkhout, B.E.N., 2000. Functional differences between the long terminal repeat transcriptional promoters of human immunodeficiency virus type 1 subtypes A through G. *Journal of virology*, *74*(8), pp.3740-3751.

Jones, K.A., Kadonaga, J.T., Luciw, P.A. and Tjian, R., 1986. Activation of the AIDS retrovirus promoter by the cellular transcription factor, Sp1. *Science*, *232*(4751), pp.755-759.

Jones, K.A. and Peterlin, M.B., 1994. Control of RNA initiation and elongation at the HIV-1 promoter. *Annual review of biochemistry*, *63*(1), pp.717-743.

Jordan, A., Bisgrove, D. and Verdin, E., 2003. HIV reproducibly establishes a latent infection after acute infection of T cells in vitro. *The EMBO journal*, *22*(8), pp.1868-1877.

Karn, J. and Stoltzfus, C.M., 2012. Transcriptional and posttranscriptional regulation of HIV-1 gene expression. *Cold Spring Harbor perspectives in medicine*, *2*(2), p.a006916.

Kent, J.R., Kang, W., Miller, C.G. and Fraser, N.W., 2003. Herpes simplex virus latency-associated transcript gene function. *Journal of neurovirology*, *9*(3), pp.285-290.

Khaitan, A. and Unutmaz, D., 2011. Revisiting immune exhaustion during HIV infection. *Current HIV/AIDS Reports*, *8*(1), pp.4-11.

Kim, Y.K., Bourgeois, C.F., Isel, C., Churcher, M.J. and Karn, J., 2002. Phosphorylation of the RNA polymerase II carboxyl-terminal domain by CDK9 is directly responsible for human immunodeficiency virus type 1 Tat-activated transcriptional elongation. *Molecular and cellular biology*, *22*(13), pp.4622-4637.

Kim, T.A., Avraham, H.K., Koh, Y.H., Jiang, S., Park, I.W. and Avraham, S., 2003. HIV-1 Tat-mediated apoptosis in human brain microvascular endothelial cells. *The Journal of Immunology*, *170*(5), pp.2629-2637.

Klaver, B. and Berkhout, B., 1994. Comparison of 5'and 3'long terminal repeat promoter function in human immunodeficiency virus. *Journal of virology*, *68*(6), pp.3830-3840.

Korin, Y.D., Brooks, D.G., Brown, S., Korotzer, A. and Zack, J.A., 2002. Effects of prostratin on T-cell activation and human immunodeficiency virus latency. *Journal of virology*, *76*(16), pp.8118-8123.

Krebs, F.C., Mehrens, D., Pomeroy, S., Goodenow, M.M. and Wigdahl, B., 1998. Human immunodeficiency virus type 1 long terminal repeat quasispecies differ in basal transcription and nuclear factor recruitment in human glial cells and lymphocytes. *Journal of biomedical science*, *5*(1), pp.31-44.

Krueger, B.J., Varzavand, K., Cooper, J.J. and Price, D.H., 2010. The mechanism of release of P-TEFb and HEXIM1 from the 7SK snRNP by viral and cellular activators includes a conformational change in 7SK. *PloS one*, *5*(8), p.e12335.

Kumar, A., Abbas, W. and Herbein, G., 2014. HIV-1 latency in monocytes/macrophages. *Viruses*, *6*(4), pp.1837-1860.

Kuppuswamy, M., Subramanian, T., Srinivasan, A. and Chinnadurai, G., 1989. Multiple functional domains of Tat, the trans-activator of HIV-1, defined by mutational analysis. *Nucleic acids research*, *17*(9), pp.3551-3561.

Lan, P., Tonomura, N., Shimizu, A., Wang, S. and Yang, Y.G., 2006. Reconstitution of a functional human immune system in immunodeficient mice through combined human fetal thymus/liver and CD34+ cell transplantation. *Blood*, *108*(2), pp.487-492.

Lassen, K.G., Ramyar, K.X., Bailey, J.R., Zhou, Y. and Siliciano, R.F., 2006. Nuclear retention of multiply spliced HIV-1 RNA in resting CD4+ T cells. *PLoS pathogens*, *2*(7), p.e68.

Liang, C. and Wainberg, M.A., 2002. The role of Tat in HIV-1 replication: an activator and/or a suppressor. *AIDS Rev*, *4*(1), pp.41-9.

Li, L., Dahiya, S., Kortagere, S., Aiamkitsumrit, B., Cunningham, D., Pirrone, V., Nonnemacher, M.R. and Wigdahl, B., 2012. Impact of Tat genetic variation on HIV-1 disease. *Advances in virology*, *2012*.

Lin, X., Irwin, D., Kanazawa, S., Huang, L., Romeo, J., Yen, T.B. and Peterlin, B.M., 2003. Transcriptional profiles of latent human immunodeficiency virus in infected individuals: effects of Tat on the host and reservoir. *Journal of virology*, *77*(15), pp.8227-8236.

Macias, M.P. and Stinski, M.F., 1993. An in vitro system for human cytomegalovirus immediate early 2 protein (IE2)-mediated site-dependent repression of transcription and direct binding of IE2 to the major immediate early promoter. *Proceedings of the National Academy of Sciences*, *90*(2), pp.707-711.

Majello, B., De Luca, P., Hagen, G., Suske, G. and Lania, L., 1994. Different members of the Sp1 multigene family exert opposite transcritional regulation of the long terminal repeat of HIV-1. *Nucleic Acids Research*, *22*(23), pp.4914-4921.

Marini, A., Harper, J.M. and Romerio, F., 2008. An in vitro system to model the establishment and reactivation of HIV-1 latency. *The Journal of Immunology*, *181*(11), pp.7713-7720.

McCune, J.M., Namikawa, R., Kaneshima, H., Shultz, L.D., Lieberman, M. and Weissman, I.L., 1988. The SCID-hu mouse: murine model for the analysis of human hematolymphoid differentiation and function. *Science*, *241*(4873), pp.1632-1639.

McNamara, L.A., Onafuwa-Nuga, A., Sebastian, N.T., Riddell IV, J., Bixby, D. and Collins, K.L., 2013. CD133+ hematopoietic progenitor cells harbor HIV genomes in a subset of optimally treated people with long-term viral suppression. *The Journal of infectious diseases*, *207*(12), pp.1807-1816.

McNamara, L.A., Ganesh, J.A. and Collins, K.L., 2012. Latent HIV-1 infection occurs in multiple subsets of hematopoietic progenitor cells and is reversed by NF-κB activation. *Journal of virology*, pp.JVI-00895.

Melkus, M.W., Estes, J.D., Padgett-Thomas, A., Gatlin, J., Denton, P.W., Othieno, F.A., Wege, A.K., Haase, A.T. and Garcia, J.V., 2006. Humanized mice mount specific adaptive and innate immune responses to EBV and TSST-1. *Nature medicine*, *12*(11), p.1316.

Messi, M., Giacchetto, I., Nagata, K., Lanzavecchia, A., Natoli, G. and Sallusto, F., 2003. Memory and flexibility of cytokine gene expression as separable properties of human T H 1 and T H 2 lymphocytes. *Nature immunology*, *4*(1), p.78.

Miller, M.D., Wang, B. and Bushman, F.D., 1995. Human immunodeficiency virus type 1 preintegration complexes containing discontinuous plus strands are competent to integrate in vitro. *Journal of virology*, *69*(6), pp.3938-3944.

Miura, Y., Koyanagi, Y. and Mizusawa, H., 2003. TNF-related apoptosis-inducing ligand (TRAIL) induces neuronal apoptosis in HIV-encephalopathy. *Journal of medical and dental sciences*, *50*(1), pp.17-25.

Moir, S., Ho, J., Malaspina, A., Wang, W., DiPoto, A.C., O'Shea, M.A., Roby, G., Kottilil, S., Arthos, J., Proschan, M.A. and Chun, T.W., 2008. Evidence for HIV-associated B cell exhaustion in a dysfunctional memory B cell compartment in HIV-infected viremic individuals. *Journal of Experimental Medicine*, *205*(8), pp.1797-1805.

Monie, D., Simmons, R.P., Nettles, R.E., Kieffer, T.L., Zhou, Y., Zhang, H., Karmon, S., Ingersoll, R., Chadwick, K., Zhang, H. and Margolick, J.B., 2005. A novel assay allows genotyping of the latent reservoir for human immunodeficiency virus type 1 in the resting CD4+ T cells of viremic patients. *Journal of virology*, *79*(8), pp.5185-5202.

Montano, M.A., Novitsky, V.A., Blackard, J.T., Cho, N.L., Katzenstein, D.A. and Essex, M., 1997. Divergent transcriptional regulation among expanding human immunodeficiency virus type 1 subtypes. *Journal of Virology*, *71*(11), pp.8657-8665.

Mousseau, G., Kessing, C.F., Fromentin, R., Trautmann, L., Chomont, N. and Valente, S.T., 2015. The Tat inhibitor didehydro-cortistatin A prevents HIV-1 reactivation from latency. *MBio*, *6*(4), pp.e00465-15.

Müller, C.W., Rey, F.A., Sodeoka, M., Verdine, G.L. and Harrison, S.C., 1995. Structure of the NF-κB p50 homodimer bound to DNA. *Nature*, *373*(6512), p.311.

Munkanta, M., Handema, R., Kasai, H., Gondwe, C., Deng, X., Yamashita, A., Asagi, T., Yamamoto, N., Ito, M., Kasolo, F. and Terunuma, H., 2005. Predominance of three NF-κB binding sites in the long terminal repeat region of HIV type 1 subtype C isolates from Zambia. *AIDS Research & Human Retroviruses*, *21*(10), pp.901-906.

Murray, J.M., Zaunders, J.J., McBride, K.L., Xu, Y., Bailey, M., Suzuki, K., Cooper, D.A., Emery, S., Kelleher, A.D. and Koelsch, K.K., 2014. HIV DNA subspecies persist in both activated and resting memory CD4+ T cells during antiretroviral therapy. *Journal of virology*, *88*(6), pp.3516-3526.

Myers, G., 1994. HIV: between past and future. *AIDS research and human retroviruses*, *10*(11), pp.1317-1324.

Nabel, G. and Baltimore, D., 1987. An inducible transcription factor activates expression of human immunodeficiency virus in T cells. *Nature*, *326*(6114), p.711.

Naghavi, M.H., Schwartz, S., Sonnerborg, A. and Vahlne, A., 1999. Long terminal repeat promoter/enhancer activity of different subtypes of HIV type 1. *AIDS research and human retroviruses*, *15*(14), pp.1293-1303.

Narasipura, S.D., Kim, S. and Al-Harthi, L., 2014. Epigenetic regulation of HIV-1 latency in astrocytes. *Journal of virology*, *88*(5), pp.3031-3038.

Nguyen, V.T., Kiss, T., Michels, A.A. and Bensaude, O., 2001. 7SK small nuclear RNA binds to and inhibits the activity of CDK9/cyclin T complexes. *Nature*, *414*(6861), p.322.

Novitsky, V., Smith, U.R., Gilbert, P., McLane, M.F., Chigwedere, P., Williamson, C., Ndung'u, T., Klein, I., Chang, S.Y., Peter, T. and Thior, I., 2002. Human immunodeficiency virus type 1 subtype C molecular phylogeny: consensus sequence for an AIDS vaccine design?. *Journal of virology*, *76*(11), pp.5435-5451.

Norton, N.J., Fun, A., Bandara, M., Wills, M.R., Mok, H.P. and Lever, A.M., 2017. Innovations in the quantitative virus outgrowth assay and its use in clinical trials. *Retrovirology*, *14*(1), p.58.

O'Doherty, U., Swiggard, W.J., Jeyakumar, D., McGain, D. and Malim, M.H., 2002. A sensitive, quantitative assay for human immunodeficiency virus type 1 integration. *Journal of virology*, *76*(21), pp.10942-10950.

Pai, A. and Weinberger, L.S., 2017. Fate-Regulating Circuits in Viruses: From Discovery to New Therapy Targets. *Annual review of virology*, *4*, pp.469-490.

Pearson, R., Kim, Y.K., Hokello, J., Lassen, K., Friedman, J., Tyagi, M. and 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. *Journal of virology*, *82*(24), pp.12291-12303.

Pereira, L.A., Bentley, K., Peeters, A., Churchill, M.J. and Deacon, N.J., 2000. SURVEY AND SUMMARY A compilation of cellular transcription factor interactions with the HIV-1 LTR promoter. *Nucleic acids research*, *28*(3), pp.663-668.

Perelson, A.S., Essunger, P., Cao, Y., Vesanen, M., Hurley, A., Saksela, K., Markowitz, M. and Ho, D.D., 1997. Decay characteristics of HIV-1-infected compartments during combination therapy. *Nature*, *387*(6629), p.188.

Pierson, T.C., Zhou, Y., Kieffer, T.L., Ruff, C.T., Buck, C. and Siliciano, R.F., 2002. Molecular characterization of preintegration latency in human immunodeficiency virus type 1 infection. *Journal of virology*, *76*(17), pp.8518-8531.

Pincus, S.H., 1996. Therapeutic potential of anti-HIV immunotoxins. *Antiviral research*, *33*(1), pp.1-9.

Pocernich, C.B., Sultana, R., Mohmmad-Abdul, H., Nath, A. and Butterfield, D.A., 2005. HIV-dementia, Tat-induced oxidative stress, and antioxidant therapeutic considerations. *Brain Research Reviews*, *50*(1), pp.14-26.

Poggi, A. and Zocchi, M.R., 2006. HIV-1 Tat triggers TGF-β production and NK cell apoptosis that is prevented by pertussis toxin B. *Journal of Immunology Research*, *13*(2-4), pp.369-372.

Pomerantz, R.J., Trono, D., Feinberg, M.B. and Baltimore, D., 1990. Cells nonproductively infected with HIV-1 exhibit an aberrant pattern of viral RNA expression: a molecular model for latency. *Cell*, *61*(7), pp.1271-1276.

Porter, D.L., Levine, B.L., Kalos, M., Bagg, A. and June, C.H., 2011. Chimeric antigen receptor–modified T cells in chronic lymphoid leukemia. *New England Journal of Medicine*, *365*(8), pp.725-733.

Prins, J.M., Jurriaans, S., van Praag, R.M., Blaak, H., van Rij, R., Schellekens, P.T.A., Ten Berge, I.J., Yong, S.L., Fox, C.H., Roos, M.T. and de Wolf, F., 1999. Immuno-activation with anti-CD3 and recombinant human IL-2 in HIV-1-infected patients on potent antiretroviral therapy. *Aids*, *13*(17), pp.2405-2410.

Procopio, F.A., Fromentin, R., Kulpa, D.A., Brehm, J.H., Bebin, A.G., Strain, M.C., Richman, D.D., O'Doherty, U., Palmer, S., Hecht, F.M. and Hoh, R., 2015. A novel assay to measure the magnitude of the inducible viral reservoir in HIV-infected individuals. *EBioMedicine*, *2*(8), pp.874-883.

Rafati, H., Parra, M., Hakre, S., Moshkin, Y., Verdin, E. and Mahmoudi, T., 2011. Repressive LTR nucleosome positioning by the BAF complex is required for HIV latency. *PLoS biology*, *9*(11), p.e1001206.

Ragoczy, T. and Miller, G., 2001. Autostimulation of the Epstein-Barr virus BRLF1 promoter is mediated through consensus Sp1 and Sp3 binding sites. *Journal of virology*, *75*(11), pp.5240-5251.

Ramanathan, Y., Rajpara, S.M., Reza, S.M., Lees, E., Shuman, S., Mathews, M.B. and Pe'ery, T., 2001. Three RNA polymerase II carboxyl-terminal domain kinases display distinct substrate preferences. *Journal of Biological Chemistry*, *276*(14), pp.10913-10920.

Ranjbar, S., Rajsbaum, R. and Goldfeld, A.E., 2006. Transactivator of transcription from HIV type 1 subtype E selectively inhibits TNF gene expression via interference with chromatin remodeling of the TNF locus. *The Journal of Immunology*, *176*(7), pp.4182-4190.

Rasmussen, T.,2014 in Special issue: abstracts from the conference on retroviruses and opportunistic infections. *Top. Antivir. Med*, (22), e-1 Abstract 438LB.

Razooky, B.S. and Weinberger, L.S., 2011. Mapping the architecture of the HIV-1 Tat circuit: A decision-making circuit that lacks bistability and exploits stochastic noise. *Methods*, *53*(1), pp.68-77.

Razooky, B.S., Pai, A., Aull, K., Rouzine, I.M. and Weinberger, L.S., 2015. A hardwired HIV latency program. *Cell*, *160*(5), pp.990-1001.

Reeder, J.E., Kwak, Y.T., McNamara, R.P., Forst, C.V. and D'Orso, I., 2015. HIV Tat controls RNA Polymerase II and the epigenetic landscape to transcriptionally reprogram target immune cells. *Elife*, *4*.

Reuse, S., Calao, M., Kabeya, K., Guiguen, A., Gatot, J.S., Quivy, V., Vanhulle, C., Lamine, A., Vaira, D., Demonte, D. and Martinelli, V., 2009. Synergistic activation of HIV-1 expression by deacetylase inhibitors and prostratin: implications for treatment of latent infection. *PloS one*, *4*(6), p.e6093.

Roizman, B., Gu, H. and Mandel, G., 2005. The first 30 minutes in the life of a virus: unREST in the nucleus. *Cell Cycle*, *4*(8), pp.1019-1021.

Roof, P., Ricci, M., Genin, P., Montano, M.A., Essex, M., Wainberg, M.A., Gatignol, A. and Hiscott, J., 2002. Differential regulation of HIV-1 clade-specific B, C, and E long terminal repeats by NF-κB and the Tat transactivator. *Virology*, *296*(1), pp.77-83.

Rosen, C.A., Sodroski, J.G. and Haseltine, W.A., 1985. The location of cis-acting regulatory sequences in the human T cell lymphotropic virus type III (HTLV-III/LAV) long terminal repeat. *Cell*, *41*(3), pp.813-823.

Roy, S., Delling, U., Chen, C.H., Rosen, C.A. and Sonenberg, N., 1990. A bulge structure in HIV-1 TAR RNA is required for Tat binding and Tat-mediated trans-activation. *Genes & development*, *4*(8), pp.1365-1373.

Ruben, S., Perkins, A., Purcell, R., Joung, K., Sia, R., Burghoff, R., Haseltine, W.A. and Rosen, C.A., 1989. Structural and functional characterization of human immunodeficiency virus tat protein. *Journal of virology*, *63*(1), pp.1-8.

Sahu, G.K., Lee, K., Ji, J., Braciale, V., Baron, S. and Cloyd, M.W., 2006. A novel in vitro system to generate and study latently HIV-infected long-lived normal CD4+ T-lymphocytes. *Virology*, *355*(2), pp.127-137.

Saleh, S., Solomon, A., Wightman, F., Xhilaga, M., Cameron, P.U. and Lewin, S.R., 2007. CCR7 ligands CCL19 and CCL21 increase permissiveness of resting memory CD4+ T cells to HIV-1 infection: a novel model of HIV-1 latency. *Blood*, *110*(13), pp.4161-4164.

Sanders, R.L., Clark, C.L., Morello, C.S. and Spector, D.H., 2008. Development of cell lines that provide tightly controlled temporal translation of the human cytomegalovirus IE2 proteins for complementation and functional analyses of growth-impaired and nonviable IE2 mutant viruses. *Journal of virology*, *82*(14), pp.7059-7077.

Sarisky, R.T., Gao, Z., Lieberman, P.M., Fixman, E.D., Hayward, G.S. and Hayward, S.D., 1996. A replication function associated with the activation domain of the Epstein-Barr virus Zta transactivator. *Journal of virology*, *70*(12), pp.8340-8347.

Scripture-Adams, D.D., Brooks, D.G., Korin, Y.D. and Zack, J.A., 2002. Interleukin-7 induces expression of latent human immunodeficiency virus type 1 with minimal effects on T-cell phenotype. *Journal of virology*, *76*(24), pp.13077-13082.

Shan, L., Deng, K., Shroff, N.S., Durand, C.M., Rabi, S.A., Yang, H.C., Zhang, H., Margolick, J.B., Blankson, J.N. and Siliciano, R.F., 2012. Stimulation of HIV-1-specific cytolytic T

lymphocytes facilitates elimination of latent viral reservoir after virus reactivation. *Immunity*, *36*(3), pp.491-501.

Shaw, J.P., Utz, P.J., Durand, D.B., Toole, J.J., Emmel, E.A. and Crabtree, G.R., 1988. Identification of a putative regulator of early T cell activation genes. *Science*, *241*(4862), pp.202-205.

Sheppard, K.A., Rose, D.W., Haque, Z.K., Kurokawa, R., McInerney, E., Westin, S., Thanos, D., Rosenfeld, M.G., Glass, C.K. and Collins, T., 1999. Transcriptional activation by NF-κB requires multiple coactivators. *Molecular and cellular biology*, *19*(9), pp.6367-6378.

Shen, A., Zink, M.C., Mankowski, J.L., Chadwick, K., Margolick, J.B., Carruth, L.M., Li, M., Clements, J.E. and Siliciano, R.F., 2003. Resting CD4+ T lymphocytes but not thymocytes provide a latent viral reservoir in a simian immunodeficiency virus-Macaca nemestrina model of human immunodeficiency virus type 1-infected patients on highly active antiretroviral therapy. *Journal of virology*, *77*(8), pp.4938-4949.

Schmitz, J.E., Kuroda, M.J., Santra, S., Sasseville, V.G., Simon, M.A., Lifton, M.A., Racz, P., Tenner-Racz, K., Dalesandro, M., Scallon, B.J. and Ghrayeb, J., 1999. Control of viremia in simian immunodeficiency virus infection by CD8+ lymphocytes. *Science*, *283*(5403), pp.857-860.

Siekevitz, M., Josephs, S.F., Dukovich, M., Peffer, N., Wong-Staal, F. and Greene, W.C., 1987. Activation of the HIV-1 LTR by T cell mitogens and the trans-activator protein of HTLV-I. *Science*, *238*(4833), pp.1575-1578.

Siliciano, J.D., Kajdas, J., Finzi, D., Quinn, T.C., Chadwick, K., Margolick, J.B., Kovacs, C., Gange, S.J. and Siliciano, R.F., 2003. Long-term follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting CD4+ T cells. *Nature medicine*, *9*(6), p.727.

Siliciano, R.F. and Greene, W.C., 2011. HIV latency. *Cold Spring Harbor perspectives in medicine*, *1*(1), p.a007096.

Singh, A., Razooky, B., Cox, C.D., Simpson, M.L. and Weinberger, L.S., 2010. Transcriptional bursting from the HIV-1 promoter is a significant source of stochastic noise in HIV-1 gene expression. *Biophysical journal*, *98*(8), pp.L32-L34.

Sobhian B, Laguette N, Yatim A, Nakamura M, Levy Y, Kiernan R, Benkirane M. HIV-1 Tat assembles a multifunctional transcription elongation complex and stably associates with the 7SK snRNP. Molecular cell. 2010 May 14;38(3):439-51.

Strain, M.C., Lada, S.M., Luong, T., Rought, S.E., Gianella, S., Terry, V.H., Spina, C.A., Woelk, C.H. and Richman, D.D., 2013. Highly precise measurement of HIV DNA by droplet digital PCR. *PloS one*, *8*(4), p.e55943.

Stinski, M.F. and Petrik, D.T., 2008. Functional roles of the human cytomegalovirus essential IE86 protein. In *Human Cytomegalovirus* (pp. 133-152). Springer Berlin Heidelberg.

Sung, J.A., Lam, S., Garrido, C., Archin, N., Rooney, C.M., Bollard, C.M. and Margolis, D.M., 2015. Expanded cytotoxic T-cell lymphocytes target the latent HIV reservoir. *The Journal of infectious diseases*, *212*(2), pp.258-263.

Swiggard, W.J., Baytop, C., Jianqing, J.Y., Dai, J., Li, C., Schretzenmair, R., Theodosopoulos, T. and O'Doherty, U., 2005. Human immunodeficiency virus type 1 can establish latent infection in resting CD4+ T cells in the absence of activating stimuli. *Journal of virology*, *79*(22), pp.14179-14188.

Teng, M.W., Bolovan-Fritts, C., Dar, R.D., Womack, A., Simpson, M.L., Shenk, T. and Weinberger, L.S., 2012. An endogenous accelerator for viral gene expression confers a fitness advantage. *Cell*, *151*(7), pp.1569-1580.

Tesmer, V.M., Rajadhyaksha, A., Babin, J. and Bina, M., 1993. NF-IL6-mediated transcriptional activation of the long terminal repeat of the human immunodeficiency virus type 1. *Proceedings of the National Academy of Sciences*, *90*(15), pp.7298-7302.

Tonomura, N., Habiro, K., Shimizu, A., Sykes, M. and Yang, Y.G., 2008. Antigen-specific human T-cell responses and T cell–dependent production of human antibodies in a humanized mouse model. *Blood*, *111*(8), pp.4293-4296.

Tyagi, M., Pearson, R.J. and Karn, J., 2010. Establishment of HIV latency in primary CD4+ cells is due to epigenetic transcriptional silencing and P-TEFb restriction. *Journal of virology*, *84*(13), pp.6425-6437.

UN Joint Programme on HIV/AIDS. MDG 6: 15 years, 15 lessons of hope from the AIDS response. Factsheet. *UNAIDS* [online], http://www.unaids.org/sites/default/files/media_asset/20150714_FS_MDG6_ Report_en.pdf (2015)

Van Lint, C., Bouchat, S. and Marcello, A., 2013. HIV-1 transcription and latency: an update. *Retrovirology*, *10*(1), p.67.

Varela-Rohena, A., Molloy, P.E., Dunn, S.M., Li, Y., Suhoski, M.M., Carroll, R.G., Milicic, A., Mahon, T., Sutton, D.H., Laugel, B. and Moysey, R., 2008. Control of HIV-1 immune escape by CD8 T cells expressing enhanced T-cell receptor. *Nature medicine*, *14*(12), p.1390.

Verdin, E.R.I.C., 1991. DNase I-hypersensitive sites are associated with both long terminal repeats and with the intragenic enhancer of integrated human immunodeficiency virus type 1. *Journal of virology*, *65*(12), pp.6790-6799.

Verdin, E., Paras, P. and Van Lint, C., 1993. Chromatin disruption in the promoter of human immunodeficiency virus type 1 during transcriptional activation. *The EMBO journal*, *12*(8), pp.3249-3259.

Vivier, E., Raulet, D.H., Moretta, A., Caligiuri, M.A., Zitvogel, L., Lanier, L.L., Yokoyama, W.M. and Ugolini, S., 2011. Innate or adaptive immunity? The example of natural killer cells. *Science*, *331*(6013), pp.44-49.

Weinberger, L.S., Burnett, J.C., Toettcher, J.E., Arkin, A.P. and Schaffer, D.V., 2005. Stochastic gene expression in a lentiviral positive-feedback loop: HIV-1 Tat fluctuations drive phenotypic diversity. *Cell*, *122*(2), pp.169-182.

Weinberger, L.S. and Shenk, T., 2006. An HIV feedback resistor: auto-regulatory circuit deactivator and noise buffer. *PLoS biology*, *5*(1), p.e9.

Weinberger, L.S., Dar, R.D. and Simpson, M.L., 2008. Transient-mediated fate determination in a transcriptional circuit of HIV. *Nature genetics*, *40*(4), p.466.

Weinberger, A.D. and Weinberger, L.S., 2013. Stochastic fate selection in HIV-infected patients. *Cell*, *155*(3), pp.497-499.

West, M.J., Lowe, A.D. and Karn, J., 2001. Activation of human immunodeficiency virus transcription in T cells revisited: NF-κB p65 stimulates transcriptional elongation. *Journal of virology*, *75*(18), pp.8524-8537.

Whitney, J.B., Hill, A.L., Sanisetty, S., Penaloza-MacMaster, P., Liu, J., Shetty, M., Parenteau, L., Cabral, C., Shields, J., Blackmore, S. and Smith, J.Y., 2014. Rapid seeding of the viral reservoir prior to SIV viraemia in rhesus monkeys. *Nature*, *512*(7512), p.74.

Wong, K., Sharma, A., Awasthi, S., Matlock, E.F., Rogers, L., Van Lint, C., Skiest, D.J., Burns, D.K. and Harrod, R., 2005. HIV-1 Tat interactions with p300 and PCAF transcriptional coactivators inhibit histone acetylation and neurotrophin signaling through CREB. *Journal of Biological Chemistry*, *280*(10), pp.9390-9399.

Xing, S. and Siliciano, R.F., 2013. Targeting HIV latency: pharmacologic strategies toward eradication. *Drug discovery today*, *18*(11-12), pp.541-551.

Yamaguchi, Y., Inukai, N., Narita, T., Wada, T. and Handa, H., 2002. Evidence that negative elongation factor represses transcription elongation through binding to a DRB sensitivity-inducing factor/RNA polymerase II complex and RNA. *Molecular and cellular biology*, *22*(9), pp.2918-2927.

Yamamoto, N., Tanaka, C., Wu, Y., Chang, M.O., Inagaki, Y., Saito, Y., Naito, T., Ogasawara, H., Sekigawa, I. and Hayashida, Y., 2006. Analysis of human immunodeficiency virus type 1 integration by using a specific, sensitive and quantitative assay based on real-time polymerase chain reaction. *Virus genes*, *32*(1), pp.105-113.

Yang, Z., Zhu, Q., Luo, K. and Zhou, Q., 2001. The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription. *Nature*, *414*(6861), p.317.

Yedavalli, V.S., Benkirane, M. and Jeang, K.T., 2003. Tat and trans-activation-responsive (TAR) RNA-independent induction of HIV-1 long terminal repeat by human and murine cyclin T1 requires Sp1. *Journal of Biological Chemistry*, *278*(8), pp.6404-6410.

Yik, J.H., Chen, R., Nishimura, R., Jennings, J.L., Link, A.J. and 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. *Molecular cell*, *12*(4), pp.971-982.

Zeichner, S.L., Kim, J.Y. and Alwine, J.C., 1991. Linker-scanning mutational analysis of the transcriptional activity of the human immunodeficiency virus type 1 long terminal repeat. *Journal of virology*, *65*(5), pp.2436-2444.

Zhang, Z., Klatt, A., Gilmour, D.S. and Henderson, A.J., 2007. Negative elongation factor NELF represses human immunodeficiency virus transcription by pausing the RNA polymerase II complex. *Journal of Biological Chemistry*, *282*(23), pp.16981-16988.

Chapter 2

Materials and Methods

2.1 Generation of three independent Jurkat cell-based models to study HIV-1C latency

We manipulated the integrity of the LTR-Tat transcription axis and the Tat intracellular stability to develop three different cellular models that would examine the molecular mechanisms governing the HIV-1 latency in the Jurkat cells. In all the experimental models, we used a minimal reporter HIV-1 vector in which the GFP expression is placed under the control of the HIV-1 LTR to represent the transcriptional activity of the viral promoter. (1) **The autonomous Tat-feedback (ATF) model:** The HIV-1 minimal reporter vector, LTR-GFP-IRES-Tat (LGIT) expresses both GFP and Tat from the LTR, thus, the LTR-Tat positive feedback loop is preserved in this model as described previously (Weinberger LS et al., 2005; Burnett JC et al., 2009). We constructed a panel of five viral vectors with the number of NF-κB sites in the viral enhancer varying from 4 to 0 to evaluate the transcriptional strength of the viral promoter as a collaborative effect of the stoichiometry of the NF-κB elements and the concomitant Tat-feedback strength. (2) **The disjoint Tat-feedback (DTF) model:** In these cells, the LTR-Tat-feedback loop is disrupted by removing Tat expression from the control of the LTR and placing it under the control of an inducible Tet-ON promoter that in turn is controlled by varying the concentration of Doxycycline. Only the GFP expression is under the LTR in the LTR-GFP (LG) reporter viral vector. In this model, the transcription profile of the viral promoters of the panel is expected to be directly proportional only to the varying number of NF-κB motifs (4 to 0 copies) at a comparable intracellular Tat concentration. (3) **The tunable Tat-feedback (TTF) model:** In the TTF model, Tat is placed under the control of the LTR, however, the strength of the Tat-feedback is amenable due to a destabilization domain (DD) introduced at the C-terminus of the Tat ORF, shortening the latter's half-life to 2 h. The Tat protein, however, can be reversibly stabilized by a small molecule ligand– Shield1 (Banaszynski LA et al., 2006) that can protect and thereby prevent the 'Tat-DD' cassette from proteasomal degradation. The TTF model is expected to offer flexibility to examine the kinetics of viral latency-establishment as a function of the Tat-feedback strength alone in the context of the fixed viral promoter based on the presence or absence of the Shield1 molecule. The construction of each of these experimental cell models is described in detail below. Despite their technical merits in offering mechanistic insights into viral latency, we acknowledge that the *in vitro* models

of latency described here are minimalistic in their approach and are limited in their ability to represent the latent infections of the replication-competent viruses.

2.1.1 Construction of the LTR-GFP-IRES-Tat (LGIT) vector and the panel of NF-κB-variant viral vectors (the ATF model)

The ATF model consists of the minimal HIV-1 vector- LTR-GFP-IRES-Tat (LGIT) encoding a single 'GFP-IRES-Tat' transcript. From the transcript, the Tat protein is expected to be translated at a 10-100 fold lower efficiency as compared to that of EGFP as the viral protein is under the control of the IRES element (Mizuguchi H et al., 2000). The LGIT construct is devoid of the several intricate gene regulatory components such as splicing and other feedback components present in a full-length virus and focuses only on the narrow feedback loop existing between the LTR and Tat (Figure 2.1). In the present study, the LGIT minimal reporter vector was slightly modified as compared to the one originally reported (Weinberger LS et al., 2005). The initial LGIT vector obtained as a kind gift from Dr. David Schaffer (University of California, USA) was subsequently engineered to substitute Tat of subtype B origin with that of subtype C (BL4-3, GenBank accession number FJ765005.1) in the first step and substitute the 3' LTR of subtype B with that of subtype C (Indie_C1, Genbank accession number AB023804) subsequently to generate the c-LTR-GFP-IRES-cTat (cLGIT) vector (Verma A et al., 2016). Throughout the thesis, the model HIV-1 vector corresponding to the ATF latency model will be referred to as LGIT. Of note, in the ATF system, the expression levels of Tat at any point in time are a combined function of the strength of the viral promoter and the robustness of the Tat-feedback circuit, autonomously determined by the model.

Using the parental LGIT backbone, we constructed a panel of five reporter viral vectors that contained varying number of NF-κB binding sites, ranging from 4 to 0 copies in the viral enhancer of the 3' LTR (p911a vector series; Figure 2.1 and Appendix-1). Using the LTR of Indie_C1, an HIV-1C molecular clone (GenBank accession No. AB023804) as template, we generated a viral LTR of the 'FHHC' configuration containing four tandem NF-κB binding sites of three different sequences (H- GGGACTTTCC, C- GGGGCGTTCC, F- GGGACTTTCT) in an overlap-extension PCR. The additional 22 bp constituting the F-κB motif are adapted from the LTR of the HIV-1 subtype C molecular clone BL42-02 (GenBank accession No. HQ202921). Inactivating point mutations were subsequently introduced into the **'FHHC' (4-κB)** LTR sequentially using the overlap-

PCR to generate NF-κB copy-number variant LTRs: **OHHC (3-κB), OOHC (2-κB), OOOC (1- κB) and OOOO (0-κB)** as depicted schematically (Figure 2.1). In the variant LTRs, the critical residues in the target NF-κB site(s) were mutated without changing the total number of the base-pairs in the viral enhancer. The mutated κB site referred to here as 'O' contains the sequence TCTACTTTTT. The various primers used in the PCR are presented (Table 2.1). The variant LTR fragments were cloned directionally between the XhoI and PmeI sites present on the outer primers N1990 FP and N1991 RP thus, substituting the original 3' LTR of LGIT. The 3' LTR sequences of all the vectors were sequence-confirmed and the expression of EGFP was confirmed in HEK293T cells.

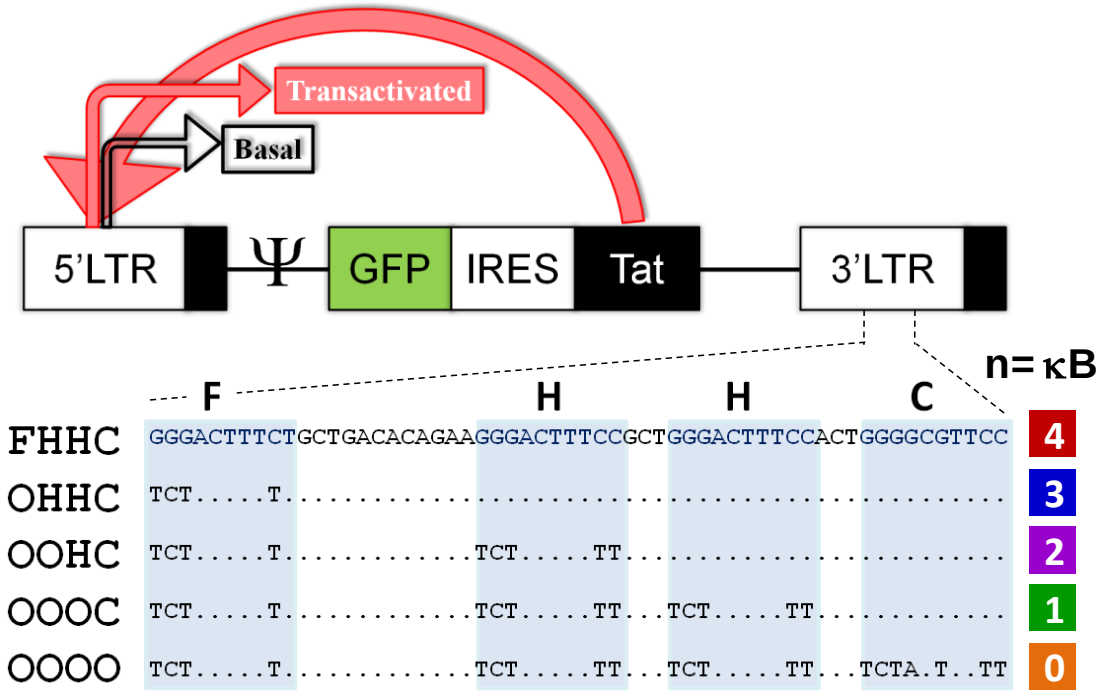


Figure 2.1: A panel of LTR-GFP-IRES-Tat (LGIT) vectors representing the 'Autonomous Tat feedback' (ATF) model of HIV-1 latency. A schematic representation of the autonomous 'LTR-Tat' transcriptional circuit in the minimal viral backbone LGIT has been depicted. In the absence or at low levels of Tat, the HIV-1 promoter (5' LTR) initiates a weak basal level transcription (thin black arrow) that doesn't lead to efficient transcription elongation. The increased levels of Tat strengthen the positive-feedback loop (thick pink arch) by rescuing the block in transcriptional elongation (~100 fold higher transcription) leading to Tat-mediated transactivation (thick pink arrow) which in turn leads to increased synthesis of Tat itself. In the LGIT minimal vector backbone, GFP and Tat are co-expressed from the same promoter, and their expression levels are expected to be proportional to the transcriptional strength of the LTR. Using the minimal LGIT design, two different panels (p911a and p911b series) each containing five LTR variants that differ in the copy-number of functional NF-κB motifs (4 to 0) have been constructed. Sequential point mutations (shaded in blue) have been introduced into NF-κB motifs to create the LTR variants, thus, maintaining the size of the viral promoter constant. The p911a series was constructed first encoding EGFP. The series p911b was constructed subsequently that encoded GFPd2 of shorter half-life (~2h) and represented the transcriptional activity of the LTR more reliably. Both the series of vectors were used in the ATF model of viral latency (see chapter 3). Vector schematic has been modified from Weinberger LS et al., 2005.

Of note, we also constructed a panel of five NF-κB-site variant viral vectors, analogous to the panel of p911a described above with the difference that a short-lived GFP reporter (GFPd2) was used in the place of EGFP (p911b vector series; Appendix-1). Later in the course of our study, we found that GFPd2 in the backbone of LTR-GFPd2-IRES-Tat (LdGIT) represented the transcriptional activity of the viral promoter more faithfully than the regular EGFP (see Figure 3.10, Chapter 3). The GFPd2 variant used in the panel contains a short half-life of 2 h instead of 48 h of EGFP (Li X et al., 1998) and serves as a rapid, reversible indicator for LTR switch-off or switch-on. To generate the p911b panel, we transferred the NF-κB variant LTRs from the p911a panel to the LdGIT vector backbone (Anjali Verma A thesis, 2015) using the PmeI and XhoI sites.

2.1.2 Construction of the LTR-GFP (LG) vector and the panel of NF-κB-variant viral vectors (the DTF model)

In the DTF model, the expression of Tat from the vector backbone LTR-GFP (LG) has been disassociated from the control of the LTR, thus, breaking the LTR-Tat-feedback loop. The expression of EGFP, however, remains under the control of the LTR. The expression of Tat is placed under the control of the chromatinized and inducible 'Tet-On-Tat' expression cassette; the levels of Tat, therefore, can be regulated by administering different doses of Doxycycline. Importantly, the DTF model, unlike the ATF model, permits the evaluation of the intrinsic transcriptional function of the viral promoters in the p912 panel (Figure 2.2 and Appendix-I) the members of which differ in the number of the NF-κB motifs, at a fixed and normalised intracellular concentration of Tat. First, the 'IRES-Tat' cassette was excised from LGIT-OHHC (LGIT-3-κB) vector using two enzyme sites, overhang ends compatible with each other, BsrGI (upstream of IRES) and BsiWI (downstream of Tat), and the digested backbone was subjected to self-ligation to generate viral vector LG-OHHC (LG-3-κB). The LG-3-κB vector was subsequently used as the parental vector to construct the four other members of the panel (4-, 2-, 1- and 0-κB) by substituting the 3' LTR between PmeI and XhoI with corresponding LTR variants from the LGIT panel (p911a series, section 2.1.1).

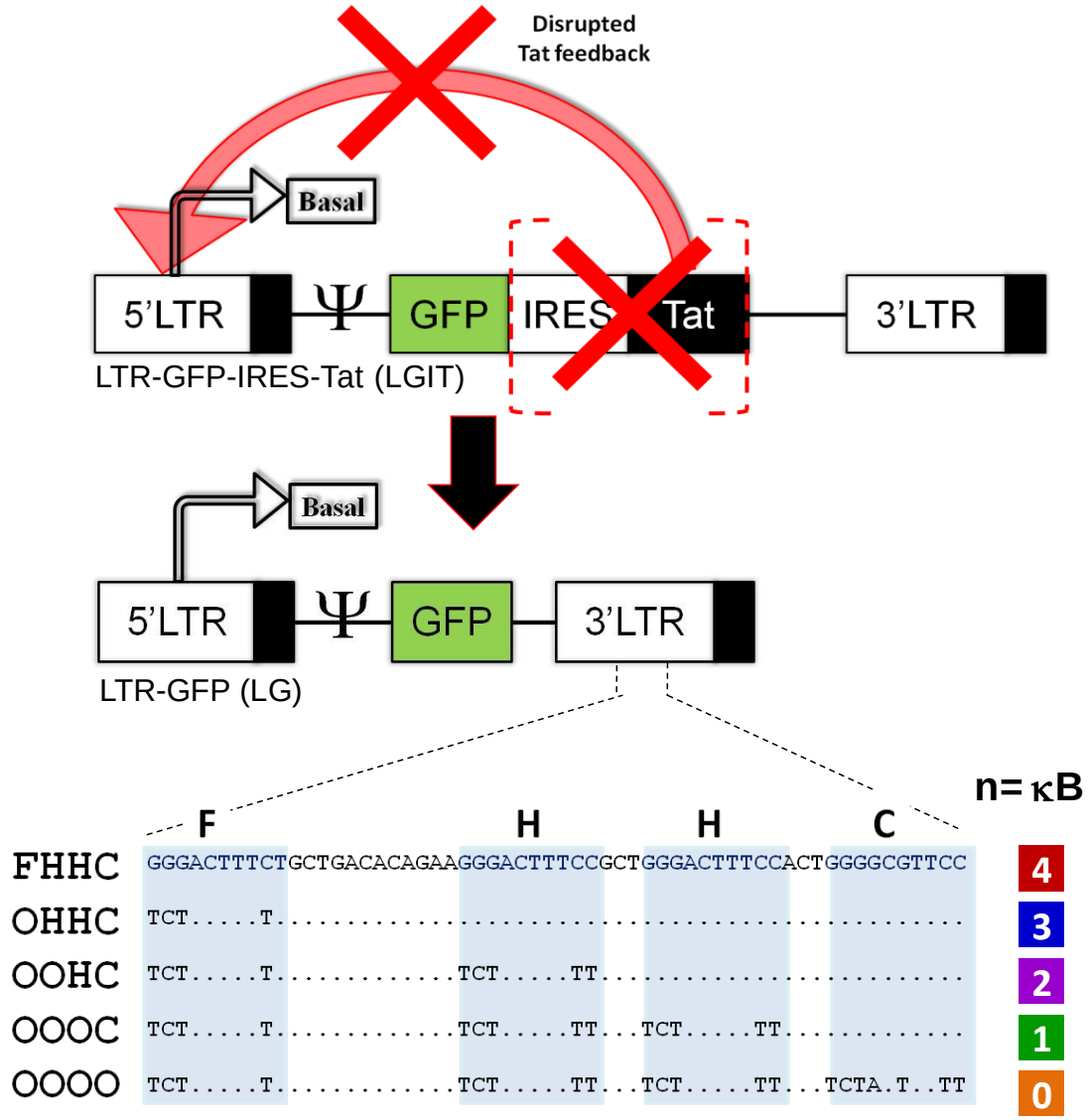


Figure 2.2: The LTR-GFP (LG) panel of vectors representing the 'Disjoint Tat feedback' (DTF) model of HIV-1 latency. A schematic representation of the disjoint 'LTR-Tat' transcriptional circuit in the minimal viral backbone LG has been presented. The LG viral backbone is derived from the LGIT backbone by deleting the 'IRES-Tat' cassette. The 5' LTR in the LG vector, therefore, drives only a basal level transcription of GFP in the absence of Tat (black arrow). The absence of Tat in the transcriptional circuit disrupts the LTR-Tat feedback axis (crossed-out pink arch) thereby, abolishing the Tat-mediated transactivation from the viral promoter. A panel of five LTR variants with varying copy numbers of NF-κB sites (4 to 0) has been created in the LG backbone to study the influence of only the LTR strength on viral latency. Of note, in the DTF model, Tat is expressed in *trans* from an inducible Tet-ON promoter (Figure 3.13, Chapter 3).

2.1.3 Construction of the LTR-GFPd2-IRES-Tat:RFP:DD (LdGITRD) vector and the panel of NF-κB-variant viral vectors (the TTF model)

It is technically challenging, if not impossible, to examine the LTR-Tat axis in determining the viral latency as the state of the cell and the host factors invariably influence the viral promoter and hence the outcome of transcription. It is, therefore, necessary to develop a system of the LTR-Tat circuitry which is amenable to

manipulation impervious to the influences of the host factors. To this end, we exploited the 'tunable protein regulation' system described previously (Banaszynski LA et al., 2006) by tagging the 12 kDa protein FK506 binding protein (FKBP) to Tat at the C-terminus. FKBP, a protein chaperone expressed in a broad range of tissues in many eukaryotes from yeasts to humans can target the associated proteins for rapid degradation. The engineering of FKBP has undergone many phases of maturation before assuming the present form (Clackson T et al., 1998; Egeler EL et al., 2011; An W et al., 2015). Endogenous proteins tagged with the FKBP domain can be reversibly rescued by a small molecule ligand Shield1 that specifically binds to FKBP preventing its degradation. In the present latency model, the c-Tat was fused with DsRed2-RFP and tagged with an FKBP destabilisation domain (DD) at its C-terminus to enable tuning of the Tat transcriptional circuit. The 5' LTR in the LTR-GFPd2-IRES-Tat:RFP:DD (LdGITRD) viral vector transcribes a single transcript encoding GFPd2 and a 1,314 bp long fusion cassette separated by the IRES element. The fusion cassette is a combination of three different ORFs- (i) a Tat expression segment (BL4-3, GenBank accession number FJ765005.1) (ii) the ORF of DsRed2-RFP, and (iii) the FKBP DD. The FKBP DD has been optimised to serve as a C-terminal tag that destabilizes the fused protein without modulating the stability of the protein in the presence of the Shield1 ligand (Clackson T et al., 1998; Egeler EL et al., 2011; An W et al., 2015). While GFPd2 represents the transcriptional activity of the viral promoter faithfully, the presence of DsRed2-RFP as a fusion protein of Tat permits the tracking of Tat in real-time. The three components of the 'Tat:RFP:DD' cassette were independently amplified using appropriate templates and primers and, finally using an overlap PCR, the fusion ORF was generated (see Table-2.1 and Appendix-I). The Tat ORF from the LGIT-3-κB vector (ATF model; Section 2.1 a) was replaced with the 'Tat:RFP:DD' ORF, thus, generating the LdGITRD-3-κB viral vector. LdGITRD-3-κB was subsequently used as the parental vector to construct the other member LdGITRD-1-κB vector of the panel p913 (Figure 2.3 and Appendix-I) by cloning the 3'LTR of LGIT-1-κB between PmeI and XhoI in the LdGITRD backbone. GFP expression from both the vectors of the panel p913 was confirmed in HEK293 cells.

We confirmed the Shield1-mediated tuning of Tat stability in an assay of transactivation using HEK293 cells. Approximately 0.6 x 10^5 cells were seeded in a 12-well dish and after a day the cells were transfected with 2 μg of purified pLdGITRD-3-κB in 1 ml of complete DMEM per well. Six hours post-transfection, the cells were washed and

replenished with 1 ml of freshly reconstituted complete DMEM supplemented with varying concentrations of Shield1 (0, 0.5, 1.0, 2.5, 4.0 or 5.0 μM). GFP and RFP images were captured using an inverted fluorescence microscope (Olympus IX-71, Olympus Corporation, Tokyo, Japan) 48 h following the transfection. A Shield1 dose dependent RFP expression was observed (Figure 3.19A, Chapter 3).

