AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Virology®

VIRUS-CELL INTERACTIONS

A Stronger Transcription Regulatory Circuit of HIV-1C Drives the Rapid Establishment of Latency with Implications for the Direct Involvement of Tat

Sutanuka Chakraborty,[a]* Manisha Kabi,[a]* Udaykumar Ranga[a]

[a]HIV-AIDS Laboratory, Molecular Biology and Genetics Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore, India

ABSTRACT The magnitude of transcription factor binding site variation emerging in HIV-1 subtype C (HIV-1C), especially the addition of NF-κB motifs by sequence duplication, makes the examination of transcriptional silence challenging. How can HIV-1 establish and maintain latency despite having a strong long terminal repeat (LTR)? We constructed panels of subgenomic reporter viral vectors with varying copy numbers of NF-κB motifs (0 to 4 copies) and examined the profile of latency establishment in Jurkat cells. Surprisingly, we found that the stronger the viral promoter, the faster the latency establishment. Importantly, at the time of commitment to latency and subsequent points, Tat levels in the cell were not limiting. Using highly sensitive strategies, we demonstrate the presence of Tat in the latent cell, recruited to the latent LTR. Our data allude, for the first time, to Tat establishing a negative feedback loop during the late phases of viral infection, leading to the rapid silencing of the viral promoter.

IMPORTANCE Over the past 10 to 15 years, HIV-1 subtype C (HIV-1C) has been evolving rapidly toward gaining stronger transcriptional activity by sequence duplication of major transcription factor binding sites. The duplication of NF-κB motifs is unique and exclusive to HIV-1C, a property not shared with any of the other eight HIV-1 genetic families. What mechanism(s) does HIV-1C employ to establish and maintain transcriptional silence despite the presence of a strong promoter and concomitant strong, positive transcriptional feedback is the primary question that we attempted to address in the present manuscript. The role that Tat plays in latency reversal is well established. Our work with the most common HIV-1 subtype, HIV-1C, offers crucial leads toward Tat possessing a dual role in serving as both a transcriptional activator and repressor at different phases of viral infection of the cell. The leads that we offer through the present work have significant implications for HIV-1 cure research.

KEYWORDS HIV-1C, LTR, NF-κB, Tat, latency, positive feedback

Postintegration HIV-1 latency is characterized by the presence of the transcriptionally silent but replication-competent provirus within host cells, a challenge for HIV-1 eradication. A significant amount of controversy surrounds HIV-1 latency following the discovery of a latent HIV-1 reservoir in resting $CD4^+$ T cells (1, 2), whether external host cell parameters or intrinsic proviral elements are deterministic in modulating HIV-1 latency. While some models consider HIV-1 latency to be an "epiphenomenon" influenced by the activation status of the host cell and host factors (3–6), others emphasize the stochastic nature of viral gene expression noise regulated by the long terminal repeat (LTR)-Tat positive feedback circuit, although the influence of host environmental factors is not disregarded (7–11). Furthermore, several groups have

Citation Chakraborty S, Kabi M, Ranga U. 2020. A stronger transcription regulatory circuit of HIV-1C drives the rapid establishment of latency with implications for the direct involvement of Tat. J Virol 94:e00503-20. https://doi.org/10.1128/JVI.00503-