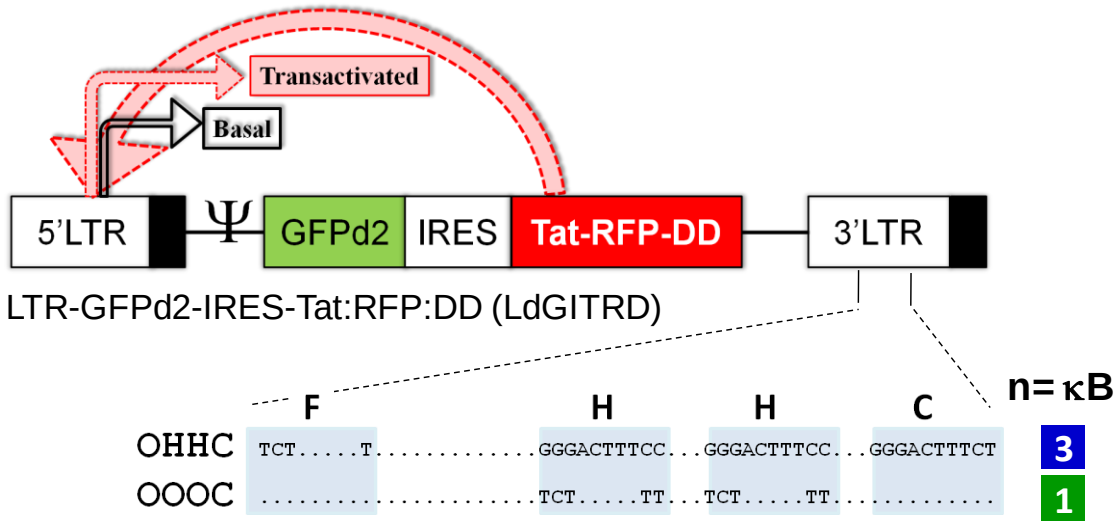


Figure 2.3: The LTR-GFPd2-IRES-Tat:RFP:DD (LdGITRD) panel of vectors representing the 'Tunable Tat feedback' (TTF) latency model. The LGIT minimal HIV-1 vector was further modified to construct the p911b series of vectors representing the ATF model. Tat in the ATF constructs was fused with DsRed2-RFP. Further, a C-terminal degradation domain (DD) of FKBP was added at the C-terminus of the fusion protein. The presence of DD permits fine tuning of the half-life of Tat-RFP by the use of Shield, a small-molecule ligand (represented as dotted pink arch). In the p913 panel of vectors, the 3- and 1-κB LTRs represented a strong and a weak promoter respectively. The TTF model was generated to examine the influence of the Tat-feedback strength on latency at a fixed LTR strength.

2.2 Generation of stable Jurkat cells with an integrated provirus

2.2.1 The ATF model (Jurkat-LGIT/Jurkat-LdGIT): Jurkat cells were individually infected with the NF-κB copy-number variant viral strains of the panel (4 – 0 copies of NF-κB motifs) and sorted by GFP expression. Pseudotyped viral stocks of the LGIT panel were added at an RIU of 0.05-0.1 to 1 x 10^6 Jurkat cells in a total volume of 2 ml of complete RPMI supplemented with 25 μg/ml DEAE-Dextran in a 35 mm tissue culture dish. After six hours of infection, cells were washed to remove DEAE-Dextran and transferred to a T-25 flask containing 5 ml of complete RPMI and maintained under standard culture conditions. The infected cell pools were expanded over a period of seven days and, then induced with a cocktail of T-cell activation agents (40 ng/ml PMA + 40 ng/ml TNFα + 200 nM TSA + 2.5 mM HMBA). GFP^{+ve} cells were sorted from each pool 18 h post-activation using a cell sorter (BD FACS AriaIII cell sorter, BD biosciences). A

small aliquot of the sorted cell population was re-analysed to confirm the purity of the GFP^{+ve} cells. The sorted cell pool with stable GFP expression represented a random population of proviral integrants of the corresponding NF-κB variant viral strain. We confirmed the integration frequency in all the cell pools using a real-time PCR.

A Taqman qPCR was performed using the genomic DNA of the stable J-LGIT cell pools to determine the mean number of proviral integrations per cell in all the stable cell pools. Genomic DNA was extracted from 1 x 10^6 stable cells using the GenElute mammalian genomic DNA kit (G1N350, Sigma-Aldrich) following manufacturer's protocol. The extracted DNA was dissolved in TE and the concentration was adjusted to 70 ng/µl. Five µl of this solution would correspond approximately to 10^5 copies of the human genome. The stock DNA solution was subjected to a 10-fold serial dilution up to a final DNA concentration of 10^1 copies/5 µl and used as template in the PCR. A 129 bp fragment spanning the R-U5 region of the HIV-1 5' LTR (+18 to +147) was amplified using the primer-probe combination N2208 FP, N2209 RP and N2210 FAM (see Table 2.2 for the primer sequences) in a Taqman real-time PCR. A standard curve was established simultaneously using the genomic DNA extracted from the J-Lat 8.4 cells that contain a single proviral copy per cell (Jordan A et al., 2003). The proviral copy number of the query samples was then estimated using the regression analysis.

To establish stable clonal lines of the viral variants, we sorted single cell per well in a 96-well plate. The wells contained a mix of 50-50 µl each of complete- and the spent-RPMI media. The cells were diluted to a cell density of 0.1 x 10^6/ml prior to the sort. The sorter was programmed to sort single cells of the highest GFP intensity (MFI ~10^5) per well of the 96-well plate. Cells of a viral variant of the panel were sorted into a single plate.

2.2.2 The DTF model (Jurkat-LG): The final Jurkat-LG (J-LG) cell line corresponding to the DTF model was generated using three different lentiviral vectors (Figure 3.13, Chapter 3) and three sequential steps of infection and selection. In the first step, the cells were stably transduced with rtTA3 and selected against puromycin. Jurkat cells were infected with pseudotyped CMV-rtTA3-IRES-Puro (p812) virus at a low RIU (< 0.1) to ensure single integration per cell. The stable cells were selected against puromycin (800 ng/ml), and individual cell lines were established by limit-dilution (Malini Menon thesis, 2017). One of the clonal lines 'J-rtTA3-Cl5' was used for the subsequent experiments.

In the second step, J-rtTA3-Cl5 cells were stably transduced with the Tat expression cassette under the control of a CMV promoter. The recombinant lentiviral vector 'pcDH-CMV-Tet-ON-Tat-EF1α-ECFP' (p815.CS.CFP), based on the inducible 'Tet-ON-Tat' cassette in the pcDH backbone (Catalog no. CD511A-1, Clonetech, California, USA), contains Tat derived from the BL4-3 HIV-1C molecular clone. The EF1α–EGFP cassette in vector p815.CS (Malini Menon thesis, 2017) was replaced by EF1α–ECFP cassette, thus, placing the Tat expression under the CMV promoter. The EF1α promoter and the ECFP ORF were amplified using p815.CS and pCFP-C1 (Addgene vector no. 6076-1) vectors as templates, respectively. Subsequently, the EF1α–ECFP cassette was generated from the two purified PCR fragments in an overlap PCR using external (N2362 FP and N2365 RP) and internal (N2363 FP and N2364 RP) primers (see Table 2.1 for the primer sequences). The EF1α–ECFP cassette was cloned between the enzyme sites SalI and XbaI in p815.CS. The cells infected with the pcDH-CMV-Tet-ON-Tat-EF1α-ECFP viral vector and selected in flow sorting for low-level fluorescent intensity (MFI range ~500-5,000). Cells were sorted into 96-well plates, 1 cell/well, and about 20 clonal cell lines (J-rtTA3-cTat lines) were expanded and monitored for stable ECFP expression for about two months.

Three clonal lines - 1a, 2c, and 4b- were selected and scored for their ability to support transcription from the LTR by Dox-induced Tat. Half a million cells of each clonal line were suspended in RPMI medium supplemented with 25 μg/ml of DEAE and 750 ng/ml Dox for 24 h. Subsequently, the cells were infected with LG-3-κB virus expressing EGFP under LTR. Six hours post-infection, the cells were washed and resuspended in complete RPMI medium supplemented with 750 ng/ml Dox. After 48 h, the clones were scored for the GFP^{+ve} expression using flow cytometry. J-rtTA3-2c cell line that showed maximal Dox-induced LTR transactivation (Figure 3.14B, Chapter 3) was considered for subsequent experiments. We also found a positive correlation between the increasing concentrations of Dox (0, 250, 500, 750 and 1,000 ng/ml Doxycycline for 48 h) and the induction in the Tat transcripts as measured in a real-time PCR normalized with the housekeeping gene GAPDH in a ΔΔCt method (Figure 3.14A, Chapter 3).

Finally, in the third step, we generated the stable Jurkat cells for each member in panel p912. Stable cell pools of J-LG panel expressing both ECFP and EGFP were also established by flow sorting using a protocol essentially similar to that of the p911a and

p911b panels described above (section 2.2.1) with a few modifications. The Jurkat cells were pre-treated with 750 ng/ml of Doxycycline (Dox; D9891, Sigma-Aldrich) for 24 h prior to the infection with the viral strains. Dox-treated cells were infected with pseudotyped LG viruses at an RIU of 0.01-0.05 in complete RPMI supplemented with 25 μg/ml of DEAE-dextran and 750 ng/ml of Dox. Six hours post-infection, the cells were washed and resuspended in 1 ml complete RPMI supplemented with Dox. The infected cells were allowed to expand for a week under standard growth conditions with the replenishment of Dox every 24 h. Seven days post-infection the cells were induced with an activation cocktail (40 ng/ml PMA + 40 ng/ml TNFα + 200 nM TSA + 2.5 mM HMBA) and 18 h post-activation, CFP^{+ve} GFP^{+ve} double positive cells representative of randomly integrated proviruses (J-LG; WT or NF-κB variant) were sorted.

2.2.3 The TTF model (Jurkat-LdGITRD): 1 x 10^6 Jurkat cells were infected with the minimal LdGITRD viral strains (3- and 1-κB variants) at an RIU of ~0.1-0.2 in 1 ml of complete RPMI supplemented with 25 μg/ml of DEAE-dextran and 1.0 μM Shield1 (#632189, Takara Clontech). Six hours post-infection, the infected cells (J-LdGITRD) were washed and replenished with 1 ml of complete RPMI supplemented with 1 μM Shield1. Post 72 h of infection, the cells were induced with an activation cocktail (40 ng/ml PMA + 40 ng/ml TNFα + 200 nM TSA + 2.5 mM HMBA) and 18 h post-activation, GFP^{+ve} cells were sorted that harboured integrated proviruses.

2.3 General techniques

2.3.1 Cell culture

Jurkat cells were maintained in RPMI 1640 medium (R4130, Sigma-Aldrich, St. Louis, USA) supplemented with 10% fetal bovine serum (RM10435, HiMedia Laboratories, Mumbai, India), 2 mM glutamine (G8540, Sigma-Aldrich), 100 units/ml penicillin G (P3032, Sigma-Aldrich) and 100 g/ml streptomycin (S9137, Sigma-Aldrich). The human embryonic kidney cell lines HEK293 and HEK293T cells were cultured in Dulbecco's modified Eagle's medium (D1152, Sigma-Aldrich) supplemented with 10% FBS. All the cells were incubated at 37^0C in the presence of 5% CO_2.

2.3.2 Generation of pseudotyped reporter virus

Pseudotyped reporter viruses pertaining to all the HIV-1 latency models (described in sections 2.1 and 2.2) were generated in HEK293T cells by transfecting each individual viral vector together with 3rd generation lentiviral packaging vectors using the standard calcium phosphate protocol (Jordan M et al., 1996). Briefly, a plasmid DNA cocktail consisting of 10 μg of the viral vector (wild-type or NF-κB mutants), 5 μg psPAX2 (#11348; NIH AIDS reagent program, USA), 3.5 μg pHEF-VSVG (#4693; NIH AIDS Reagent program) and 1.5 μg pCMV-rev (#1443; NIH AIDS Reagent program) was transfected in a 100 mm dish seeded with HEK293T at 30% cell confluence. pCMV-RFP (0.2 μg) was used as an internal control for transfection. Six hours post-transfection, the medium was replenished with complete DMEM. Culture supernatants were harvested at 48 h post-transfection, filtered using 0.22 μ filter and stored in 1 ml aliquots in a deep freezer for future use.

2.3.3 Estimation of relative infectious units (RIU) of the NF-κB-variant, minimal reporter viruses in each panel

The RIU was quantified in Jurkat T-cells by monitoring GFP expression by flow cytometry. Precisely, 3 x 10^4 Jurkat cells in each well of a 12-well tissue culture plate were infected with viral stocks serially diluted 2-fold (from 10 xd to 80 xd) in a total volume of 1 ml of 10% RPMI containing 25 μg/ml of DEAE-Dextran. Six hours post-infection, the cells were washed and replenished with 1 ml of complete RPMI. Post 48 h, the cells were activated with a combination of 40 ng/ml PMA (P8139, Sigma Aldrich), 40 ng/ml TNFα (T0157, Sigma-Aldrich), 200 nM TSA (T8552, Sigma Aldrich) and 2.5 mM HMBA (224235, Sigma-Aldrich) for 18 h, following which percent GFP^{+ve} cells were analysed by a flow cytometer (BD FACSAriaIII sorter, BD biosciences, New Jersey, USA). Following this, we constructed titration curves and determined 5-10% infectivity of the cells by regression analysis which would correspond to ~0.05-0.1 RIU. All the flow cytometry data were analysed using FCS express 4 and 6 versions (De Novo Software, Los Angeles, CA)

2.4 Molecular techniques

2.4.1 Tat-transcript estimation in the stable Jurkat cells

Transcription from the viral promoter was monitored using a real-time PCR for Tat, in addition to GFP expression, as indicative of the functioning of the LTR during the course of latency-establishment, or activation of a latent provirus. Total mRNA was extracted at specified time points from 0.5 x 10^6 cells using a single-step RNA isolation reagent- TRI reagent (T9424, Sigma-Aldrich). 250-1,000 ng of extracted RNA was converted to cDNA using random hexamer primers in a reaction volume of 20 µl using the Tetro cDNA synthesis kit (BIO-65043, Bioline, London, UK). The cDNA was then amplified using a SYBR green RT-PCR kit (06924204001, Roche Products, Mumbai, India) for a 139 bp region in Tat exon-1 using the primer pair N1783-N1784 while GAPDH was used as the housekeeping gene control (primer details in Table 2.2). The relative gene expression at each time point was calculated using the ΔΔ Ct method.

2.4.2 Chromatin immunoprecipitation assay

Chromatin preparation equivalent of 2 x 10^6 cells was used per immunoprecipitation assay following the protocol described previously (Verma A et al., 2016). Briefly, 2 x 10^6 Jurkat cells collected in a 1.5 ml vial were washed with 1X PBS, resuspended in 1 ml of RPMI supplemented with 1% formaldehyde and incubated with gentle agitation for 10 min at room temperature. The cross-linking reaction was quenched by incubating the cells with 0.125 M glycine for 5 min with mild agitation at room temperature. The cells were then centrifuged at 3,000 rpm for 5 min at 4^0C followed by a wash with 1X PBS (containing 0.01X protease inhibitor cocktail or PIC; #11836170001, Roche Applied Science, Indianapolis, USA). The PBS was carefully removed and the cells were resuspended in 100 µl of ice-chilled lysis buffer (1% SDS, 50 mM Tris buffer, pH 8.0, 10 mM EDTA). The lysis reaction was incubated on ice for 20 min with occasional mixing of the lysate using a wide-bore tip. The lysate in each vial was then subjected to 22 cycles of sonication at the high mode, using 30-second-ON followed by 30-second-OFF pulse scheme in the Bioruptor plus sonicator (UCD-300, Diagenode, Liege, Belgium) containing pre-chilled water. The sonicated lysate was centrifuged at 12,000 rpm for 10 min at 4^0C to remove any cellular debris; the clear supernatant was transferred to a fresh 1.5 ml vial, and stored at -80^0C until use. One-tenth of the lysate (10 µl) was used to

confirm the magnitude of chromatin shearing for the generation of the fragment sizes ranging from 200 to 500 bp. Each immunoprecipitation reaction comprised of 100 µl of lysate and 2 µg of an antigen-specific antibody against p50 (ab7971, Abcam, Massachusetts, USA), p65 (ab7970, Abcam), NFAT1 (ab2722, Abcam), NFAT2 (ab2796, Abcam), HIV-1 Tat (ab43014, Abcam; #7377, NIH AIDS reagent program; #4374, NIH AIDS reagent program), RNA Pol II CTD phospho S2 (ab5095, Abcam), or H3K9 Tri Meth (ab8898, Abcam). The immunoprecipitated DNA was amplified using the primer pair N1054 FP (5'-GATCTGAGCC(T/C)GGGAGCTCTCTG-3') and N1056 RP (5'-TCTGAGGGATCTCTAGTTACCAGAGTC-3') that spans a 240 bp region comprising the enhancer-core promoter region in the LTR. The amplified DNA fragments were subjected to agarose gel electrophoresis and the band intensities were normalised using the percent-input method to compare differential recruitment of each transcription factor at the active vs. latent promoter. To enhance the sensitivity of the assay, TaqMan qPCR was performed on the ChIP-DNA using the primer-probe combination- N2493 FP, N2215 RP and N2492 Hex (refer to Table 2.2). The final data were evaluated using the percent input method.

2.5 Microscopy techniques

2.5.1 Live-cell, time-lapse microscopy

Jurkat cells, stably integrated with the NF-κB variant, LdGIT viral strains (J-LdGIT) were subjected to live-cell, fluorescent, time-lapse microscopy to measure the dynamics of the promoter reactivation at the level of a single-cell. Approximately 30,000 stable J-LdGIT cells were used per assay. A 35mm glass-bottom cell-culture dish was coated for 45 min at room temperature with a thin film of the BD Cell-Tak reagent, an immobilizing agent comprising of protein formulations extracted from the marine mussels (#354240, Corning, New York, USA). The dish was washed twice with sterile water and air dried for 5 min. Cells suspended in 160 µl complete RPMI were deposited on the coated dish to form a uniform monolayer and allowed to stand for 15 min. An appropriate microscopic field with optimum number of immobilized cells was selected and subjected to uninterrupted imaging of GFP fluorescence for 24 h using an inverted fluorescent microscope (Zeiss LSM 880) containing a motorized stage. During imaging, the cells were maintained under optimal culture conditions consisting of 5% CO2, 70-90% humidity, and 37°C temperature. The cells were excited and images captured every 15

min using a 10 mW 488 nm argon laser and a Plan-Apochromat X20/0.8 objective operated at 0.6% power. To monitor reactivation of the latent proviral promoter, GFP^{-ve} cells were sorted from the stable cell pool, the sorted cells were activated with TNFα (40 ng/ml) and selected cells were tracked for GFP expression every 15 min for 24 h. In each assay, we collected the GFP trajectories of 25-30 individual cells as a function of time and constructed the average GFP trajectories to monitor the function of the viral promoter. The microscope hardware and software were adjusted to correct for any drift in the focal plane during the assay. Simultaneously, using a 561 nm argon laser, we also monitored the stable expression of RFP under the control of an EF1α promoter from every cell at all the time points as a normalization control for the varying expression levels of GFP under the LTR. The GFP data from only those cells that showed a uniform RFP expression throughout the course of imaging were considered and analysed. Further, values of 'Hill coefficient' were computed from the GFP trajectories of each variant LTR following the protocol mentioned in Razooky BS et al., 2011.

2.5.2 Indirect immunofluorescence of Tat

Immunofluorescence staining was performed to trace the presence of Tat in different sub-cellular compartments at multiple time points during the establishment of viral latency. Stable J-LdGIT-3-κB cells characterised by strong GFP fluorescence (MFI range 5,000-50,000) were sorted and used as the D-0 sample. Approximately, 3 x 10^6 cells were collected in a 1.5 ml vial, washed once with 1X PBS and fixed with 2% paraformaldehyde in PBS for 10 min at room temperature with mild rocking. Fixed cells were washed once again followed by permeabilization with 0.2% Triton-X-100 in PBS for 10 min with gentle intermittent vortexing. Cells were washed once and blocked with 4% BSA in PBS for 30 min at room temperature with mild rocking. The cells were incubated with Rabbit anti-Tat antibody (ab43014, Abcam) at 1: 250 dilution for 1h at room temperature followed by two PBS washes. This was followed by incubation with a 1:500 dilution of Goat anti-rabbit Alexa Fluor 568 (A-11010, Molecular Probes, Thermo Fisher Scientific) for 20 min in the dark at room temperature followed by a PBS wash. For the nuclear staining, 4 μg/ml of DAPI was used for 20 min in the dark at room temperature. Cells were washed twice and mounted with 70% glycerol on coverslips for the confocal imaging. Images were acquired with a Zeiss LSM 880 confocal laser

scanning microscope with Airyscan using a Plan Apochromat X63/1.4- oil immersion objective and analyzed using the ZEN 2.1 software.

2.5.3 The proximity ligation assay

To identify the cellular proteins differentially interacting with Tat in the context of the active vs latent viral promoter, an in situ proximity ligation assay (PLA) was performed using the stable HEK293 cells infected with the LdGIT-4-κB reporter virus. A commercial kit (Duolink In Situ Red Starter kit Mouse/Rabbit, #DUO92101, Sigma-Aldrich) was used to perform the assay following the instructions of the manufacturer. HEK293 cells harbouring active or latent LdGIT-4-κB virus, marked by the presence or absence of green fluorescence, respectively, were seeded on glass coverslips and allowed to grow to 60-70% confluence. The cells, evenly spread on the coverslip were fixed with 4% paraformaldehyde for 20 min at room temperature, permeabilized with 0.1% Triton-X-100 for 10 min at room temperature and washed thrice with 1X PBS. The cells were blocked for one hour with the blocking reagent supplied in the kit. The blocked cells were then treated with the rabbit polyclonal anti-Tat antibody at 1: 250 dilution (Catalog no. ab43014, Abcam) in combination with either (1) mouse monoclonal anti-Tat antibody at 1: 250 dilution (Catalog no. 7377, NIH AIDS reagent program); or (2) mouse monoclonal anti-p65 antibody at 1: 300 dilution (Catalog no. 6956S, Cell Signaling Technology, MA, USA). The cells were incubated with a pair of probes (the PLA probe Anti-Mouse MINUS; DUO92004 and PLA probe Anti-Rabbit PLUS; DUO92002) in a 40 μl reaction volume, for one hour at 37^0C followed by washing twice with 500 μl of wash buffer A for 5 min each. The ligation and amplification reactions were performed as per manufacturer's instructions using the Duolink In Situ Detection reagents Red (Catalog no. DUO92008). The DAPI-supplemented mounting medium (Catalog no. DUO82040, supplied in the PLA kit) was used for mounting the cells. Imaging of the cells was performed using a Zeiss LSM 880 confocal laser scanning microscope with Airyscan fitted with a Plan Apochromat 63X/1.4-oil immersion objective. Signal intensities from PLA positive spots were quantitated manually using the Image J software.

Table 2.1 Cloning primers used in the generation of the three latency models

Primers used for the introduction of point mutations to create NF-κB copy-number variant LTRs (section 2.1.1)			
LTR-Variant	**Primer pair**	**Description**	**Sequence of primers (5'-3')**
	N1990 FP N1991 RP	Outer primers	GCGTACCTCGAGTGGAAGGGTTAATTTACTCCAAGAAAAGGC TATGTCGTTTAAACCTGCTAGAGATTTTCCACACTACCAAAAGGGTCTGAG
FHHC (4-κB)	N1992 FP N1993 RP	Inner primers	TGACACAGAAGGGACTTTCTGCTGACACAGAAGGGACTTTCCGCTGGGACTTTCCACTGGGGCGTTCC AAGTCCCAGCGGAAAGTCCCTTCTGTGTCAGCAGAAAGTCCCTTCTGTGTCAGCAGTCTTTGTAAAACTCCG
OHHC (3-κB)	N1994 FP N1995 RP		TGACACAGAATCTACTTTTTGCTGACACAGAAGGGACTTTCCGCTGGGACTTTCCACTGGGGCGTTCC AAGTCCCAGCGGAAAGTCCCTTCTGTGTCAGCAAAAAGTAGATTCTGTGTCAGCAGTCTTTGTAAAACTCCG
OOHC (2-κB)	N1996 FP N1997 RP		TGACACAGAATCTACTTTTTGCTGACACAGAATCTACTTTTTGCTGGGACTTTCCACTGGGGCGTTCC AAGTCCCAGCAAAAAGTAGATTCTGTGTCAGCAAAAAGTAGATTCTGTGTCAGCAGTCTTTGTAAAACTCCG
OOOC (1-κB)	N1998 FP N1999 RP		TGACACAGAATCTACTTTTTGCTGACACAGAATCTACTTTTTGCTTCTACTTTTTACTGGGGCGTTCC AAGTAGAAGCAAAAAGTAGATTCTGTGTCAGCAAAAAGTAGATTCTGTGTCAGCAGTCTTTGTAAAACTCCG
OOOO (0-κB)	N2000 FP N2001 RP		TGACACAGAATCTACTTTTTGCTGACACAGAATCTACTTTTTGCTTCTACTTTTTACTTCTACTTTTTAGG AAGTAGAAGCAAAAAGTAGATTCTGTGTCAGCAAAAAGTAGATTCTGTGTCAGCAGTCTTTGTAAAACTCCG
Primers used for cloning the LdGITRD vector backbone (p913 series) (section 2.1.3)			
Amplicon	**Amplicon length**	**Primer pair**	**Sequence of primers (5'- 3')**
GFPd2-IRES-Tat	1779 bp	N2720 FP N2724 RP	TTTCTTCCATTGCGGCCGCGCCGCCACCATGGCCTCCTCCGAGAACGTC GGCCATTTCGAAGTCGAAGGGGTCT
DsRed2-RFP	629 bp	N2723 FP N2726 RP	ACTTCGAAATGGCCTCCTCCGAGAACG TCCGATATCCAGGAACAGGTGGTGGC
FKBP DD	353 bp	N2725 FP N2722 RP	TGTTCCTGGATATCGGAGTGCAGGTGGAAACCATC CGTACGCGGCGCGCCTCATTCCAGTTCTAGAAGCTCC

Primers used for cloning the pcDH-CMV-Tet-ON-Tat-EF1α-ECFP vector backbone (p815.CS.CFP) (section 2.2.2)			
Amplicon	**Amplicon length**	**Primer pair**	**Sequence of primers (5'- 3')**
EF1a-ECFP	1271 bp	N2362 FP	AAATGGATCCGCGGCCGCAAGG
		N2365 RP (outer)	CCTCTAGTCGACTTACTTGTACAGCTCGT
		N2364 FP	CGGCGCCTACTCTAGAATGGTGAGCAAGGGCGAGG
		N2363 RP (inner)	CCTCTAGTCGACTTACTTGTACAGCTCGT

Table 2.2 Details of primers used in various qPCR Assays

Target region	**Primer pair/ primer-probe combination**	**Primer/probe sequence (5'-3')**
Primers used to determine the proviral integration frequency in the NF-κB-variant, J-LGIT stable cells (section 2.2.1)		
Strong-stop DNA	N2208 FP	GATCTGAGCC(T/C)GGGAGCTCTCTG
	N2209 RP	TCTGAGGGATCTCTAGTTACCAGAGTC
	N2210 probe	FAM-CTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCT-TAMRA
Primers used to estimate Tat-transcript levels in stable Jurkat cells (section 2.4.1)		
Tat Exon1	N1783 FP	GGAATCATCCAGGAAGTCAGCCCGAAAC
	N1784 RP	CTTCGTCGCTGTCTCCGCTTCTTCCTG
GAPDH	N2232 FP	GAGCTGAACGGGAAGCTCACTG
	N2233 RP	GCTTCACCACCTTCTTGATGTCA
Primers used in Chromatin immunoprecipitation assay (section 2.4.2)		
LTR-enhancer	N2493 FP	CCGGAGT(A/T)TTACAAAGACTGCTG
	N2215 RP	CTGCTTATATGCAGCATCTGAGG
	N2492 probe	HEX- CACTGGGGCGTTCCAGG(G/A)GG(A/T)GT -BHQ

2.6 Bibliography

An, W., Jackson, R.E., Hunter, P., Gögel, S., Van Diepen, M., Liu, K., Meyer, M.P. and Eickholt, B.J., 2015. Engineering FKBP-based destabilizing domains to build sophisticated protein regulation systems. *PloS one*, *10*(12), p.e0145783.

Banaszynski, L.A., Chen, L.C., Maynard-Smith, L.A., Ooi, A.L. and Wandless, T.J., 2006. A rapid, reversible, and tunable method to regulate protein function in living cells using synthetic small molecules. *Cell*, *126*(5), pp.995-1004.

Burnett, J.C., Miller-Jensen, K., Shah, P.S., Arkin, A.P. and Schaffer, D.V., 2009. Control of stochastic gene expression by host factors at the HIV promoter. *PLoS pathogens*, *5*(1), p.e1000260.

Clackson, T., Yang, W., Rozamus, L.W., Hatada, M., Amara, J.F., Rollins, C.T., Stevenson, L.F., Magari, S.R., Wood, S.A., Courage, N.L. and Lu, X., 1998. Redesigning an FKBP–ligand interface to generate chemical dimerizers with novel specificity. *Proceedings of the National Academy of Sciences*, *95*(18), pp.10437-10442.

Egeler, E.L., Urner, L.M., Rakhit, R., Liu, C.W. and Wandless, T.J., 2011. Ligand-switchable substrates for a ubiquitin-proteasome system. *Journal of Biological Chemistry*, *286*(36), pp.31328-31336.

Jordan, A., Bisgrove, D. and Verdin, E., 2003. HIV reproducibly establishes a latent infection after acute infection of T cells in vitro. *The EMBO journal*, *22*(8), pp.1868-1877.

Jordan, M., Schallhorn, A. and Wurm, F.M., 1996. Transfecting mammalian cells: optimization of critical parameters affecting calcium-phosphate precipitate formation. *Nucleic acids research*, *24*(4), pp.596-601.

Li, X., Zhao, X., Fang, Y., Jiang, X., Duong, T., Fan, C., Huang, C.C. and Kain, S.R., 1998. Generation of destabilized green fluorescent protein as a transcription reporter. *Journal of Biological Chemistry*, *273*(52), pp.34970-34975.

Mizuguchi, H., Xu, Z., Ishii-Watabe, A., Uchida, E. and Hayakawa, T., 2000. IRES-dependent second gene expression is significantly lower than cap-dependent first gene expression in a bicistronic vector. *Molecular Therapy*, *1*(4), pp.376-382.

Razooky, B.S., Pai, A., Aull, K., Rouzine, I.M. and Weinberger, L.S., 2015. A hardwired HIV latency program. *Cell*, *160*(5), pp.990-1001.

Verma, A., Rajagopalan, P., Lotke, R., Varghese, R., Selvam, D., Kundu, T.K. and Ranga, U., 2016. Functional Incompatibility between the Generic NF-κB Motif and a Subtype-Specific Sp1III Element Drives the Formation of the HIV-1 Subtype C Viral Promoter. *Journal of virology*, *90*(16), pp.7046-7065.

Weinberger, L.S., Burnett, J.C., Toettcher, J.E., Arkin, A.P. and Schaffer, D.V., 2005. Stochastic gene expression in a lentiviral positive-feedback loop: HIV-1 Tat fluctuations drive phenotypic diversity. *Cell*, *122*(2), pp.169-182.

Weinberger, L.S., Dar, R.D. and Simpson, M.L., 2008. Transient-mediated fate determination in a transcriptional circuit of HIV. *Nature genetics*, *40*(4), p.466.

Chapter 3

HIV-1C LTR with an intact Tat positive-feedback undergoes rapid silencing with progressive increase in the copies of NF-κB motifs

3.1 Introduction

A single model cannot perfectly replicate all the aspects of HIV-1 latency; the choice of an appropriate model depends on the problem being investigated. Human clinical trials undoubtedly furnish the closest natural setting to the study of HIV latency. Given the ethical issues and limitations of therapeutic interventions, in vivo animal models such as the SIV-Macaque model or the humanized mice models stand the next best in the physiological relevance to natural infection (Archin NM et al., 2014). The ex vivo cell-line models represent relatively simple models permitting a near-perfect recapitulation of viral latency. These models can also aid in the initial screening of latency-reversal agents (LRAs). Among the cell-line models, the J-Lat model shed significant light on the chromatin-remodelling mechanisms and the associated transcription factors involved in HIV-1 latency (Mahmoudi T et al., 2006; Williams SA et al., 2006). This latency model also contributed significantly to the understanding of latency-reactivation (Karn J et al., 2011) and the mechanisms of the Tat-mediated stochastic regulation of HIV-1 latency (Weinberger LS et al., 2005).

In the absence of an appropriate marker for HIV latency, the fluorescently labelled HIV-1 vectors have made a remarkable contribution to examine the mechanisms of viral latency in the recent years. The engineering of a fluorescent protein into a full-length HIV-1 molecular clone could often prove challenging because of the viral packaging restrictions, the risk of perturbing the functional overlap of the ORFs in the viral genomes and additional intrinsic difficulties. The pseudotyped reporter HIV-1 vectors, therefore, have been popular as the substitutes of the replication-competent viral strains as model vectors to study basic mechanisms of viral latency. The minimal reporter viral constructs consisting of only the viral LTR and Tat have been instrumental in discerning the basic mechanism of transcriptional silencing and reversal processes. One of the most widely accepted reporter vectors to study HIV-1 latency is the LTR-GFP-IRES-Tat (LGIT) vector, first documented by Weinberger et al who used it to identify the influence of the Tat mediated positive auto-regulation in deciding between an active and latent phenotypes in subtype B (Weinberger LS et al., 2005).

3.2 The study rationale

The examination of the mechanisms governing latency in HIV-1C is technically more challenging as compared to that of HIV-1B for two different reasons. First, HIV-1C contains a larger number of NF-κB elements- three or four, in the enhancer as opposed to only two or one such elements in the other HIV-1 families (Bachu M et al., 2012a; Boullosa J et al., 2014). This might appear paradoxical since, a progressive gain in additional NF-κB motifs should disfavour the virus to establish and maintain latency. Second, the additional NF-κB binding elements are genetically diverse. Further, the genetically distinct, acquired NF-κB motifs unique to HIV-1C are spatially conserved with respect to other TFBS in the LTR implicating possible mechanistic differences in signal transduction through these sites. The impact of such genetic diversity on the establishment and maintenance of transcriptional silence has not been examined previously. HIV-1B on the other hand, possesses two genetically identical conserved NF-κB motifs, conserved across all the HIV-1 subtypes. In this backdrop, the present study is an attempt to examine the influence of the number of NF-κB binding elements in the C-LTR. **Importantly, the focus of the present study is on the copy-number difference of the NF-κB binding sites, therefore, on the overall strength of transcription and its influence on viral latency.** The present study did not aim to examine the impact of genetic diversity of any specific NF-κB binding motif on viral latency.

We employed three different cell models to examine HIV-1C latency using Jurkat T-cells and pseudotyped reporter viruses expressing GFP (EGFP or GFPd2 variant). The three experimental models differ in the strategy by which the LTR-Tat-feedback axis is regulated (detailed account of the three models in Chapter 2; section 2.1). (1) **The autonomous Tat-feedback (ATF) model**: The natural LTR-Tat positive feedback loop is preserved in this model (Burnett JC et al., 2009; Weinberger LS et al., 2005). Tat as well as the reporter gene EGFP/GFPd2 are co-expressed from the viral LTR in the vector LTR-GFP-IRES-Tat (LGIT). The concentration the Tat protein in this model is expected to be proportional to the transcription strength of the viral promoter. (2) **The disjoint Tat-feedback (DTF) model**: The LTR-Tat-feedback loop has been disrupted in this model by removing Tat expression from the LTR-control and placing it under a chromatinized, inducible Tet-ON-Tat cassette. The HIV-1 vector expressing EGFP alone

under the LTR is called LTR-GFP (LG) reporter vector. In this model, the intracellular concentration of Tat is expected to be independent of the transcriptional strength of the LTR. (3) **The tunable Tat-feedback (TTF) model**: In the TTF model, Tat is placed under the control of the LTR like in the ATF model. Tat, however, contains a destabilization domain (DD) at the C-terminus of the Tat ORF, using which the half-life and consequently the intracellular concentration of the viral protein Tat could be regulated using the small molecule ligand– Shield1 (Banaszynski LA et al., 2006). **Using these three different experiment systems, we examined transcriptional activation and silencing from a panel of HIV-1C LTRs that vary in the number of functional NF-κB binding sites.** (Figure 3.1)

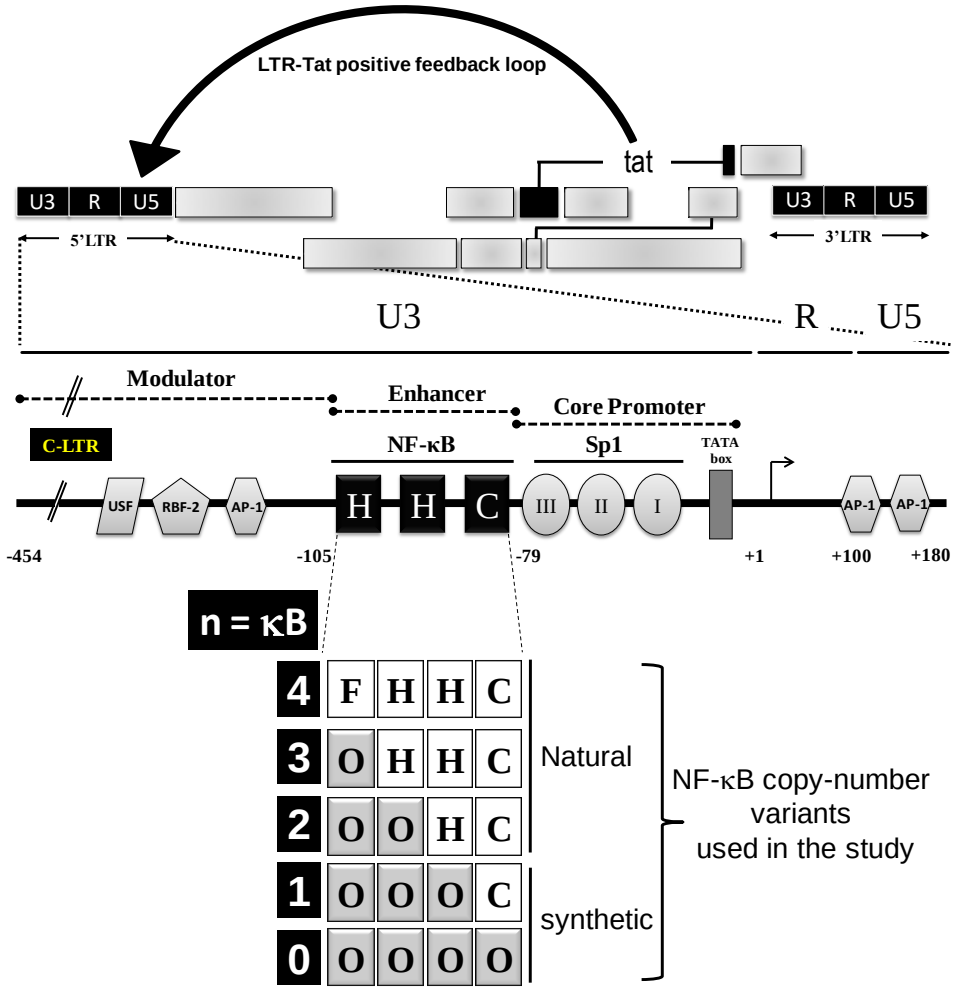


Figure 3.1: A high level TFBS variation in HIV-1C makes the question of viral latency more complex. A typical HIV-1C LTR containing three NF-κB motifs is presented. The distinct regulatory regions of the viral promoter (U3, R, U5, the Modulator, the Enhancer and the Core promoter) and the important transcription factor binding sites (TFBS) have been depicted. Although subtype-specific variations are pervasive among all the regulatory elements, the present study focuses on NF-κB motif variation of HIV-1C. The enhancer region (-105 to -79) comprises exclusively of multiple NF-κB binding motifs arranged in a tandem fashion. Unlike other genetic subtypes of HIV-1, HIV-1C naturally possesses two (rare strains; Hanna LE et al., 2014), three (canonical strains) or four (newer strains, Bachu M

et al., 2012a; Boullosa J et al., 2014) functional copies of NF-κB sites in the enhancer. The NF-κB copy-number variation in HIV-1C is expected to concomitantly alter the transcriptional strength of the LTR and that of the Tat feedback loop, thus, making the phenomenon of viral latency more complex and diverse in HIV-1C. Shown above is the panel of five LTRs that contain a variable number of NF-κB elements. 'H' represents the canonical H-κB motif (GGGACTTTCC) universally present in all the HIV-1 subtypes; 'C' stands for the subtype C unique NF-κB element (GGGGCGTTCC), and 'F' denotes the fourth -κB element (GGGACTTTCT) found in the emerging strains of HIV-1C (Bachu M et al., 2012a). 'O' indicates the mutated -κB site.

3.3 The 'Autonomous Tat-feedback' (ATF) model: The transcriptional strength of the LTR is directly proportional to the number of functional NF-κB binding elements.

The 'Autonomous Tat-feedback' (ATF) model (described in details in Chapter 2; sections 2.1.1 and 2.2.1) comprising the minimalist presence of the two major viral factors, the LTR and Tat, and an association between these two factors represents the natural context of the virus as reported previously (Weinberger LS et al., 2005; Burnett JC et al., 2009). Several groups have adopted the ATF model to elucidate the molecular mechanisms involved in governing viral latency. The minimalist vector design consists of the co-expression of EGFP and Tat form the viral LTR: LTR-GFP-IRES-Tat (LGIT) in the absence of all the other viral factors (Figure 2.1). The GFP expression from the virus serves as readout for the transactivation from the viral promoter; the expression of Tat regulates transactivation from the promoter as in the natural context. Although GFP and Tat are co-transcribed in a single viral mRNA, the translation of Tat is expected to be at a relatively lower level (~10-100 fold reduced) as compared to that of EGFP since an IRES element controls the translation of the viral protein Tat (Mizuguchi H et al., 2000). The viral promoter can function in two different modes, in the absence of Tat and in its presence. In the initial phases of infection, in the absence of Tat, a low basal transcription is primarily regulated by NF-κB that does not lead to efficient transcriptional elongation. The gradual accumulation of Tat generated through the basal transcription effectively promotes the release of the stalled RNA Pol II thereby accelerating viral RNA elongation by ~100 fold, a process referred to as the transactivated transcription. In this process, Tat itself being a product of the viral transcription initiates a positive feedback of transcriptional cascades at the promoter. It might thus be logical to assume that any component, *cis* or *trans*, that enhances the strength of transcription from the viral promoter is likely to accelerate the Tat-feedback circuit positively. Given the natural propensity of HIV-1C to contain more copies of NF-κB binding elements in the enhancer,

three copies typically and upto four frequently (Bachu M et al., 2012a), we constructed a panel of reporter viruses with varying number of NF-κB motifs to examine viral latency (p911a and p911b series; Appendix-I). The Jurkat based ATF latency-model follows the same working principles of latency (Figure 3.2) as the J-Lat model described previously (Jordan A et al., 2003, Weinberger LS et al., 2005). Stable proviral integrants in infected Jurkat cells may enter latency within a few hours of infection, a phenomenon termed early-latency. Alternatively, a combination of host and viral factors and/or stochastic events might turn off the active promoter at a later time point in late-latency. The inactive promoter can be reactivated using agents that activate T-cell signalling pathways or mediate chromatin remodelling. Of note, the reactivating agents can activate only a fraction of the latent viruses that have a potential for reactivation. Why a large fraction of the inducible latent proviruses is refractory to transcriptional activation is not understood.

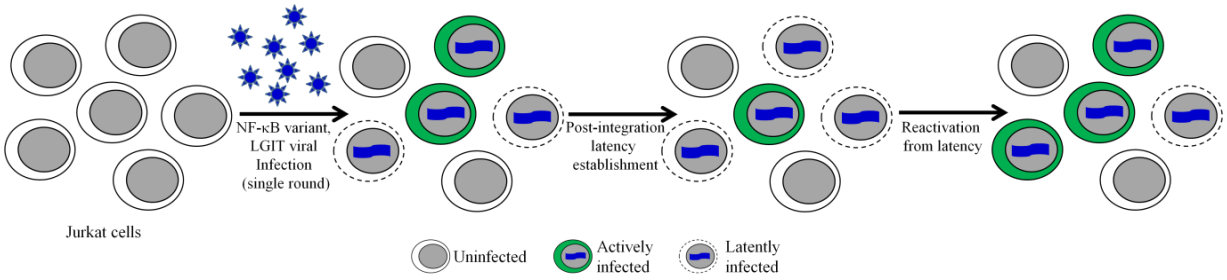


Figure 3.2: The 'Autonomous Tat feedback' (ATF) model of HIV-1 latency. Jurkat cells are infected with a panel of minimal-reporter HIV-1 strains (LGIT/LdGIT; p911a/b series) pseudotyped with VSV-G envelope. Following proviral integration, a fraction of the infected cells may harbour transcriptionally active proviruses while in other cells the proviruses may undergo transcriptional silencing. Actively transcribing proviruses in all or a fraction of cells may switch off at a later time point (late latency); depending on several silencing factors such as the site of integration, epigenetic modulation at the promoter, and the activation status of the host cell. The silent provirus can be reactivated using several host-cell activating agents.

As described above, among all the genetic families of HIV-1, only in subtype C a variation in the copy-number of NF-κB binding sites and a genetic variation within the additional copies of the NF-κB binding sites are manifested. Previous work from our laboratory generated convincing experimental evidence that the transcriptional strength of HIV-1 C viral promoter is proportional to the number of functional copies of NF-κB binding sites (Bachu M et al., 2012b). Importantly, the influence of the transcription strength of the viral promoter on viral latency has not been examined previously. In this backdrop, we constructed a panel of NF-κB-motif copy-number variant LTRs using the LGIT vector-backbone (the p911a/b vector-series, as described in Chapter 2, section 2.1.1). We used a prototype subtype C LTR containing four functional NF-κB binding

sites (FHHC, see section 2.1.1 for the construction of the 'FHHC' LTR) and introduced inactivating point mutations into the TFBS from 5' end to 3' end of the enhancer to reduce progressively the number of functional NF-κB motifs from 4 copies to 0 copies. The panel of five reporter viruses contained 4, 3, 2, 1, and 0 copies of functional NF-κB binding sites in the viral enhancer (Chapter 2, Section 2.1.1). Using this panel of reporter viruses, we infected Jurkat cells and examined the kinetics of latency-establishment and maintenance.

The viral strains of the panel were pseudotyped with VSV-G envelope in HEK293T cells and the relative infectious units (RIU) of the stocks were determined in Jurkat cells using GFP fluorescence (Chapter 2, section 2.3.3). Using the titer information, an appropriate quantity of each viral stock was used to infect approximately 50% of Jurkat cells and the assay was conducted as outlined (Figure 3.3A). Briefly, Jurkat cells were infected with the viral variants using an approximate RIU and three days following infection, the cells were activated using a cocktail of global T-cell activators (40 ng/ml PMA + 40 ng/ml TNFα + 200 nM TSA + 2.5 mM HMBA). Twenty four hours following activation, the GFP fluorescence levels of control and activated cells were determined using flow cytometry. It is evident from the comparative GFP-histogram profiles (Figure 3.3B), HIV-1 transcription varied in an NF-κB dose-responsive manner. The differences in the GFP-intensities between the LTR-variants were more pronounced following cell activation (filled, grey histogram) compared to the un-induced fractions (hollow, black histogram).

The transcription-efficiency in the form of GFP expression was examined further from the LTR-variant reporter viruses. We found the GFP-MFI to be proportional to the number of functional NF-κB motifs in the LTR (Figure 3.3C, left panel) although the levels of viral infectivity were comparable (Figure 3.3C, middle panel). The LTR containing four NF-κB motifs (FHHC; 4-κB) demonstrated the highest magnitude of fluorescence expression with (82,917.51 ± 825.7 RFU) and without (12,365.13 ± 179.3 RFU) activation. In contrast, the LTR in which all the four NF-κB motifs have been mutated (OOOO; 0-κB) demonstrated the lowest levels of the reporter expression with (22,190.38 ± 668.1 RFU) and without (6,083.36 ± 290.5 RFU) activation; whereas the other three LTRs containing 3 (OHHC; 3-κB), 2 (OOHC; 2-κB), or 1 (OOOC; 1-κB)

functional NF-κB motifs remained between the two extremes demonstrating progressively decreasing promoter activity in that order (Figure 3.3C, left panel). The fold transactivation was directly proportional to the number of functional NF-κB motifs in the LTR (Figure 3.3C, top, right panel) with a linear correlation (r = 0.98) between the promoter activity and the functional NF-κB motifs in the LTR (Figure 3.3C, bottom, right panel).

Additionally, we also quantitated the Tat-transcript level from all the variant promoters with or without activation using reverse-transcription PCR. In the RT-PCR, uninfected Jurkat cells served as the untreated control and GAPDH gene expression was used for normalisation. The relative Tat expression profile was determined using the ΔΔCt method. Similar to the expression of GFP, the level of Tat-transcript expression was directly proportional to the NF-κB copy-number in the LTRs of the reporter viral strains with or without activation (Figure 3.3D, left panel) and the fold transactivation progressively decreased from the 4-κB to the 0-κB LTR (Figure 3.3D, top, right panel) with a perfect correlation between the fold transactivation and NF-κB copy-number (Figure 3.3D, bottom, right panel). It is evident from the expression profile that a perfect correlation exists between the NF-κB motifs in the promoter and the level of gene expression from the promoter. Additionally, from the expression profile of the two genes from the viral promoter, the expression of GFP can be used as a surrogate marker for the expression of Tat, since a perfect correlation exists between the two genes co-expressed from the viral promoter. In the subsequent assays of viral latency, we have routinely used the expression of GFP as a measure of the transcriptional activity of the viral promoter with a frequent confirmation of Tat expression.

Importantly, the GFP^{+ve} cells in the experiment could be further classified into two distinct populations based on the fluorescence intensity – low or basal and high or transactivated GFP phenotypes (Figure 3.3B). Of the two GFP^{+ve} cell fractions in the assay, the GFP^{High} fraction represents a stronger Tat feedback circuit thus, denoting the 'Tat-transactivated' population.

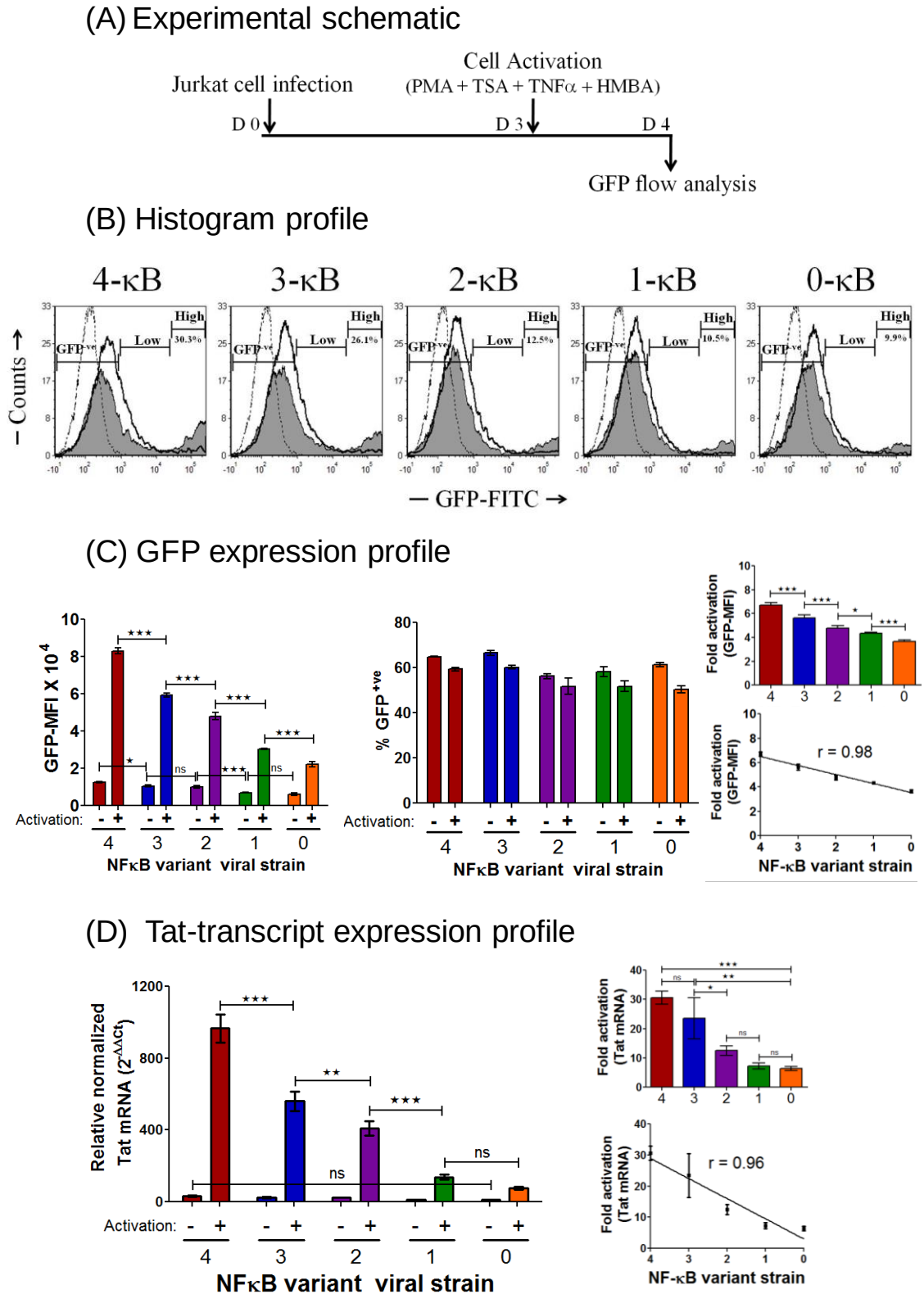


Figure 3.3: Gene expression profile from a panel of LTR-variants representing the ATF model (A) Schematic representation of the experimental strategy. One million Jurkat cells were infected at ~0.5-0.6 RIU independently with each of the viruses of the panel as shown. After 72 h of infection, half of the infected cells were activated for 24 h followed by flow analysis of the GFP expression. **(B)** Histogram profiles to compare the GFP intensities among the variants. The black dotted histogram represents Jurkat cells neither infected nor activated; the black hollow histogram represents cells infected but not activated. The solid grey histogram represents cells infected and activated. The intensity ranges of GFP^{-ve}, GFP^{Low} and GFP^{High} are marked. The GFP^{High} population (MFI $>10^4$ RFU) represents the

transactivated population with an intact LTR-Tat transcription feedback. As evident, the peak-height of the transactivated population gradually declined as the NF-κB copies reduced from 4 to 0, thus demonstrating a positive correlation between the numbers of NF-κB binding sites in the viral enhancer and Tat-mediated transactivation. **(C)** Viral gene expression manifests a linear positive correlation with the NF-κB copy-number. The mean GFP-MFI values from experimental quadruplicates ± SD of the infected cells with or without activation (left panel) and the percent infected cells (middle panel) are presented. The fold enhancement in the GFP expression from the variant LTRs (top, right panel) and its correlation with the number of NF-κB copies (bottom, right panel) are also presented. The data are representative of two independent experiments. Two-way ANOVA with Bonferroni post-test correction was used for the statistical evaluation (*p<0.05, **p<0.01, ***p<0.001 and ns – non-significant). **(D)** The Tat-transcript expression is directly proportional to the NF-κB copy number. Total mRNA was extracted from 0.5 million cells corresponding to control and activated populations. Relative Tat expression was determined in a RT-PCR using the ΔΔCt method. GAPDH gene expression was used as the normalization control. Mean values of the relative Tat expression from three independent experiments ± SEM are plotted. Two-way ANOVA with Bonferroni post-test correction was used for the statistical evaluation.

Interestingly, although the transcriptional strength of the viral promoter influenced the percent of cells in each of the three populations- GFP^{-ve}, GFP^{Low} and GFP^{High}, the most pronounced impact on the GFP expression was manifested on the GFP^{High} fraction (Table-3.1). On day-4 following cell infection, the percentage of the GFP^{High} fraction was directly proportional to the number of NF-κB motifs in the LTR. The viral promoters could be categorised into three groups based on the percentage of cells in the GFP^{High} fraction. The 4-κB LTR demonstrated the largest number of cell percentage, 32.30 ± 0.4 to be GFP^{High} and the LTRs containing 0, 1, and 2 copies of NF-κB the lowest percentage, while the 3-κB LTR occupied an intermediate position, with 22.80 ± 1.7% cells being positive for GFP^{High} expression. Importantly, the promoter strength was negatively correlated with the percentage of cells in the GFP^{low} and GFP^{-ve} fractions suggesting that the transcriptional strength of the viral promoter could have a significant impact on viral gene expression and latency.

Table 3.1 GFP expression profile of the NF-κB-variant viral strains

NF-κB copies	% GFP^{-ve}	% GFP^{+ve} cells		
		GFP^{Low}	GFP^{High}	Total
4	40.75 ± 0.33	27.70 ± 0.80	32.30 ± 0.45	60.00 ± 1.05
3	40.89 ± 1.15	36.32 ± 2.36	22.80 ± 1.70	59.11 ± 1.15
2	47.88 ± 1.54	38.61 ± 1.08	13.52 ± 0.46	52.13 ± 1.54
1	48.21 ± 1.15	41.29 ± 1.21	10.48 ± 0.15	51.76 ± 1.12
0	50.28 ± 0.73	39.47 ± 0.63	10.25 ± 0.15	49.72 ± 0.73

Data presented here represent the experimental results of Figure 3.3B. Two-way ANOVA was used for statistical evaluation. Mean values from experimental quadruplicates ± SD are indicated. Data are representative of two independent experiments.

Collectively, the data of the ATF model confirm that the expression of both GFP and Tat from the LTR is a function of the transcriptional strength of the viral promoter which in turn is directly proportional to the number of the functional NF-κB motifs present in the enhancer. Next, we asked how the transcriptional strength of the viral promoter influences the kinetics of viral latency-establishment.

3.4 In the ATF model, the stronger the viral promoter, the faster the latency-establishment.

The results of the above analysis unequivocally established that the transcriptional strength of HIV-1C LTR is directly proportional to the number of the functional NF-κB motifs in the enhancer (Figure 3.3 and Table-3.1). It is rather paradoxical that a virus that must establish transcriptional silence should possesses a strong promoter especially when the other genetic families of HIV-1 do not employ such a strategy. To understand this paradox, we determined the kinetics of transcriptional silence of the panel of ATF reporter viral strains that differ in the number of functional NF-κB motifs in the enhancer ranging from 4 to 0. The basic theme of the assay consisted of infecting the Jurkat cells at a low RIU, expansion of the infected cells, activation of the proviruses with a cocktail of global activators, sorting all the GFP^{+ve} cells, and examining the kinetics of latency-establishment as a function of the NF-κB copy-number (Figure 3.5A). A representative strategy for the cell-sorting of the GFP^{+ve} cells is presented (Figure 3.4).

Jurkat cells were infected individually with each NF-κB copy-number variant viral strain at an RIU of 0.1-0.2 to ensure single integration event per cell. The infected cells were expanded for a week in complete RPMI medium under standard growth conditions followed by treating the cells with a cocktail of global activators (as previously mentioned in section 3.2) to recover any latent provirus that may have switched-off during the early phases of the infection. All the GFP^{+ve} cells were sorted following 24 h post-stimulation and the percentage of cells that continued to express GFP was monitored using a flow cytometer every four days for 16 days. Simultaneously, we quantitated the expression of Tat transcripts on D0, D8 and D16 in an RT-PCR.

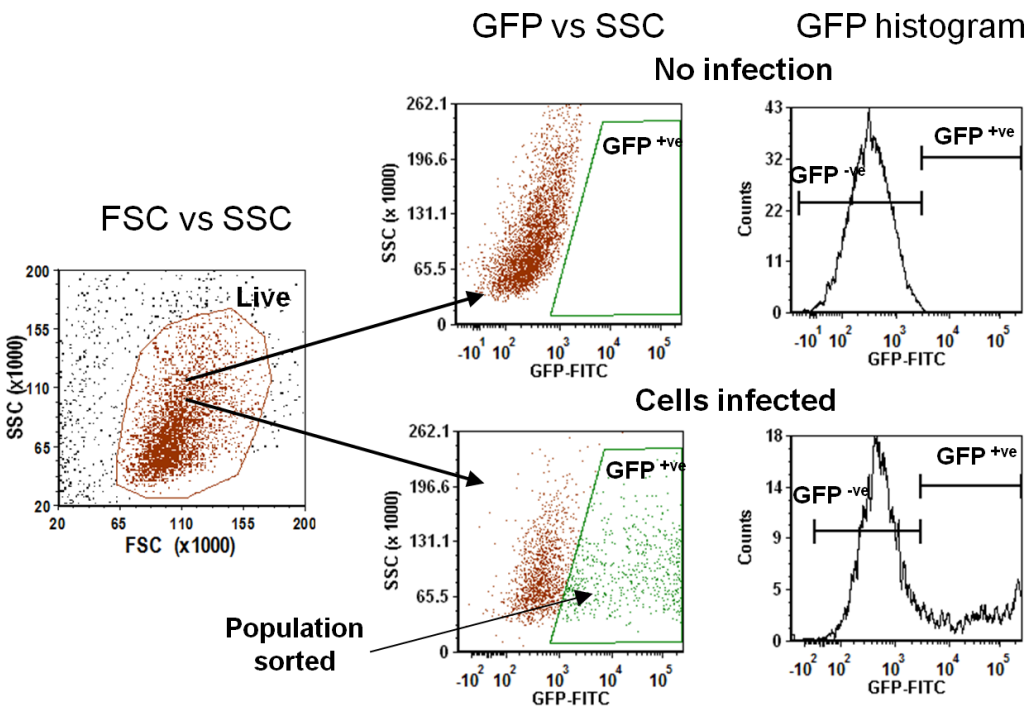


Figure 3.4: Representative gating strategy to sort a pool of GFP^{+ve} cells for the analysis of latency-establishment. The enriched GFP^{+ve} population represents a pool of cells with heterogeneous levels of GFP expression, the fluorescent intensities ranging from 10^3 to 10^5 RFU. LTR-silencing events in all the variant strains were scored both by estimating the percent GFP^{+ve} to GFP^{-ve} conversion as well as the reduction in the GFP-MFI over time.

The representative histogram profiles of GFP expression of all the five viral strains have been presented (Figure 3.5B). The kinetic curves of latency-establishment for each variant promoter were constructed by plotting the percentage of GFP^{+ve} cells against the day post-sorting (Figure 3.5C, bottom, left panel) along with the determination of MFI at each time point (Figure 3.5C,top, left panel). The analysis identified many interesting features. The expression of GFP from all the five LTRs reduced progressively as a function of time, up to 16 days. Surprisingly, the rate of GFP switch-off was faster from the 3- and 4-κB LTRs as compared to the other three promoters (0-, 1- and 2-κB LTRs) in the panel. The kinetics of latency-establishment was the slowest from the 0-κB LTR followed by 1- and 2-κB LTRs suggesting that the presence of the NF-κB binding motifs plays a direct role in establishing HIV-1 latency. Importantly, although both 3- and 4-κB LTRs demonstrated a rapid GFP downregulation, the 3-κB LTR established viral latency at a faster rate and the difference between the two promoters was highly reproducible and significant.

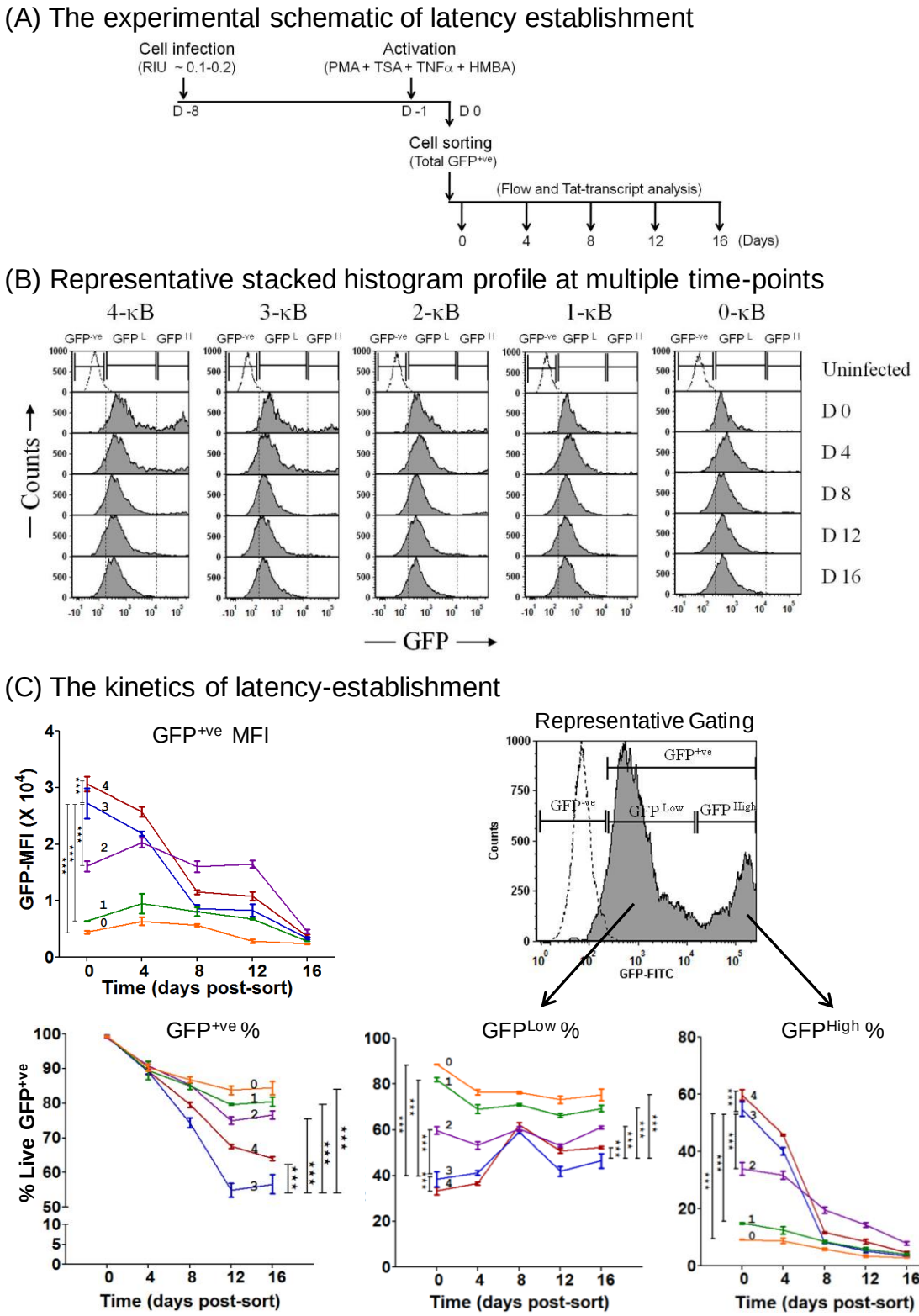


Figure 3.5: The kinetics of viral latency-establishment in the ATF model. (A) A schematic of the experimental layout. One million Jurkat cells were infected with each reporter viral strain of the panel at a low infectious titer (~0.1-0.2 RIU). The infected cells were maintained in culture for seven days following which they were treated with a cocktail of global T-cell activators (PMA+TNFα+TSA+HMBA) to activate the proviral pool. After 24 h of activation, all the GFP^{+ve} cells (harbouring active provirus) were sorted. The sorted cells were then maintained in culture and the GFP expression was monitored by flow cytometry every four days and that of Tat transcripts on days 0, 8 and 16. **(B)** GFP expression histogram profile of the LTR panel**.** The sorted GFP^{+ve} cells exhibit two distinct populations, the

GFP^{Low} peak (GFP^{Low}, MFI ~10^2 –10^4 RFU) representing cells with basal-level transcription and the GFP^{High} peak (GFP^{High}, MFI >10^4 RFU) comprising of the Tat-mediated transactivated cells. As is evident, the stronger the viral promoter the higher the number of transactivated cells. Importantly, a significantly faster reduction in the transactivated peak is evident for the strong promoters (4- and 3-κB LTRs) but not for the other three (2-, 1- and 0-κB LTRs). **(C)** A two-phase latency-establishment by the strong viral promoters. The percentage of cells positive for GFP expression (Bottom, left panel) and the MFI values of all the GFP^{+ve} cells (Top, left panel) were monitored every four days. Mean values from experimental triplicates ± SD are plotted. Data are representative of three independent experiments. Two-way ANOVA with Bonferroni post-test correction was used for the statistical evaluation (*p<0.05, **p<0.01, ***p<0.001 and ns – non-significant). A representative post-sort histogram profile of GFP^{+ve} cells is shown (Top, right panel). Two cell-pools showing the GFP^{Low} and GFP^{High} phenotypes are evident in the histogram. In the GFP^{Low} (Bottom, middle panel) and GFP^{High} (Bottom, right panel) cell pools, the percentages of cells positive for GFP expression at multiple time points are shown.

Of note, the process of latency-establishment was not complete from any of the promoters that even for the 3-κB LTR that demonstrated the fastest GFP switch-off kinetics, approximately 56.6% cells remained GFP^{+ve} at day-16 indicating that only 43.4% cells became latent. The percentage of cells becoming latent was significantly smaller for the other three viral promoters (0-, 1- and 2-κB LTRs). Multiple reasons could explain why GFP downregulation was not complete in this experimental model. For technical reasons, a T-cell line such as Jurkat was used in these assays and the Jurkat cells do not recapitulate the physiologically inactive state of the primary T-cells. Additionally, the prolonged half-life of the EGFP, approximately 48 h, is not likely to represent the true transcriptional activity of the viral promoter in real-time. The cells, therefore, may continue to be scored as GFP^{+ve} for a significant period of time even after the LTR is switched off. In the subsequent sections, we replaced EGFP with GFPd2 that contains a half-life of a significantly shorter period (2 vs 48 h) and GFPd2 faithfully represented LTR transcriptional status unlike EGFP (see Figure 3.10) where complete transcriptional silence could be achieved within a week.

The mean fluorescence intensities of the LTRs, in contrast, were proportional to the NF-κB number in the panel and ranged from the highest for the 4-κB LTR to the lowest for the 0-κB LTR at the 0 h time point and broadly at the subsequent time points (Figure 3.5C, top, left panel). On day-16, the MFI values for all the five promoters were comparable as the GFP down modulation was the most rapid for the stronger LTRs and proportionately slower for the other LTRs in the panel. For instance, the GFP intensity of 4-κB LTR reduced approximately 8-fold from a value of 30,631.64 ± 1,278.3 on D0 to 3,771.06 ± 245.2 on D16 whereas the corresponding values for the weakest 0-κB LTR were 4,455.11 ± 258.9 and 2,371.98 ± 59.3, respectively within the same time-frame that amounted to a modest 2-fold drop.

In the above analysis, we considered all the GFP^{+ve} cells as a homogeneous population although two different cell pools, GFP^{High} and GFP^{Low}, were evident in each histogram profile (Figure 3.5C, top, right panel). When the GFP^{High} and GFP^{Low} pools were gated and analysed independently at all the time points, the reduction of GFP-MFI observed in the total GFP^{+ve} cell pool (Figure 3.5C, top, left panel) corresponded perfectly, only with the $\%GFP^{High}$ cells (Figure 3.5C, bottom, right panel), but not with the $\%GFP^{Low}$ cell pool (Figure 3.5C, bottom, middle panel). Of note, the switch-off of the LTRs especially from the GFP^{High} cell pool demonstrated two important features. First, the stronger LTRs (4- and 3-κB) produced the highest percentage of GFP^{High} cell pool at the 0 h time point. In contrast, the weak LTRs (1- and 0-κB) generated the lowest percentage whereas the 2-κB LTR occupied an intermediary position between these two extremes (Figure 3.5C, bottom, right panel). Second, the LTR switch-off phenomenon of the stronger LTRs (4- and 3-κB) was manifested in two distinct phases: the most rapid reduction of GFP expression between days 0 and 8 and a slower rate of reduction after D8. The three weak LTRs in the panel (2-, 1- and 0-κB) do not appear to demonstrate this phenomenon. The rapid fall of GFP expression of the stronger LTRs between days 0 and 8 resulted in a significant rise in the GFP^{Low} cell pool percentage peaking on D8 (Figure 3.5C, bottom, middle panel). Such a phenomenon is not evident for the weak LTRs. The profile of Tat-transcript expression from all the five viral promoters was determined at D0, D8 and D16 using an RT-PCR (Figure 3.6A). Expression profile of the Tat-transcripts resembled that of GFP expression- the stronger LTRs (4- and 3-κB) clearly segregated from the rest of the three (2-, 1- and 0-κB). A two-phase reduction in Tat expression appears to be manifested for both the strong viral promoters with a faster drop in gene expression between D0 and D8. The 4-κB LTR showed the highest level of Tat expression 92.94 ± 5.4 at D0 which dropped to 12.02 ± 0.8 at D8 and to 2.3 ± 0.01 at D16. The corresponding values for the 3-κB LTR are 71.76 ± 2.5, 12.8 ± 0.73 and 1.15 ± 0.1 respectively.

To rule out the possibility that a difference in the number of proviral integration events among the reporter viral strains influenced the assay results, we took a precaution of infecting the Jurkat cells at a low RIU to avoid multiple proviral integration. Additionally, we quantitated the mean number of the proviral integration in all the stable cell pools of all the five reporter viral strains using a PCR that amplified the LTR strong-stop region

(experimental details in Chapter 2, section 2.9A). We used the J-Lat 8.4 clonal line that contains a single provirus per cell (Spina CA et al., 2013) to generate a standard curve (Figure 3.6B, left and middle panels) and used regression analysis to determine the number of integration events in our stable Jurkat cell pools. The analysis confirmed a single integration event per cell in all the five stable cell pools (Figure 3.6B, right panel).

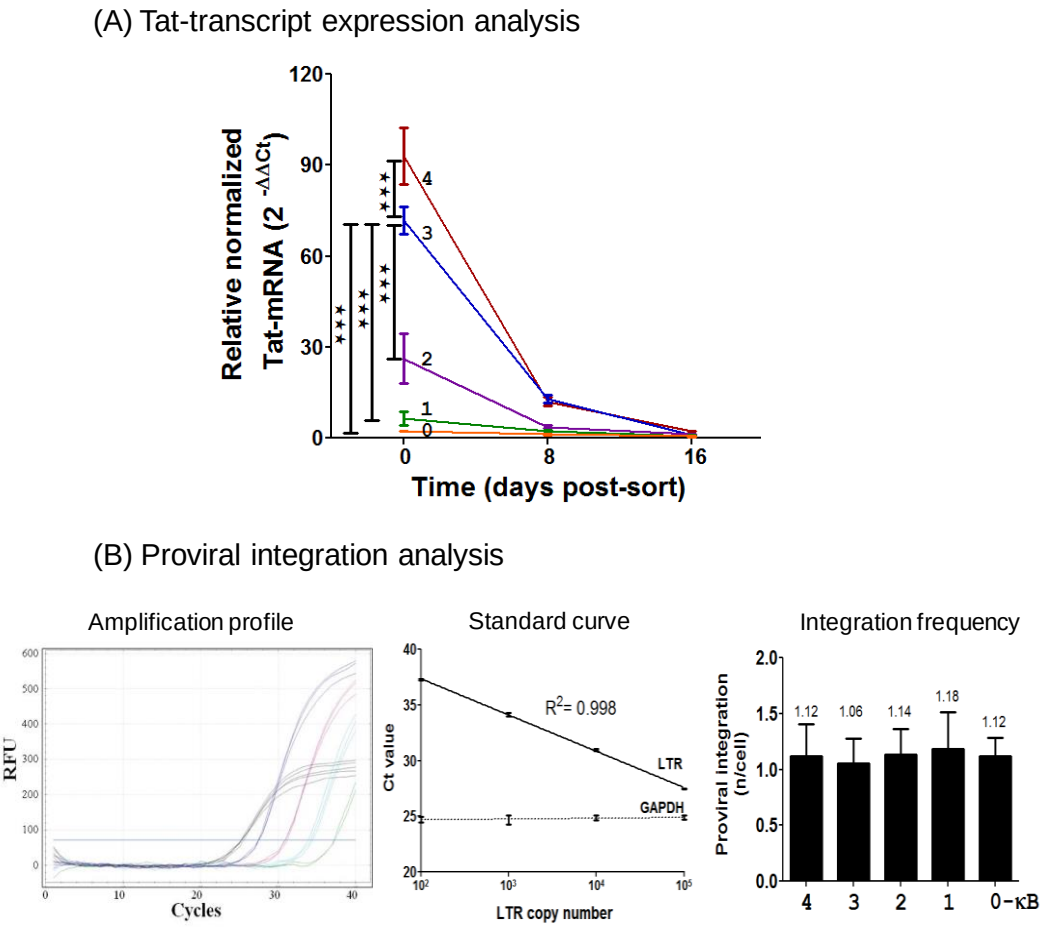


Figure 3.6: Tat-transcript profile and frequency of proviral integration in the LTR-variants during the establishment of latency in the ATF model. **(A)** The kinetics of Tat-mRNA expression. Cellular RNA was extracted at 0, 8 and 16 days post-sorting from 0.5 million cells from all the NF-κB-variants. Relative Tat expression was quantified in RT-PCR using the ΔΔCt method. GAPDH gene expression was used as the housekeeping control. Mean values from experimental triplicates ± SD values are plotted. Data are representative of two independent experiments. For all the experiments, Two-way ANOVA with Bonferroni post-test was used for statistical evaluation (*** $p<0.001$). **(B)** Proviral copy levels among the LTR-variants. Genomic DNA isolated from two million J-Lat 8.4 cells was serially diluted ten-fold and a 130 bp LTR sequence was amplified in a Taqman qPCR to construct the standard curve (left and middle panels). GAPDH was used as the reference gene. Integration events for the LTR-variants were estimated on the D0 samples using regression analysis and found to be ~1.0 per cell (right panel). Mean values from experimental triplicates ± SD values are plotted.

In summary, our data are suggestive that the transcriptional strength of HIV-1C LTR is an important parameter regulating viral latency. Additionally, the transcriptional silencing of the strong promoters (4- and 3-κB LTRs) may be regulated differently at the molecular

level as compared to the low-strength viral promoters (2-, 1- and 0-κB). A positive correlation between the Tat-transcript levels and the rapidity of GFP switch-off in the stronger viral promoters is also indicative of the Tat-mediated positive feedback playing a critical role in regulating viral latency.

3.5 GFP downregulation in the GFP^{High} cells transits via the GFP^{Low} compartment.

In the ATF model, downregulation of GFP expression from GFP^{High} cells appeared to have transited via the transient phase of GFP^{Low} cells. The transient phase could not be distinctly identified as we sorted all GFP^{+ve} cells together, both GFP^{High} and GFP^{Low} cells included (Figure 3.5). In the subsequent assay, we sorted only the GFP^{High} cells, ignoring the GFP^{Low} cells (see Figure 3.7 for the sorting strategy).

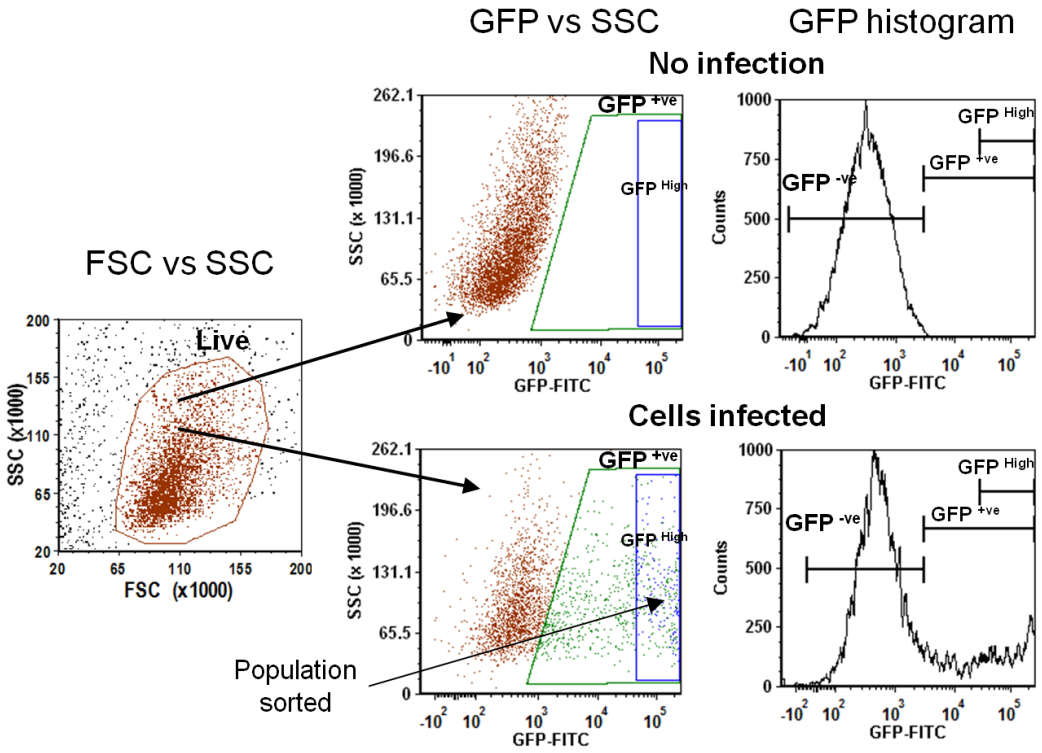


Figure 3.7: Gating strategy for the GFP^{High} cell-sorting. The GFP^{High} pool (GFP-MFI >10^4 RFU) was enriched from all the LTR-variants in the ATF panel to compare the kinetic profiles of latency. The GFP^{+ve} pool here was a more homogeneous cluster compared to the previous experiment (Figures 3.4) with respect to the GFP expression.

We monitored GFP downregulation every four days for 24 days (Figure 3.8A).The stacked histograms representing the GFP expression profile from all the variant LTRs of the panels are presented (Figure 3.8B). The D0 time point represents the validation of the cells soon after cell sorting. Of note, although only the GFP^{High} cells were sorted, some of these cells have already initiated the process of GFP downregulation during the short gap between the sort and validation and as a consequence, the GFP^{Low} cells are readily visible (See D0 profiles, Figure 3.8B). Importantly, at D0, there were almost no GFP^{-ve} cells even under the strong LTRs. The data of the GFP^{High} only cell sorting clearly demonstrated, as speculated above, the transition of the GFP^{High} cells via the GFP^{Low} compartment to the GFP^{-ve} pool which manifested in a spike of percentage increase on day-8 of the GFP^{Low} cell profile (Figure 3.8C, bottom, middle panel). As demonstrated above, the downregulation of GFP expression was not complete for any of the viral promoters. The biphasic mode of latency-establishment rather than a gradual and monophasic decline in the GFP peak is indicative that not all the cells in the population attained latency simultaneously. Thus, post-integration latency in the ATF model was found to be non-synchronous and probabilistic. Furthermore, depending on the MFI values of GFP expression (Figure 3.8C, top, left panel), or the percentage of cells expressing GFP (Figure 3.8C, bottom profiles) it is evident that the viral promoter could be classified into two categories where GFP downregulation was significantly faster for the stronger LTRs (4- and 3-LTRs) as compared to the other three (2-, 1-, and -0 κB LTRs).

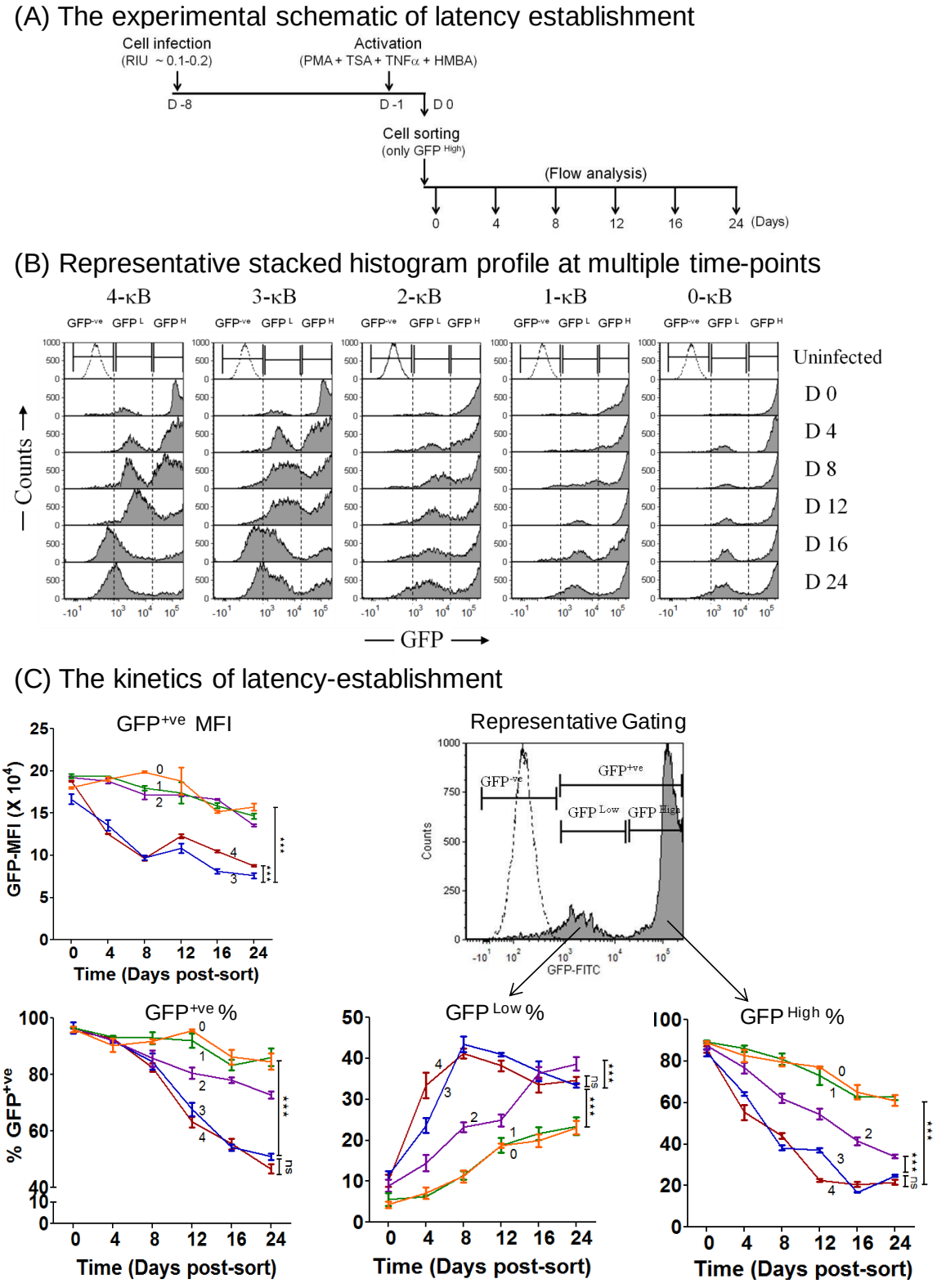


Figure 3.8: The kinetics of viral latency-establishment in the transactivated cells (GFP^{High}) of the ATF panel. (A) A schematic of the experimental layout. One million Jurkat cells were infected with individual LTR-variant viral strains at a low infectious titer (~0.1-0.2 RIU), maintained in culture for seven days, treated with the global T-cell activators (PMA+TNFα+TSA+HMBA) and sorted for GFP^{High} cells after 24 h of activation. GFP expression from the sorted cells was then monitored by flow cytometry every four days to compare the rate of LTR silencing among the variants specifically in the transactivated proviral pool. **(B)** GFP expression histogram profile of the LTR panel**.** The GFP^{High} cells in all the LTR-variants demonstrated a binary mode of GFP downregulation, although the stronger promoters (4-

and 3-κB LTRs) exhibited a faster kinetics than the others (2-, 1- and 0-κB LTRs). Similar to Figure 3.5B, the regions for GFP^{-ve} (MFI $<10^3$) GFP^{Low} (GFP^{L}; MFI $\sim10^2 - 10^4$ RFU) and GFP^{High} (GFP^{H}; MFI $>10^4$ RFU) phenotypes are indicated in stacked-histograms. **(C)** A combination of strong Tat-positive feedback and enhanced LTR-strength accelerates latency-establishment. The percentage of cells positive for high GFP expression (Bottom, left panel) and their mean intensities (Top, left panel) at multiple time points are presented. Mean values from experimental quadruplicates ± SD are plotted. Data are representative of three independent experiments. Two-way ANOVA with Bonferroni post-test correction was used for the statistical evaluation (*$p<0.05$, **$p<0.01$, ***$p<0.001$ and ns – non-significant). A representative post-sort histogram profile of GFP^{High} cells is shown (Top, right panel). The kinetic profiles of the percent GFP^{+ve} cell pools corresponding to GFP^{Low} (Bottom, middle panel) and GFP^{High} (Bottom, right panel) phenotypes are indicated.

3.6 Transactivated LTRs follow a rapid, biphasic mode of silencing with an increased frequency of occurrence in higher -κB promoters

One significant difference that we noted in the above analysis, as compared to the previously reported work based on HIV-1B (Weinberger LS et al., 2005; Burnett JC et al., 2009) is the biphasic transcriptional silencing of the GFP^{High} cells in the HIV-1C model. In an attempt to examine the nature of HIV-1 latency, Weinberger LS et al. previously identified in the Jurkat cell model that the 'Dim-sort' GFP cells, but not the Bright-sort or Mid-sort GFP cells demonstrated the phenomenon of phenotypic bifurcation (PheB). Only the Dim-sort GFP cells moved into high-GFP or negative-GFP expression regardless of external environmental factors or chromatin integration site differences (Weinberger LS et al., 2005). The authors coined the term PheB to describe this phenomenon ascribing this to gene noise and stochastic nature in the decision making that can have more than one outcome. The reporter virus used in the assays, LGIT, represented the minimal structural design of HIV-1 by retaining the LTR-Tat positive feedback axis alone in the absence of any other viral factors. The LTR and Tat both were derived from HIV-1 subtype B in the assay. We substituted both of these elements with the counterparts of HIV-1C. Our prime objective was to compare the rate of transcriptional-silencing among the promoter-variant strains that possessed varying copies of functional NF-κB sites. Hence GFP^{High} cells were considered as the starting point of the relaxation kinetics assay that would enable us to assess the silencing ability of the transactivated population in the variant strains and also ensure uniformity in the initial levels of GFP and Tat expression of the experimental samples. Thus, there exist important differences between the vectors used by Weinberger LS et al. and our laboratory with respect to the subtype origin of the viral promoter and Tat as well as the latency-kinetic profiles obtained particularly in the Tat-transactivated cells.

A molecular feature common to both the vectors is the presence of the EGFP as the reporter gene co-expressed from the viral strains along with Tat. The mean half-life of EGFP is ~48 h which probably is too long to represent the kinetics of the transcriptional activity of the viral promoter faithfully in real-time. A cell may continue to be scored positive for GFP expression for several hours even though the viral promoter has been transcriptionally silent, thus, skewing the data towards false positivity. Additionally, the strong GFP signal of the GFP^{High} cells pushed these cells towards the right extreme of the square in our analyses (Figure 3.8B). To circumvent these problems and to establish a real-time association between the reporter gene expression and the transcriptional activity of the viral promoter, we substituted EGFP in the reporter virus with GFPd2, a variant form of GFP containing a mean half-life of only 2 h (Li X et al., 1998). We constructed a new panel of reporter viral strains (LdGIT; p911b series) with NF-κB motif copy-number variation that is analogous to the panel used above. Using the new panel of the reporter viruses, we established the latency profiles in Jurkat cell infections (see Figure 3.9 for sorting strategy) essentially as described above using only the GFP^{High} sorted cells (Figure 3.10A).

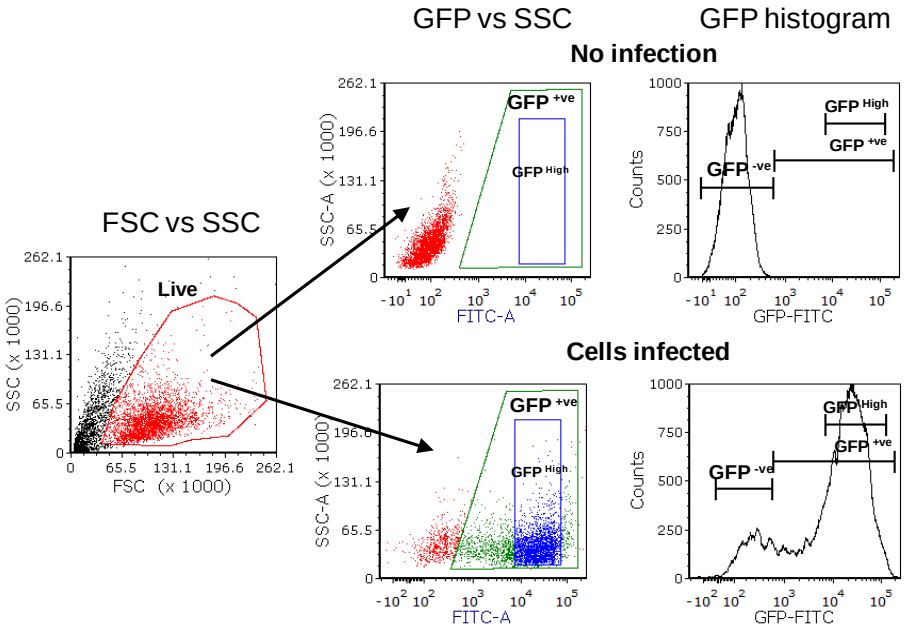


Figure 3.9: The gating scheme for sorting of GFP^{High}cells. The GFP^{High}pools (GFP-MFI ~10^3-10^4 RFU) corresponding to the LdGIT panel of variant viral strains were subjected to latency-establishment as described forthe LGIT strains (Figures 3.7 and 3.8). Note that the fluorescent intensities reduced approximately ten-fold with the substitution of EGFP with GFPd2.

The stacked GFP histogram profiles of all the five reporter viral strains are presented (Figure 3.10B). Several differences are readily evident. The kinetics of the viral latency is significantly faster for all the five viral strains regardless of the copy-number difference of the NF-κB motif. Unlike in the previous model LGIT that used EGFP with a significantly longer half-life, all the viral strains of the LdGIT panel successfully established a near complete viral latency. The MFI values of all the viral strains reduced to the baseline within 96 h following activation and sorting (Figure 3.10C, top, left panel). The percentage of cells downregulating GFP expression was directly proportional to the number of the NF-κB motifs whether total number of cells (Figure 3.10C, bottom, left panel) or only the GFP^{High} cells (Figure 3.10C, bottom, right panel) were considered.

The time required for the loss of fluorescence in half of the cells (FL_{50}) varied among the panel members such that the stronger LTR down regulated the fluorescence at a faster rate. The FL_{50} values for the 4-, 3-, 2-, 1-, and 0-κB viral strains were 23.3, 22.1, 24.64, 32.9 and 48 h respectively. Thus, there was a direct correlation between the transcriptional strength of the viral promoter and the rate of latency-establishment as earlier observed in the EGFP model. The rate of latency-establishment was significantly faster for 4- and 3-κB LTRs as compared to the other three but not between the two promoters themselves (Figure 3.10C, bottom, left panel). From the trajectory of fluorescence downregulation in the LdGIT model, it was quite evident that GFP^{High} cells transited via the GFP^{Low} compartment to latency. The movement of the GFP^{High} cells to the GFP^{Low} compartment manifested in a spike of fluorescence at 18 h for the stronger LTRs (4- and 3-κB LTRs) and at a later time point for the other three LTRs. Collectively, our data are assertive that the transcriptional strength of HIV-1 promoter is an important parameter regulating viral latency. The stronger the transcriptional strength the faster the rate of latency-establishment. GFPd2, a variant form of GFP containing a shorter mean half-life represents the kinetics of the viral transcription more faithfully than the regular EGFP mostly used in the assays. Latency in J-LdGIT was achieved much earlier than the EGFP counterparts (within 7 days) while the trend in GFP downregulation remained the same as in the J-LGIT. The latency kinetics in the transactivated population (GFP^{High} cells) substantiated the NF-κB-site dependent latency mechanism as observed in the previous experiment of the ATF model (section 3.3). In this regard, a biphasic mode of promoter-silencing by the transactivated cells was an interesting observation.

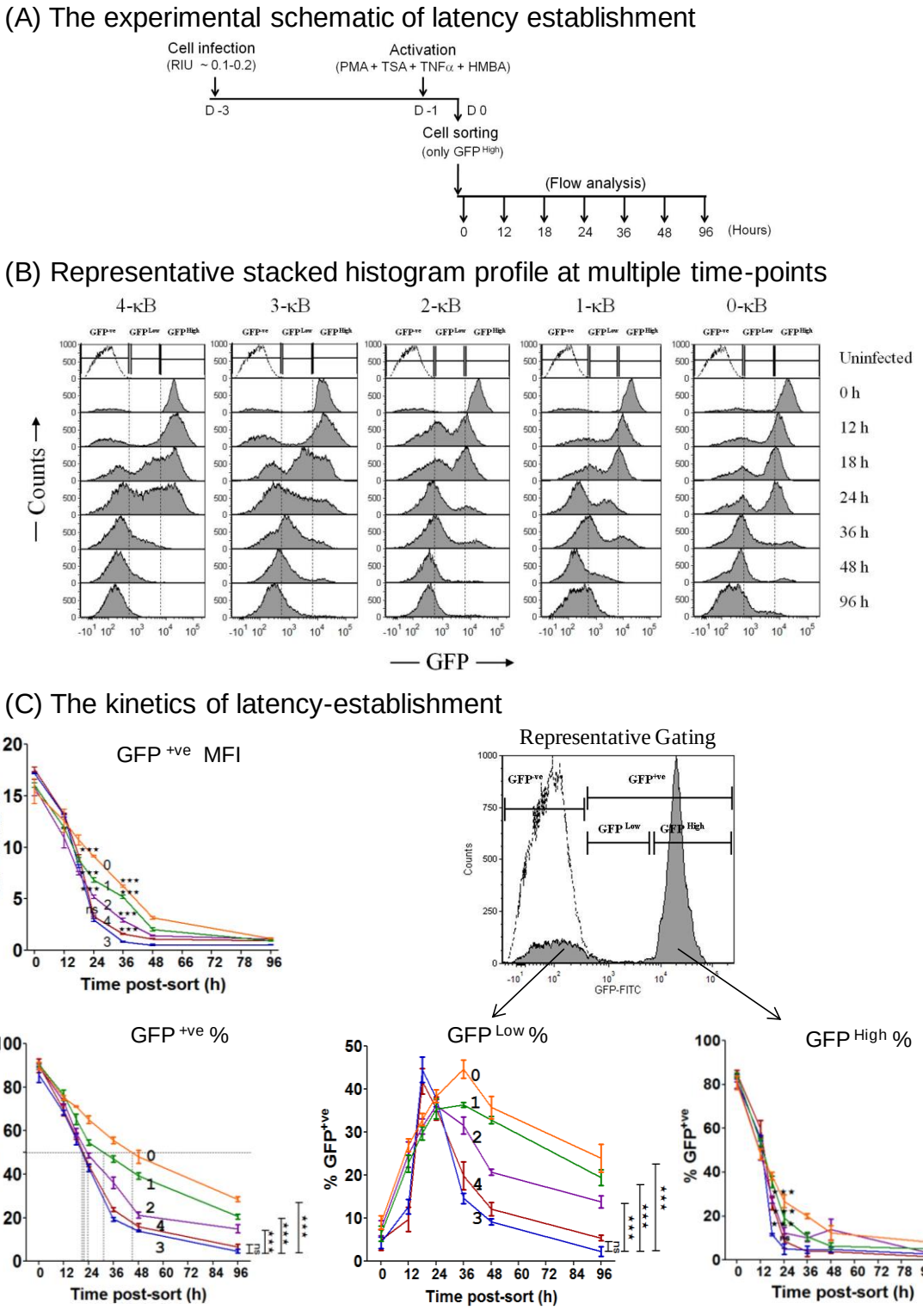


Figure 3.10: The kinetics of viral latency-establishment in the transactivated cells of the LdGIT variants. (A) The schematic of the experimental set-up. The parameters and experimental conditions used in the present assay were similar to that of the previous experiment (Figure 3.8A) with the exception that the GFP-readout was monitored at a shorter interval, every six hours as opposed to the four-day time points. The GFPd2 reporter also enabled the completion of the assay within four days as opposed to 24 days in the case of EGFP vectors (Figure 3.8C). **(B)** The histogram profile of GFP expression. The regions of GFP^{-ve} (MFI $<10^3$ RFU), GFP^{Low} (GFP^{L}; MFI ~10^2–10^4 RFU), and GFP^{High} (GFP^{H}; MFI $>10^4$ RFU) phenotypes are demarcated in the stacked-histograms. Of note, the half-life of GFPd2 being only 2h, the stability of the GFP^{Low} phenotype was extremely transient, hence the present system lacked a distinct

GFP^{Low} cluster at any time point unlike in the EGFP system (Figures 3.5B and 3.8B). **(C)** Promoter-silencing was favoured by the cumulative effect of the strong Tat-positive feedback and enhanced LTR-strength. Unlike for the LGIT panel of vectors, a complete loss of GFP expression was observed with the LdGIT panel of vectors within four days post-sorting. The percent GFP^{+ve} cells (Bottom, left panel) and their mean intensities (Top, left panel) at multiple time points are indicated. The time required for the loss of fluorescence in half of the cells (FL_{50}) for each variant is indicated using the dotted lines. Mean values from experimental triplicates ± SD are plotted. Data are representative of three independent experiments. Two-way ANOVA with Bonferroni post-test correction was used for the statistical evaluation (*p<0.05, **p<0.01, ***p<0.001 and ns – non-significant). A representative post-sort histogram profile of GFP^{High} cells is shown (Top, right panel). The kinetic profiles of the percent GFP^{+ve} cell-pools corresponding to GFP^{Low} (Bottom, middle panel) and GFP^{High} (Bottom, right panel) phenotypes are shown.

3.7 The stronger viral promoters demonstrate a phenotypic bifurcation in the live-cell tracking assay.

'Genetic noise' refers to the fluctuations in the levels of gene expression in a biological system, such as the mammalian cells of homogeneous population with an identical genetic constitution. Genetic noise becomes particularly apparent when a molecular species such as a protein, DNA or RNA molecule is limiting in any biochemical reaction; and slight fluctuations in the physiological levels of this species contributes to diverse phenotypic outcomes. Noisy gene expression could influence population variability (Raser JM et al., 2005). Fluctuation in gene expression is a common phenomenon in all forms of life ranging from single-cell prokaryotes to multicellular eukaryotes. Pathogenic organisms could exploit the phenomenon of genetic noise to randomly switch the surface receptors to increase virulence, acquire antibiotic resistance or evade immune responses. In the bacteriophage λ, stochastic fluctuations in the key regulatory proteins- Cro and CI and the genetic circuits involved have been extensively studied as the source of the two phenotypic outcomes- lysis or lysogeny (Arkin A et al., 1998). In complex gene-regulatory networks, DNA-protein feed-forward loops have been recognised as important modules contributing to differentiation, developmental processes and behavioural outcomes of an organism. These are auto-regulatory circuits (positive or negative) where the output acts as the regulatory input. A positive-feedback loop increases genetic noise by manifesting a bistable or bimodal phenotype (Isaacs FJ et al., 2003; To TL et al., 2003).

Using a combined approach of mathematical modelling and time-lapse, live-cell microscopy of individual cells expressing GFP under HIV-1 LTR, Razooky BS et al. demonstrated that in HIV-1B, the autonomous LTR-Tat-feedback loop is the 'decision-making' circuit for a transcriptional 'ON' or 'OFF' state (Razooky BS et al., 2011). Thus,

the tracking of live cells offers a powerful strategy to examine genetic noise and more appropriately to investigate the decision making in HIV-1 latency. Since the data described above using populations of virus-infected Jurkat cells demonstrated explicitly that the transcriptional strength of the viral promoter plays an important role in influencing viral latency, we explored if the genetic noise could also be significantly modulated by the strength of the viral promoter.

To this end, we first generated stable Jurkat cells constitutively expressing DsRed2-RFP under the EF1α promoter. The resulting Jurkat-RFP stable cells were used to infect individual NF-κB copy-number variant (LdGIT) viral strains. In the live cell tracking, the expression of RFP from the cells is intended to function as an internal control to normalize the varying expression levels of GFP under the viral LTRs (Figure 3.11). The live-cell assay for qualitative investigation of phenotypic variability was performed according to the protocol mentioned in Chapter 2, section 2.5.1 and presented schematically (Figure 3.11A). Approximately 0.5 million Jurkat-RFP stable cells were infected independently with LdGIT variant strains at a low infectious titre (RIU ~0.1).Two days after infection, the cells were activated with a cocktail of global activators, and the cells double positive for high intensity of GFP (GFP^{High}) and RFP were sorted. The cells were allowed to relax for seven days and RFP^{+ve} GFP^{-ve} cells were sorted, thus, selecting the latently infected cells. Finally, approximately 30,000 sorted cells were immobilised on a glass-bottom in 35 mm dish and reactivated only with TNFα (40 ng/ml) to activate the NF-κB pathway and specifically to avoid global activation. Representative microscopy images showing temporal GFP expression for each promoter are depicted in Figure 3.11B. Individual cells were tracked for GFP and RFP expression every 30 min for 24 h using a confocal microscope with an auto-corrected focal drift. The detailed experimental protocol is outlined in Chapter 2, section 2.5.1. GFP expression from all the five promoter-variant viral strains increased progressively during the period of observation while the expression of RFP remained constant (Figure 3.11B).

It is evident from the GFP-tracking of individual cells that the fluorescence intensity of the cells is directly correlated with the number of functional κB-motifs in the viral promoter; the results being perfectly consistent with the data of the LTR-transactivation assay using Jurkat cell population (Figure 3.3).Thus, the transcriptional strength of the

viral promoter is an important parameter that can modulate the kinetics of viral latency regardless of whether the cells are examined individually or as a population.

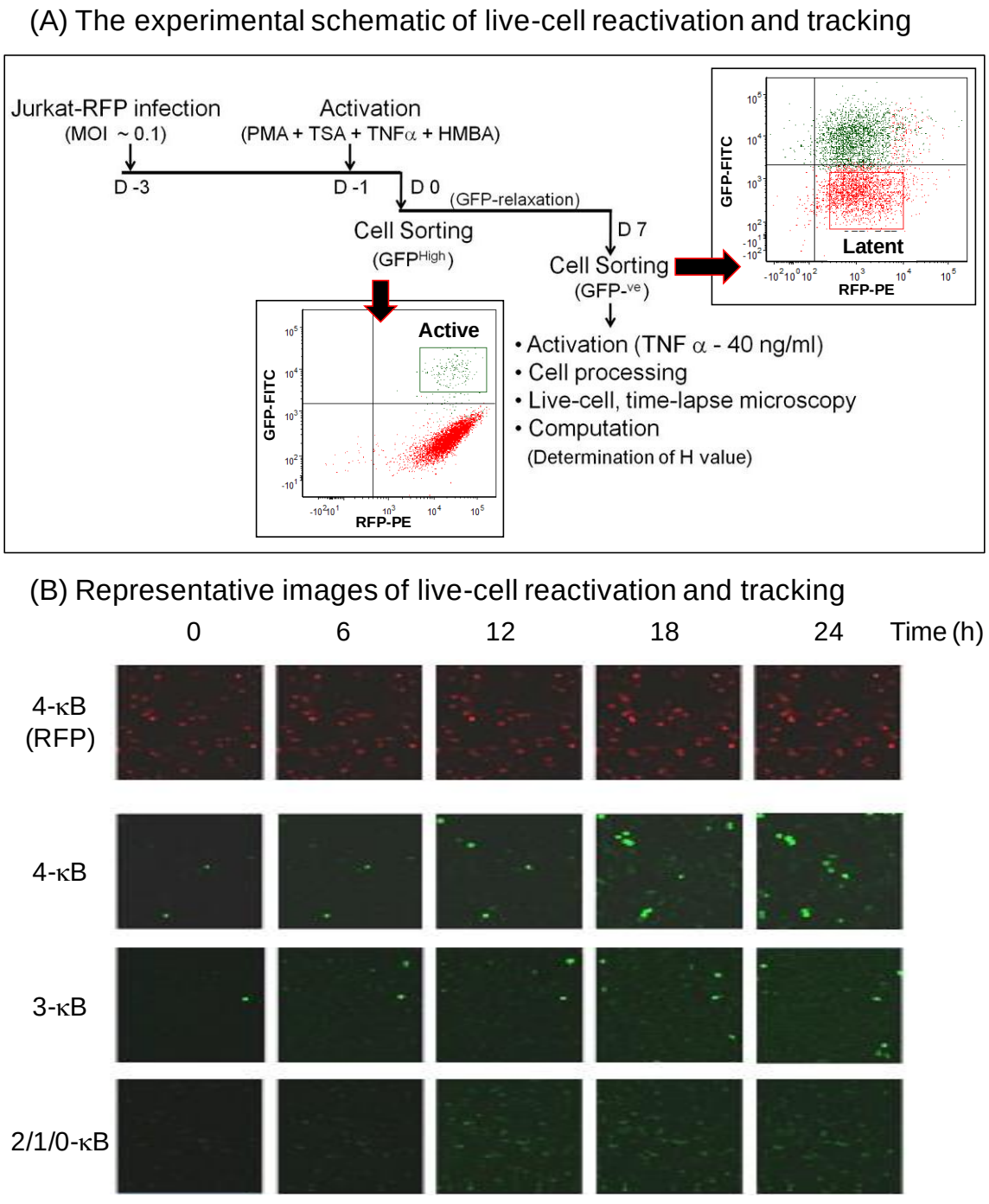


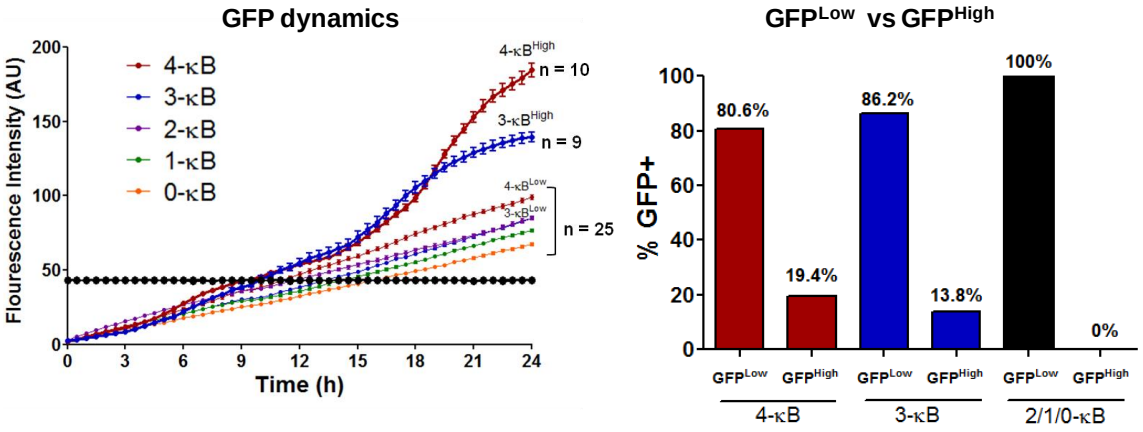
Figure 3.11: Live and single-cell reactivation kinetics of the NF-κB variant LdGIT strains. (A) Schematic representation of the experimental strategy. Jurkat cells were engineered for stable, constitutive RFP expression downstream of an EF1α promoter to differentiate LTR-activity (GFP) from the intrinsic noise levels (RFP) and to track unbiased cells. One million Jurkat-RFP cells were infected individually with the variant viral strains of the LdGIT panel, activated with the global T-cell activators (PMA+TNFα+TSA+HMBA), and GFP^{High} cells were sorted after 24 h. The active proviral LTRs were then allowed to relax for a week followed by a second sorting to enrich the GFP^{-ve} cells harboring transcriptionally silent proviruses. Gating profiles for the active and latent proviral sorts are indicated. Approximately 30,000 GFP^{-ve} cells sorted from each LTR variant were immobilized on the glass-bottom of a 35 mm

petri dish, treated with 40 ng/ml TNFα, and individual live-cells were monitored for GFP reactivation by capturing images for 24 h every 30 min using a confocal microscope. Uniform RFP expression was noted throughout the experiment. Values of the Hill coefficient (H) were determined from the GFP trajectories for all the LTR variants. **(B)** Representative GFP and RFP images at multiple time points of the LTR-reactivation. The RFP images of a single variant are presented as a representation (top panel; 4-κB) and the expression levels were found consistent at all the time points.

Of note, the two stronger viral LTRs containing 4- and 3-κB motifs manifested a unique phenomenon following activation unlike the other three (2-, 1-, and 0-κB LTRs). The individual cells of the two strong viral promoters displayed two characteristically distinct fluorescent intensities (Figure 3.12A left panel). While the majority of the cell population displayed a GFP^{Low} phenotype (80.6% and 86.2% for 4- and 3-κB LTRs, respectively; Figure 3.12A, right panel) similar to the cells of the other three LTRs, a minor population demonstrated the GFP^{High} phenotype (19.4% and 13.8% for 4- and 3-κB LTRs, respectively). The GFP^{High} phenotype is not manifested by the 2-, 1-, and 0-κB LTRs. We measured GFP and RFP fluorescence from 25-35 individual cells for each variant LTR from two independent experiments and the mean intensities ± SEM are presented (Figure 3.12A left panel). We further attempted kinetic modelling with ordinary differential equation (ODE) from the GFP-fluorescence trajectories of the GFP^{High} cells following the protocol developed by Razooky BS et al. (Figure 4; Razooky BS et al., 2011) to derive the Hill coefficient (H) values for the 4- and 3-κB positive Tat-feedback circuits. Since, the fluorescence-bifurcation phenomenon upon TNFα stimulation into GFP^{Low} and GFP^{High} phenotypes was noted only in the 3- and 4-κB LTRs, we first obtained the difference between the mean fluorescent intensities of the GFP^{High} and GFP^{Low} cells for the strong LTR variants at all time-points and the velocity vs enhanced fluorescence curves were then obtained. The curves fitted to the increasing phase of the data in the velocity vs florescence plot using H =1, 2 or 3 are shown in Figure 3.12B. The 4-κB LTR-Tat circuit clearly demonstrated H >1, indicating a cooperative nature of Tat in the $4\text{-}\kappa B^{High}$ cells although the exact value (H =2 or 3) could not be discerned due to the noise in the individual cell-trajectories (Figure 3.12B). The data need to be further enriched using closely-spaced time intervals as well counting larger number of cells for each variant. We are presently exploring the molecular mechanisms underlying 'the phenotypic bifurcation' manifested by only the stronger viral promoters. The higher intracellular concentration of Tat and/or subtype-specific differences in the LTR-Tat

positive feedback axis possibly might underlie this phenomenon. Of note, the 4- and 3-κB LTRs both are seen naturally in natural infection of HIV-1C.

(A) The kinetics of single-cell reactivation



(B) Cooperativity of Tat in the 4-κB LTR-Tat circuit

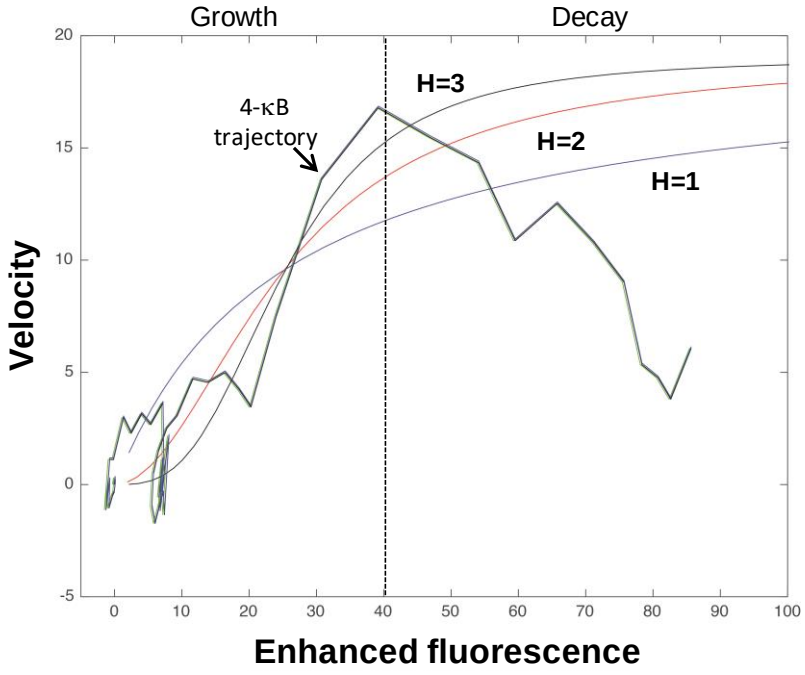


Figure 3.12: A positive cooperativity in Tat function of HIV-1C. (A) The kinetic profiles of reactivation of the LTR-variant viral strains. GFP trajectory curves (arbitrary fluorescent units vs time) were constructed for 25 individual GFP^{Low} cells from each variant strain using the ImageJ software and the mean values ± SEM are presented (left panel). Additionally, nine cells from the 3-κB and ten cells from the 4-κB variants in the GFP^{High} category were analysed. Data from two independent sets of experiments are presented collectively. The frequencies of GFP^{Low} and GFP^{High} cells in the population (right panel). The number of GFP^{Low} and GFP^{High} cells was counted in two independent microscopic fields for all the five LTR-variants and the percent values are indicated. **(B)** The 4-$κB^{High}$ cells display an H>1 reactivation kinetics. At each time point, the difference in the mean green fluorescent intensities of the GFP^{High} and GFP^{Low} cells of the 4-κB variant was estimated and termed as the enhanced fluorescence. This was followed by the

generation of a velocity vs enhanced fluorescence trajectory. Of note, the trajectory displayed a growing and a decaying phase, and the H values were obtained only for the growing phase using the parameters mentioned in Razooky BS et al., 2011; Figure 4. Evidently, the experimental trajectory does not fit the H=1 simulation curve.

3.8 The 'Disjoint Tat-feedback' (DTF) circuit abrogates the impact of NF-κB copy-number variation on HIV-1 transcription.

The transcription-output from the HIV-1 LTR is a coupled event contributed by the promoter-architecture as well as the Tat-mediated positive feedback. In accordance with the circuit design of the ATF model (Chapter 2; Figure 2.1), as demonstrated earlier, the variation in the functional copies of NF-κB motifs in the LTR affected the transcription-efficiency as well as the Tat-feedback strength. In other words, a stronger viral promoter not only contained a larger number of NF-κB motifs but also generated more intracellular Tat protein (Figure 3.3C and 3.3D), thus, making it difficult to discern the influence of either of the factors on viral latency. To discern the influence of only the NF-κB-number difference on viral latency in the absence of an interference from the Tat-mediated auto-regulation, it was essential to disrupt the LTR-Tat-feedback axis. To this end, we generated a new panel of five κB-variant reporter viral vectors, analogous to the ATF model described above (Section 3.3) and established stable Jurkat cells. The new viral expression system which we call the 'Disjoint Tat-feedback' (DTF) model, comprises of the variant LTRs expressing only EGFP (LG), but not Tat (Figures 2.2, 3.13). In the DTF model, the Tat-mediated positive feedback circuit is abrogated even though the LG vectors continue to express GFP albeit at a low, basal level. In the DTF model, Tat is expressed ectopically under an inducible 'Tet-ON' promoter (Das AT et al., 2016), thus, de-linking its expression from the viral genetic circuit. A stable Tat-expressing Jurkat cell line was used for the generation of the variant promoter panel to ensure uniform Tat expression that can be tuned further by Dox. The working principle of the DTF system is depicted in Figure 3.13.

The panel of promoter-variant Jurkat cells was generated in three sequential engineering steps using three different lentiviral constructs as described (Chapter 2, section 2.2.2). Jurkat cells transduced with the first lentiviral vector expressing rtTA3 under the CMV-promoter essential for the Dox-mediated induction of the 'Tet-ON' promoter, and selected using puromycin. The Jurkat-rtTA3 stable cell pool (Parent-I) was subsequently transduced with the second lentiviral vector containing the 'Tet-ON-Tat' expression

cassette and an independent 'EF1α-ECFP' expression cassette; the latter for stable selection of the CFP^{+ve} cells with an integrated 'Tet-ON-Tat' component; thus, generating the Jurkat-rtTA3-Tet-ON-Tat stable cells (Parent-II).

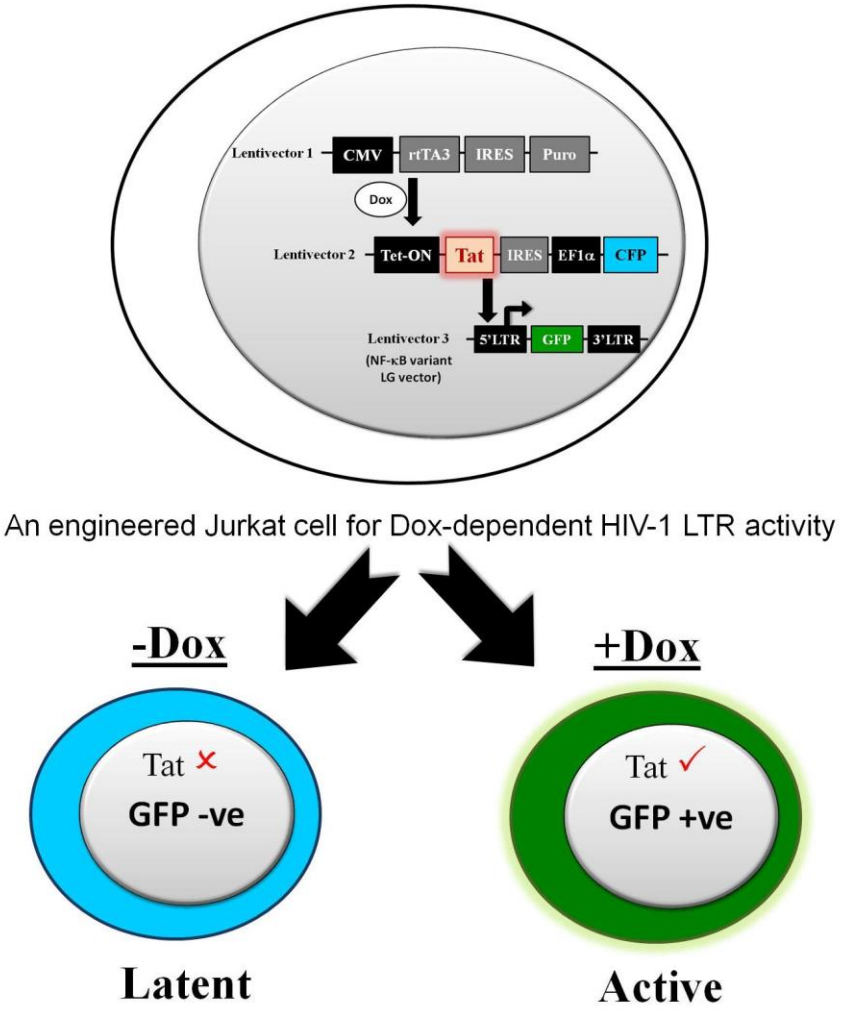


Figure 3.13: The 'Disjoint Tat feedback' (DTF) model of HIV-1 latency. The Tat ORF is de-linked from its natural LTR-circuit and expressed *in trans* in a Dox-inducible fashion. A panel of NF-κB-variant viral strains was generated in the HIV-1 vector backbone LTR-GFP (LG) (p912 series; Appendix-I). Jurkat cells were engineered using a three-step-infection approach to generate the stable Jurkat-LTR-GFP (Jurkat-LG) cells. First, stable Jurkat-rtTA3 (Parent-I) clonal lines were generated from Jurkat cells by infecting a CMV driven tetracycline transactivator (rtTA3) lentiviral construct followed by puromycin selection. Next, a Parent-I clone was stably transduced with the second lentivirus containing an inducible 'Tet-On-Tat' cassette and a constitutive 'EF1α-ECFP' cassette (Parent-II; Jurkat-rtTA3-Tet-ON-Tat); the latter for stable sorting of single CFP^{+ve} cells. A suitable Parent-II clone was chosen and infected with the viral strains of the LG panel to obtain the variant Jurkat-LG stable cells. In the absence of Dox, the Jurkat-LG cells fail to express GFP (CFP^{+ve} GFP^{-ve} phenotype; transcriptionally silent) and in the presence of Dox both Tat GFP are expressed (CFP^{+ve} GFP^{+ve}; transcriptionally active).

Stable cell clones of 'Parent-II' cells were generated by single-cell sorting and several stable cell lines were established. Cells expressing low-level CFP were sorted into individual wells of 96-well plates, and approximately 20 clonal lines were recovered and

expanded. Dox-responsiveness of the clones was examined using one of the representative clones '2c' in the presence of increasing concentrations of Dox (0, 250, 500, 750 and 1000 ng/ml) and using the protocol as described (Chapter 2, section 2.2.2). Tat-transcript levels increased in a Dox dose-dependent fashion. The highest level of Tat transcripts were induced at a Dox concentration of 750 ng/ml as quantitated in an RT-PCR reaction (Figure 3.14A).Under these experimental conditions, 94.87 ± 9.4 fold enhancement in the levels of Tat transcripts was detected following 750 ng/ml Dox-induction. Subsequently, Tat-transcript expression was compared in three different cell clones (1a, 2c, and 4b) in the presence of 750 ng/ml of Dox and clone 2c was selected for subsequent experiment based on Tat induction profile (Figure 3.14B). Clone 2c generated significantly higher levels of Tat at 24 h and 48 h following Dox induction, 98.6 and 135.3 fold, respectively.

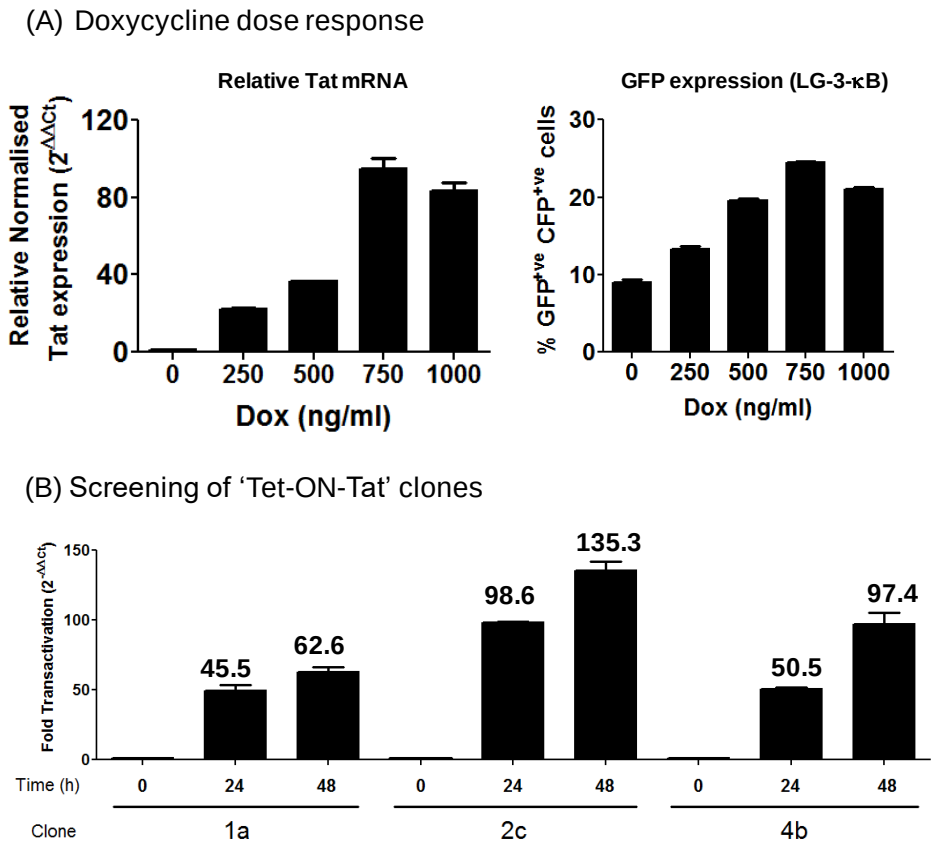


Figure 3.14: Validation of the Dox-mediated HIV-1 transcription in the DTF model. (A) Optimization of Dox concentration. The clonal line 2c (J-rtTA3-2c) was used to determine the optimal Dox concentration for the maximal Tat and GFP induction. Approximately two million cells were separately treated with varying concentrations of Dox for 24h as indicated. The treated cells were harvested, the total RNA extracted, the cDNA synthesized, and RT-PCR was performed for Tat transcripts. GAPDH was used as the reference gene control. Mean values of relative Tat transcripts (ΔΔCt method) from experimental triplicates ± SD are plotted. Data are representative of two independent experiments. A parallel experiment to determine percent GFP^{+ve} cells was conducted by infecting 0.3 million cells with the 3-κB

viral strain (LG-OHHC; 10 ng/ml p24 equivalent) and treating with varying doses of Dox. Mean values ± SD are plotted. Optimal Dox concentration was found to be 750 ng/ml. Data are representative of two independent experiments. **(B)** Selection of an appropriate Jurkat-rtTA3-Tet-ON-Tat (Parent-II) clone. Following the infection of Jurkat-rtTA3 (Parent I) cells with the 'Tet-ON-Tat-EF1α-ECFP' lentivirus at a low infectious titer (RIU ~0.01), a stable pool of low CFP^{+ve} cells was initially sorted and expanded followed by single-cell sorting of the pool. Twenty clones were established. Three of the clonal cell lines 1a, 2c and 4b were induced with 750 ng/ml of Dox and Tat transcripts were quantitated in an RT-PCR at 24 and 48 h following induction. The clonal line 2c (J-rtTA3-Tet-ON-Tat) was selected for the subsequent experiments.

Clone 2c was selected to generate the promoter-variant LG viral panel consisting of the five viral strains (p912 vector-series; Appendix-I). Jurkat-rtTA3-2c (J-rtTA3-2c) cells were infected with each LG-viral strain at ~0.5 RIU and three days following the infection, half of the infected cells were stimulated with a cocktail of global T-cell activators to compare GFP expression profiles between induced and un-induced promoters (Figure 3.15A). The cells were maintained at 750 ng/ml of Dox throughout the assay to maintain uniform levels of Tat among the cell lines, thus making Tat concentration not a variable among the cell lines (Figure 3.15B).The Tat-transcript expression among the variant stable cell pools was determined using an RT-PCR 24 h following the activation or in the absence of activation. Importantly, at a fixed Dox concentration (750 ng/ml), the five promoter-variant viral strains of the panel expressed comparable levels of Tat-transcript regardless of the κB-motif copy-number difference (Figure 3.15B). Thus, in the DTF model, the transcriptional strength of the viral promoter is uncoupled from the intracellular differences of Tat concentration among the panel members.

The stable cells of the Jurkat-LG (J-LG) panel thus generated would be positive for both the fluorescent proteins (CFP^{+ve} GFP^{+ve}) in the presence of Dox induction. Withdrawal of Dox will silence the 'Tet-ON' promoter and limit the synthesis of Tat thus expressing CFP but not GFP. Having confirmed uniform Tat expression among the panel of J-LG cell lines, we determined the GFP histogram profiles for each LG-variant with or without global activation but in the presence of a constant (750 ng/ml) Dox induction (Figure 3.15C).

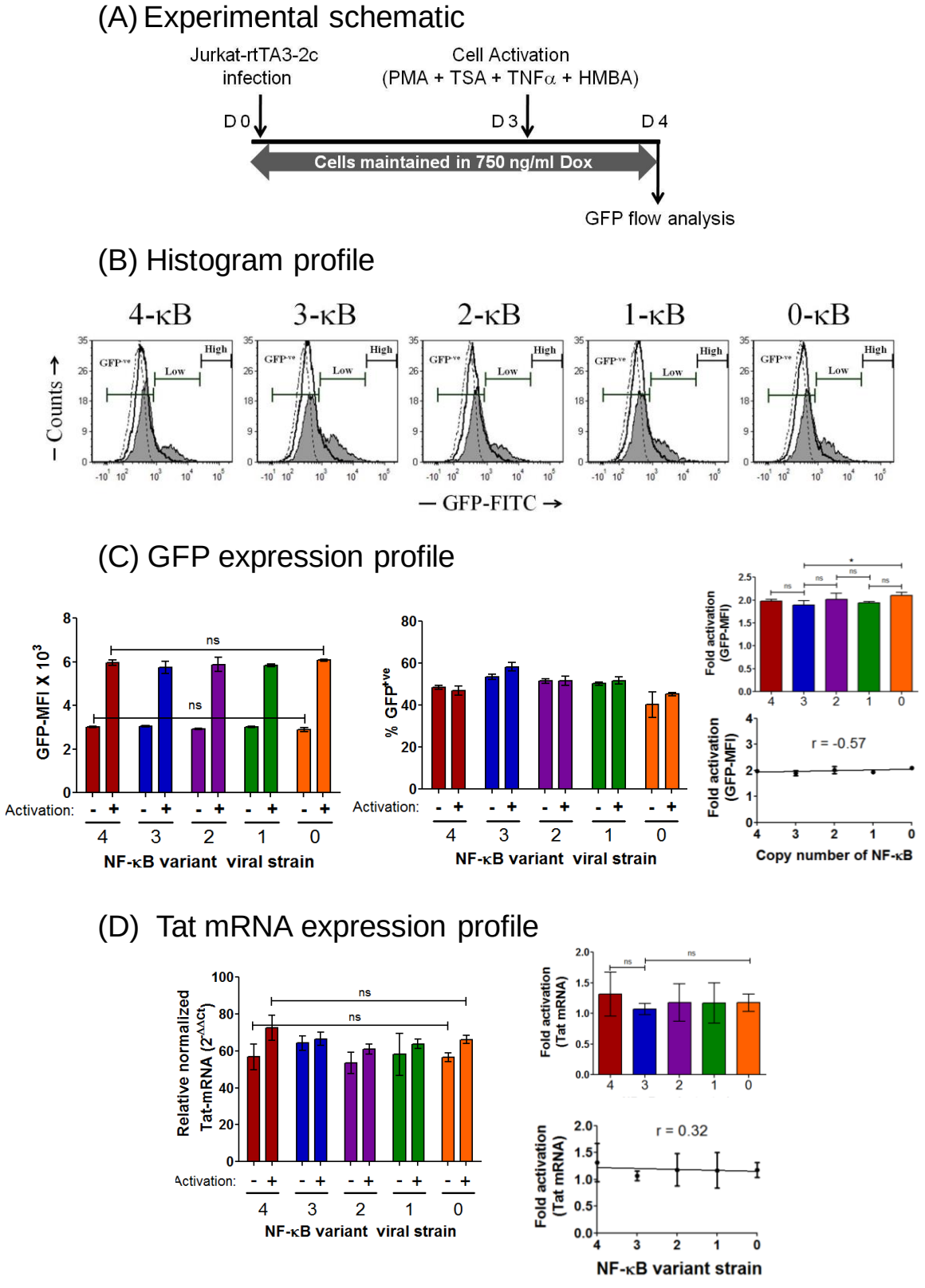


Figure 3.15: The DTF system fails to demonstrate Tat-transactivated population (A) Schematic representation of the experimental strategy. One million Jurkat-rtTA3-2c cells were pre-treated with 750 ng/ml Dox for 24 h and subsequently infected at ~0.5 RIU independently with each of the viruses of the panel as shown. After 72 h of infection, half of the infected cells were activated for 24 h followed by GFP expression analysis using a flow cytometer. The cells were maintained at 750 ng/ml Dox throughout the experiment. **(B)** Histogram profiles of GFP intensity among the

variants. The black dotted histogram represents Jurkat cells neither infected nor activated; the black hollow histogram represents cells infected but not activated. The solid grey histogram represents cells infected and activated. The intensity ranges for GFP^{-ve} and GFP^{Low}, and GFP^{High} cells are marked. Unlike in the ATF model (Figure 3.3), none of the NF-κB-variant strains in the DTF system displayed the Tat-feedback mediated transactivated population (GFP-MFI $>10^4$RFU). **(C)** Viral gene expression fails to correlate with the NF-κB copy number in the DTF model. The mean values of GFP-MFI from experimental quadruplicates ± SD both with or without activation (left panel) and the percent infected cells (middle panel) are presented. The fold transactivation of the variant LTRs (top, right panel) and the correlation between the NF-κB copy number and the fold transactivation (bottom, right panel) are presented. The data are representative of two independent experiments. Two-way ANOVA with Bonferroni post-test correction was used for the statistical evaluation (*$p<0.05$ and ns – non-significant). **(D)** Tat-transcript levels remain constant despite a variation in the copy number of functional NF-κB sites. Total mRNA was extracted from 0.5 million cells corresponding to control and activated populations. Relative Tat expression was determined in an RT-PCR using the ΔΔCt method. GAPDH gene expression was used as the normalization control. Mean values of the relative Tat expression from three independent experiments ± SEM are plotted. Two-way ANOVA with Bonferroni post-test correction was used for the statistical evaluation.The present data showing uniform levels of Tat mRNA in the cells infected with LTR-variant strains were further validated in the DTF model engineered to express Tat from an external inducible 'Tet-ON' promoter without any influence from the LTR.

Two different features are quite evident. First, the expression pattern of GFP with or without activation was indistinguishable among all the five cell lines regardless of the κB-motif copy-number difference. The GFP expression was comparable among all the J-LG cell lines with respect to percent expression, the MFI values, or fold activation (Figure 3.15D). The MFI values of all the five variant viral strains enhanced approximately by two-fold as compared to the respective controls without activation. As one might expect from the GFP-MFI data, LTR-inducibility among the variants were highly consistent (Figure 3.15C; top, right panel) and unrelated to the copy-number of NF-κB motifs (Figure 3.15.C bottom, right panel). The low correlation coefficient value ($r = -0.57$) is indicative of the absence of an association between the number of NF-κB sites and fold activation (Figure 3.15D). The second prominent feature of the DTF model is the conspicuous absence of the GFP^{High} cluster (GFP-MFI $>10^4$ RFU) unlike the ATF model. All the five J-LG cell lines, including that of 3- and 4-κB promoters, demonstrated the GFP^{Low} and GFP^{-ve} clusters but not the GFP^{High} cluster (Figure 3.15C). A comparison of the relative intracellular Tat-transcript levels between the two models is suggestive that the low-level Tat expression in the DTF model probably underlies the absence of the GFP^{High} cluster.

3.9 The kinetic profile of latency-establishment in the DTF model is conserved across the NF-κB-variants

Having confirmed uniform Tat expression at a fixed Dox concentration, we examined the kinetics of latency establishment using the panel of J-LG stable cells following the

experimental schematic depicted (Figure 3.17A). The cells were maintained for one day in the presence of 750 ng/ml of Dox and infected with the variant LG viral strains at an RIU of 0.1-0.2 to infect approximately 10-20% of cells. The cells were maintained for a week in complete RPMI medium supplemented with 750 ng/ml Dox, replenished every 24 h, and then, stimulated with the cocktail of global activators. Twenty four hours following activation, the CFP^{+ve} GFP^{+ve} cells representing active proviruses were sorted. The sorted cells of each variant virus were maintained in the presence (750 ng/ml) or absence of Doxycycline and the GFP expression profile of the cells of each viral variant was monitored every 24 h using a flow cytometer. The representative gating strategy for cell sorting is presented (Figure 3.16).

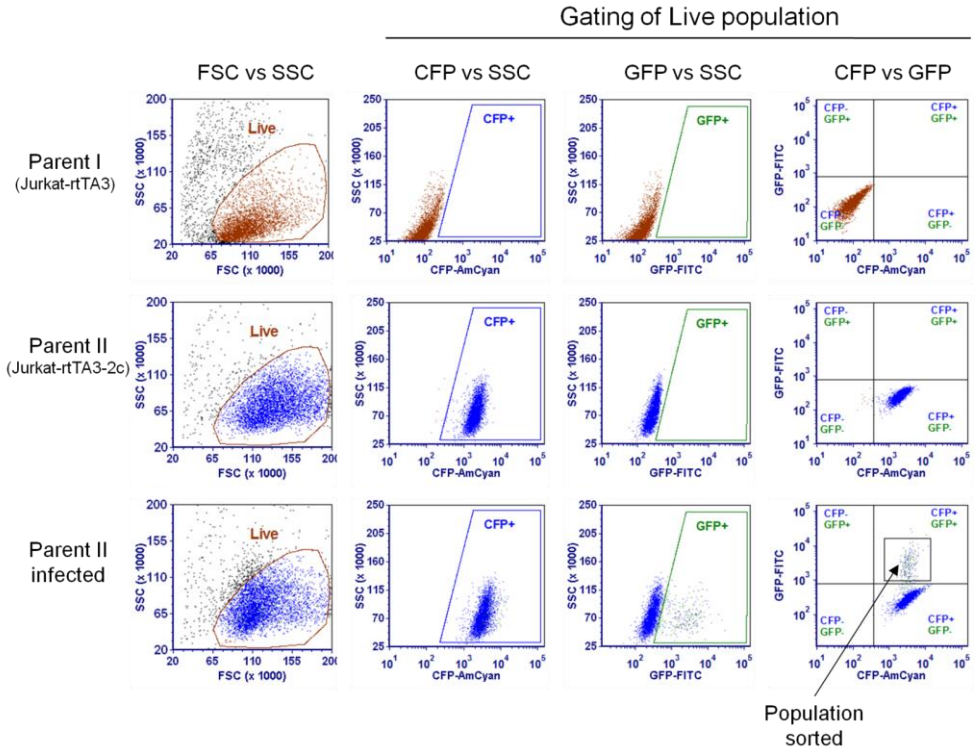


Figure 3.16: The gating scheme for the sorting of the Jurkat-LG-GFP^{+ve} cells. The three panels in the figure correspond to the three stages of the stable-cell generation in the DTF model (Figure 3.13). The parent-I in the first panel expresses rtTA3 from the CMV promoter and is devoid of any fluorescent marker. Parent-II (Panel 2) refers to the cell line: Jurkat-rtTA3-2c harbours the 'Tet-ON-Tat' cassette and constitutively expresses CFP but not GFP. A defined Parent-II cell line 2c was selected and used for the infection with the LG-variant viral strains (Panel 3). For the kinetic analysis of latency-establishment, the CFP^{+ve} GFP^{+ve} cells were sorted as indicated.

The stacked histogram profiles of GFP expression representing the viral strains are depicted in the absence (Figure 3.17B, left panel) or presence (Figure 3.17C, left panel) of Dox. Several interesting observations were made. Contrary to the ATF model where

GFP^{High} cells (GFP-MFI $>10^4$ RFU) were prominent, the GFP-histograms in the DTF system were devoid of the transactivated cells. At D0, immediately after sorting, a single GFP peak corresponding to the basal transcription (GFP-MFI 10^2-10^4) was noted for all the promoter variant viral strains in the absence of Dox (Figure 3.17B) or in its presence (Figure 3.17C). The GFP expression dropped sharply both in the absence (Figure 3.17B; top, right panel) or presence (Figure 3.17C top, right panel) of Dox from all the promoters. The GFP switch off was particularly sharper in the absence of Dox-mediated Tat expression. Nearly all the cells stopped expressing GFP by day 6 in the absence of Dox but a few cells continued to express in the presence of Dox at the end of day 7. The MFI values too decreased precipitately.

Tat expression was uniform among all the five variant viral strains at day 0 and at the subsequent time points in the presence of Dox. No Tat transcripts were seen in the cells in the absence of Dox supplementation (Figure 3.17D). Collectively, the data demonstrated a rapid latency establishment regardless of the difference in the κB-motif copies when the LTR-Tat-feedback axis is disrupted in the DTF model unlike that of the ATF model where the LTR-Tat axis is preserved. One technical limitation of the DTF model is the significantly low level of Tat expression even under the most optimal conditions. In the DTF model, the fold activation of the Tat transcript levels could only be mildly enhanced by the Dox induction (750 ng/ml) compared to the saturating Tat-transcript levels generated in the ATF model. Thus, the Tet-ON promoter system offered the control of Tat expression but could not produce sufficient quantities of Tat to recapitulate the physiologically relevant concentrations of Tat.

In summary, the results of the ATF and DTF models collectively point at Tat and the intracellular concentration of Tat playing a critical and diametrically opposite functions in controlling HIV latency reversal and establishment. Our data are assertive that the HIV-1 LTR containing a larger number of functional NF-κB sites generates a higher quantity of intracellular Tat which in turn regulates viral latency by causing a stronger LTR-Tat-feedback loop.

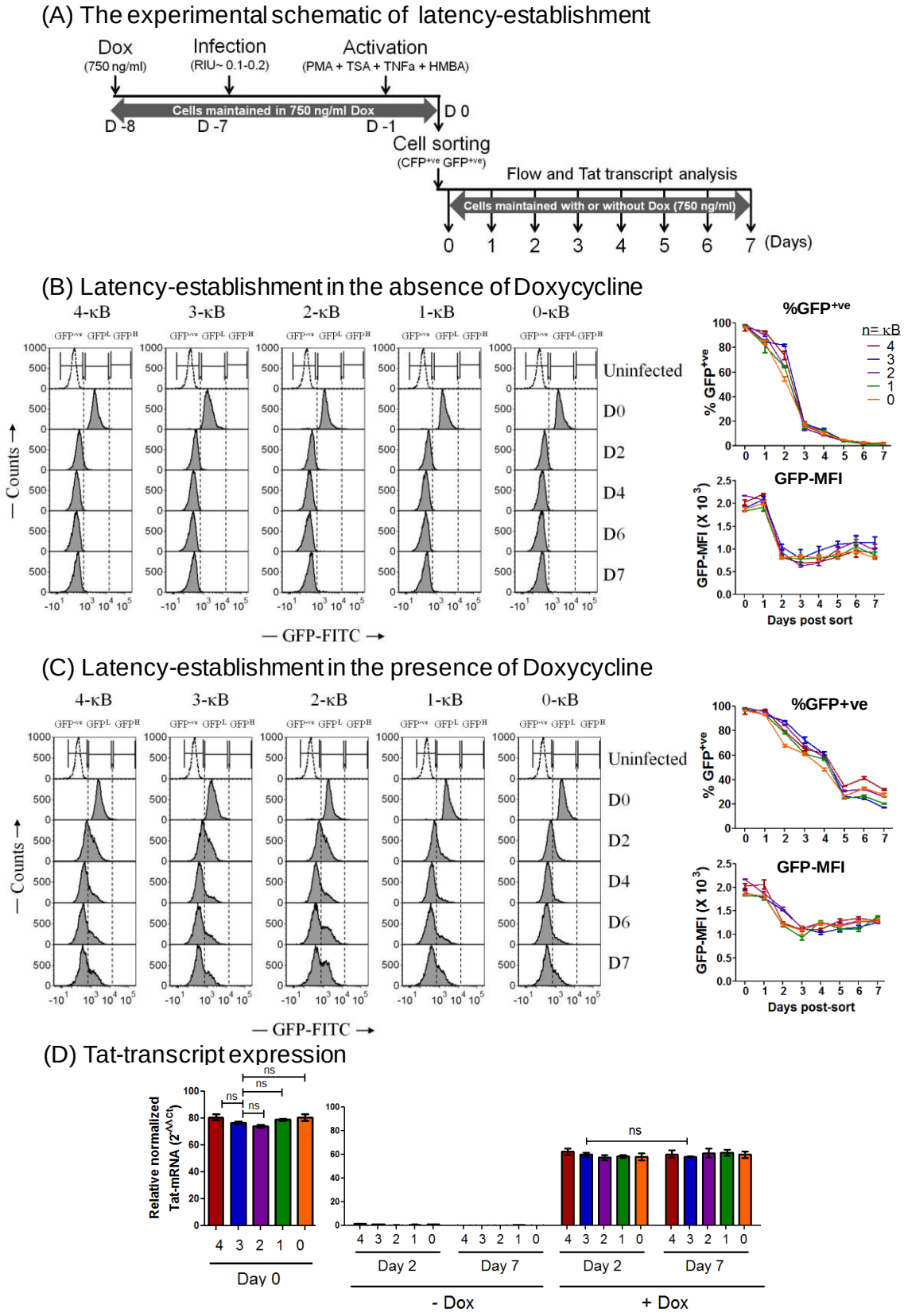


Figure 3.17: The kinetics of viral latency-establishment in the DTF model (A) A schematic representation of the experimental layout. One million Jurkat-rtTA3-2c cells were pretreated with Dox for 24 h and infected independently with each reporter LG viral strain at a low infectious titer (~0.1-0.2 RIU). Following activation (PMA+TNFα+TSA+HMBA), the infected cells were cultured for seven days in the continued presence of Dox. After

24 of activation, the CFP^{+ve} GFP^{+ve} cells were sorted from each viral variant. The sorted cells were maintained either in the absence or presence of Dox (750 ng/ml) and the GFP expression was monitored using a flow cytometer every day for a week. Tat-transcript levels were assessed at three different time points (days 0, 2 and 7) at both the Dox concentrations. **(B)** Latency kinetics in the continued presence of Dox. Stacked-histogram profiles for the variant strains are shown (left panel). As evident from the percent GFP (top, right panel) and the GFP-MFI (bottom, right panel) profiles, upon de-linking the LTR-Tat axis, the DTF model failed to produce a significant difference in the kinetic profile of latency-establishment among the NF-κB-variant strains. **(C)** Latency kinetics in the absence of Dox. The stacked histograms (left panel), the percent GFP^{+ve} (top, right panel) and GFP-MFI (bottom, right panel) are depicted. Mean values from experimental quadruplicates ± SD are plotted. Two-way ANOVA with Bonferroni post-test correction was used for the statistical evaluation (ns – non-significant). Data represent three independent experiments. (**D**) The Tat-transcript expression profile. Cellular RNA was extracted at 0, 2 and 7 days post-sorting from 0.5 million cells at each time point from all the NF-κB-variants in both the +Dox and –Dox conditions. Relative Tat expression was quantified in RT-PCR using the ΔΔCt method. GAPDH was used as the reference gene. Mean values from three independent experiments ± SEM are plotted. Two-way ANOVA with Bonferroni post correction was used for statistical evaluation (ns - non-significant).

3.10 The 'Tunable Tat-feedback' (TTF) latency model: LTR-transactivation is directly proportional to the strength of the LTR-Tat-feedback circuit.

We have demonstrated that the pattern of silencing of HIV-1C differs between the ATF and the DTF models despite the caveat that not enough Tat was made in the DTF model. Nevertheless, the common theme that emerged from both the models is that both the transcriptional strength of the LTR and the presence of a functional Tat-feedback circuit are necessary for the initiation of viral latency. In the ATF model that preserves the natural LTR-Tat transcription circuit, the rapidity of viral latency establishment was directly proportional to the number of NF-κB elements in the enhancer. The DTF model emphasised the significance of an intact LTR-Tat-feedback axis in the absence of which the kinetic profiles of transcriptional silence did not vary significantly, despite the difference in the NF-κB copies. However, a technical limitation of the DTF model was the inability to achieve Tat levels at par with that of the ATF model. To circumvent the low level Tat expression in the DTF model, we devised a different model to regulate Tat levels keeping the transcriptional strength of the viral promoter constant i.e. a fixed number of NF-κB motifs in the LTR.

In the 'Tunable Tat-feedback' (TTF) model, the LTR-Tat-feedback circuit was modulated by regulating the half-life of Tat based on the principle of the 'tunable proteolytic degradation' (Banaszynski LA et al., 2006). The Tat-RFP fusion protein in this model was tagged with the C-terminal degradation domain (DD) of FK506 binding protein (FKBP).

The presence of the DD motif at the C-end prompts the 'Tat:RFP:DD' fusion protein for rapid degradation through the proteasome machinery. The fusion protein, however, can be reversibly rescued and stabilised by the use of a small molecule ligand 'Shield' which binds and protects the protein from rapid processing. Thus, by the selective addition of Shield1 to the culture medium, the 'Tat:RFP:DD' fusion protein can be physiologically sustained for longer periods in the presence of Shield1 or exposed to rapid degradation in its absence (Figure 3.18). Furthermore, a dose response is evident between the concentration of Shield1 in the medium and RFP expression, thus, the model would permit a quantitative analysis of the effect of Tat on viral latency (Figure 3.19).

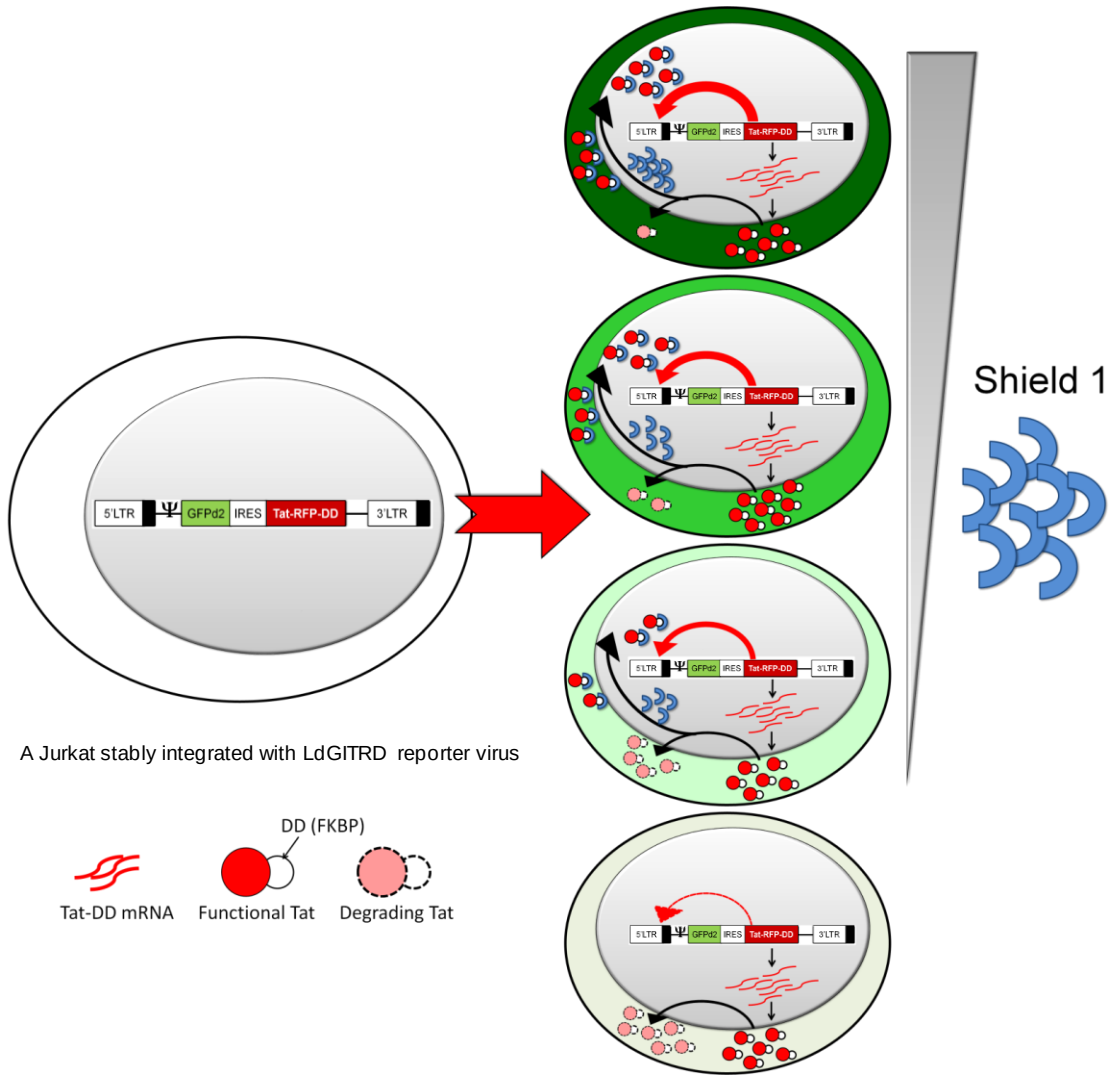


Figure 3.18: The 'Tunable Tat feedback' (TTF) model of HIV-1 latency. The design of the minimal HIV-1 vector LdGITRD as has been discussed in Figure 2.3. A panel of LTR-variant viral vectors (p913 series; Appendix-I) has been generated (Figure 2.3). DD in 'Tat:RFP:DD' targets the fusion protein to proteasomal degradation which can be reversed using Shield1- a small-molecule ligand, thus, enabling the modulation of the physiological levels of the Tat fusion protein.

In the minimal HIV-1 reporter vector- LTR-GFPd2-IRES-Tat:RFP:DD or the LdGITRD, (Figure 3.18, see Chapter 2, section 2.1.3 for the vector design and construction) GFPd2, a variant form of EGFP with low half-life (2h, Li X et al., 1998) is directly expressed under the LTR thus, faithfully reflecting the transcriptional activity of the viral promoter. The 'Tat:RFP:DD' cassette was also co-expressed from the LTR but under the translational control of an IRES element. Importantly, the fusion of Tat with DsRed2-RFP not only offers the advantage of tracing the presence of Tat but also permits the categorization of the active viral promoter (GFP^{+ve} cells) into Tat-dependent ($Tat\text{-}RFP^{+ve}$) and Tat-independent ($Tat\text{-}RFP^{-ve}$) transactivation. We constructed a full panel of LTR-variant reporter viral strains with the number of functional NF-κB motifs varying from 4 to 0 (the LdGITRD viral strains, the p913 series) and established stable Jurkat cells. We, however, used only two different viral strains one each representing the stronger (3-κB LTR) and the weaker (1-κB LTR) viral promoters for the subsequent analysis. At the level of validation, we examined the expression of GFP and RFP in HEK293T cells (Figure 3.19A). Approximately, 0.6 million HEK239T cells were transfected with 1μg LdGITRD-3-κB vector in each well of a 12-well culture dish and treated with Shield1 concentrations ranging from 0 to 5 μM as indicated. After 48h of transfection, the expression of both RFP and GFP was recorded using a fluorescent microscope. While the expression of GFP was largely independent of the presence of Shield1, we observed a perfect correlation between the concentration of Shield1 and RFP expression. It was evident that the expression of GFP was several-fold stronger than that of RFP. Unlike GFP, the expression of RFP was under the translational control of an IRES element which imposes a cost of 10-100 fold inferior gene expression (Mizuguchi H et al., 2000). Furthermore, the slower maturation time of DsRed2 (24h) as compared to that of GFPd2 (6 h) could also be an additional factor that underlies the intensity difference between the fluorescent proteins (Piatkevich KD and Verkhusha VV., 2011).

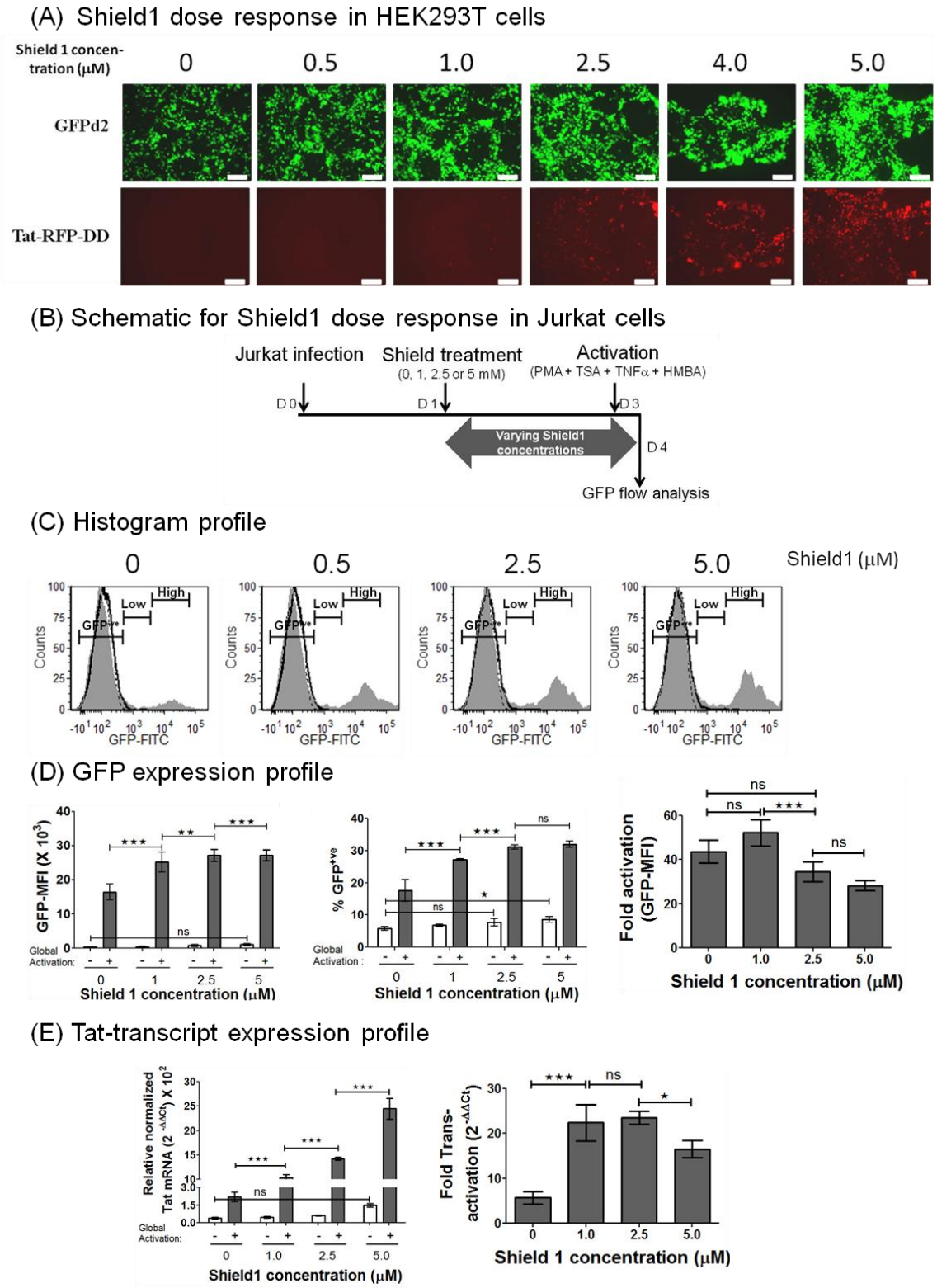


Figure 3.19: Validation of the TTF model of HIV-1 latency. (A) Shield1 dose-response assay in HEK293T cells. Approximately, 0.6 million HEK293T cells were transfected with 1 μg of the LdGITRD-3-κB vector in the presence of varying concentrations of Shield1 as indicated. GFP and RFP images were captured 48 h post-transfection. The experiment was repeated twice with comparable results. **(B)** The experimental schematic for the Shield1 dose response in Jurkat cells. Approximately, 0.3 million Jurkat cells were infected with the LdGITRD-3-κB minimal virus (20 ng/ml; p24 equivalent) and post 24 h, the infected cells were split into four fractions and each treated with a different concentration of Shield1 ranging from 0 to 5 μM. After 48 h of infection, half of the cells from each fraction were

activated for 24 h followed by the quantitation of the GFP and Tat mRNA expression levels for both the induced and uninduced fractions. **(C)**The representative GFP histogram profiles at a varying concentrations of Shield1. **(D)** Shield1 dose-dependent LTR-transactivation. **(E)** Shield1 dose-dependent Tat-mRNA expression. Both the GFP-MFI and percent GFP^{+ve} values (3.19 C left and central panels), as well as the Tat mRNA levels (3.19 D left panel) responded to Shield1 dose. The optimal Shield1dose for fold activation was obtained between 1.0 to 2.5 μM (3.19 C right panel and 3.19 D right panel). Mean values from experimental triplicates ± SD are plotted. Relative Tat expression was quantified (by ΔΔCt method) against the housekeeping control GAPDH. Mean values from three independent experiments ± SEM are plotted. Two-way ANOVA with Bonferroni post-test correction was used for the statistical evaluation (*$p<0.05$, **$p<0.01$, ***$p<0.001$ and ns - non-significant).

The 3-κB viral strain produced in HEK293T cells was examined for infectivity and fluorescent protein expression in Jurkat T-cells in the absence or presence of Shield1 at varying concentrations (0, 1.0, 2.5, and 5.0 μM) (Figure 3.19B). Approximately, 0.3 million Jurkat Cells were infected with 20 ng/ml of p24-equivalent of the viral stock and 24 h later the cells were divided into four fractions. The fractions were treated with varying concentrations of Shield1 (0, 1.0, 2.5 or 5.0 μM) for 48 h. Following this, the cells were activated with the cocktail of global activators for 24 h and the GFP expression was estimated in flow cytometry. Interestingly, the GFP histogram profile with varying concentrations of Shield1 clearly indicated a dose-dependent increment in the peak of GFP^{High} cells indicating the direct influence of Shield1 on Tat-transactivation (Figure 3.19C). The magnitude of LTR-transactivation or GFP expression in the infected Jurkat cells was directly proportional to the concentration of Shield1 as indicated by the GFP-MFI values (Figure 3.19D, left panel). An increase in the Shield1 concentration from 0 to 5 μM enhanced GFP-MFI values from 375.98 ± 12.2 to 1,047.67 ± 79 in the absence of global activation and from 25,412.16 ± 359.9 to 29,429.04 ± 238.1 when the cells were activated. Of note, although in our protocol the cell infection was a common event, the GFP expression demonstrated dose-response proportional to the Shield1 concentration even though the GFP itself does not contain the DD of FKBP (Figure 3.19D, central panel). This was possibly an indirect effect of the progressively enhanced LTR-Tat-feedback loop due to the increased Tat:RFP:DD stabilization proportional to the increased Shield1 concentration. Importantly, the quantitation of the Tat-transcripts also demonstrated a perfect dose response proportional to the Shield1concentration (Figure 3.19E, left panel). The fold activations of the expression of GFP (Figure 3.19D, right panel) and Tat-transcripts (Figure 3.19E, right panel) were found maximum at 1 μM and 2.5 μM, respectively. In the subsequent experiments therefore, we used Shield1 in the range of 0 to 3 μM.

3.11 The TTF model identifies two distinct modes of latency-establishment in HIV-1C and detectable quantities of Tat were observed in the LTR-'OFF' state.

Having validated the TTF model, we selected two viral strains (3- and 1-κB LTRs) as representatives of the stronger (3- and 4-κB LTRs) and weaker (0-, 1- and 2-κB LTRs) viral promoter variants and determined the kinetics of latency establishment in the presence of varying concentrations of Shield1, as depicted schematically (Figure 3.21A). Jurkat cells were infected with 3- or 1-κB viral strains at an RIU of ~0.1-0.2 in the presence of 1 μM Shield1 and expanded for a week in the presence of Shield1. Subsequently, the cells were activated with a cocktail of global activators for 24 h following which GFP^{High} cells (GFP-MFI ~10^4 RFU) were sorted. The sorted cells were divided into four batches and maintained separately at different concentrations of Shield1 (0, 0.5, 1.0 and 3.0 μM). The expression of GFP and RFP was monitored by flow every 24 h. Shield1 was supplemented at the respective concentrations every 24 h. The representative gating strategy for the sorting of GFP^{High} cells is presented (Figure 3.20).

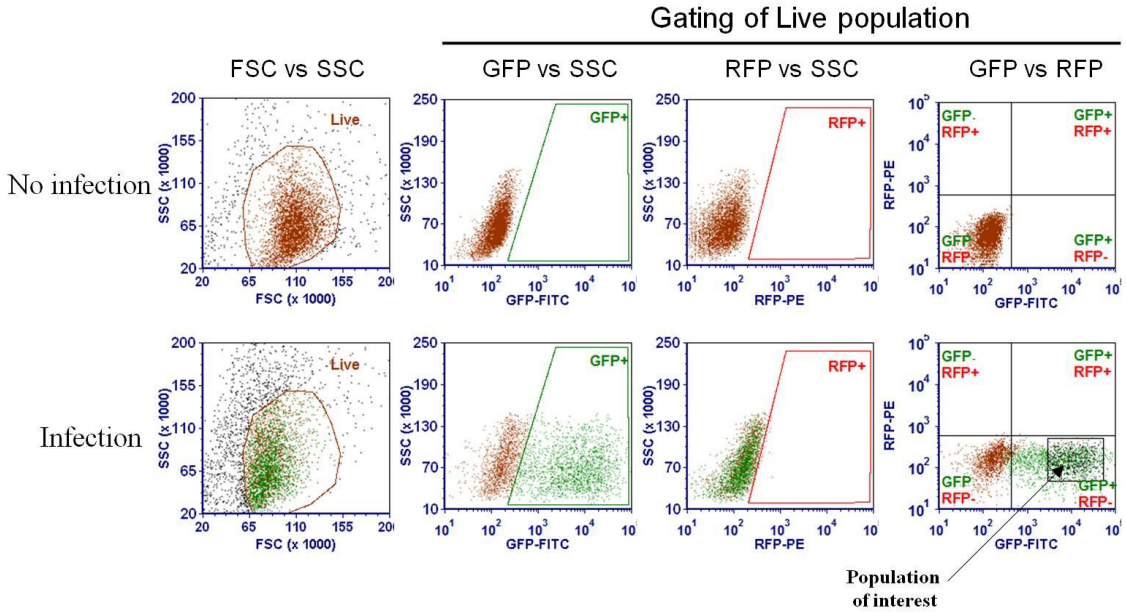


Figure 3.20: A schematic of cell-sorting in the TTF model. In the TTF model, two distinct fluorescent markers were used, GFPd2 (for the tracking of the LTR activity) and DsRed2-RFP (the real-time expression kinetics of Tat).

The observations of the latency assay in the TTF model were multifaceted (Figures 3.21 and 3.22). Of primary importance was the identification of distinct temporal cascades of

GFP and Tat-RFP expressions in the course of latency establishment. Secondly, the temporal profiles significantly differed between the two variant promoters used in the study. Lastly, two unique modes of viral silencing were recognised, a different mode predominant in each promoter-type (Figure 3.21B and C).

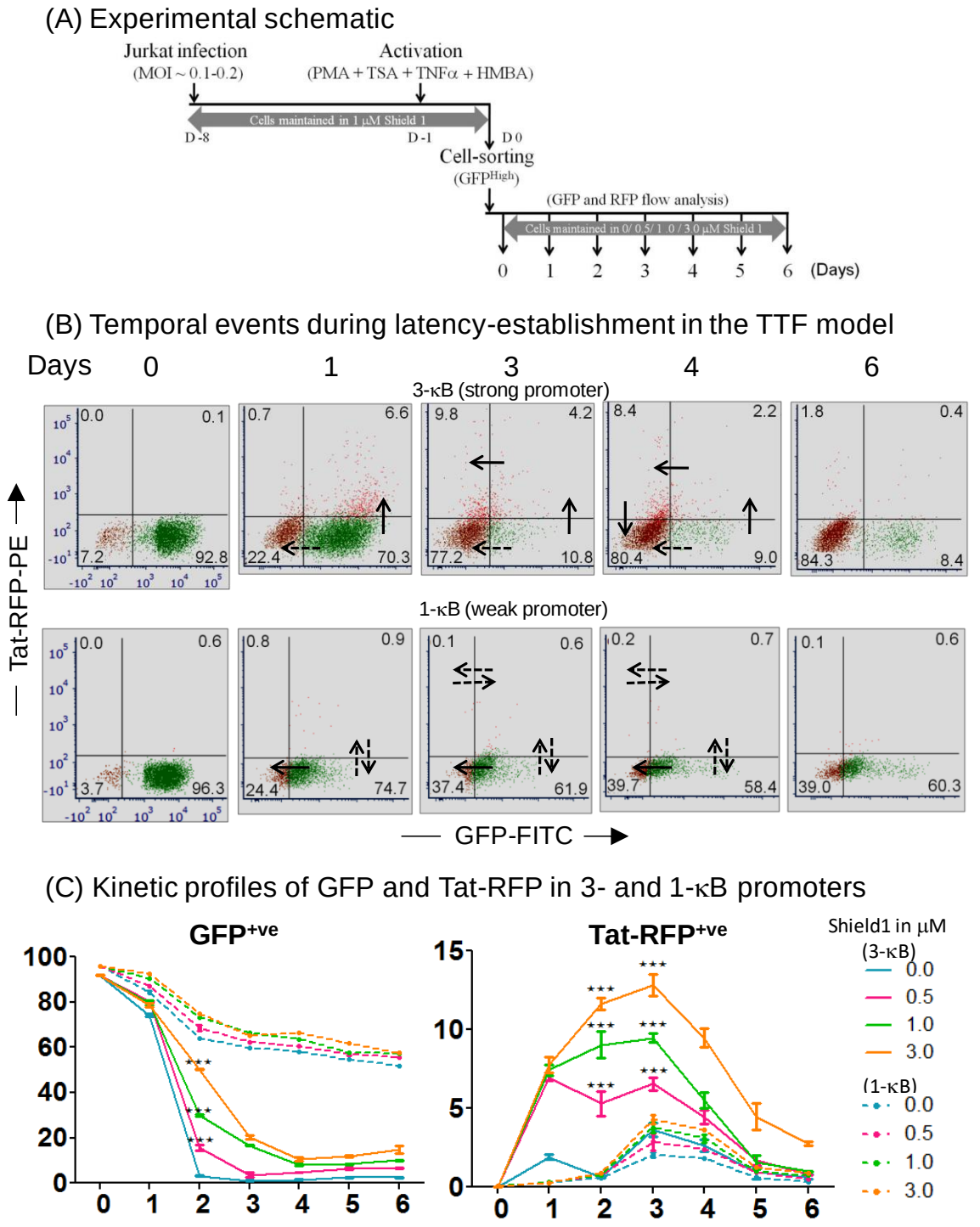


Figure 3.21: The kinetics of viral latency-establishment in the TTF model. (A) The time schematic of the assay. One million Jurkat cells were independently infected at a low infectious titer (RIU~0.1-0.2) with one of the two NF-κB-variant LdGITRD strains 3- and 1-κB, representing the strong and the weak promoters, respectively. The infected cells

were maintained for seven days in culture supplemented with 1 μM of Shield1, treated with global T-cell activators for 24 h and GFP^{High} cells were sorted. The sorted cells were split into four separate fractions and each maintained at a different concentration of Shield1 (0, 0.5, 1.0 or 3.0 μM). The cells were analyzed every 24 h for temporal pattern of GFP and RFP expression using flow cytometry. **(B)** Two distinct modes of latency-establishment in HIV-1C. Based on the GFP and RFP expression profile, the strong (3-κB) and the weak (1-κB) LTRs followed different trajectories of transcriptional silence. The four quadrants of the GFP vs Tat-RFP dot plots represent distinct phases of the LTR-activity and Tat expression for 3 μM Shield1. The weak LTR predominantly follows a Tat-independent silencing route (lower panel) while the strong LTR additionally demonstrates a Tat-dependent latency mechanism (upper panel). Interestingly, the Tat-dependent latency path in the strong promoter routes the cells via the GFP^{-ve} Tat-RFP^{+ve} quadrant (upper, left quadrant) back to latency indicating the presence of Tat in latent infection (see Figure 3.22 for quantitative analysis). **(C)** Kinetic profiles of %GFP^{+ve} and %Tat-RFP^{+ve} cells. The frequency of GFP^{+ve} to GFP^{-ve} switch demonstrated Shield1 dose dependence within each LTR-variant and an over-all faster kinetics in the strong LTR compared to the weak LTR. A prominent, Shield1 dose responsive Tat-RFP expression was also noted in the 3-κB LTR but not in the 1-κB LTR. Importantly, %Tat-RFP^{+ve} cells in the 3-κB LTR peaked on day-3, the time-point at which most of the GFP^{+ve} cells had transited the GFP^{-ve} phenotype. The solid and dotted lines denote the kinetic curves corresponding to 3- and 1-κB LTRs, respectively. Mean values from experimental triplicates ± SD are plotted. Data represent three independent experiments. Two-way ANOVA with Bonferroni post-test was used for the statistical evaluation (*$p<0.05$, ***$p<0.001$ and ns - non-significant).

In the TTF system, unlike in the DTF model described previously, the expression kinetics of Tat could be directly interrogated using RFP. Thus, the transcription activity of the viral promoter and the expression profile of Tat could be simultaneously visualised using GFP and RFP, respectively. Taking advantage of the expression of the two fluorescent proteins, we monitored viral latency using flow cytometry every 24 hours for six days in Jurkat cells- infected, activated and sorted for GFP^{High} cells (Figure 3.21A). The data of the TTF model, which are highly reproducible, not only appear to be broadly consistent with those of the ATF model but also provided more information regarding the nature of HIV-1 latency establishment. Prior to sorting, only less than 1% cells were positive for RFP expression even though 10-20% cells were GFP^{+ve}. The increased molecular size of the 'Tat:RFP:DD' fusion protein, the slow maturation of DsRed2, and compromised translation efficiency due to the IRES element control, all may have contributed to the observed difference between the GFP and RFP expression profile.

We primarily focused on the cells positive for GFP expression since these cells indicate the transcriptionally active viral promoter. The cells were analysed for the expression of both the fluorescent proteins in Jurkat cells infected with 3- or 1-κB viral strains, and this analysis was continued through the following six days (Figures 3.21 and 3.22). Soon after sorting, at the day-0 time point, all the cells were GFP^{+ve} representing a transcriptionally active viral promoter (Figure 3.21B; GFP^{+ve} Tat-RFP^{-ve} cells and Figure 3.22; bottom, right panel). The difference in the transcriptional strength between the two viral promoters had no bearing on this since the cells were selected for GFP expression. At this

time point, barely any cells expressed Tat-RFP suggesting that the transactivation activity was Tat-independent. The demarcation in the behaviour of the two viral promoters in the establishment of viral latency was evident from day-1 onwards. The stronger viral promoter (3-κB LTR) rapidly upregulated Tat-RFP expression within 24 hours owing to a strong and rapid Tat-transcriptional circuit such that RFP expression reached a peak by day-2 (Figure 3.21B; GFP^{+ve} $Tat\text{-}RFP^{+ve}$ cells and Figure 3.22; top, right panel). These cells however, represented only a minority, approximately 7.5% of the GFP^{+ve} cell population, whereas the bulk of the cells at day-2 were GFP^{+ve} $Tat\text{-}RFP^{-ve}$. Thus, at day-2 the 3-κB LTR demonstrated two independent modes of transactivation: Tat-independent and Tat-dependent. Importantly, at the peak of Tat-RFP expression at day-2, the process of transcriptional silencing had already begun. By day-3 and on the subsequent days, the dual-positive cells rapidly down regulated the expression of GFP, indicating the LTR switch-off, to become GFP^{-ve} $Tat\text{-}RFP^{+ve}$ cells (Figure 3.21B and Figure 3.22; top, left panel). These GFP^{-ve} $Tat\text{-}RFP^{+ve}$ cells rapidly extinguished RFP expression to return to the double negative state. In parallel, the GFP^{+ve} $Tat\text{-}RFP^{-ve}$ cell population also started to down regulate GFP expression and rapidly returned to the double-negative state (Figure 3.21B; GFP^{-ve} $Tat\text{-}RFP^{-ve}$ cells; and Figure 3.22; bottom, left panel). Thus, while the LTRs activated by Tat-independent transactivation (GFP^{+ve} $Tat\text{-}RFP^{-ve}$ cells) returned to latency directly, the LTRs activated by Tat-dependent transactivation (GFP^{+ve} $Tat\text{-}RFP^{+ve}$ cells) established latency via GFP^{-ve} $Tat\text{-}RFP^{+ve}$ phenotype (Figure 3.22). The GFP^{-ve} $Tat\text{-}RFP^{+ve}$ cells are unique because in these cells the LTR was switched off despite the presence of significant quantities of Tat (Tat-RFP).

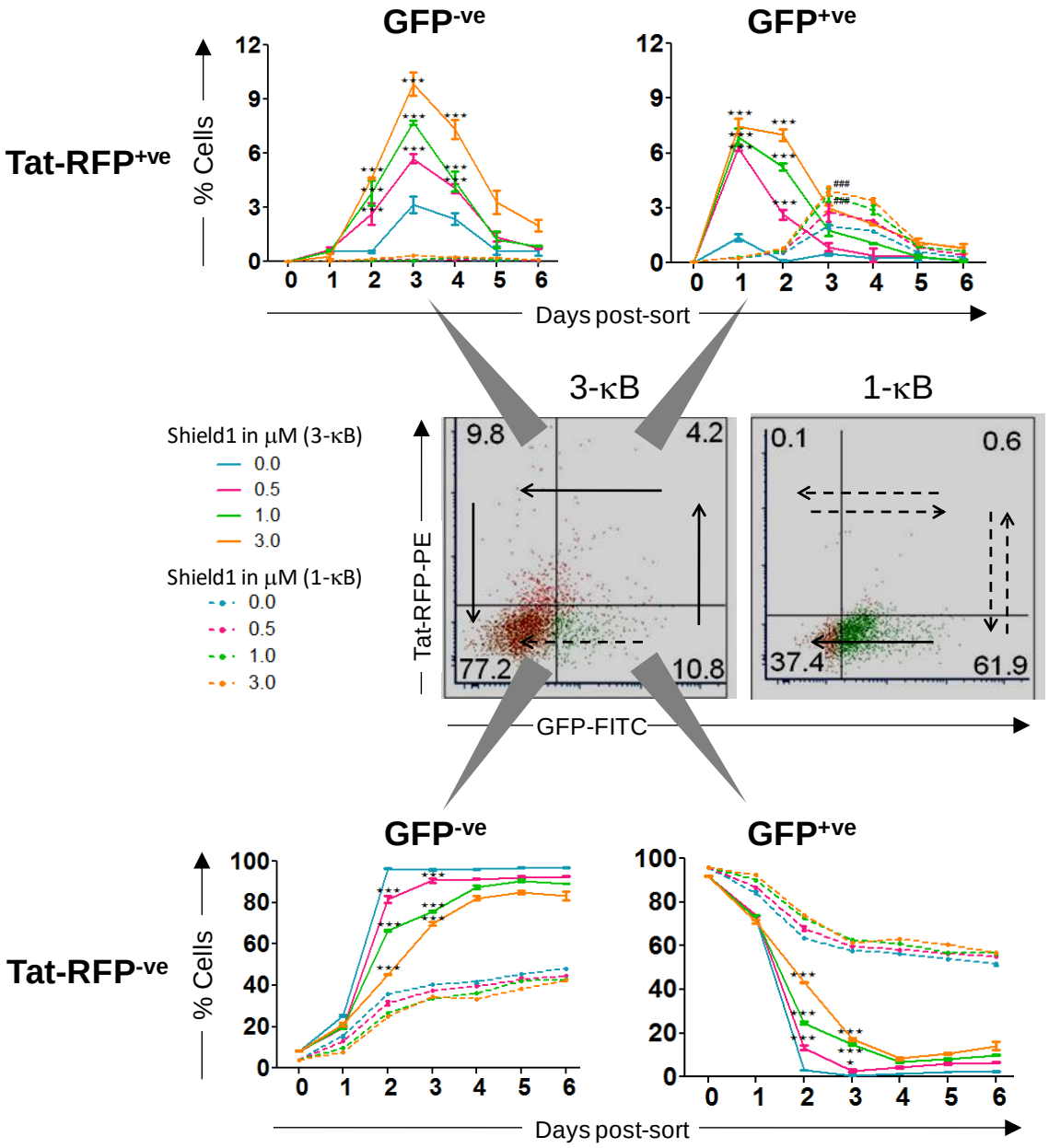


Figure 3.22: The Tat-dependent and Tat-independent modes of viral latency-establishment. The strength of the promoter determines the silencing route in the TTF model to be either Tat-dependent (Tat-RFP^{+ve}, upper panel) or Tat-independent (Tat-RFP^{-ve}, lower panel). A snapshot of cellular distribution in the four quadrants of the GFP vs Tat-RFP dot plot has been presented for both the strong (3-κB) and the weak (1-κB) LTRs on day-3 and for 3 µM Shield1 concentration (central panel; see Figure 3.21B for detailed temporal events). While the Tat-independent silencing pathway involves the simple transition from GFP^{+ve} Tat-RFP^{+ve} (lower right panel) to GFP^{-ve} Tat-RFP^{+ve} cells (lower left panel), the Tat-dependent route involves the transition through two additional phases- GFP^{+ve} Tat^{+ve} (upper right panel) and GFP^{-ve} Tat-RFP^{+ve} (upper left panel). A comparison of the kinetic profiles for all the four phenotypes between the strong and the weak promoter variants has been presented. The solid and dotted lines denote the kinetic curves corresponding to 3- and 1-κB LTRs, respectively. Mean values from experimental triplicates ± SD are plotted. Data represent three independent experiments. Two-way ANOVA with Bonferroni post-test was used for the statistical evaluation (*p<0.05, ***p<0.001 and ns - non-significant).

The kinetics of latency establishment of the 1-κB LTR appeared to differ considerably from that of the stronger viral promoter (3-κB) both qualitatively and quantitatively. The 1-κB LTR predominantly displayed the Tat-independent transactivation (lower panel, Figure 3.21B). Although approximately 4% of these cells expressed RFP at a Shield1 concentration of 3 μM (Figure 3.22; top, right panel), the Tat-RFP expression was delayed by 24 h, as compared to that of the 3-κB LTR, with the RFP expression reaching a peak only on day-3. Importantly, despite the presence of Tat, these dual-positive cells of 1-κB LTR returned to the GFP^{+ve} RFP^{-ve} quadrant, but not to the GFP^{-ve} $Tat\text{-}RFP^{+ve}$ compartment, unlike the cells of 3-κB LTR. The progression of the dual-positive cells of 1-κB LTR to the GFP^{-ve} $Tat\text{-}RFP^{+ve}$ quadrant would have manifested the presence of RFP^{+ve} cells in this quadrant on day-4 or later, however, no such cells were seen (Figure 3.22, top, left panel). Thus, the weak HIV-1 LTR appears to function primarily through Tat-independent transactivation and is incapable of effectuating sustained levels of Tat-dependent transactivation unlike the stronger LTR (compare the differences in the pattern of fluorescent cell migration, (Figure 3.21B).

Furthermore, a clear demarcation in the profiles of the weak and stronger LTRs is evident at the level of Tat-independent transactivation– the RFP negative cell populations (Figure 3.22, bottom panels). The GFP^{+ve} $Tat\text{-}RFP^{-ve}$ cells of the 3-κB LTR down regulated GFP at all the concentrations of Shield1 by day-4 (Figure 3.21B, bottom, right panel) and rapidly migrated to the dual-negative quadrant. In contrast, latency-establishment in the GFP^{+ve} $Tat\text{-}RFP^{-ve}$ cells of the 1-κB LTR was incomplete and nearly half of these cells remained GFP positive on day-6 (Figure 3.22, bottom, right panel). This was primarily because a subset of the GFP^{+ve} $Tat\text{-}RFP^{-ve}$ cells at the later time points (day-1 and beyond) followed the Tat-dependent route to latency in the case of the stronger 3-κB, but not the weaker 1-κB, promoter. Therefore, from the data of the TTF model it appears that a Tat-dependent transactivation can silence the promoter at a faster rate as compared to that of the Tat-independent pathway. Further, the kinetics of percent GFP^{+ve} to GFP^{-ve} transition, irrespective of the Tat-RFP expression (GFP^{+ve} $Tat\text{-}RFP^{-ve}$ and GFP^{+ve} $Tat\text{-}RFP^{+ve}$ cells combined) demonstrated an identical pattern of promoter silencing when compared with the ATF model (Figure 3.21C; left panel). At all the concentrations of Shield1, the stronger 3-κB LTR switched off faster than the weaker 1-κB LTR. Thus, the data obtained from the TTF model are strongly suggestive that the transcriptional strength of

the HIV-1 LTR plays a critical role in controlling viral latency as a validation of the ATF model. A stronger LTR is not only faster in establishing viral latency but also is rapid in revival kinetics from latency, whereas a weak viral promoter appears to behave in a diametrically opposite fashion.

Of note, the remarkable merit of the TTF model lies in its ability to differentiate between the individual contributions made towards viral latency by the transcriptional strength of the viral promoter and the transactivation strength of Tat that are intricately interdependent. A positive feedback loop established between the LTR and Tat precludes attempts to differentiate between the individual function of the LTR and Tat to evaluate which of these two elements plays the cardinal role in regulating transcription and latency. In the TTF model, it was possible to separate these two functions and appreciate their independent contributions towards viral latency. The kinetics of the establishment of viral latency was compared between two viral promoters (1-κB LTR vs 3- κB LTR) that differed remarkably in the transcriptional strength and in the presence of varying intracellular concentration of Tat. A comparison of the Tat-RFP kinetic profiles between the two LTRs at different Shield1 concentrations is presented in Figure 3.21C right panel. While the cells carrying the 3-κB variant virus and treated with 3 μM Shield1 manifested ~12% Tat-RFP^{+ve} phenotype on day-3, the 1-κB variant could attain only a maximum of ~4% cells expressing Tat-RFP^{+ve} at the same time-point, for the same dose of Shield1. Although under identical Shield1 concentration, the intracellular stability of Tat-RFP is normalized, the production rate of the same is governed by the LTR-strength and the assisted functional rate of the LTR-Tat feedback component. Strikingly, on day-3, for 3 μM Shield1 treatment, the GFP^{+ve} phenotype declined to ~20% in the 3-κB variant, whereas, ~65% of GFP^{+ve} cells were still noted in the 1-κB variant (Figure 3.21C left panel) despite limited intracellular Tat-RFP levels compared to the stronger promoter variant. The results indirectly hint at an accessory role for Tat in silencing the strong HIV-1 promoters, which is the theme of Chapter 4.

Since the Tat-RFP fusion protein was molecularly linked to the DD domain of FKBP, the intracellular levels of Tat could be fine-tuned by varying the concentration of Shield1 in the medium (Figure 3.19) and the molecular levels of Tat further contributed in manipulating the transcriptional feedback strength. Hence, as an additional feature, the LTR-Tat-feedback strength in the TTF model could be modulated without altering the

genetic architecture of the LTR such as varying the copy-numbers of NF-κB elements which was not feasible in either of the two earlier models- ATF and DTF. As described above in great detail, Jurkat cells infected with the viral strain containing a stronger viral promoter (3-κB LTR) manifested the higher levels of Tat-dependent transactivation within 24 h following sorting and rapidly initiated the process of viral latency almost without a time delay. The weak viral promoter (1-κB LTR) in contrast, demonstrated a delayed transcriptional activity lagging by 48 hours as that of the stronger viral promoter and failed to sustain Tat-dependent transactivation. Thus, there is a significant lag in the establishment of viral latency by the weak promoter. Importantly, such temporal differences in the kinetics of latency establishment were not manifested by either of the viral promoters at varying concentrations of Tat. For instance, regardless of the concentration of Shield1, the 3-κB LTR or the 1-κB LTR demonstrated the peak transcriptional activity at the same time – day-1 and day-3 following sorting, respectively (Figure 3.22, top, right panel). Likewise, the peak of viral latency establishment was seen on the same day-3 for the 3-κB LTR viral strain regardless of the difference in the Shield1 concentration (Figure 3.22, top, left panel). Thus, the temporal kinetics of latency establishment is a direct function of the transcriptional strength of the viral promoter. Importantly, when the transcriptional strength of the viral LTR is a constant factor, the difference in Tat concentration doesn't appear to influence the time of latency establishment provided the Tat levels are above a certain threshold.

3.12 Summary

In summary, the data of the ATF and TTF models collectively appear to suggest, for the first time, that in the presence of an intact LTR-Tat positive feedback loop, the transcriptional strength of the viral promoter plays the cardinal role in regulating viral latency. However, the Tat-feedback circuit being intricately hardwired to the transcriptional strength of the LTR, a dominant role of a stronger viral promoter appears to increase the molecular levels of Tat and favour a Tat-dependent transactivation over the Tat-independent phenomenon. The Tat-mediated transactivated cells would then undergo rapid silencing over the Tat-independent transactivated promoters as evident in the data of the TTF model (Figure 3.21B, bottom, right panel). The data are also the first

to suggest that in addition to its primary transactivation function, saturating levels of Tat could also negatively regulate HIV-1 transcription at least in the ex vivo context. The insights from the three Jurkat cell-based models have been tabulated below (Table-3.2).

Table 3.2. The salient features of the Jurkat cell-based latency models to study latency in HIV-1C

Model	LTR-Tat-feedback	Vector backbone	Fluorescent marker(s)	Parameter influencing latency	Results
Autonoumous Tat-feedback (ATF)	Preserved	LTR-GFP-IRES-Tat (LGIT)	EGFP or GFPd2	Synergistic action of LTR- and Tat-feedback strengths	LTR-silencing is directly proportional to the combined LTR-Tat-feedback strength
Disjoint Tat-feedback (DTF)	Disrupted	LTR-GFP (LG)	EGFP	LTR-strength alone	LTR-silencing is unrelated to promoter-strength alone
Tuneable Tat-feedback (TTF)	Preserved but manipulable	LTR-GFPd2-IRES-Tat:RFP:DD (LdGITRD)	GFPd2 (LTR activity) DsRed2-RFP (Tat)	Tat-feedback strength alone	LTR-silencing is inversely proportional to Tat-feedback strength alone

3.13 Bibliography

Arkin, A., Ross, J. and McAdams, H.H., 1998. Stochastic kinetic analysis of developmental pathway bifurcation in phage λ-infected Escherichia coli cells. *Genetics*, *149*(4), pp.1633-1648.

Archin, N.M., Sung, J.M., Garrido, C., Soriano-Sarabia, N. and Margolis, D.M., 2014. Eradicating HIV-1 infection: seeking to clear a persistent pathogen. *Nature Reviews Microbiology*, *12*(11), p.750.

Bachu, M., Mukthey, A.B., Murali, R.V., Cheedarla, N., Mahadevan, A., Shankar, S.K., Satish, K.S., Kundu, T.K. and Ranga, U., 2012a. Sequence insertions in the HIV type 1 subtype c viral promoter predominantly generate an additional NF-κB binding site. *AIDS research and human retroviruses*, *28*(10), pp.1362-1368.

Bachu, M., Yalla, S., Asokan, M., Verma, A., Neogi, U., Sharma, S., Murali, R.V., Mukthey, A.B., Bhatt, R., Chatterjee, S. and Rajan, R.E., 2012b. Multiple NF-κB sites in HIV-1 subtype C long terminal repeat confer superior magnitude of transcription and thereby the enhanced viral predominance. *Journal of Biological Chemistry*, *287*(53), pp.44714-44735.

Banaszynski, L.A., Chen, L.C., Maynard-Smith, L.A., Ooi, A.L. and Wandless, T.J., 2006. A rapid, reversible, and tunable method to regulate protein function in living cells using synthetic small molecules. *Cell*, *126*(5), pp.995-1004.

Boullosa, J., Bachu, M., Bila, D., Ranga, U., Süffert, T., Sasazawa, T. and Tanuri, A., 2014. Genetic diversity in HIV-1 subtype C LTR from Brazil and Mozambique generates new transcription factor-binding sites. *Viruses*, *6*(6), pp.2495-2504.

Burnett, J.C., Miller-Jensen, K., Shah, P.S., Arkin, A.P. and Schaffer, D.V., 2009. Control of stochastic gene expression by host factors at the HIV promoter. *PLoS pathogens*, *5*(1), p.e1000260.

T Das, A., Tenenbaum, L. and Berkhout, B., 2016. Tet-On systems for doxycycline-inducible gene expression. *Current gene therapy*, *16*(3), pp.156-167.

Hanna, L.E., Neogi, U., Ranga, U., Swaminathan, S. and Prasad, V.R., 2014. Phylogenetic Characterization of Six Full-Length HIV-1 Subtype C Molecular Clones from Three Patients: Identification of Rare Subtype C Strains Containing Two NF-κB Motifs in the Long Terminal Repeat. *AIDS research and human retroviruses*, *30*(6), pp.586-591.

Isaacs, F.J., Hasty, J., Cantor, C.R. and Collins, J.J., 2003. Prediction and measurement of an autoregulatory genetic module. *Proceedings of the National Academy of Sciences*, *100*(13), pp.7714-7719.

Jordan, A., Bisgrove, D. and Verdin, E., 2003. HIV reproducibly establishes a latent infection after acute infection of T cells in vitro. *The EMBO journal*, *22*(8), pp.1868-1877.

Karn, J., 2011. The molecular biology of HIV latency: breaking and restoring the Tat-dependent transcriptional circuit. *Current opinion in HIV and AIDS*, *6*(1), p.4.

Li, X., Zhao, X., Fang, Y., Jiang, X., Duong, T., Fan, C., Huang, C.C. and Kain, S.R., 1998. Generation of destabilized green fluorescent protein as a transcription reporter. *Journal of Biological Chemistry*, *273*(52), pp.34970-34975.

Mahmoudi, T., Parra, M., Vries, R.G., Kauder, S.E., Verrijzer, C.P., Ott, M. and Verdin, E., 2006. The SWI/SNF chromatin-remodeling complex is a cofactor for Tat transactivation of the HIV promoter. *Journal of Biological Chemistry*, *281*(29), pp.19960-19968.

Mizuguchi, H., Xu, Z., Ishii-Watabe, A., Uchida, E. and Hayakawa, T., 2000. IRES-dependent second gene expression is significantly lower than cap-dependent first gene expression in a bicistronic vector. *Molecular Therapy*, *1*(4), pp.376-382.

Piatkevich, K.D. and Verkhusha, V.V., 2011. Guide to red fluorescent proteins and biosensors for flow cytometry. In *Methods in cell biology* (Vol. 102, pp. 431-461). Academic Press.

Raser, J.M. and O'shea, E.K., 2005. Noise in gene expression: origins, consequences, and control. *Science*, *309*(5743), pp.2010-2013.

Razooky, B.S. and Weinberger, L.S., 2011. Mapping the architecture of the HIV-1 Tat circuit: A decision-making circuit that lacks bistability and exploits stochastic noise. *Methods*, *53*(1), pp.68-77.

Spina, C.A., Anderson, J., Archin, N.M., Bosque, A., Chan, J., Famiglietti, M., Greene, W.C., Kashuba, A., Lewin, S.R., Margolis, D.M. and Mau, M., 2013. An in-depth comparison of latent HIV-1 reactivation in multiple cell model systems and resting CD4+ T cells from aviremic patients. *PLoS pathogens*, *9*(12), p.e1003834.

To, T.L. and Maheshri, N., 2010. Noise can induce bimodality in positive transcriptional feedback loops without bistability. *Science*, *327*(5969), pp.1142-1145.

Weinberger, L.S., Burnett, J.C., Toettcher, J.E., Arkin, A.P. and Schaffer, D.V., 2005. Stochastic

gene expression in a lentiviral positive-feedback loop: HIV-1 Tat fluctuations drive phenotypic diversity. *Cell*, *122*(2), pp.169-182.

Williams, S.A., Chen, L.F., Kwon, H., Ruiz-Jarabo, C.M., Verdin, E. and Greene, W.C., 2006. NF-κB p50 promotes HIV latency through HDAC recruitment and repression of transcriptional initiation. *The EMBO journal*, *25*(1), pp.139-149.

Chapter 4

A putative role for Tat in the rapid transition from active to latent phenotype in HIV-1C

4.1 Introduction

In HIV-1, the influence of the Tat auto-regulatory circuit in deciding viral transcription and latency has emerged as a prominent area of investigation. Previous research on the HIV-1B transcriptional circuit has revolutionised our understanding of HIV-1 latency (Weinberger LS et al., 2005; Weinberger LS and Shenk T., 2006; Weinberger LS et al., 2008., Razooky BS et al., 2011; Razooky BS et al., 2015). The HIV-1 transactivator Tat induces gene expression from HIV-1 LTR by enhancing transcription elongation and thus, increases the production of its own mRNA by ~50-100 folds (Feinberg MB et al., 1991). Flow cytometry analysis of isoclonal Jurkat cells harbouring pseudotyped GFP viruses with an intact Tat positive circuit, revealed a molecular phenomenon called 'phenotypic bifurcation' especially when the intracellular Tat concentration is limiting. At a low Tat concentration (characterized by the low-GFP cells or the dim population), a fraction of cells exhibits high GFP expression (the ON state) while the remainder of the cells demonstrate no viral expression at all (the OFF state) (Weinberger LS et al., 2005). Interestingly, the high-GFP or the bright population neither undergoes transcriptional silencing within the duration of the assay nor a bimodal GFP expression pattern is displayed. Further, the Hill-coefficient analysis (the dose-response curve in a 'LTR-activity vs time' plot of the single-cell time-lapse microscopy) demonstrates that Tat essentially functions as a monomer and a dimer formation is not manifested (Razooky BS et al., 2011). A Hill value of ≥ 1 implies self-cooperativity while a value of ≤ 1 signifies the lack of cooperativity as observed in HIV-1B. Based on these data, Tat was proposed to function as a monomer in the HIV-1 B circuit. How the silencing of HIV-1 promoter is inevitable in the face of a positive feedback loop, induced by a monostable transactivator (Tat)? – This question still remains elusive.

Unlike other genetic subtypes of HIV-1, subtype C demonstrates extensive TFBS polymorphism consisting of NF-κB, NFAT, AP-1, RBEIII sites, both genetic as well as number variation, that modulates the transcriptional strength of the viral promoter significantly. Given the positive feedback nature of the LTR-Tat transcription circuit, a change in the transcriptional strength of the viral promoter should affect the functioning of the LTR-Tat circuit and consequently the qualities of HIV-1 latency. The data presented in the previous chapter of the thesis clearly demonstrated that NF-κB number

variation in the C-LTR not only modulated transcriptional strength of the promoter but also the associated Tat-feedback strength (Figure 3.3). To evaluate the individual contributions made by the transcriptional strength of the LTR and the transactivation strength of the Tat-feedback circuit towards viral transcriptional silencing, it was necessary to engineer the viral genetic circuits to permit the analysis of viral latency as a function of promoter strength alone or the Tat-feedback strength alone or a combination of the two. To this end, we devised three distinct models in Jurkat cells and the data of these assays have been presented in Chapter 3. The data from the three cell models collectively suggested that the strength of the LTR and that of the Tat-feedback circuit collectively, but not individually, can effectuate transcriptional silencing. Additionally, we found many important differences in the nature of viral latency between the published reports that used HIV-1B and the data we generated using HIV-1C, both using Jurkat cells. Interestingly, in contrast to the HIV-1B model where the GFP^{High} (bright) cells failed to establish transcriptional silence within the period of the assay (Figure 3D, Weinberger LS et al., 2005), the GFP^{High} cells of HIV-1C LTR downregulated the expression of the reporter gene. Importantly, the rate of transcriptional silence was directly proportional to the transcriptional strength of the C-LTR i.e., a stronger promoter established a faster transcriptional silence. Additionally, the transcriptional silence followed a biphasic trajectory instead of a gradual shut down process (Figures 3.8, 3.10). The most interesting observation emerged from the TTF model (Figure 3.21), which was the presence of Tat in the cells in the promoter-OFF condition (See Figure 3.21B). Of note, while a stronger viral promoter (3-κB LTR) downregulated GFP expression at a significantly faster rate in a Tat-dependent fashion, a weaker viral promoter (1-κB LTR) in contrast was unable to express sustained Tat expression and established latency in a Tat-independent pathway.

4.2 The study rationale

In the present chapter, we plan to examine the kinetics of reporter gene expression from variant viral promoters in the context of the ATF model, at a resolution of single cell. The nature of the biological events, such as the gene expression analyses, is masked or averaged when pools of cells are used in the analyses. We use clonal cell lines to

circumvent these technical limitations. We also attempt to delineate the nature of binding of various transcription factors to the viral promoter between the ON and OFF states.

4.3 The single transactivated J-LGIT cells of the ATF system established latency in a bimodal fashion (ON or OFF mode) with no significant intermediate phenotype

To examine the kinetics of latency-establishment and to study the cell-to-cell variability among clonal cell variants in the context of a single LTR as well as across all the five LTR variants, we tracked GFP expression in single J-LGIT cells sorted from the GFP^{High} zone. The experimental set-up was comparable to the one described in section 3.4, Chapter 3 with a slight modification. Following the infection of Jurkat cells with the NF-κB-variant LGIT strains and stimulating the infected cells with the global activation cocktail, single GFP^{High} cells were sorted into individual wells of a 96-well culture plate (Figure 4.1 A). The sorted cells were allowed to grow for 3-4 weeks following which GFP-expression profiles were assessed using the BD FACSAriaIII flow cytometer.

We recovered 16-25 clones from each NF-κB variant, 3 weeks following sorting and the first 16 clones from each variant were considered for the GFP expression analysis. Based on the GFP expression pattern, the clones can be categorised into three distinct types- (1) **the persistent type**: All the individual cells of a clone sustain the GFP expression high and comparable to that of the parental cell through the observation period, up to day 28 or even longer indicating actively transcribing provirus in all the daughter cells, (2) **the latent type**: All the daughter cells of the clonal cell lines switch-off GFP expression completely, and (3) **the bi-modal type**: the distinctive feature of the clone type was the simultaneous demonstration of both the phenotypes. Although all the cells in the cluster were derived from the same GFP^{High} parental cell, one subset of the cells maintains high GFP expression whereas the other subset down regulates the reporter gene completely without the manifestation of an intermediate phenotype. Representative fluorescent images of each clonal phenotype are depicted (Figure 4.1 B).

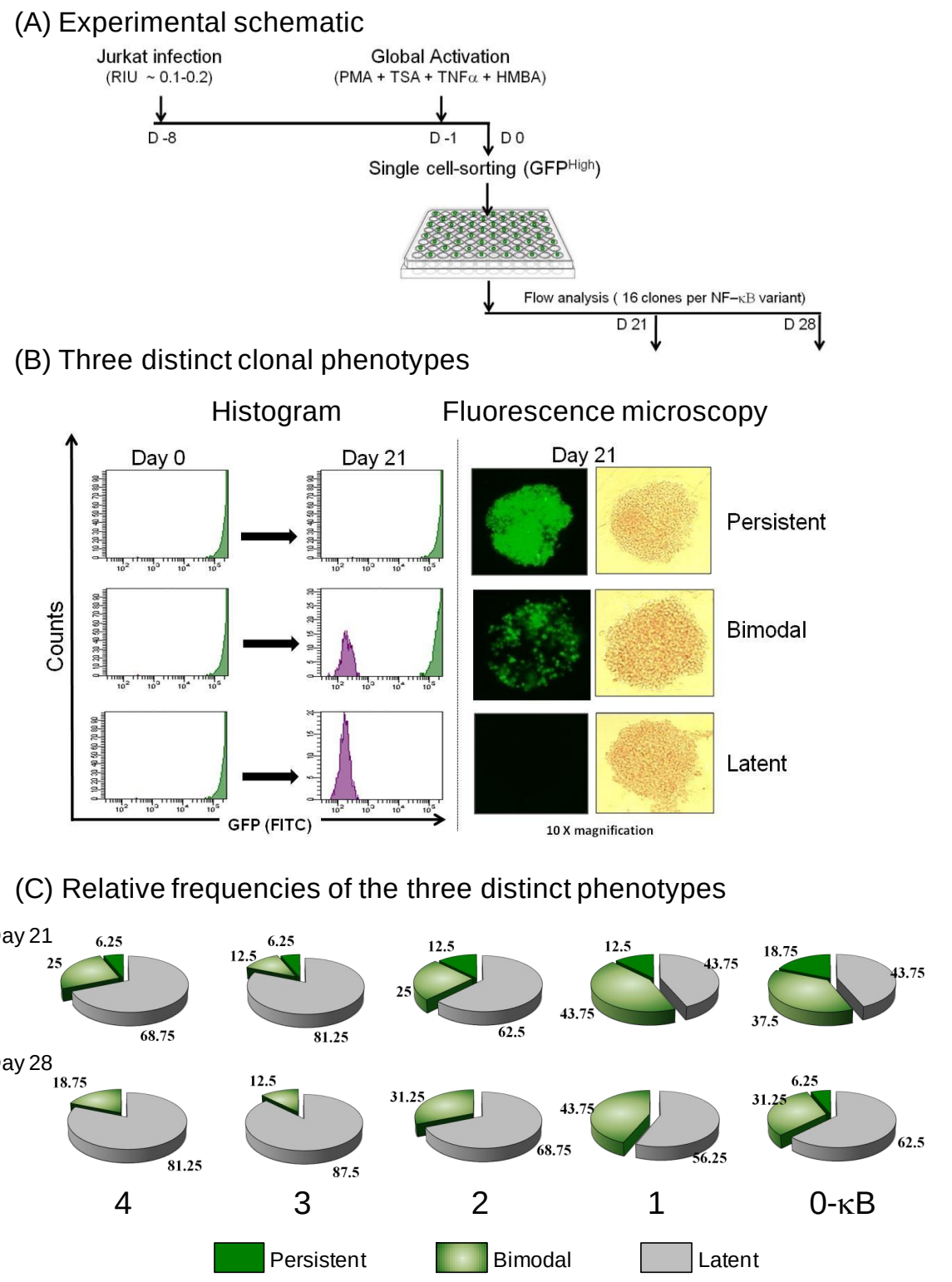


Figure 4.1: The kinetics of latency-establishment in single cells of the ATF model. (A) The schematic representation of the experimental procedure. Approximately one million Jurkat cells were infected at a low infectious titer (~0.1-0.2 RIU) with each of the five LGIT, LTR-variant viral strains. Seven days after infection, the cells were activated with global T-cell activators for 24 h and sorted for single GFP^{High} cells (GFP-MFI $>10^4$ RFU). Single GFP^{High} cells were sorted into individual wells of 96-well culture dishes (one dish for each viral strain). Importantly, alternate wells were assigned for the collection of single cells to prevent cross contamination. The sorted cells were allowed to expand in culture for three-four weeks under standard growth conditions. The first 16 clones corresponding to each variant strain were analyzed on days 21 and 28 post-sorting for GFP expression by flow cytometry and images were captured using a fluorescence microscope. **(B)** Three distinct categories of clones were identified based on the GFP expression profile. The sorted, single GFP^{High} cells from the five viral strains expanded into one of the following classes

(1) the persistent clones: all the daughter cells continue to express high GFP (GFP-MFI $>10^4$ RFU) even 21 days after the sorting; (2) the latent clones: all the daughter cells switched off GFP expression (GFP^{-ve}; MFI $<10^3$), and (3) the bimodal clones: some daughter cells expressed high levels of GFP while the others do not down regulate GFP completely without an intermediate phenotype, thus pointing to the stochastic nature of silencing of the LTR. **(C)** The proportion of the three phenotypes among the LTR-variants across time. Two sets of pies corresponding to days 21 and 28 are shown. Each pie represents a viral variant at a particular time point. The solid green, shaded green and grey fractions represent the relative percentages of the persistent, bimodal and latent clones, respectively, of the total of 16 clonal cells representing each of the 5 viral strains of the panel.

4.4 The frequencies of the three clonal phenotypes differ significantly among the NF-κB-variant strains

The profiles of the three phenotypes described above have been determined for each of the five LTR-variant viral strains using flow cytometry at two different time points- day 21 and day 28. Each LTR-variant was represented by 16 different clonal cell lines. We could not analyse the cells at an earlier time point as the clones needed expansion to provide enough cells for flow cytometry. For each clonal population, the numbers of cells representing the three phenotypes have been determined and presented as percent values in a pie chart format (Figure 4.1C). The kinetics of latency-establishment of the 16 clonal lines representing a specific NF-κB-variant LTR did not exhibit an identical GFP phenotype. Regardless of the transcriptional strength of the LTR, all the five viral promoters demonstrated the three different GFP expression phenotypes: the persistent, bimodal, and latent phenotypes. The profile of the three phenotypes differed among the LTR variants and appeared to be associated with the transcriptional strength of the LTR. All the five variant LTRs downregulated GFP expression progressively between days 21 and 28. Based on the rate of silencing of the reporter gene the variant LTRs could be classified into two different categories– the stronger (4- and 3-κB LTRs) and the weaker (1- and 0-κB LTRs) viral promoters whereas the 2-κB LTR appears to fall between these two classes. The stronger LTRs contained fewer persistent and bimodal GFP clones at day 21 as compared to the other three LTRs. By day 28, the stronger LTRs did not contain any persistent phenotype and smaller bimodal clones. The stronger viral promoters downregulated GFP expression, both persistent and bimodal phenotypes, at a significantly faster rate to establish latency. Thus, the rate of transcriptional silence was a function of the transcriptional strength of the LTR. Importantly, the data of the clonal cell line assays are perfectly consistent with the results of the population-level studies (see Figures 3.5, 3.8 and 3.10) faithfully recapitulating the influence of the transcriptional strength of the LTR on the kinetics of viral latency. Both the experimental strategies

demonstrated a direct correlation between the transcriptional strength of the LTR and the rapid rate of latency-establishment. Further, both the experimental strategies are consistent with each other in demonstrating bimodal nature, but not a gradual, shut off of viral gene expression.

Of note, the complete absence of a phenotype representing an intermediate GFP expression in the bimodal clones is suggestive of an inherent stochastic trigger for the onset of promoter silencing. It was particularly intriguing to note that proviral GFP^{High} Jurkat clonal cells harbouring B-LTR did not downregulate GFP expression even after considerable number of days post-sorting (Weinberger LS et al., 2005) and all the clonal lines demonstrated the sustained phenotype. This difference between B- and C-LTRs could possibly arise from subtype-specific phenomena unique to the LTR-Tat circuits. Further, our data particularly in the context of the clonal cells are suggestive of a predominantly cumulative role of multiple NF-κB sites mediating an actively repressive role. When the GFP expression levels are normalised, a context signifying comparable levels of intracellular Tat among the variant cell lines, the stronger viral promoters (3- or 4-κB LTRs) favoured efficient silencing as opposed to the weaker viral promoters (0-, 1- or 2-κB LTRs) possibly by the recruitment of repressive factors by the NF-κB sites. To this end, we examined the nature of a few important transcription factors and chromatin modulators recruited to the active and latent viral promoters using the bimodal clonal cells in the chromatin immunoprecipitation assay.

4.5 Molecular complexes differentially binding to the active and silent LTRs govern the alternate phenotypic outcomes in the GFP^{-ve} and GFP^{High} fractions of the bimodal clones.

The HIV-1 LTR encompasses a plethora of overlapping transcription factor binding elements (Pereira LA et al., 2000). Recruitment of appropriate active or repressive complexes to the LTR depends on multiple factors including the activation status of the host cell, the nature of diverse cellular signals, site of integration, and many others. An active HIV-1 chromatin is characterised by the acetylated histone marks at lysine 9 and 14 (H3K9Ac and H3K14Ac) (Klichko V et al., 2006) while deacetylation of histones by HDACs (such as HDAC1 and HDAC3) and methylation of histones such as H3K9Me3,

H3K27Me3 or H3K9Me2 are characteristic repressive marks for a silent HIV-1 provirus (Williams SA et al., 2006; Mbonye U and Karn J., 2011). Furthermore, alternate combinations of Rel family members have been shown to modulate the HIV-1 transcription differentially. While the p50-p65 heterodimer promotes active transcription from the HIV-1 LTR, the p50 homodimer predominantly functions as a transcription suppressor (Stroud JC et al., 2009; Burnett JC et al., 2009; Williams SA et al., 2006). Given the larger number and genetic variation of the additional NF-κB motifs, the situation is further complicated in the case of HIV-1C LTR, augmenting the probability of recruiting a much wider range of transcription factors. Additionally, the NF-κB motifs function as overlapping sites with other transcription factor families, thus, creating a context of competition for the same sites. For instance, members of the NF-κB and NFAT families sharing overlapping sites might compete for the binding under varying circumstances in the HIV-1 LTR (Pessler F et al., 2004). The crystal structures of NFAT1 in a form bound to the κB-sites in the HIV-1 LTR has been solved (Bates DL et al., 2008; Giffin MJ et al., 2003). Although NFAT2 has been identified as a potential activator of HIV-1 transcription, reports for the influence of NFAT1 are controversial. Seminal work by Macian F et al. demonstrated the negative regulation of NFAT1 on HIV-1 transcription via binding to the NF-κB sites and by physically interacting with the HIV-1 Tat protein (Macian F et al., 1999). The group further showed that in contrast to NFAT2, NFAT1 outcompeted p50 binding to the site, inhibited p65 mediated transcription, and exerted a suppressive role on HIV-1 LTR and the promoters of other cytokine genes.

A previous report from our laboratory demonstrated the recruitment of NFAT1, 2 and 5 proteins to the C-κB site, a variant NF-κB motif unique for HIV-1C, with an affinity comparable to that of the H-κB site, the canonical NF-κB motif, in the EMSA assay (Verma A et al., 2016). The kinetics of NFAT occupancy of HIV-1C LTR, in the context of active and latent phenotypes, however, remains unexplored. One of the objectives of the present work, therefore, was to compare the nature of the transcription factors and other host factors binding to the viral promoter between the active and suppressed states under identical experimental conditions. The clonal cell lines that display the bimodal GFP phenotype offer an excellent experimental model as these cells demonstrate two contrasting phenotypes (GFP^{High} and GFP^{-ve} expression) despite an identical viral genotype, chromatin background and, host-cell activation (Figure 4.1B). We selected two

clonal cell lines, 3c and 8c, representative of the stronger viral promoters, 4-κB and 3-κB LTRs, respectively, for the analysis.

Prior to assessing the transcription factor complexes recruited to the active and silent promoters, we confirmed that the levels of proviral integration between the two subpopulations (GFP^{High} and GFP^{-ve}) of each bi-modal clone were comparable. The clonal lines were expanded, globally activated, and allowed to relax for a week until the bimodal GFP expression pattern was displayed (Figure 4.2A and 4.2B, left panels). GFP^{High} and GFP^{-ve} sub-fractions were sorted separately from the expanded clonal cells, genomic DNA was isolated from one million cells each, and the frequency of the proviral integration per cell was determined using a real-time PCR as described earlier (Section 3.4 and Figure 3.6, Chapter 3). The analysis confirmed a comparable frequency of proviral integration between the two cell sub-fractions (GFP^{High} and GFP^{-ve}) which was close to one (Figure 4.2A and 4.2B, top, right panels), in both the clonal cell lines (4-κB and 3-κB LTRs) thus, ruling out the possibility of integration frequency differences underlying the bimodal phenotype. Importantly, despite comparable frequencies of integration, the expression of the Tat trasnscript quantitated using a Tat RT-PCR as described previously (section 3.3, and Figure 3.3, Chapter 3) was significantly different between the cell fractions in both the clonal cell lines (Figure 4.2A and 4.2B, bottom, right panels). As expected, the Tat-transcript levels significantly reduced in the GFP^{-ve} fraction compared to the GFP^{+ve} fraction; approximately 112 folds for 4-κB (Figure 4.2A right panel) and 80 folds 3-κB clonal cell lines, (Figure 4.2B right panel) respectively. Of note, the persistence of a small but detectable quantity of Tat-mRNA species in the GFP^{-ve} fractions of both the clones was nevertheless intriguing. Compared to the uninfected Jurkat cells (the no-infection control used for the estimation of the relative Tat-transcripts using the ΔΔCt method), the GFP^{-ve} cells of 4-κB and 3-κB demonstrated approximately 21.63 ± 1.6 and 12.41 ± 1.3 folds higher Tat mRNA expression, respectively. The presence of the Tat-transcripts in the GFP^{-ve} cells could be ascribed to a few residual copies of Tat mRNA from the proviral promoter that was on the verge of silencing. Alternatively, the transcripts may have originated from leaky transcription from a promoter not fully repressed.

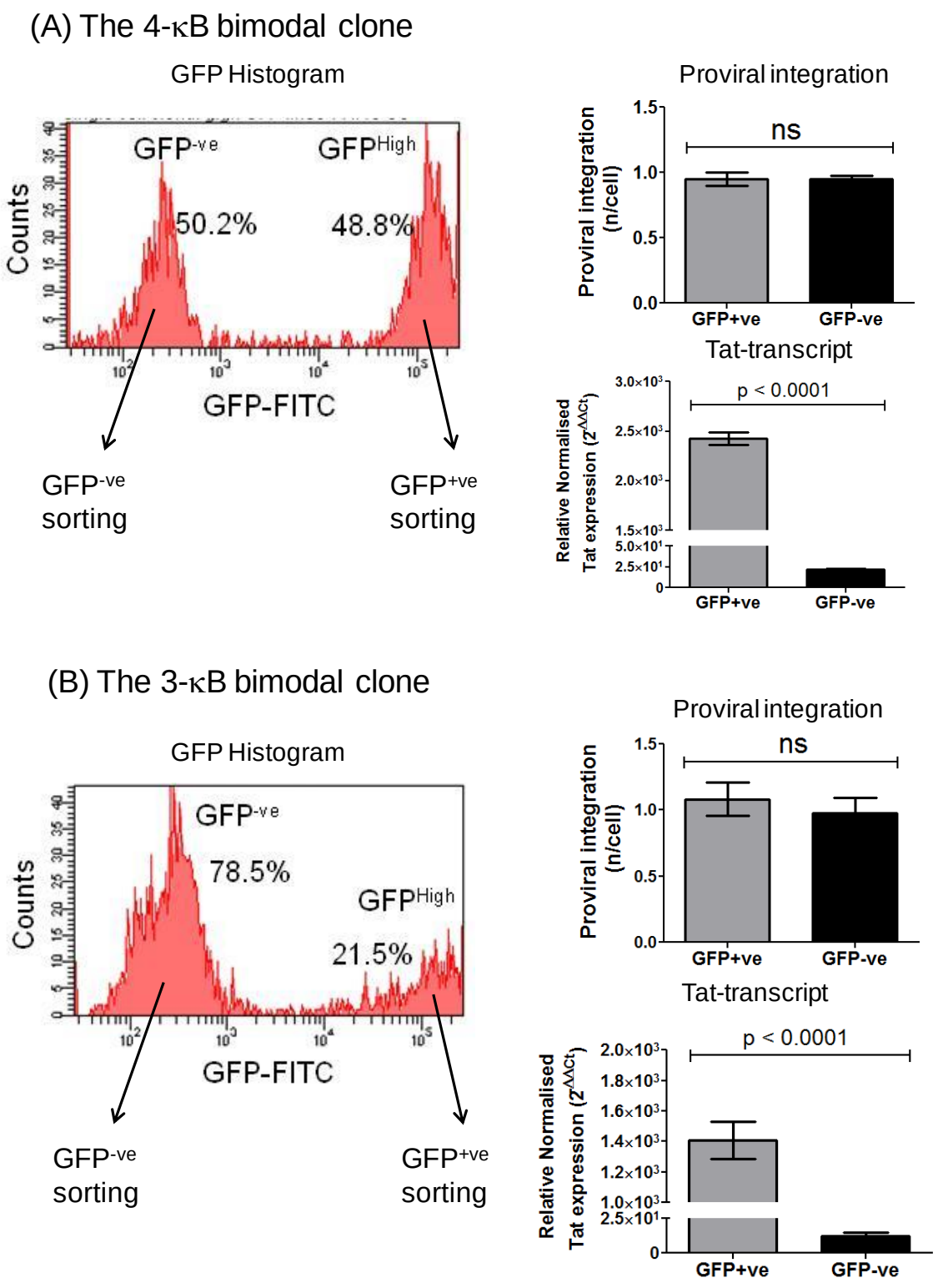


Figure 4.2: Integration frequency and Tat-expression quantification in the bimodal clones of the strong LTRs. Two cell clones representing **(A)** 4- and (**B**) 3-κB LTRs were used in the analysis. The percentages of cells representing the GFP^{-ve} and GFP^{High} fractions prior to sorting are depicted and the two fractions were sorted. A Taqman qPCR amplifying a sequence in the proviral LTR showed the integration frequency to be comparable (~1.0 provirus per cell) between the GFP^{-ve} and GFP^{High} fractions of both the LTRs (Top, right panels). J-Lat cells containing a single copy of provirus were used to generate the standard curve (as in Figure 3.6.). A quantitative real-time PCR for the Tat-transcripts demonstrated significantly higher levels of Tat-transcripts in the GFP^{High} fractions. Mean values from three independent experiments ± SEM are plotted. A two-tailed unpaired t-test was used for the statistical evaluation.

Importantly, the presence of the Tat-transcript in the latent cell is consistent with our previous data of Tat protein analysis. In the Jurkat TTF model, the cells infected with the 3-κB LTR viral strain and harbouring the latent viral promoter (negative for GFP expression) prominently expressed the Tat-RFP fusion protein suggesting the presence of the Tat protein associated with the latent viral promoter (Figure 3.21). In summary, we for the first time could demonstrate the presence of Tat protein and Tat mRNA in cells that contain the viral LTR in the 'OFF' state.

Having demonstrated the presence of Tat in the cells containing the latent provirus, we asked if Tat in these cells is recruited to the latent viral promoter. Taking advantage of the clonal cell lines that displayed the phenotypic bifurcation and using the chromatin immunoprecipitation assay, we examined the presence of several important host factors (Rel members- p50, p65; NFAT1 and NFAT2; RNA polymerase Ser2 phosphorylation, and Histone 3 lysine 9 trimethylation- a mark for repressive chromatin), and Tat, in the chromatin preparations of the active and latent cells of the bimodal clonal cell lines. The experimental protocol for ChIP was essentially as described in the materials and methods section (Chapter 2, section 2.4.2). The GFP^{High} and GFP^{-ve} cells of the bimodal clonal cell lines- 3c (4-κB LTR) and 8c (3-κB LTR) were sorted using a flow sorter. The chromatin binding complexes were precipitated from the cell-lysates equivalent of two million cells and 2 μg of each antibody specific to the target protein was used for every immunoprecipitation reaction. A semi-quantitative PCR that amplified a 240 bp fragment spanning the enhancer-core promoter region in the C-LTR was performed (Figure 4.3A). The image-densitometry data was then calculated from the quadruplicate values for each transcription factor and normalised to the 10% input signals. Additionally, we also performed a Taqman probe-based qPCR, amplifying a 127 bp region spanning the C-κB and Sp1 sites for a precise quantitation of the proviral DNA copy numbers in the chromatin samples.

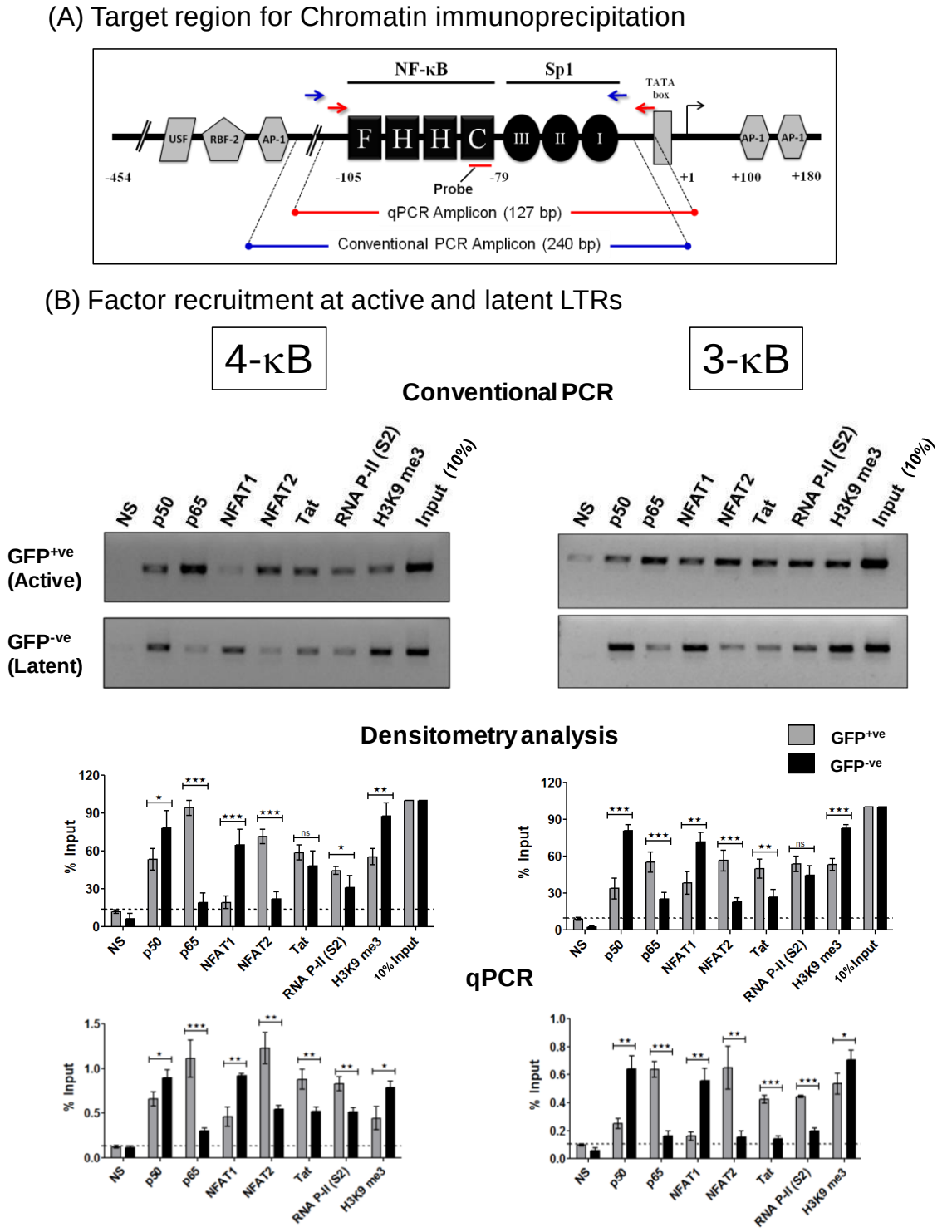


Figure 4.3: Cellular factor and Tat recruitment to the active and latent LTRs in the Chromatin immunoprecipitation (ChIP) analysis. (A) Schematic representation of the LTR sequence and the location and direction of the primers used for the amplification by conventional PCR (240 bp, blue) and qPCR (127 bp, red). **(B)** ChIP of the active (GFP^{+ve}) and latent (GFP^{-ve}) promoters of the two bimodal clones representing the 4-κB and 3-κB proviruses as depicted. Cell-lysate from two million cells (GFP^{+ve} or GFP^{-ve}) and 2 µg of respective antibody were used for individual IP reactions. One-tenth of the IP chromatin was used as the input control. Conventional PCR (top panel) and the corresponding densitometry analyses of the gel images (middle panel) are presented. Data for each band are normalized to the input. Taqman qPCR (bottom panel) for the enhancer region was performed and data for each IP was calculated using the percent-input method. Mean values from experimental quadruplicates ± SD are plotted. A two-tailed unpaired t-test was used for statistical analysis (*$p<0.05$, ***$p<0.001$ and ns- non-significant).

The profiles of the host factors bound to the active versus latent viral promoters were consistent with the transcription promoting factors being associated with the active LTR and the transcription suppressing factors binding to the silent LTR (Figure 4.3B). Higher levels of RNA Pol II S2 and the repressive mark H3K9Me3 were found associated with the active and silent promoters, respectively, as expected. Furthermore, the Rel protein p65 and NFAT2 were preferentially associated with the active LTR as opposed to p50 and NFAT1 that were found associated with the latent promoter at a significantly higher level. That the p50-p65 heterodimer is transcription-promoting, the presence of a significantly higher concentration of p65 at the active promoter is expected (Barbeau B et al., 1997; Chen-Park FE et al., 2002; Stroud JC et al., 2009). On the other hand, the preferential association of p50 with the latent promoter is suggestive of the formation of the p50 homodimer, a known transcription suppressor (Williams SA et al., 2006). Similarly, our data are also in agreement with the previous reports regarding the transcription suppressive and supportive functions of NFAT1 and NFAT2, respectively (Kinoshita S et al., 1997; Kinoshita S et al., 1998; Macián F et al., 1999).

The most crucial finding of the present study is the detection of the association of the Tat protein with the latent LTR. The results were reproducible and consistent between the two strong viral promoters (Figures 4.3B; left and right panels). The data were also consistent between the conventional PCR and the quantitative real-time PCR performed following immunoprecipitation. The Tat protein was found associated with the active 4-κB and 3-κB promoters at 1.7- and 3-folds higher, respectively, as compared to their latent counterparts. To the best of our knowledge, the present study is the first one to demonstrate the association of Tat with the latent LTR, albeit at a lower intensity as compared to the active promoter.

We wished to validate the ChIP results further for Tat, particularly at the latent promoter. Hence, we performed additional ChIP experiments using two other anti-Tat antibodies to enhance the credibility of the signal at the latent viral promoters (Figure 4.4). Similar to the previous experiment, cell-lysates equivalent of two million cells were used for individual IP reactions for the two sub-populations (GFP^{-ve} and GFP^{High}) from each of the bimodal clones (3- and 4-κB). In addition to the rabbit-polyclonal anti-Tat antibody previously used, two mouse-monoclonal anti-Tat antibodies, raised against diverse amino

acid sequences were used to immuno pull-down the Tat-bound chromatin complexes from the cell lysates. The ChIPed DNA was quantified using the Taqman probe-based qPCR amplifying the 127 bp LTR sequence as mentioned earlier. All the three antibodies furnished positive ChIP signals for Tat at both the latent viral promoters (3- and 4-κB), over and above the respective IgG-isotyped controls. A general trend of higher association of Tat with the active over the latent promoters was also noted for each of the anti-Tat antibodies, thereby authenticating the previous ChIP results. Of note, the fold differences in the qPCR signals for the active and the latent promoters were not uniform across the three antibodies, possibly due to their differences in the clonality, affinity for the epitopes, the extent of epitope masking or a combination of all the factors. The pattern however, was reproducible in the 3- and 4-κB LTRs for each of the individual antibodies.

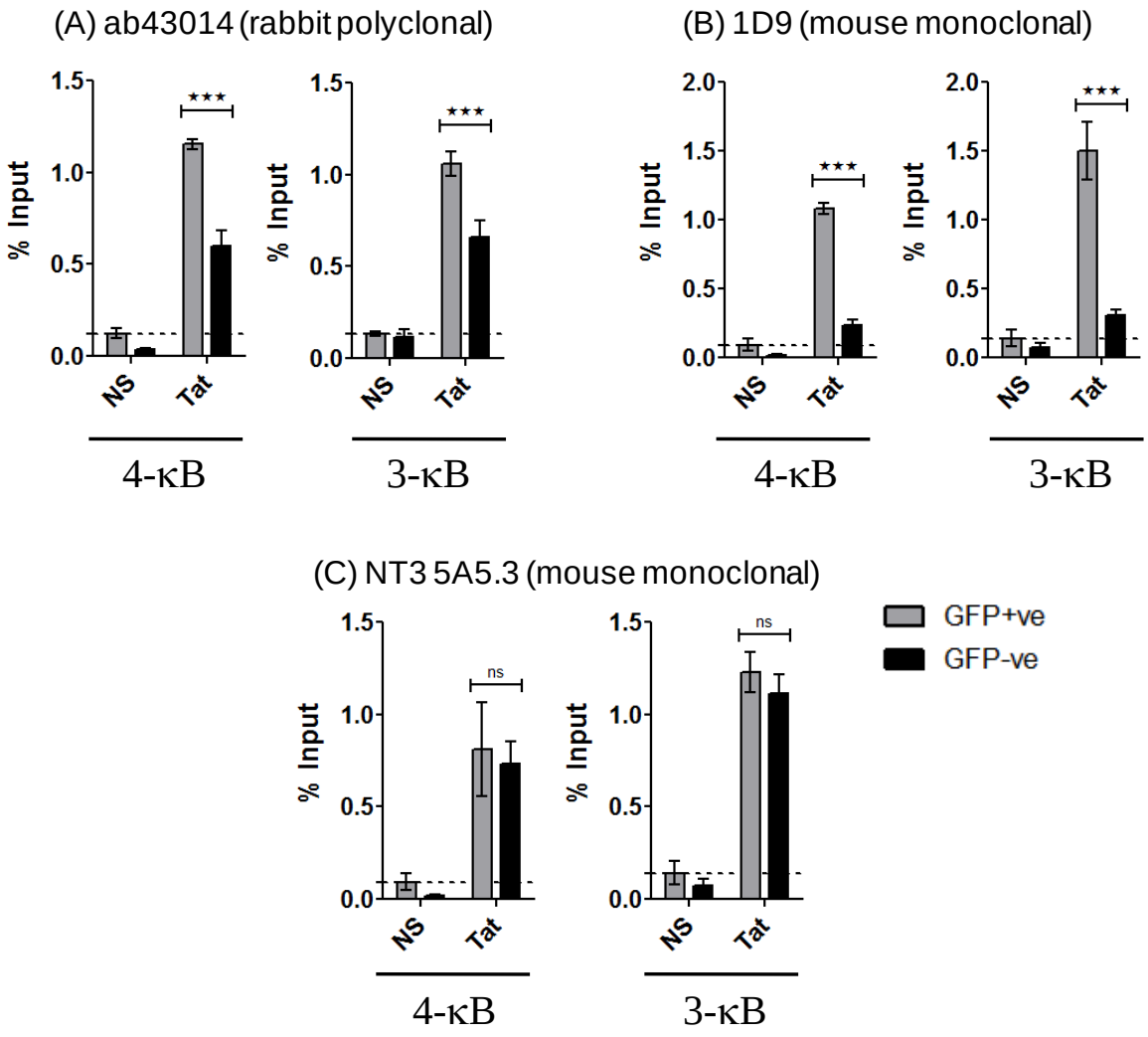


Figure 4.4: ChIP analysis to validate the recruitment of Tat at the latent promoter using diverse anti-Tat antibodies. ChIP of the active (GFP^{+ve}) and latent (GFP^{-ve}) promoters of the two bimodal clones representing the 4-κB and 3-κB proviruses was performed as described previously (Figure 4.3). Three independent anti-Tat antibodies were used for the ChIP assay- ab43014 (rabbit polyclonal), 1D9 (mouse monoclonal; binds to a Tat peptide spanning amino

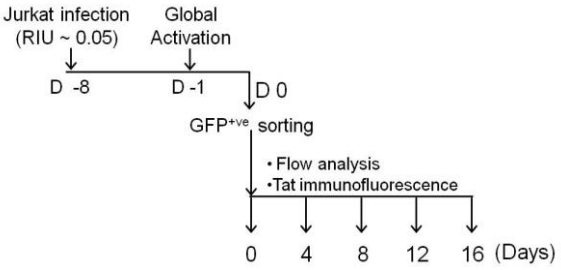
acids 1-20) and NT3 5A5.3 (mouse monoclonal; raised against synthetic Tat peptide spanning amino acid 73-86). One-tenth of the IP chromatin was used as the input control. Taqman qPCR for the enhancer region was performed and data for each IP was calculated using the percent-input method. Mean values from experimental triplicates ± SD are plotted. A two-tailed unpaired t-test was used for statistical analysis (***p<0.001 and ns- non-significant).

4.6 A sustained presence of Tat in the nucleus long after the switch-off of the LTR

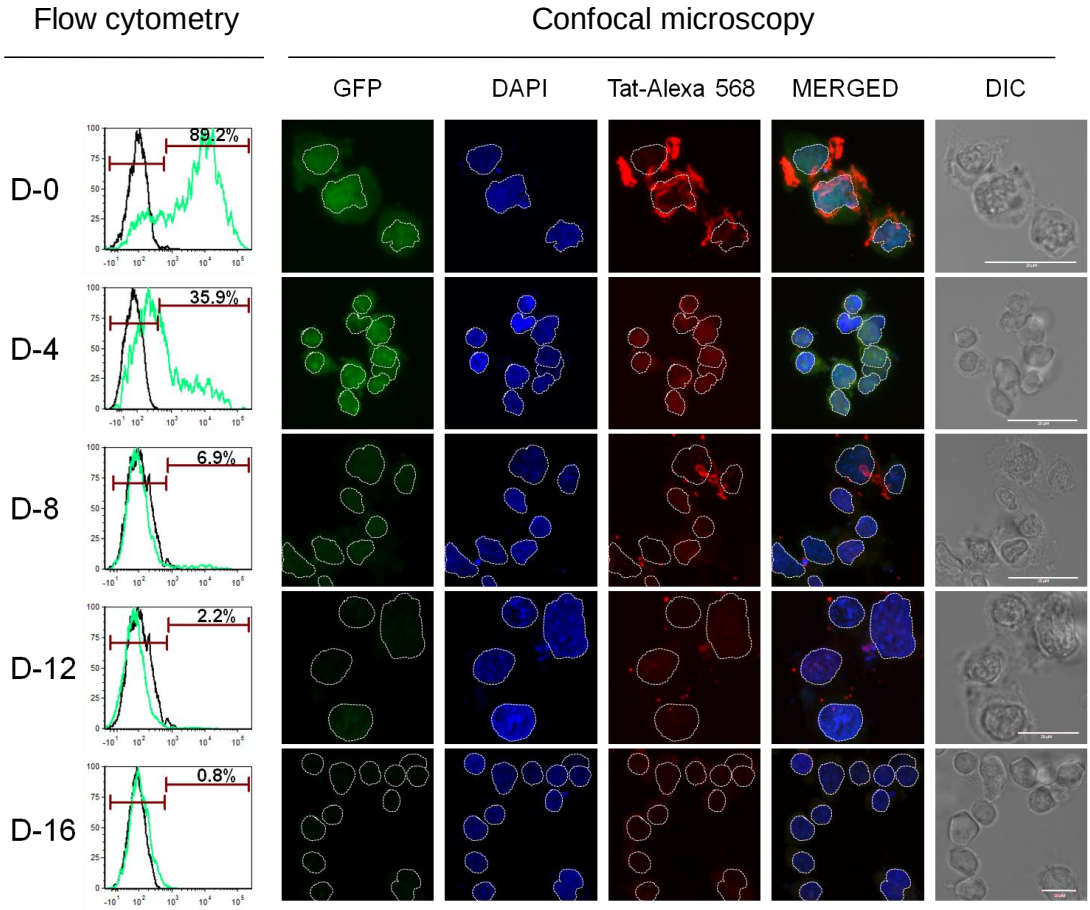
The data presented in Chapter 3 using the ATF model of latency demonstrated a direct correlation between the transcriptional strength of the viral promoter and the rapidity of viral latency (Figures 3.5, 3.8, 3.10). These data also alluded to the possibility that the intracellular Tat concentration might be directly involved in the repression of the viral promoter (Figure 3.6). Furthermore, we could detect the presence of Tat at the transcriptionally silent promoter using the ChIP assay (Figures 4.3 and 4.4). Therefore, it was necessary to understand the relative distribution and the physiological levels of Tat in the cell during the process of latency establishment. To this end, we tracked Tat expression in Jurkat cells using indirect immunofluorescence during the transition from the 'LTR-ON' to 'LTR-OFF' state.

Jurkat cells infected with the 3-κB (J-LdGIT-3-κB) viral strain were used in the assay. The expression of GFPd2 was monitored temporally at different time points using flow cytometry to measure the transcriptional activity of the LTR (Figure 4.5B, left panel). The cells were sorted for GFP^{High} expression on day-0 and subjected to indirect immunofluorescence of Tat using a rabbit polyclonal anti-Tat antibody at different stages of transcriptional silencing (D0, D4, D8, D12, D14 and D16) (Figure 4.5A). Tat was first stained with a rabbit anti-Tat primary antibody and this was followed by the staining with an anti-rabbit secondary antibody conjugated to Alexa 568. Both the antibodies were titrated prior to the experiment. DAPI was used to stain the nuclei. The experimental procedure is described in details (Section 2.5.2, Chapter 2).

(A) Experimental scheme



(B) Relative distribution of Tat during latency-establishment



(C) Estimation of Tat intensities in the nuclear and extra-nuclar compartments

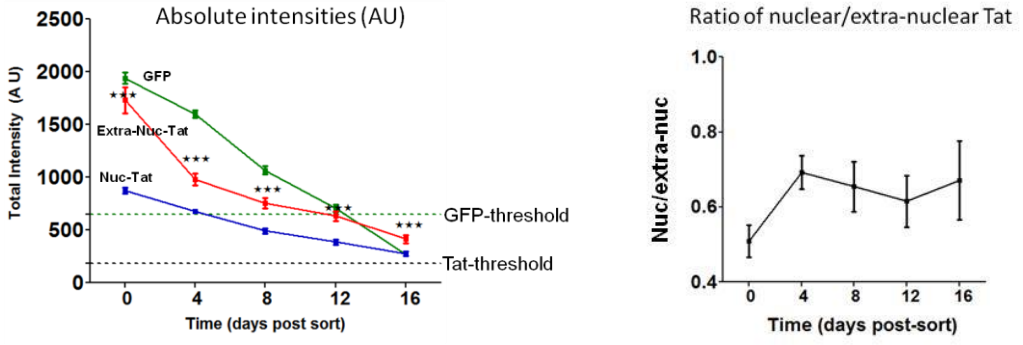


Figure 4.5: Indirect immunofluorescence identifies persistent levels of nuclear Tat long after the silencing of the LTR. (A) A schematic of the experimental layout. One million Jurkat cells were infected with the 3-κB viral strain at an infectious titer of ~0.05 RIU. The cells were induced with the global T-cell activators after a week post-infection and

sorted for GFP^{High} cells after 24 h post-activation. The GFP^{High} sorted cells were subjected to either flow quantitation of GFP expression or indirect immunofluorescence of Tat using a rabbit anti-Tat primary antibody and an anti-rabbit Alexa-568 conjugated secondary antibody. Tat expression in the nuclear and extra-nuclear compartments was quantitated. DAPI was used as the nuclear stain. **(B)** GFP-flow profiles (left panels) and representative confocal images of Tat and GFP-expression at multiple time-points (right panels) are presented. A white dotted line demarcates the nuclear from the extra-nuclear compartments in each cell. Scale bar = 20μM (**C**) Quantitative estimation of Tat-expression in the nuclear and extra-nuclear compartments during latency-establishment. The total cellular GFP intensity and the intensities of Tat in the nuclear and extra-nuclear compartments (Alexa 568 signal) were measured in arbitrary units from individual cells and plotted against the days post-sorting (left panel). The threshold values for GFP-intensity and Tat-intensity were obtained from uninfected Jurkat cells and infected, unstained cells (no-anti-Tat control) respectively. Data are collected from 150 individual cells for each time point and from three independent experiments. Mean values ± SEM are plotted. One-way ANOVA was used for statistical evaluation (***$p<0.001$). The ratio of the mean intensities of Tat-Alexa 568 in the nuclear and the extra-nuclear compartments is presented as a function of time (right panel).

The cells progressively downregulated GFP expression and by D8 only 6.9% of cells were positive for GFP by flow cytometry and by D16 less than 1% cells were GFP positive confirming LTR silencing (Figure 4.5B; left panel). The GFP-expression pattern of individual cells monitored by confocal microscopy was perfectly consistent with that of the flow analysis that no visible fluorescence could be detected at D8 or beyond (Figure 4.5B; right panel). Cellular distribution of Tat as demonstrated by the indirect immunofluorescence (Alexa-568) followed distinct temporal patterns in the nuclear and extra-nuclear compartments (DAPI overlap) as transcriptional silence established progressively (Figure 4.5B; right panel and Figure 4.5C; left panel). On days D0, D4 and also D8 where the transcriptional activity of the viral promoter was the highest, overall Tat protein expression was also high (red fluorescent signal) that decreased progressively with time. With the progressive silencing of the LTR, the physiological levels of Tat dropped considerably at D8, D12 and D16, due to the attenuation of the concomitant LTR-Tat feedback circuit.

Importantly, a dramatic reduction in the Tat levels was noted from D0 to D4 in the extra-nuclear compartment (~4.2 fold) while the drop in the nuclear Tat signal was relatively milder (~3.1 fold) during the same period (Figure 4.5C; left panel). The slopes of reduction in the Tat intensities during the initial phase (D0 to D4) of latency-establishment were estimated to be -74.54 ± 16.8 and -37.28 ± 3.2, respectively, for the extra-nuclear and nuclear Tat. At the later time-points (D8, D12 and D16), there was only a moderate reduction in the Tat levels both in the extra-nuclear and nuclear compartments although the Tat concentration at both the locations was above the background level as late as D16 during which the GFP expression of the LTR completely extinguished (Figure 4.5C; left panel). The data thus, suggested that the stability of Tat in the nucleus could be

higher. Further, the ratios of nuclear and extra-nuclear Tat signals were compared at every time point. This analysis found that the intracellular Tat levels are relatively stable during the process of LTR silencing alluding to a direct role for Tat in HIV-1 latency (Figure 4.5C; right panel). The data of Tat-immunofluorescence are in perfect agreement with the observation of the 'GFP^{-ve} Tat-RFP^{+ve}' cells emerging at the later stages of promoter-silencing in the TTF model (Figures 3.21 and 3.22). To establish a quantitative measure of the Tat distribution profile among the two cellular compartments, we monitored the fluorescence profile of 150 individual cells using confocal microscopy as the cells downregulated GFP expression. The total fluorescent intensities of Tat-Alexa 568 were determined independently for the nuclear and the extra-nuclear compartments and the threshold for background was set from the values of a no-primary control for Tat intensity (n = 10), and uninfected Jurkat control for GFP (n = 10).

4.7 The highly sensitive in situ proximity ligation assay (PLA) detects the presence of Tat in the latently infected cells.

The Tat-immunofluorescence described above (section 4.6) not only detected Tat in the latently infected cells as late as D16 post-sorting, but also demonstrated a rapid loss of Tat from the extra-nuclear compartment while its relative stability in the nucleus. However, the over-all intensity of the Tat signals at D12, D14 and D16 in both the compartments was only marginally above the background level. To increase the sensitivity of Tat detection in the 'LTR OFF' cells and to gain valuable insights into the Tat-interacting host cellular factors, we used the highly sensitive proximity ligation assay (PLA).

The PLA conjugates immunostaining with the rolling-circle replication and outperforms the traditional immune assays in sensitivity to detect trace amounts of endogenous proteins at concentrations as low as a single copy. PLA not only permits the co-localisation of two different proteins in the same cellular compartment but also identifies physically interacting partners within a 40 nm territory with a precision unimaginable with the standard immunofluorescence protocols (Figure 4.6; Gustafsdottir SM et al., 2005; Soderberg O et al., 2006).

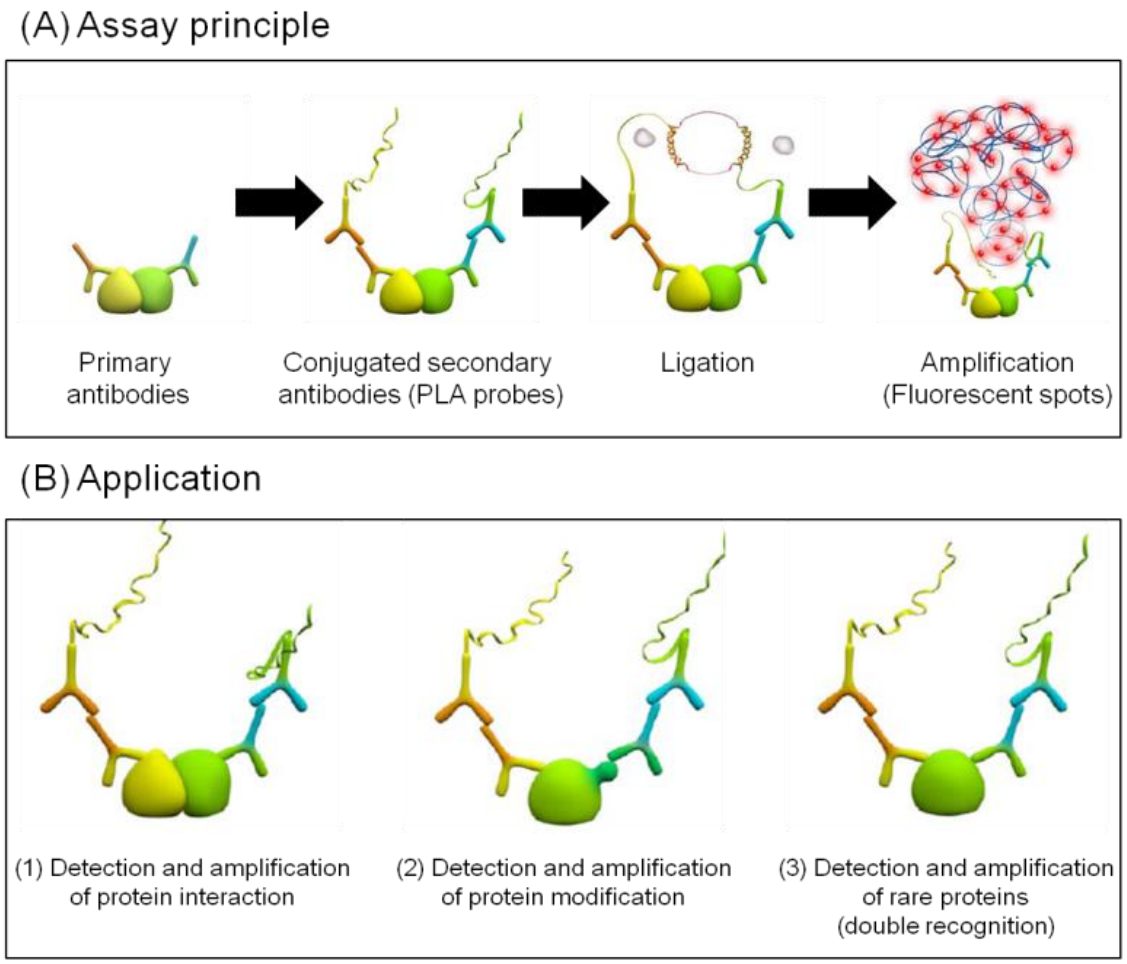


Figure 4.6: Schematic representation of the in-situ proximity ligation assay (PLA). In PLA, a single rare protein, definite post-translational modifications of proteins or two physically interacting proteins (distance ≤ 40 nm) can be evaluated with a higher sensitivity and precision than the traditional confocal techniques. A detailed account of PLA is provided in doi:10.1007/978-1-4939-2742-5_15. The figure has been modified from the product manual, Cat# DUO12901, Sigma-Aldrich.

We optimised Tat-PLA in HEK293T cells using a pair of anti-Tat primary antibodies raised in different hosts (rabbit and mouse), and the experimental strategy as described (Section 2.5.3, Chapter 2). Since PLA does not work well in non-adherent cells, and our attempts to adapt the protocol to the Jurkat cells were not successful, we used HEK293/HEK293T cells in this assay. Initially, we performed a set of control experiments with the above pair of primary antibodies to validate specific PLA signals from exogenous Tat protein in HEK293T cells. To this end, HEK293T cells were transfected with either of the two miniviral vectors- pLGIT expressing B-Tat or different doses of pcLGIT expressing C-Tat as indicated (Figure 4.7; left panel). Following 36 h of transfection, the cells were stained with the anti-Tat antibodies (at concentrations previously titrated for immunofluorescence staining) and PLA was performed following manufacturer's instructions. Distinct PLA signals (white dots) could be detected both for

B-Tat as well as C-Tat (Figure 4.7; left panel) using the antibody pair (Lanes 3-6) as opposed to sparse dots in single antibody controls (Lanes 1 and 2). Further, in the case of C-Tat (pcLGIT), progressive increase in both GFP-MFI and the number of PLA dots were visible in a plasmid dose-responsive manner confirming the specific detection of C-Tat in PLA by the chosen set of antibodies (Figure 4.7; right panel).

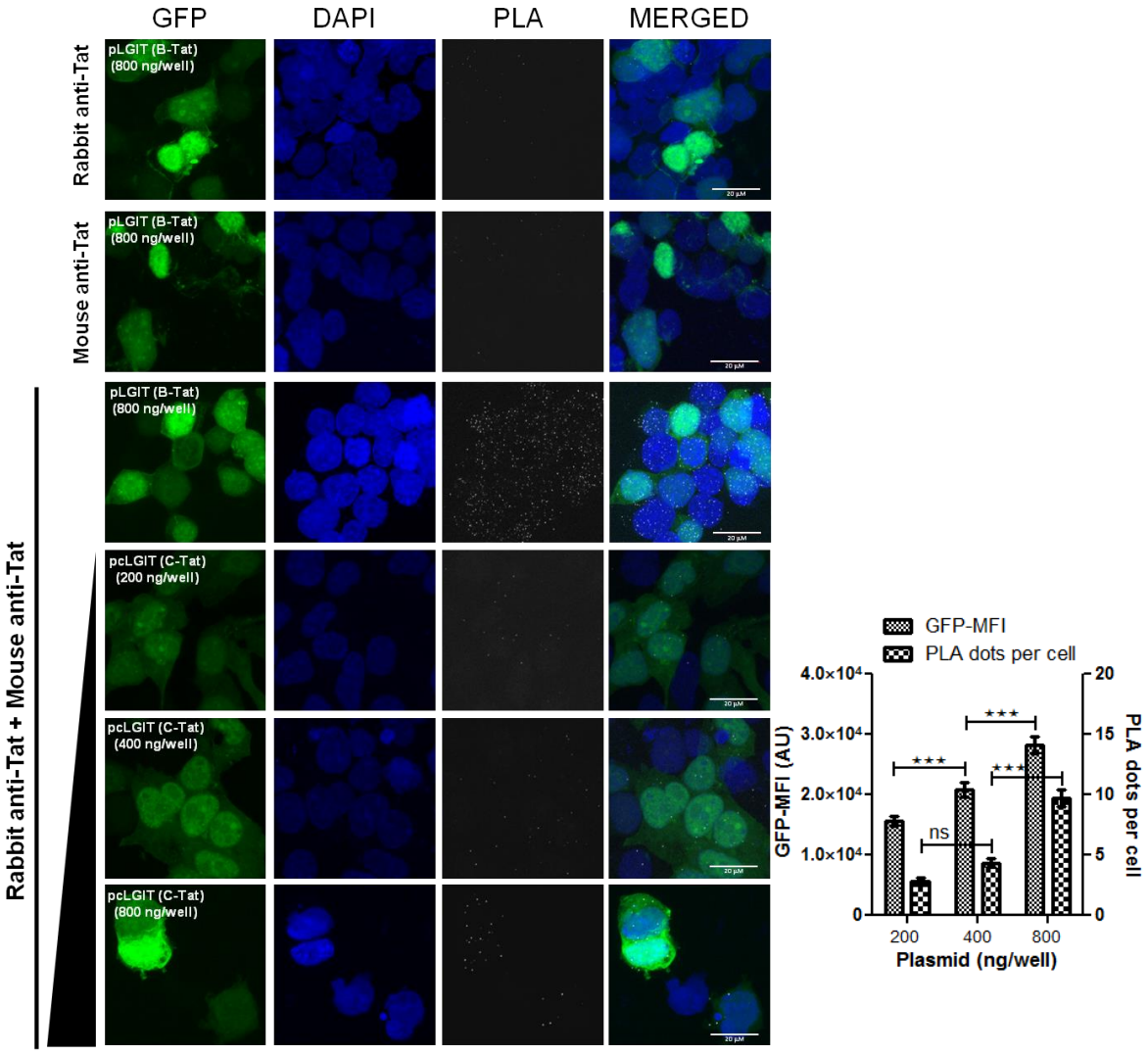


Figure 4.7: Validation of anti-Tat primary antibodies in the in-situ proximity ligation assay (PLA). A pair of anti-Tat primary antibodies (rabbit-polyclonal; Cat # ab43014 and mouse-monoclonal; Cat # 1D9) was tested for their specific detection of Tat protein in PLA in HEK293 cells. Approximately 0.5 million HEK293T cells were transfected on poly-L-lysine coated coverslips in each well of a 12-well culture dish with either pLGIT (800 ng) or pcLGIT (200, 400, 800 ng) as indicated. Following immunostaining of Tat using the above pair of primary antibodies, PLA was performed according to manufacturer's protocol (Cat# DUO12901, Sigma-Aldrich). Representative fields of confocal images (left panel). Lane1: Rabbit ant-Tat single antibody control; Lane 2: Mouse anti-Tat single antibody control; Lane 3: Tat-PLA using Rabbit and Mouse anti-Tat antibodies to stain exogeneous B-Tat expressed from the vector pLGIT. Lane 4-6: Tat-PLA using Rabbit and Mouse anti-Tat antibodies to stain exogeneous C-Tat expressed from increasing concentrations of the vector pcLGIT. Distinct Tat-PLA dots (white) indicated the presence of both B- and C-Tat proteins, as opposed to sparse signals in the single antibody controls. Quantitative estimation of Tat-PLA dots revealed a DNA dose-dependent increase in signal (right panel). Mean GFP intensities and the average number of dots per cell were determined for each dose of the vector pcLGIT (ng/well) transfected. Mean values from three independent experiments ± SEM are plotted. Total numbers of cells counted for each plasmid dose were 25. A one-way ANOVA with Bonferroni's multiple comparison post-test was used for statistical analyses (***$p<0.001$).

Using the optimised PLA protocol for Tat, we asked if the presence of Tat could be detected in latent cells. Following the experimental schematic as depicted (Figure 4.8A), HEK293 cells were stably infected with 4-κB variant of the LdGIT panel of the ATF model and the cells expressing high levels of GFP were sorted to isolate infected cells. After a week of incubation following the enrichment, approximately 50% of the cells expressed GFP, thus, the cell pool contained both active (GFP^{+ve}) and latent (GFP^{-ve}) cells. The cell pool was stained for Tat using the pair of anti-Tat antibodies and the optimised PLA protocol. The cells stained with either of the antibodies alone did not show any Tat-specific signals confirming the specificity of the assay (Figure 4.8B). Tat-specific staining was evident only in the presence of both the antibodies not only in the GFP^{+ve} cells but also in the GFP^{-ve} cells (Figure 4.8B). The mean Tat staining intensity was determined in a total of 85 GFP^{+ve} cells and 119 GFP^{-ve} cells comprising of three independent experiments (Figure 4.8C). These values were found to be 2.91 ± 2.5 and 2.34 ± 1.9 for GFP^{+ve} and GFP^{-ve} cells respectively, although the difference was not significant statistically. The data of PLA staining of Tat in HEK293 cells collectively confirmed the presence of Tat in the latent cells at a concentration comparable to that of active viral transcription.

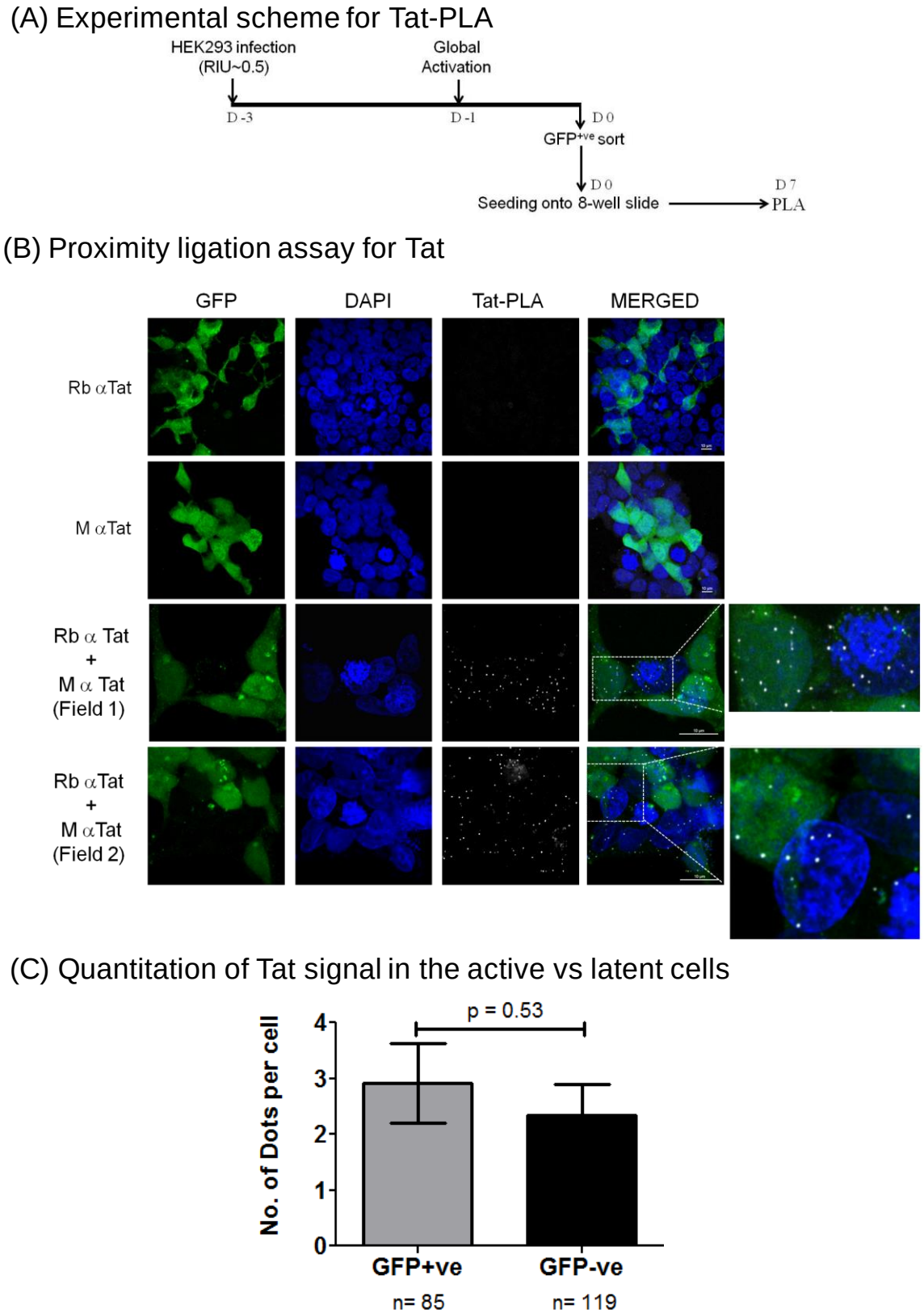


Figure 4.8: The proximity ligation assay detects Tat in the latent cells. (A) The schematic representation for the Tat-PLA assay in HEK293 cells. Approximately one million cells were infected with the 4-κB strain of the LdGIT panel of the ATF model (~ 0.5 RIU). After three days of infection, the infected cells were stimulated with the global activators and the GFP^{+ve} cells were sorted. The sorted cells with active provirus were relaxed for a week to arrive at a mixed population of active and latent cells (~50% each of GFP^{+ve} and GFP^{-ve}). Approximately 50,000 mixed GFP cells were

seeded on an 8-well slide chamber, allowed to adhere to the glass surface and stained for Tat using a pair of primary antibodies raised in rabbit and mouse. PLA was performed according to the manufacturer's protocol (Cat# DUO12901, Sigma-Aldrich). **(B)** Representative fields of confocal images depicting mixed GFP cells. Lane1: Rabbit-αTat single antibody control; Lane 2: Mouse-αTat single antibody control; Lanes 3 and 4: Tat-PLA using Rabbit-αTat and Mouse-αTat antibodies. Distinct Tat-PLA dots (white) were noticed in both the active and latent phenotypes (Lanes 3 and 4), while the single antibody controls were devoid of the signal. Two representative fields have been presented and a sub-field from each containing the Tat signals in both GFP^{+ve} (active infection) or GFP^{-ve} (latent infection) cells have been enlarged for clarity. **(C)** Quantitative estimation of Tat in the active and latent cells. Individual dots were counted and the dot density per cell was determined independently for GFP^{+ve} as well as GFP^{-ve} cells. Mean values from three independent experiments ± SEM are plotted. Total numbers of cells counted were 85 from GFP^{+ve} and 119 from GFP^{-ve} categories. A two-tailed, unpaired t-test was used for statistical analyses. Importantly, the analysis indicated no significant difference in the Tat signal densities between the active and latent phenotypes.

Using PLA, we next asked if Tat is in proximity to the transcription factors, such as p50 or p65, known to be recruited to the LTR in active and latent cells. We used p65 to address this question as Tat is known to directly interact with multiple components of the NF-κB signalling pathway. Nuclear export of Tat via physical interaction with the NF-κB inhibitor IκB-α is an established mechanism to negatively regulate HIV-1 gene expression (Puca A et al., 2007; Vitagliano L et al., 2011). In a subsequent study by the same group, Tat was found to compete for the p65 binding-site on the IκB-α further suggesting a direct interaction of Tat with p65 (Fiume G et al., 2011). Prior to performing the Tat-p65 PLA in the HEK293 cells, stably infected with the LdGIT-4-κB virus, we carried out a control experiment to validate the anti-p65 primary antibodies for specific detection of p65 in PLA (Figure 4.9). Towards this, we used a pair of anti-p65 primary antibodies raised in rabbit and mouse. HEK293T cells were co-transfected with the p65 expression vector- pCMV-3xFLAG-p65 and pCMV-sLuc-IRES-GFP (pCMV-LIG; transfection control) and post 36 h of transfection, immunostaining and subsequently PLA were performed using the above anti-p65 primary antibodies (previously optimised for immunofluorescence). Bright white PLA dots overlapping with the GFP expression (Figure 4.9; Lane 3) as opposed to relatively scarce dots in the single antibody controls (Figure 4.9; Lanes 1 and 2) confirmed the specific detection of p65 in PLA using the above pair of antibodies.

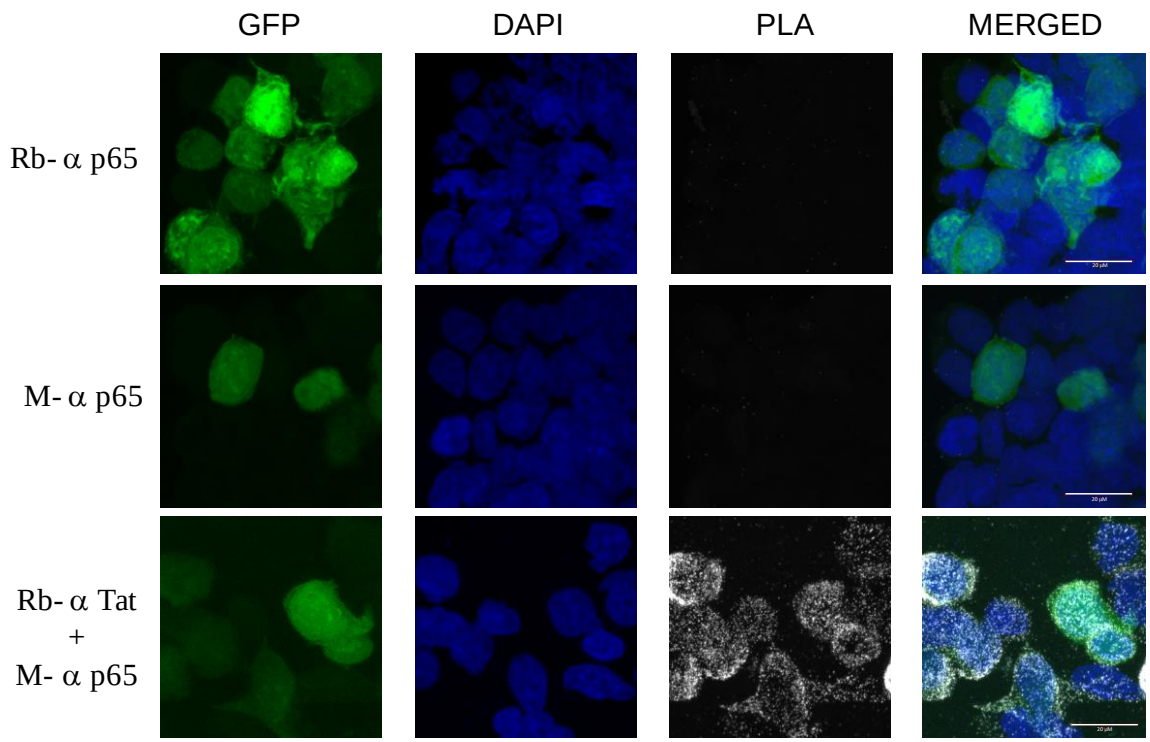


Figure 4.9: Validation of anti-p65 primary antibodies in the in-situ proximity ligation assay (PLA). A pair of anti-p65 primary antibodies (rabbit-polyclonal; Cat # ab16502 and mouse-monoclonal; Cat # CST 6956) was tested for their specific detection of p65 in PLA in the HEK293T cells. Approximately 0.5 million HEK293T cells were co-transfected with 600 ng of pCMV-3xFLAG-p65 and 200 ng pCMV-LIG on poly-L-lysine coated coverslips in each well of a 12-well culture dish. Following the usual staining procedure for IF with the above antibodies, PLA was performed according to manufacturer's protocol (Cat# DUO12901, Sigma-Aldrich). Representative fields of confocal images (left panel). Lane1: Rabbit anti-p65 single antibody control; Lane 2: Mouse anti-p65 single antibody control; Lane 3: p65-PLA using Rabbit and Mouse anti-p65 antibodies. Saturating p65-PLA signals represented by bright white dots as opposed to scarce dots in the single antibody controls validated specific detection of p65 in PLA using the above pair of antibodies.

The PLA-validated rabbit anti-Tat (Figure 4.7) and mouse anti-p65 (Figure 4.9) antibodies were then used to compare the physical proximity of Tat and p65 in active vs latent infections. A stable pool of HEK293 cells harbouring the LdGIT-4-κB virus and comprising of both active (GFP^{+ve}) and latent (GFP^{-ve}) phenotypes was chosen for the assay. The experimental schematic was identical to the previous PLA experiment to detect and compare the presence of Tat in the active and latent cells (Figure 4.8A). In the present experiment, distinct Tat-p65 interaction was evident from the cells only in the presence of both the antibodies but not when either of the antibodies was used alone (Figure 4.10A). The mean number of PLA dots representing Tat-p65 interaction was found to be 3.17 ± 0.5 and 1.52 ± 0.2 for the GFP^{+ve} and (actively infected) GFP^{-ve} (latently infected) cells by counting 166 and 194 cells, respectively (Figure 4.10B). The difference in the signal intensity between the transcriptionally active and latent proviruses was statistically significant.

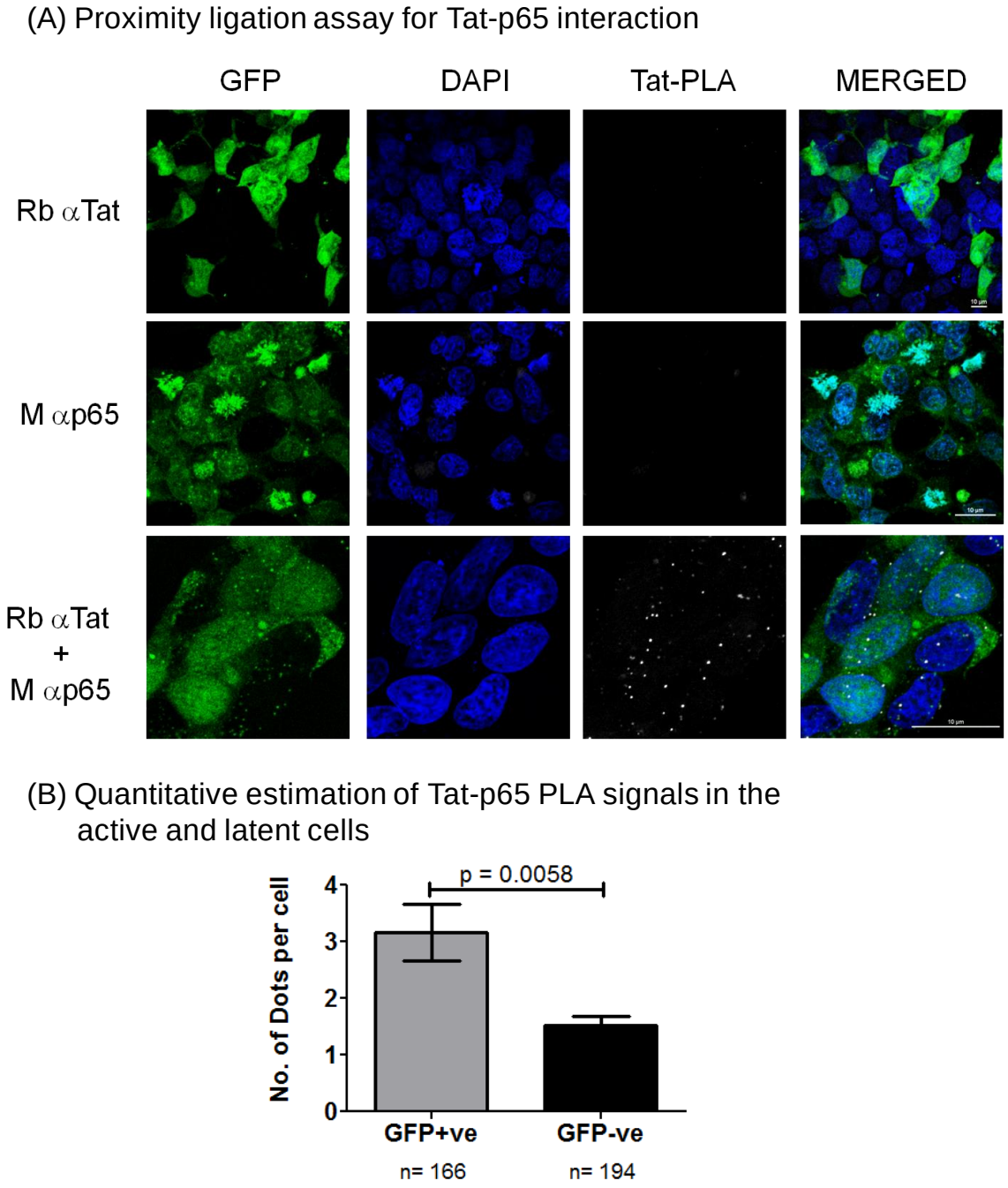


Figure 4.10: Enhanced Tat-p65 interaction in the actively infected cells. (A) Representative confocal images of the Tat-p65 interaction in the active and latent cells. A mixed GFP population of stable HEK293 cells harbouring the 4-κB strain of the LdGIT panel (the ATF model) was probed for the proximal association of Tat and p65 using a rabbit anti-Tat and a mouse anti-p65 antibodies following the manufacturer's protocol (Cat# DUO12901, Sigma-Aldrich). Lane 1: Rabbit anti-Tat single antibody control; Lane 2: Mouse anti-p65 single antibody control; Lane 3: Samples stained for PLA using rabbit anti-Tat and mouse anti-p65antibodies. **(B)** Quantitation of Tat-p65 PLA signals in the active and latent cells. Individual dots were counted and the Tat-density per cell was determined independently for GFP^{+ve} as well as GFP^{-ve} cells. Mean values (dots per cell) from three different experiments ± SEM are plotted. A total of 166 and 194 cells respectively from GFP^{+ve} GFP^{-ve} category were counted. A two-tailed, unpaired t-test was used for statistical analyses. The association of Tat-p65 was found to be ~2-fold higher in the active cells as compared to latent cells.

4.8 Summary

In summary, the data of Tat-PLA confirm the presence of Tat in the latently infected cells and permitted the examination of the molecular complexes differentially recruited to the active versus latent promoter and the interaction of Tat with some of these factors at the single-cell resolution. For instance a preferred physical proximity of Tat with p65 was noted in the case of active transcription over latent infection. These results are also consistent with our previous data that suggested a positive correlation between the transcription strength of the LTR and the Tat-feedback strength which favoured latency establishment. Importantly, we identified a reciprocal binding of essential transcription factors such as p50, p65, NFAT1 and NFAT2 to active versus latent LTRs. The presence of Tat at the latent promoter as identified by ChIP is suggestive of the viral factor playing a critical role in initiating the process of transcriptional silence and its regulation.

4.9 Bibliography

Barbeau, B., Bernier, R., Dumais, N., Briand, G., Olivier, M., Faure, R., Posner, B.I. and Tremblay, M., 1997. Activation of HIV-1 long terminal repeat transcription and virus replication via NF-κB-dependent and-independent pathways by potent phosphotyrosine phosphatase inhibitors, the peroxovanadium compounds. *Journal of Biological Chemistry*, *272*(20), pp.12968-12977.

Bates, D.L., Barthel, K.K., Wu, Y., Kalhor, R., Stroud, J.C., Giffin, M.J. and Chen, L., 2008. Crystal structure of NFAT bound to the HIV-1 LTR tandem κB enhancer element. *Structure*, *16*(5), pp.684-694.

Burnett, J.C., Miller-Jensen, K., Shah, P.S., Arkin, A.P. and Schaffer, D.V., 2009. Control of stochastic gene expression by host factors at the HIV promoter. *PLoS pathogens*, *5*(1), p.e1000260.

Chang, H.C., Samaniego, F., Nair, B.C., Buonaguro, L. and Ensoli, B., 1997. HIV-1 Tat protein exits from cells via a leaderless secretory pathway and binds to extracellular matrix-associated heparan sulfate proteoglycans through its basic region. *Aids*, *11*(12), pp.1421-1431.

Chen-Park, F.E., Huang, D.B., Noro, B., Thanos, D. and Ghosh, G., 2002. The κB DNA sequence from the HIV long terminal repeat functions as an allosteric regulator of HIV transcription. *Journal of Biological Chemistry*, *277*(27), pp.24701-24708.

Ensoli, B., Barillari, G., Salahuddin, S.Z., Gallo, R.C. and Wong-Staal, F., 1990. Tat protein of HIV-1 stimulates growth of cells derived from Kaposi's sarcoma lesions of AIDS patients. *Nature*, *345*(6270), p.84.

Feinberg, M.B., Baltimore, D. and Frankel, A.D., 1991. The role of Tat in the human immunodeficiency virus life cycle indicates a primary effect on transcriptional elongation. *Proceedings of the National Academy of Sciences*, *88*(9), pp.4045-4049.

Fiume, G., Vecchio, E., De Laurentiis, A., Trimboli, F., Palmieri, C., Pisano, A., Falcone, C., Pontoriero, M., Rossi, A., Scialdone, A. and Fasanella Masci, F., 2011. Human immunodeficiency virus-1 Tat activates NF-κB via physical interaction with IκB-α and p65. *Nucleic acids research*, *40*(8), pp.3548-3562.

Giffin, M.J., Stroud, J.C., Bates, D.L., von Koenig, K.D., Hardin, J. and Chen, L., 2003. Structure of NFAT1 bound as a dimer to the HIV-1 LTR κB element. *Nature Structural and Molecular Biology*, *10*(10), p.800.

Gustafsdottir, S.M., Schallmeiner, E., Fredriksson, S., Gullberg, M., Söderberg, O., Jarvius, M., Jarvius, J., Howell, M. and Landegren, U., 2005. Proximity ligation assays for sensitive and specific protein analyses. *Analytical biochemistry*, *345*(1), pp.2-9.

Kinoshita, S., Su, L., Amano, M., Timmerman, L.A., Kaneshima, H. and Nolan, G.P., 1997. The T cell activation factor NF-ATc positively regulates HIV-1 replication and gene expression in T cells. *Immunity*, *6*(3), pp.235-244.

Klichko, V., Archin, N., Kaur, R., Lehrman, G. and Margolis, D., 2006. Hexamethylbisacetamide remodels the human immunodeficiency virus type 1 (HIV-1) promoter and induces Tat-independent HIV-1 expression but blunts cell activation. *Journal of virology*, *80*(9), pp.4570-4579.

Lieb, M., 1953. The establishment of lysogenicity in Escherichia coli. *Journal of bacteriology*, *65*(6), p.642.

Macián, F. and Rao, A., 1999. Reciprocal modulatory interaction between human immunodeficiency virus type 1 Tat and transcription factor NFAT1. *Molecular and Cellular Biology*, *19*(5), pp.3645-3653.

Mbonye, U. and Karn, J., 2011. Control of HIV latency by epigenetic and non-epigenetic mechanisms. *Current HIV research*, *9*(8), pp.554-567.

Pereira, L.A., Bentley, K., Peeters, A., Churchill, M.J. and Deacon, N.J., 2000. SURVEY AND SUMMARY A compilation of cellular transcription factor interactions with the HIV-1 LTR promoter. *Nucleic acids research*, *28*(3), pp.663-668.

Pessler, F. and Cron, R.Q., 2004. Reciprocal regulation of the nuclear factor of activated T cells and HIV-1. *Genes and immunity*, *5*(3), p.158.

Puca, A., Fiume, G., Palmieri, C., Trimboli, F., Olimpico, F., Scala, G. and Quinto, I., 2007. IκB-α represses the transcriptional activity of the HIV-1 Tat transactivator by promoting its nuclear export. *Journal of Biological Chemistry*, *282*(51), pp.37146-37157.

Razooky, B.S. and Weinberger, L.S., 2011. Mapping the architecture of the HIV-1 Tat circuit: A decision-making circuit that lacks bistability and exploits stochastic noise. *Methods*, *53*(1), pp.68-77.

Razooky, B.S., Pai, A., Aull, K., Rouzine, I.M. and Weinberger, L.S., 2015. A hardwired HIV latency program. *Cell*, *160*(5), pp.990-1001.

Söderberg, O., Gullberg, M., Jarvius, M., Ridderstråle, K., Leuchowius, K.J., Jarvius, J., Wester, K., Hydbring, P., Bahram, F., Larsson, L.G. and Landegren, U., 2006. Direct observation of individual endogenous protein complexes in situ by proximity ligation. *Nature methods*, *3*(12), p.995.

Stroud, J.C., Oltman, A., Han, A., Bates, D.L. and Chen, L., 2009. Structural basis of HIV-1 activation by NF-κB—A higher-order complex of p50: Rela bound to the HIV-1 LTR. *Journal of molecular biology*, *393*(1), pp.98-112.

Süel, G.M., Garcia-Ojalvo, J., Liberman, L.M. and Elowitz, M.B., 2006. An excitable gene regulatory circuit induces transient cellular differentiation. *Nature*, *440*(7083), p.545.

Verma, A., Rajagopalan, P., Lotke, R., Varghese, R., Selvam, D., Kundu, T.K. and Ranga, U., 2016. Functional Incompatibility between the Generic NF-κB Motif and a Subtype-Specific Sp1III Element Drives the Formation of the HIV-1 Subtype C Viral Promoter. *Journal of virology*, *90*(16), pp.7046-7065.

Vitagliano, L., Fiume, G., Scognamiglio, P.L., Doti, N., Cannavò, R., Puca, A., Pedone, C., Scala, G., Quinto, I. and Marasco, D., 2011. Structural and functional insights into IκB-α/HIV-1 Tat interaction. *Biochimie*, *93*(9), pp.1592-1600.

Wei, P., Garber, M.E., Fang, S.M., Fischer, W.H. and Jones, K.A., 1998. A novel CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its high-affinity, loop-specific binding to TAR RNA. *Cell*, *92*(4), pp.451-462.

Weinberger, L.S., Burnett, J.C., Toettcher, J.E., Arkin, A.P. and Schaffer, D.V., 2005. Stochastic gene expression in a lentiviral positive-feedback loop: HIV-1 Tat fluctuations drive phenotypic diversity. *Cell*, *122*(2), pp.169-182.

Weinberger, L.S. and Shenk, T., 2006. An HIV feedback resistor: auto-regulatory circuit deactivator and noise buffer. *PLoS biology*, *5*(1), p.e9.

Weinberger, L.S., Dar, R.D. and Simpson, M.L., 2008. Transient-mediated fate determination in a transcriptional circuit of HIV. *Nature genetics*, *40*(4), p.466.

Williams, S.A., Chen, L.F., Kwon, H., Ruiz-Jarabo, C.M., Verdin, E. and Greene, W.C., 2006. NF-κB p50 promotes HIV latency through HDAC recruitment and repression of transcriptional initiation. *The EMBO journal*, *25*(1), pp.139-149.

Chapter 5

Discussion

5.1 Genetic feedback circuits have evolved as molecular switches to decide between alternate fates in all the domains of life.

Phenotype-switching is a fundamental strategy employed by all life forms- viruses, bacteria, yeasts, lower metazoans and mammals to survive hostile environmental conditions and attain evolutionary fitness. Genetically identical cells exposed to the same environmental conditions may manifest diametrically opposite and heritable phenotypes. The phenomenon is widely investigated in bacteria. A landmark research led by Jacob and Monod on the lac operon in *E. coli* identified bacterial cells that could manifest alternate phenotypes under identical genetic and environmental settings based on their lactose metabolizing property (Jacob F and Monod J., 1961). Recently, *B. subtilis* has emerged as an exemplary tool to study multiple switching phenotypes during 'spore-formation' and 'bacterial competence' (Galhardo RS et al., 2007; Suel GM et al., 2006). Among eukaryotes, the simplest model yeast (*S. cerevisiae)* was used to examine the galactose-uptake decision-making pathway (Biggar SR and Crabtree, 2001). Likewise, apoptosis, cell-cycle, and stem-cell differentiation are a few of the physiological processes that require switching mechanisms in higher eukaryotes and mammals. Interestingly, all cell-fate switching mechanisms use one or more feedback circuits in their gene-regulatory networks.

5.2 Several viral families including HIV-1 exploit feedback circuits to switch between active and silent replication modes.

HIV-1 latency is a reversible phenomenon by which a replication-competent virus stays transcriptionally silent as a stable chromatin integrant within the host cells. Latency is the classic example of phenotype-switching in viruses and one of the many survival strategies to evade the host-defense mechanisms for sustenance and successful propagation. Alternately latency could be the resting phase of the virus in an inactive host cell until the latter attains an optimal physiological state to trigger active viral proliferation. For instance, an appropriate threshold for the host transcription factors is mandatory for cellular as well as viral transcription. Therefore, viral latency could either be an inadvertent consequence of the inactivated state of the host cell or it may be an intrinsic, viral-driven phenomenon; a matter presently of considerable debate. Based on whether

the host cell or the infected virus primarily determines the replication-fate of the virus, two distinct models of latency have been proposed- the 'epiphenomenon' or deterministic model and the 'viral circuitry' or the intrinsic model, respectively.

The developmental bifurcation in bacteriophage λ was first observed by Lieb M et al. Virions manifested either the lytic or latent phenotypes from a genetically identical pool of *E. coli* (Lieb M et al., 1953). A double-negative feedback loop involving two inducible proteins Cro and λ-repressor (CI) was later identified at the core of the regulatory mechanism. Excess Cro and CI proteins mutually inhibited each other's synthesis from a single bi-directional promoter-operator region, thus locking the viral fate in either of the two states: high Cro and low CI (favouring lysis) or high CI and low Cro (favouring lysogeny) (Arkin A et al., 1998). Subsequently, cooperativity or oligomerization of the λ-repressor was found essential to establish the 'bistable' circuit (Arkin A et al., 1998; Dodd IB et al., 2001, Gottesman M et al., 1999, Ptashne M et al., 2005).

Latency in eukaryotic viruses fundamentally differs from that of the prokaryotic counterparts in the absence of (i) a negative factor (repressor) and (ii) a negative-feedback loop in the viral circuitry. In fact, all eukaryotic viruses with a potential to establish latency such as EBV (Sarisky RT et al., 1996; Ragoczy T et al., 1998; Ragoczy T et al., 2001), HSV-1 (Cai W et al., 1993), CMV (Cherrington JM et al., 1989; Malone CL et al., 1990) and HIV-1 exploit one or more excitatory feedback circuits that synthesize positive transactivator proteins to drive promoter silencing. The establishment of latency despite a positive auto-regulation seems contradictory in these viruses. The paradox is resolved to a great extent through the 'cooperative effect' of the transactivator proteins. Active viral transcription occurs when the transactivators multimerize at the activation site, while, fail to support viral transcription below a certain threshold due to inefficient oligomerization. Thus, concentration-dependent oligomerization of transactivator proteins in these viruses account for the bistable (all or none) mode of transcription regulation.

5.3 The HIV-1 transcription circuit fundamentally differs from the rest of the latency-establishing viruses.

HIV-1 unlike the bacteriophage λ does not possess a repressor mediated negative-feedback circuit nor is known to exhibit cooperativity in the transcriptional circuit (Razooky BS et al., 2011) unlike other eukaryotic viruses. How does HIV-1 then attain viral latency- a stable, transcriptionally OFF state? This intriguing question forms the basis of the present study. Two different hypotheses have been proposed to explain how HIV-1 can establish a stable OFF state despite lacking bistability in its LTR-Tat transcription-circuit. The first mechanism represents the existence of a 'feedback-resistor' module within the LTR-Tat auto-regulatory circuit that comprises of an enzymatic switch between an active (acetylated) and inactive (deacetylated) forms of Tat (Ott M et al., 2004; Weinberger LS and Shenk T., 2006). Since Tat-acetylation follows a slower kinetics than the deacetylation reaction, the former acts as the rate-limiting step for the Tat-mediated transactivation. The deacetylation step thus serves as the 'resistor' or dissipater of the positive Tat-feedback loop and ensures a stable latent state. This model however, is limited in offering a convincing explanation as to the nature of the driving force that can mediate a viral rebound in spite of the viral circuit predisposed to maintain only the 'OFF' state. The second hypothesis suggests that the positive Tat-feedback loop generates long-lasting transcription pulses and additionally, that the half-life of Tat exceeds the doubling time of the infected cell itself (Weinberger LS et al., 2008). The extended stability of Tat accounts for the physiological persistence of a few acetylated residues sufficient for successful Tat-TAR interaction and a transcriptional 'ON' state. The provirus can prevail in the active state until the cell tilts to favor latency by which time the positive Tat-feedback would generate more Tat molecules. In this way, the virus persists in the active transcription mode for several generations. The hypotheses described above, however, have been proposed primarily based on computational studies using the HIV-1B circuit with little experimental validation from natural infection. In summary, empirical evidence towards understanding the mechanisms regulating viral latency in HIV-1 is scanty and lacking totally in genetic subtypes other than HIV-1B.

5.4 The copy-number variation of NF-κB motifs in HIV-1C offers an opportunity to discern the independent contributions of the LTR and Tat towards latency.

The LTR-Tat transcription-circuit in HIV-1C is similar to that in HIV-1B. Nevertheless, the C-LTR demonstrates extensive variation with respect to the sequence and copy-number of transcription-factor binding sites (TFBS) particularly the NF-κB sites, and Tat displays subtype-specific amino acid variations. HIV-1C alone among all the strains of HIV-1, HIV-2 and SIV contains two, three or four NF-κB motifs sites naturally, yet is capable of establishing latency. The biological significance of the LTR and Tat variations has not been examined. Importantly, the additional copies of NF-κB motifs are expected not only to modulate the transcriptional strength of C-LTR but also the associated LTR-Tat feedback strength. Furthermore, the genetically diverse NF-κB motifs may significantly modulate the quality of the Tat feedback loop by mediating the assembly of diverse host-factor complexes (activators or repressors) on the viral promoter thereby causing phenotypic outcomes of diverse nature. Thus, the copy-number variation in the NF-κB sites may set HIV-1C apart from all the other HIV-1 subtypes in the ability to switch between the active and latent phenotypes dynamically. Although it was experimentally verified that the acquisition of additional (the fourth) NF-κB motif endowed the newer HIV-1C strains with superior replication fitness (Bachu M et al., 2012a; Bachu M et al., 2012b), the influence of the enhanced copies of the NF-κB motifs on latency has not been examined. To this end, the present study was an attempt to address the question of the varying number of NF-κB motifs on viral latency in HIV-1C. In the present study, the genetic diversity of the NF-κB motifs was not considered. Additionally, we explored if the physiological concentrations of HIV-1C Tat, a direct consequence of the NF-κB copy-number variation, played a prominent role in latency.

5.5 The two components of the HIV-1C transcriptional circuit, the LTR and Tat, must function synergistically to establish latency.

The present study, to the best of our knowledge, is the first one to successfully uncouple the *cis*-element LTR and the *trans*-factor Tat from the transcriptional circuit to discern

their independent as well as combined influences of these factors on latency in the context of HIV-1C. As the only subtype with naturally occurring NF-κB copy-number variation, HIV-1C offers an ideal model to investigate transcriptional strength as an independent parameter affecting viral latency, a condition not explored previously. Three Jurkat cell-based latency models having distinct LTR-Tat feedback configurations in a minimal HIV-1C backbone permitted the above investigation. The details of the three models in the present study have been summarized (Table-3.2, Chapter 3). In the context of all the three models, we generated analogous panels of NF-κB copy-number variant LTRs ranging from 4 to 0 copies.

The ATF model, as expected, demonstrated a perfect positive correlation between the functional copies of NF-κB sites and the viral transcriptional strength (Figure 3.3). Both GFP (flow cytometry) and Tat-transcript expression (RT-PCR) from the NF-κB variant viral strains demonstrated high levels of correlation coefficient values ($r = 0.98$ and 0.96 for GFP and Tat-mRNA, respectively) proportional to the copy-numbers of NF-κB sites. The observation largely supported the previous report from our laboratory where the 4-κB molecular strain demonstrated high magnitude of replication competence than the 3-κB strain (Bachu M et al., 2012b). The data resulted in the identification of several novel concepts regarding latency in HIV-1– some of the observations have not been reported previously and some others differed at the mechanistic level as compared to non-HIV-1C strains.

Firstly, a novel finding of the present work is the positive correlation identified between the NF-κB copy-number and the faster kinetics of latency establishment (Figure 3.5). A progressive increase in the NF-κB motifs not only correlated positively with the transcriptional strength of the LTR, but also faster LTR-silencing; an observation not been identified previously probably due to the absence of TFBS polymorphism in non-HIV-1C LTRs. Importantly, of the two GFP populations that can be distinguished based on the fluorescence intensity, only the GFP^{High} cells (GFP-MFI $>10^4$ RFU), but not the GFP^{Low} cells, contributed significantly to both NF-κB dependent transcription as well as latency-establishment (Figures 3.3B and 3.5B). Since the Tat-feedback strength in the ATF circuit is a direct corollary of the NF-κB copy-number variation, it might be reasonable to assume that the proviruses of the GFP^{High} cells represent the Tat-

transactivated cells with a strong and operational Tat-feedback circuit while those of the GFP^{Low} cells with little or no LTR-Tat-feedback. It is tempting to propose that the autonomous Tat-feedback loop initially favors the steady accumulation of GFP and Tat molecules but subsequently switches to LTR silencing. The stronger promoters (3- and 4-κB LTRs) characterized by a stronger Tat-feedback, therefore, also initiate a rapid transcriptional silence as compared to the weak promoters (2-, 1- and 0-κB LTRs) (Figures 3.5C top, left and bottom, right panels, Figure 3.6A). The kinetics of latency-establishment were consistent especially in the GFP^{High} fraction regardless of whether the EGFP reporter (Figure 3.8) or a shorter half-life version GFPd2 (Figure 3.10) was used in the study; transcriptional silence was favored by the stronger LTRs with rapid kinetics. Thus, our data unequivocally prove that the transcriptional strength of the viral promoter combined with the LTR-Tat-feedback strength collectively determine the rapidity of latency-establishment. The ATF model, however, could not demarcate which of the two components plays a dominant role in latency. We, therefore, devised the DTF and TTF models for an independent evaluation of the transcriptional and Tat-feedback strengths, respectively.

Secondly, the ATF circuit of HIV-1C displayed a distinct phenotype related to the PheB phenomenon (all or none GFP expression by individual cells) as compared to HIV-1B. In HIV-1B, the PheB phenotype was manifested by the GFP^{Dim} clones (analogous to our GFP^{Low} cells) but not GFP^{High} cells even as late as one month post-sorting (Weinberger LS et al., 2005, Figure 3D). In HIV-1C, in contrast, the PheB phenotype was observed in the GFP^{High} cells and in all the NF-κB-variant strains (Figure 3.8B), and importantly, by clonal cell lines (Figure 4.1B). Whether the distinct behavior of the GFP^{High} cells of HIV-1C could be ascribed to subtype-specific differences or individual LTRs needs additional investigation.

Lastly, the question of Tat cooperativity appears to be distinct in HIV-1C. Razooky BS et al experimentally demonstrated a Hill coefficient value H =1 in HIV-1B, thus, ruling out the possibility of Tat cooperativity in LTR transactivation (Razooky BS et al., 2011). In contrast, we identified two distinct modes of reactivation of the latent cells, the slow (GFP^{Low}) and fast (GFP^{High}) reactivation kinetics although this phenomenon was noted only with the stronger (3- and 4-κB) LTRs. Using essentially the experimental strategy

described by Razooky BS et al, we found a Hill coefficient of H>1 with the 4-κB LTR implying a positive cooperativity in Tat function of HIV-1C (Figure 3.11). Although the data are quite noisy, a comparable trend was evident with the 3-κB LTR. The weak viral promoters (2-, 1-, and 0-κB LTRs), in contrast, manifested only the slow reactivation kinetics. Importantly, the percentages of GFP^{High} cells were proportional to the number of NF-κB motifs with 4- and 3-κB LTRs displaying 19.4% and 13.8% GFP^{High} cells, respectively. Thus, Tat cooperativity appears to be a function of the transcriptional strength of the LTR. The phenomenon of Tat cooperativity, therefore, was not identified in HIV-1B possibly given the fewer NF-κB motifs present in the viral promoter. Work is presently in progress in our laboratory to ask if the Tat-cooperativity identified in HIV-1C could be attributed to Tat oligomerization, heteromeric interaction with other transcription-factors or post-translation modifications of Tat. We also plan to understand why and how Tat-cooperativity is associated with the transcriptional strength of the LTR and the implications of this phenomenon on latency.

The DTF model differed from the ATF model in the absence of a functional LTR-Tat-feedback circuit. Tat was supplied in *trans* from an inducible Tet-ON promoter instead of being controlled by the LTR itself (Figure 3.13). Thus unlike in the ATF model, the physiological Tat levels could be normalized across the LTR-variants (Jurkat-LG). However, Tat-expression from the Tet-ON promoter was much lower than the saturating Tat levels in the ATF model such that even the highest dose of Dox failed to produce adequate molecules of Tat to generate the GFP^{High} cells in the strong LTRs (3- and 4-κB) (Figure 3.15). Nevertheless, when compared across the variant LTRs, the variation in the NF-κB copies failed to produce any difference in the transcriptional strength or in the kinetics of latency-establishment from the GFP^{Low} population (Figure 3.15 and 3.17).

The TTF model is analogous to the ATF model in preserving the LTR and Tat feedback loop intact. However, the ability to fine tune the intracellular concentrations of Tat by manipulating the Shield1 concentration in the medium offered the advantage of regulating the LTR Tat feedback loop with a fixed number of NF-κB motifs in the LTR. Importantly, the fusion of Tat with the DsRed2-RFP permitted the direct visualization of Tat (Figure 3.19A). The DsRed2-RFP tag with the Tat-ORF permitted the simultaneous tracking of the LTR activity (GFPd2 expression) as well as the Tat-expression levels in a

temporal fashion. In the TTF model, the 3-κB variant attained latency with significantly faster kinetics than the 1-κB LTR thus reproducing the observation of the ATF model- that stronger the promoter, faster the latency-establishment. Importantly, the ability to track Tat in the TTF model offered more information regarding the mechanism of transcriptional silencing that was not visible through the conventional ATF model. With the help of the TTF model, we could identify that the stronger promoter (3-κB) employs two independent modes of transcriptional silencing- Tat-dependent and Tat-independent whereas the weak 1-κB LTR displayed only the Tat-independent pathway to latency. In the Tat-independent route, the single GFP^{+ve} cells (GFP^{+ve} Tat-RFP^{-ve}) of 3-κB LTR turned to latency (GFP^{-ve} Tat-RFP^{-ve}) by directly switching off the expression of Tat (Figure 3.21). In contrast, in the Tat-dependent mode, the single GFP^{+ve} cells (GFP^{+ve} Tat-RFP^{-ve}) of 3-κB LTR, transited to latency via Tat-RFP expression successively passing through the phases of double positivity (GFP^{+ve} Tat-RFP^{+ve}) to single positivity (GFP^{-ve} Tat-RFP^{+ve}) to double negativity (GFP^{-ve} Tat-RFP^{-ve}; Figures 3.21 and 3.22). The GFP^{-ve} Tat-RFP^{+ve} cells are of great interest since the sustained presence of Tat in a promoter-silent condition has not been demonstrated previously. Furthermore, the presence of Tat in latent cells is strongly suggestive of the negative regulatory function of Tat contributing towards latency. Furthermore, the magnitude of Tat-dependent latency increased proportionately with the concentration of Shield1 hence with the strength of the Tat-feedback. The 1-κB promoter failed to take the Tat-dependent route presumably due to the absence of sufficient Tat concentration owing to a feedback of suboptimal strength. Thus, the TTF model directly revealed the significance of the Tat-feedback circuit in driving the Tat-dependent latency.

5.6 The multi-κB configuration in the C-LTR permits reciprocal binding of host-factors at the active and latent promoters; a switching advantage.

Gene expression is the outcome of multiple layers of regulatory events. At the fundamental level, gene regulation is affected by the nature of network motifs (a positive feedback in the case of HIV-1) and the sequence of the *cis*-acting TFBS that control the magnitude and timing of transcription factor occupancy at the promoter. Higher levels of regulation are attributed to the *trans*-acting chromatin remodelers, epigenetic marks and a

cross-talk between a wide array of proteins that ultimately bring about the diverse phenotypic outcomes. Host-epigenetics at the HIV-1 promoter and a cross-talk between the host and viral factors play a critical role to accomplish dynamic switching between the active and latent states (Pearson R et al., 2008, Mahmoudi T et al., 2012). The first study to analyze the recruitment of host-factors differentially at the active and latent promoters in the chromatinized context was conducted by Dr. Schaffer's group. B-LTR variants in the LGIT backbone were created by introducing point mutations at the individual NF-κB and Sp1 elements, and ChIP assay performed for transcription factor recruitment at the active vs latent promoters in stable Jurkat cells. The two HIV-1 NF-κB sites, despite being genetically identical played non-overlapping roles in viral transcription; the κB site I played a preferential role in activation while the κB site II favored p50-p50 homodimer recruitment and the establishment of latency (Figures 6B, 6E and S3, Burnett JC et al., 2009). Importantly, the assays were performed using PheB cells derived from the GFP^{Mid} parental cells (analogous to GFP^{Dim} in Weinberger LS et al., 2005 and GFP^{Low} in the present study) and the cells manifested a bimodal GFP distribution.

Unlike the above HIV-1B studies, we generated clonal cell lines from single GFP^{High} cells of the strong promoters (4- and 3-κB) and identified clones that manifested the PheB phenotype. In spite of the common integration site, a proportion of the proviruses undergo transcriptional silence while others maintain transcription, thus, manifesting the PheB phenotype. The GFP positive and negative cells derived from the single PheB clones offered a great technical advantage of the normalization of the inherent experimental parameters in the analysis of the transcription factor recruitment by ChIP assay. A single bimodal clone representing each of the 4- and 3-κB LTRs was selected; GFP^{High} and GFP^{-ve} fractions were sorted and assessed for factor recruitment at the active and latent LTRs using ChIP. The preferential binding of p50 and p65 (RelA) at the latent and active promoters ascertained the repressive and inducing functions of p50-p50 homodimer and p50-RelA heterodimer respectively, of HIV-1 transcription (Figure 4.3) (Stroud JC et al., 2009; Burnett JC et al., 2009; Williams SA et al., 2006). Unlike the NF-κB proteins, the impact of individual NFAT members on HIV-1 latency has not been examined in great detail. To the best of the knowledge, the present study is the first to demonstrate a reciprocal binding pattern for NFAT1 and NFAT2 at the active and latent promoters respectively, in the context of clonal cells. Since, NF-κB and NFAT factors share

overlapping sites (Pessler F et al., 2004, Bates DL et al., 2008, Giffin MJ et al., 2003), an increase in the copies of NF-κB motifs will also have increased concomitantly the binding sites for NFAT in C-LTR. Furthermore, the NF-κB sites in the C-LTR (F, H and C-κB sites) are genetically different adding to the multitude of possible combinations. Targeted inactivation of each κB site, one at a time, followed by ChIP may provide meaningful insights into the contribution of each κB sequence to diverse signaling pathways and HIV-1C latency.

5.7 Recruitment of Tat to the latent LTR

Tat predominantly is an RNA-binding protein that specifically forms a quaternary complex with pTEFb and the nascent TAR-mRNA (Fujinaga K et al., 1998; Wei P et al., 1998; Bieniasz PD., 1999) to enhance the processivity of RNA Pol II (Feinberg MB et al., 1991). Few studies examined the DNA binding properties of Tat. Southgate CD et al. first reported that Tat can effectively drive promoter activity when bound to upstream promoter elements and this activity requires cooperation with other transcription factors (Southgate CD et al., 1991). A single group thus far demonstrated the direct binding of Tat to the NF-κB sequences in the NL4-3 strain in both *in vitro* as well as *in vivo* assays (Dandekar JM et al., 2004). The group subsequently demonstrated a repressive role for Tat at the c-Rel promoter implying its negative regulation of cellular genes (Dhamija N et al., 2015).

The flow analysis of Tat-RFP (Figures 3.21 and 3.22) and the ChIP data (Figures 4.3 and 4.4) offered a strong experimental evidence for the presence of Tat in the latent cell and recruited to the latent LTR. Using ChIP, we could ascertain the recruitment of Tat to the latent LTR using three different anti-Tat antibodies although the signal intensities were less than or equal to those of the active promoter (Figures 4.3 and 4.4). Furthermore, a weak but positive signal of the Tat-transcripts were discernible in the latent fractions (~80 to 112 fold lower compared to active cells) of the bimodal clones of both the LTRs used in the assay (Figure 4.2). Although, these data are not compellingly assertive of a negative role for Tat promoting latency, they are indicative of such a possibility.

To confirm the presence of Tat in the latently infected cell, we performed two other experiments, indirect immunofluorescence for Tat through the successive stages of latency in Jurkat cells and a proximity ligation assay for Tat. Jurkat cells stably infected with the 3-κB LdGIT strain (ATF model) were analysed for intra- and extra-nuclear Tat levels by indirect immunofluorescence during the LTR-ON to LTR-OFF transition (Figure 4.5). A significant reduction in the overall Tat levels from D0 to D4 and minute reduction at the subsequent time-points implied the progressive cessation of the LTR-Tat feedback circuit. Importantly, positive Tat signal could be detected in both the compartments even on D16 post-sorting while a parallel flow analysis of GFP indicated a near-complete silencing of LTR by D8 (Figure 4.5). There was a significant loss of the Tat signal (~4.2 fold) in the extra-nuclear compartment, particularly from the membrane while the drop in the intra-nuclear Tat was relatively milder (~3.1 fold). If the intra-nuclear Tat is predominantly the transactivating form, then how is it detected above the threshold in the LTR-OFF cells on days 8, 12 and 16? Secondly, what makes this form of Tat comparatively more stable than the extra-nuclear form? These observations reinforce the speculation that Tat may undergo differential PTMs in the active and latent cells and the latent form of Tat is possibly more stable compared to the transactivating form. Intracellular compartmentalization of Tat to the plasma membrane, cytoplasm, nucleoplasm and nucleolus has been well documented (Ranki A et al., 1994; Marasco WA et al., 1994). Tat translocation to the nucleus is necessary for transactivation and the basic domain of Tat has been shown to be both essential and sufficient for nuclear translocation (Siomi H et al., 1990). Here, we attempted to examine the preferential compartmentalization of Tat during the process of latency establishment. The nature of the PTMs in the active and latent forms of Tat and the mechanism(s) for their inter-conversion must be deciphered further.

Tat interacts with multiple cytosolic partners such as Tubulin (Chen D et al., 2002, Battaglia PA et al., 2005), protein phosphatase-1 or PP1 (Ammosova T et al., 2005) and IκB-α (Puca A et al., 2007). The physical interaction of Tat with IκB-α is an essential mechanism for the nuclear export and sequestration of Tat into the cytoplasm thereby preventing Tat mediated transactivation and also aiding in the repression of viral genes (Puca A et al., 2007). Having detected Tat in the latent cells, we were particularly interested to identify differential interacting partners of Tat in the active and latent cells.

Since we already observed reciprocal binding for NF-κB and NFAT factors at the active and latent LTRs (Figure 4.3), we first examined the NF-κB factors as possible partners of Tat. The confocal signal for Tat in the latent cells was only marginally above the background (D16, Figure 4.5). To amplify the latent form of Tat, we resorted to the highly sensitive PLA. A phenotypic stage where the GFP^{+ve} and GFP^{-ve} cells were equally represented was chosen for all the PLA experiments in HEK293 cells stably infected with the 4-κB strain of the LdGIT panel. Initially we tried to capture Tat alone in the opposing phenotypes using two different Tat primary antibodies (Figure 4.8). Consistent with our previous data, we could detect Tat signals (PLA dots) both in the active and latent cells and importantly at comparable levels. Next, association between Tat and p65 was noted in both the fractions but with a ~2 fold higher incidence in the active compared to the latent cells (Figure 4.10). Using in vitro, transient transfection and single-round infection strategies, Fiume G et al. demonstrated that Tat physically interacts with p65 and the interaction enhances its DNA binding ability and transcriptional efficiency (Fiume G et al., 2012). Our PLA data validate these reports.

5.8 A hypothetical model for the Tat-mediated negative feedback mechanism to regulate HIV-1C latency.

We propose a novel model for the transcriptional repression of HIV-1 through a Tat-negative feedback mechanism previously unexplored. The attenuation of the Tat-positive feedback has been proposed to cause the LTR silencing, triggered by an extracellular cue (deterministic model) or limiting Tat levels probabilistically (stochastic model) (Weinberger LS et al., 2005; Weinberger LS and Shenk T., 2006; Weinberger LS et al., 2008; Burnett JC et al., 2009). In either case, Tat concentration gradually falls below a threshold for self-renewal or successful transcriptional elongation. Our data allude to a concentration-dependent inter-conversion of the active form of Tat to a repressive form, the latter competing with the former strengthening a negative-feedback circuit leading to the rapid silencing of the promoter. This hypothesis can explain why the stronger promoters (3- and 4-κB LTRs) can cause a faster silencing of viral transcription due to the higher accumulation of Tat, and a greater probability of the concentration-dependent Tat modification from the active to repressive form (data from the ATF and TTF models). We do not know if the phenomenon of transcriptional strength influencing latency is

applicable to the other genetic subtypes of HIV-1. The duplication of NF-κB binding motifs is a phenomenon unique to HIV-1C not manifested by any other genetic subtype of HIV-1 (Bachu M et al., 2012a; Boullosa J et al., 2014).

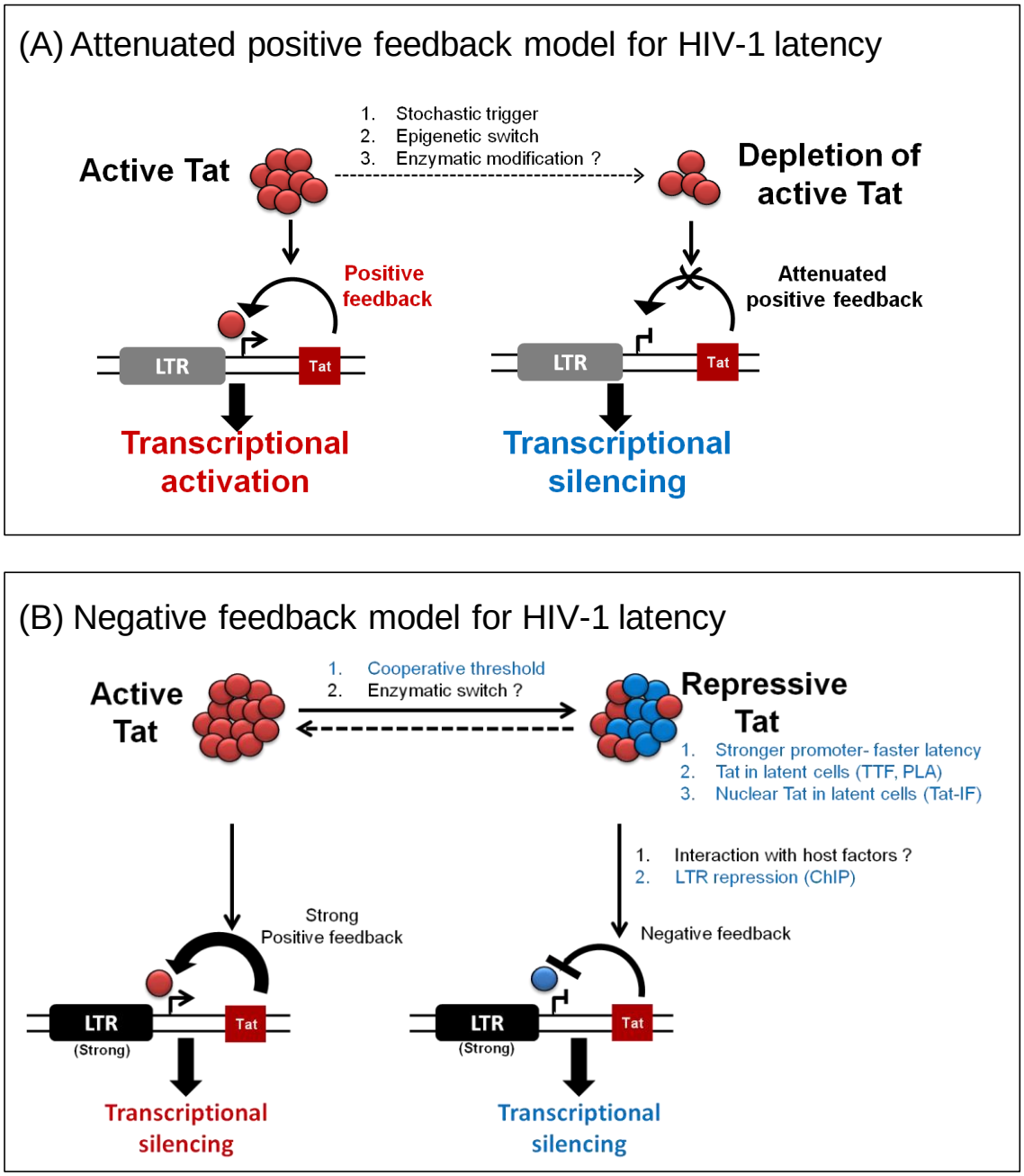


Figure 5.1: A hypothetical model of Tat negative-feedback circuit for HIV-1 latency. (A) The prevailing mode of transcriptional repression in HIV-1 (Weinberger LS et al., 2005; Weinberger LS and Shenk T., 2006; Weinberger LS et al., 2008; Burnett JC et al., 2009). The positive Tat-feedback is attenuated by a cue from an external stimulus or an intrinsic viral program. Transcriptional silence is due to Tat falling below a suboptimal threshold. **(B)** Tat PTM causing HIV-1 repression (the present work). Tat switches between active and repressive forms depending on the PTM. The repressive form of Tat initiates a negative-feedback circuit to silence the LTR. Importantly, our data provides assertive evidence for the positive cooperativity of Tat especially encoded by for stronger LTRs. Additionally, we identified the presence of Tat in the latent cells using flow cytometry (TTF model), PLA (ATF model), and, importantly, recruited to the latent LTR in ChIP (ATF model).

Our work raises several important questions pertaining to HIV-1C latency which were beyond the scope of the present study. We have demonstrated for the first time that transcriptional strength is an important parameter to drive HIV-1 latency. Is there still a window to acquire additional copies of NF-κB motifs in the C-LTR? Of note, unpublished data from our laboratory (Bhange D et al) demonstrate a recent trend of the substitution of the 4-κB strains of HIV-1C in India with a new variant 3-κB viral strain. Secondly, if transcriptional strength was the only parameter to initiate a negative-feedback circuit, why did HIV-1C acquire genetically variant NF-κB motifs? Are the variant motifs responsive to differential signaling pathways? What are the different PTMs of Tat and how these Tat forms regulate viral transcription and latency? Is the Tat negative-feedback circuit unique to HIV-1C alone or is the circuit common to all the subtypes? Work is in progress in our laboratory to address some of these questions. Answers to these questions will shed light on the mechanism of HIV-1 latency and likely to help design novel therapeutic strategies to purge HIV infection.

5.9 Bibliography

Ammosova, T., Washington, K., Debebe, Z., Brady, J. and Nekhai, S., 2005. Dephosphorylation of CDK9 by protein phosphatase 2A and protein phosphatase-1 in Tat-activated HIV-1 transcription. *Retrovirology*, *2*(1), p.47.

Arkin, A., Ross, J. and McAdams, H.H., 1998. Stochastic kinetic analysis of developmental pathway bifurcation in phage λ-infected Escherichia coli cells. *Genetics*, *149*(4), pp.1633-1648.

Bachu, M., Mukthey, A.B., Murali, R.V., Cheedarla, N., Mahadevan, A., Shankar, S.K., Satish, K.S., Kundu, T.K. and Ranga, U., 2012a. Sequence insertions in the HIV type 1 subtype c viral promoter predominantly generate an additional NF-κB binding site. *AIDS research and human retroviruses*, *28*(10), pp.1362-1368.

Bachu, M., Yalla, S., Asokan, M., Verma, A., Neogi, U., Sharma, S., Murali, R.V., Mukthey, A.B., Bhatt, R., Chatterjee, S. and Rajan, R.E., 2012b. Multiple NF-κB sites in HIV-1 subtype C long terminal repeat confer superior magnitude of transcription and thereby the enhanced viral predominance. *Journal of Biological Chemistry*, *287*(53), pp.44714-44735.

Battaglia, P.A., Ponti, D., Naim, V., Venanzi, S., Psaila, R. and Gigliani, F., 2005. The HIV-Tat protein induces chromosome number aberrations by affecting mitosis. *Cytoskeleton*, *61*(3), pp.129-136.

Bates, D.L., Barthel, K.K., Wu, Y., Kalhor, R., Stroud, J.C., Giffin, M.J. and Chen, L., 2008. Crystal structure of NFAT bound to the HIV-1 LTR tandem κB enhancer element. *Structure*, *16*(5), pp.684-694.

Bieniasz, P.D., Grdina, T.A., Bogerd, H.P. and Cullen, B.R., 1999. Analysis of the effect of natural sequence variation in Tat and in cyclin T on the formation and RNA binding properties of Tat-cyclin T complexes. *Journal of virology*, *73*(7), pp.5777-5786.

Biggar, S.R. and Crabtree, G.R., 2001. Cell signaling can direct either binary or graded transcriptional responses. *The EMBO journal*, *20*(12), pp.3167-3176.

Boullosa, J., Bachu, M., Bila, D., Ranga, U., Süffert, T., Sasazawa, T. and Tanuri, A., 2014. Genetic diversity in HIV-1 subtype C LTR from Brazil and Mozambique generates new transcription factor-binding sites. *Viruses*, *6*(6), pp.2495-2504.

Burnett, J.C., Miller-Jensen, K., Shah, P.S., Arkin, A.P. and Schaffer, D.V., 2009. Control of stochastic gene expression by host factors at the HIV promoter. *PLoS pathogens*, *5*(1), p.e1000260.

Cai, W.E.I.Z.H.O.N.G., Astor, T.L., Liptak, L.M., Cho, C., Coen, D.M. and Schaffer, P.A., 1993. The herpes simplex virus type 1 regulatory protein ICP0 enhances virus replication during acute infection and reactivation from latency. *Journal of virology*, *67*(12), pp.7501-7512.

Chen, D., Wang, M., Zhou, S. and Zhou, Q., 2002. HIV-1 Tat targets microtubules to induce apoptosis, a process promoted by the pro-apoptotic Bcl-2 relative Bim. *The EMBO journal*, *21*(24), pp.6801-6810.

Cherrington, J.M. and Mocarski, E.S., 1989. Human cytomegalovirus ie1 transactivates the alpha promoter-enhancer via an 18-base-pair repeat element. *Journal of virology*, *63*(3), pp.1435-1440.

Dandekar, D.H., Ganesh, K.N. and Mitra, D., 2004. HIV-1 Tat directly binds to NFκB enhancer sequence: role in viral and cellular gene expression. *Nucleic acids research*, *32*(4), pp.1270-1278.

Dhamija, N., Choudhary, D., Ladha, J.S., Pillai, B. and Mitra, D., 2015. Tat predominantly associates with host promoter elements in HIV-1-infected T-cells–regulatory basis of transcriptional repression of c-Rel. *The FEBS journal*, *282*(3), pp.595-610.

Dodd, I.B., Perkins, A.J., Tsemitsidis, D. and Egan, J.B., 2001. Octamerization of λ CI repressor is needed for effective repression of P RM and efficient switching from lysogeny. *Genes & development*, *15*(22), pp.3013-3022.

Feinberg, M.B., Baltimore, D. and Frankel, A.D., 1991. The role of Tat in the human immunodeficiency virus life cycle indicates a primary effect on transcriptional elongation. *Proceedings of the National Academy of Sciences*, *88*(9), pp.4045-4049.

Fiume, G., Vecchio, E., De Laurentiis, A., Trimboli, F., Palmieri, C., Pisano, A., Falcone, C., Pontoriero, M., Rossi, A., Scialdone, A. and FasanellaMasci, F., 2011. Human immunodeficiency virus-1 Tat activates NF-κB via physical interaction with IκB-α and p65. *Nucleic acids research*, *40*(8), pp.3548-3562.

Fujinaga, K., Cujec, T.P., Peng, J., Garriga, J., Price, D.H., Graña, X. and Peterlin, B.M., 1998. The ability of positive transcription elongation factor B to transactivate human immunodeficiency virus transcription depends on a functional kinase domain, cyclin T1, and Tat. *Journal of Virology*, *72*(9), pp.7154-7159.

Galhardo, R.S., Hastings, P.J. and Rosenberg, S.M., 2007. Mutation as a stress response and the regulation of evolvability. *Critical reviews in biochemistry and molecular biology*, *42*(5), pp.399-435.

Giffin, M.J., Stroud, J.C., Bates, D.L., von Koenig, K.D., Hardin, J. and Chen, L., 2003. Structure of NFAT1 bound as a dimer to the HIV-1 LTR κB element. *Nature Structural and Molecular Biology*, *10*(10), p.800.

Gottesman, M., 1999. Bacteriophage λ: the untold story. *Journal of molecular biology*, *293*(2), 177-180.

Jacob, F. and Monod, J., 1961, January. On the regulation of gene activity. In *Cold Spring Harbor Symposia on Quantitative Biology* (Vol. 26, pp. 193-211). Cold Spring Harbor Laboratory Press.

Lieb, M., 1953. The establishment of lysogenicity in Escherichia coli. *Journal of bacteriology*, *65*(6), p.642.

Mahmoudi, T., 2012. The BAF complex and HIV latency. *Transcription*, *3*(4), pp.171-176.

Malone, C.L., Vesole, D.H. and Stinski, M.F., 1990. Transactivation of a human cytomegalovirus early promoter by gene products from the immediate-early gene IE2 and augmentation by IE1: mutational analysis of the viral proteins. *Journal of virology*, *64*(4), pp.1498-1506.

Marasco, W.A., Szilvay, A.M., Kalland, K.H., Helland, D.G., Reyes, H.M. and Walter, R.J., 1994. Spatial association of HIV-1tat protein and the nucleolar transport protein B23 in stably transfected Jurkat T-cells. *Archives of virology*, *139*(1-2), pp.133-154.

Ott, M., Dorr, A., Hetzer-Egger, C., Kaehlcke, K., Schnolzer, M., Henklein, P., Cole, P., Zhou, M.M. and Verdin, E., 2004, June. Tat acetylation: a regulatory switch between early and late phases in HIV transcription elongation. In *Reversible Protein Acetylation: Novartis Foundation Symposium 259* (pp. 182-196). John Wiley & Sons, Ltd.

Pearson, R., Kim, Y.K., Hokello, J., Lassen, K., Friedman, J., Tyagi, M. and 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. *Journal of virology*, *82*(24), pp.12291-12303.

Pessler, F. and Cron, R.Q., 2004. Reciprocal regulation of the nuclear factor of activated T cells and HIV-1. *Genes and immunity*, *5*(3), p.158.

Puca, A., Fiume, G., Palmieri, C., Trimboli, F., Olimpico, F., Scala, G. and Quinto, I., 2007. IκB-α represses the transcriptional activity of the HIV-1 Tat transactivator by promoting its nuclear export. *Journal of Biological Chemistry*, *282*(51), pp.37146-37157.

Ptashne, M., 2005. Regulation of transcription: from lambda to eukaryotes. *Trends in biochemical sciences*, *30*(6), 275-279.

Ragoczy, T., Heston, L. and Miller, G., 1998. The Epstein-Barr virus Rta protein activates lytic cycle genes and can disrupt latency in B lymphocytes. *Journal of virology*, *72*(10), pp.7978-7984.

Ragoczy, T. and Miller, G., 2001. Autostimulation of the Epstein-Barr virus BRLF1 promoter is mediated through consensus Sp1 and Sp3 binding sites. *Journal of virology*, *75*(11), pp.5240-5251.

Ranki, A., Lagerstedt, A., Ovod, V., Aavik, E. and Krohn, K.J.E., 1994. Expression kinetics and subcellular localization of HIV-1 regulatory proteins Nef, Tat and Rev in acutely and chronically infected lymphoid cell lines. *Archives of virology*, *139*(3-4), pp.365-378.

Razooky, B.S. and Weinberger, L.S., 2011. Mapping the architecture of the HIV-1 Tat circuit: A decision-making circuit that lacks bistability and exploits stochastic noise. *Methods*, *53*(1), pp.68-77.

Sarisky, R.T., Gao, Z., Lieberman, P.M., Fixman, E.D., Hayward, G.S. and Hayward, S.D., 1996. A replication function associated with the activation domain of the Epstein-Barr virus Ztatransactivator. *Journal of virology*, *70*(12), pp.8340-8347.

Siomi, H., Shida, H., Maki, M. and Hatanaka, M., 1990. Effects of a highly basic region of human immunodeficiency virus Tat protein on nucleolar localization. *Journal of virology*, *64*(4), pp.1803-1807.

Southgate, C.D. and Green, M.R., 1991. The HIV-1 Tat protein activates transcription from an upstream DNA-binding site: implications for Tat function. *Genes & Development*, *5*(12b), pp.2496-2507.

Stroud, J.C., Oltman, A., Han, A., Bates, D.L. and Chen, L., 2009. Structural basis of HIV-1 activation by NF-κB—A higher-order complex of p50: Rela bound to the HIV-1 LTR. *Journal of molecular biology*, *393*(1), pp.98-112.

Suel, G.M., Garcia-Ojalvo, J., Liberman, L.M. and Elowitz, M.B., 2006. An excitable gene regulatory circuit induces transient cellular differentiation. *Nature*, *440*(7083), pp.545-551.

Wei, P., Garber, M.E., Fang, S.M., Fischer, W.H. and Jones, K.A., 1998. A novel CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its high-affinity, loop-specific binding to TAR RNA. *Cell*, *92*(4), pp.451-462.

Weinberger, L.S., Burnett, J.C., Toettcher, J.E., Arkin, A.P. and Schaffer, D.V., 2005. Stochastic gene expression in a lentiviral positive-feedback loop: HIV-1 Tat fluctuations drive phenotypic diversity. *Cell*, *122*(2), pp.169-182.

Weinberger, L.S. and Shenk, T., 2006. An HIV feedback resistor: auto-regulatory circuit deactivator and noise buffer. *PLoS biology*, *5*(1), p.e9.

Weinberger, L.S., Dar, R.D. and Simpson, M.L., 2008. Transient-mediated fate determination in a transcriptional circuit of HIV. *Nature genetics*, *40*(4), p.466.

Williams, S.A., Chen, L.F., Kwon, H., Ruiz-Jarabo, C.M., Verdin, E. and Greene, W.C., 2006. NF-κB p50 promotes HIV latency through HDAC recruitment and repression of transcriptional initiation. *The EMBO journal*, *25*(1), pp.139-149.

Appendix-I

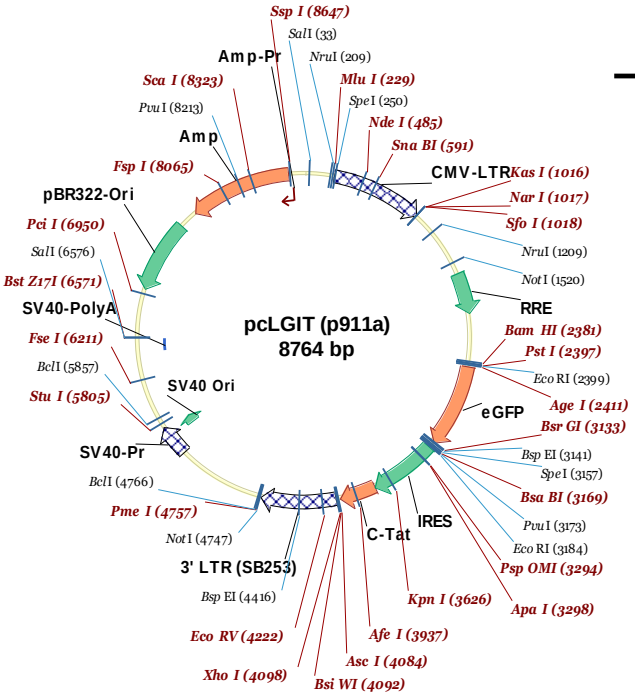
(Vector maps)

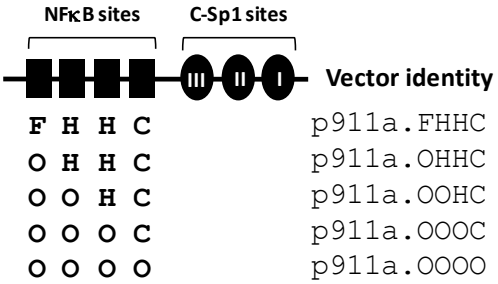
Isogenic NF-κB variants of pcLGIT vector (p911a)

Plasmid No. - p911 and variants
Lab - Prof. Udaykumar Ranga, HIV-AIDS Lab, MBGU
Backbone origin - Sutanuka Chakraborty
Construction date - Jan 2014

Key words
Pseudotyped viral vectors, reporter mini-HIV, EGFP, C-Tat, NF-κB, latency

Plasmid diagnosis
Enzyme : EagI; Fragments : 4018, 3227, 1519 bp





Feature elements	Map sites
5' LTR	236-1012
RRE	1615-1966
eGFP	2423-3139
IRES	3191-3765
C-Tat	3766-4071
3' LTR	4108-4744
pBR322-Ori	6995-7614
Amp	7769-8629
Amp Pr	8671-8699

Notes

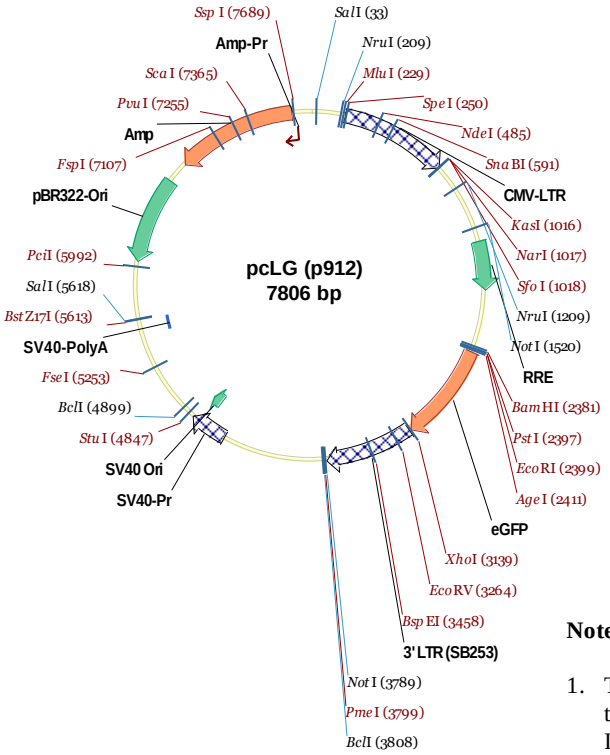
1. The present set of pseudotyped HIV-1 vectors expresses a reporter mini-virus with EGFP and C-Tat under the control of subtype C isogenic LTR variants. Subtype C-LTR from the HIV-1 molecular clone-pIndie C-1 (Gen bank accession number AB023804) was used as the template to first generate the variant 'FHHC' by overlap PCR, where, the F-κB sequence was obtained from the LTR of the HIV-1 subtype C molecular clone BL43 (GenBank accession no. FJ765005.1).
2. p911a series of vectors constitute isogenic 3'-LTR variants with NF-κB copy number variation in the 3'LTR. Variant LTRs were generated by inactivating NF-κB sites in 'FHHC' sequentially through point mutations. The resulting LTRs were cloned between XhoI and PmeI sites present on the outer primers N1990 FP and N1991 RP respectively.
3. An analogous series, p911b was generated by replacing the EGFP with the shorter half-life version GFPd2.
4. The NF-κB copy-number variant vectors of both the p911a and p911b series was used to study HIV-1 latency in the **autonomous Tat feedback (ATF)** model.

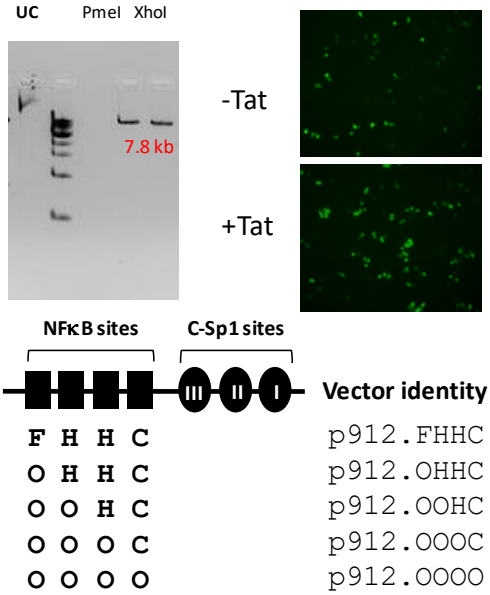
Isogenic, NF-κB variants of pcLG vector (p912)

Plasmid No. - p912 and variants
Lab - Prof. Udaykumar Ranga, HIV-AIDS Lab, MBGU
Backbone origin - Sutanuka Chakraborty
Construction date - July 2014

Key words
Pseudotyped viral vectors, reporter mini-HIV, EGFP, C-Tat, NF-κB, latency

Plasmid diagnosis and GFP expression (HEK 293T)
Enzyme : PmeI- 7806 bp Linear
XhoI – 7806 bp Linear





Feature elements	Map sites
5' LTR	236-1012
RRE	1615-1966
eGFP	2423-3139
3' LTR	4108-4744
pBR322-Ori	6197-6816
AmpR	6971-7831
Amp Pr	7831-7901

Notes

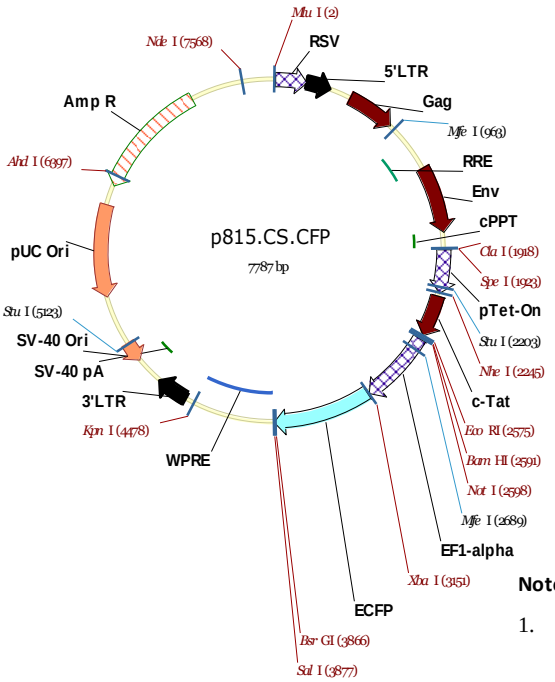
1. The p912 series of pseudotyped, HIV-1 mini virus is a truncated version of p911a series (pcLGIT backbone). It consists of a panel that reports EGFP under the viral promoter. The pcLG backbone was created by removing IRES and Tat from the pcLGIT backbone.
2. p911a.OHHC (pcLGIT-OHHC) was digested with the two enzymes with compatible 5' cohesive ends- BsrGI (upstream of IRES) and BsWI (downstream of Tat) and religated to generate the parent vector pcLG-OHHC. Importantly, this engineering has deleted the last amino acid Lys_{238} from the original EGFP coding sequence without any effect in the GFP expression.
3. pcLG-OHHC vector was used as the parent for cloning the remaining NF-κB variant LTRs between PmeI and XhoI sites at the 3'end. The p912 vector series was used to study HIV-1 latency in the **disjoint Tat feedback (DTF) model**.

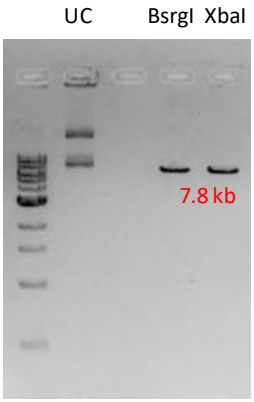
pCMV-Tet-ON-Tat-EF1α–ECFP vector backbone (p815.CS.CFP)

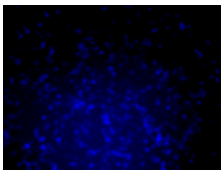
Plasmid No. - p815.CS.CFP
Lab - Prof. Udaykumar Ranga, HIV-AIDS Lab, MBGU
Backbone origin - Sutanuka Das
Construction date - July 2014

Key words
Pseudotyped viral vectors, reporter mini-HIV, EGFP, C-Tat, NF-κB, latency

Plasmid diagnosis and CFP expression (HEK 293T)
Enzyme : PmeI- 7806 bp Linear; XhoI – 7806 bp Linear







Feature elements	Map sites
RSV	7-234
5' LTR	235-414
Gag	567-919
RRE	1076-1308
Env	1309-1797
Tet-ON pro	1928-2243
c-Tat	2268-2593
EF1-α	2604-3149
ECFP	3159-3875
3' LTR	4544-4777
pUC Ori	5505-6178
AmpR	6323-7183

Notes

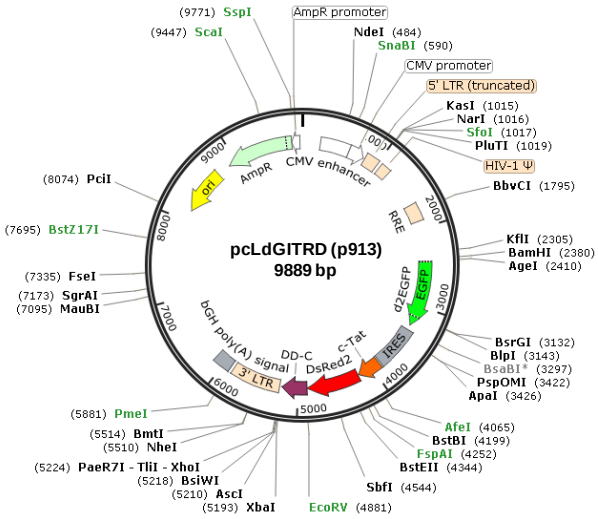
1. This plasmid is derived from p815.CS by replacing cop GFP with ECFP.
2. The EF1α and ECFP sequences were individually amplified from p815.CS and the Addgene vector pECFP-C1 (cat no. 6076-1) respectively using the primer pairs –N2362 FP (BamHI) + N2363 RP for EFI-a and N2364 FP and N2365 RP (SalI) for ECFP sequences. The resulting amplicons were used as templates for the overlap PCR product EF1-a-ECFP cassette using the external primers N2362 and N2363. The EF1-a- ECFP cassette was cloned into the p815.CS backbone by replacing the EF1α-cop GFP cassette between the enzyme sites BamHI and SalI. In the cloning process a unique enzyme site XbaI was introduced between EF1α and ECFP sequences making it possible in future to clone any gene of interest under EF1α promoter by replacing ECFP between XbaI and SalI.
3. This vector was constructed to generate a second round stable cell line (Parent-II) in the Jurkat-rtTA3 (Parent-I) using ECFP as a marker for sorting the positive cells. The primary objective is to study viral latency associated with copy-number variation of NF-κB elements in a disjoint LTR-Tat feedback context within an inducible Tet-On-Tat expression system (The disjoint Tat-feedback or DTF model)

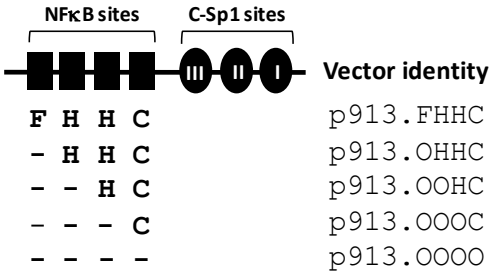
Isogenic, NF-κB variants of pcLdGITRD vector (p913)

Plasmid No. - p913 and variants
Lab - Prof. Udaykumar Ranga, HIV-AIDS Lab, MBGU
Backbone origin - pcLdGIT
Constructed by - Sutanuka Chalraborty
Construction date - Dec 2016

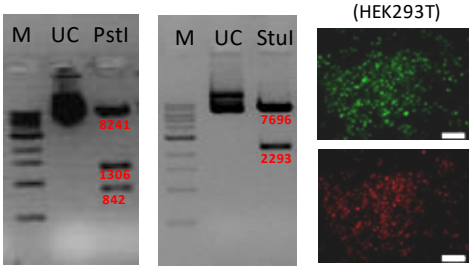
Key words
Pseudotyped viral vectors, reporter mini-HIV, GFPd2, DsRED2, C-Tat, NF-κB, FKBP, latency

Plasmid diagnosis
PstI- 8241, 1306, 842 bps; StuI- 7696, 2293 bps





48 h post transfection (HEK293T)



Notes

1. p913 series is a set of pseudotyped mini HIV-1 constructs containing NF-κB copy number variation at the 3'LTR (pIndie C-1; **Accession No**-AB023804).
2. The HIV-1 LTR in the dual-reporter virus expresses a) GFPd2 and b) a fusion protein 'cTat:DsRED2-RFP:FKBP-DD' linked by an IRES element.
3. The vector backbone used for p913 variants originated from pcLdGIT (p911b series). The fusion protein cassette 'cTat:DsRED2-RFP:FKBP-DD' was generated by overlap PCR followed by three-way ligation in three steps.

   Step 1- Three independent PCRs were performed to amplify the fragments as follows:-

Fragment	Primer pair	Template	Amplicon size
GFPd2-IRES-Tat	N2720 + N2724	pcLdGIT (p911b)	1779 bp
DsRED2	N2723 + N2726	pEF1α-DsRED2	629 bp
FKBP-DD	N2725 + N2722	pBMN-FKBP (Addgene; Catalog no. 31766)	353 bp

   Step 2- Overlap PCR was performed to obatain 'DsRED2-RFP:FKBP-DD' fragment
   Step 3- Three way ligation of 'GFPd2:IRES:cTat, 'DsRED2-RFP:FKBP-DD' and pcLdGIT vector backbone.
4. The isogenic NF-κB-variant 3' LTRs were subcloned from the p911b (pcLdGIT) series between XhoI and PmeI sites in the pcLdGITRD backbone.
5. The present set of mini reporter virus was used in the **tunable Tat feedback model (TTF)** to study HIV-1C latency by varying the Tat-feedback strength in a fixed LTR context. The Shield1 mediated 'tunable proteolytic degradation' was used to maipulate the stability of Tat and hence the Tat-feedback strength. Importantly, only two LTR-variants, OHHC (3-κB) and OOOC (1-κB) representing a strong and a weak promoter respectively were used in the study.

Publications

1. Mamata Mishra, Rebu K. Varghese, Anjali Verma, **Sutanuka Das**, Renato Santana Aguiar, Amilcar Tanuri, Anita Mahadevan, Susarla K. Shankar, Parthasarathy Satishchandra, and Udaykumar Ranga. "Genetic diversity and proviral DNA load in different neural compartments of HIV-1 subtype C infection." *Journal of neurovirology* 21, no. 4 (2015): 399-414.

2. **Chakraborty, S** et al., The subtype C of HIV-1 exploits multiple NF-κB sites in the enhancer to promote enhanced transcription as well as drive rapid viral latency (Manuscript under preparation).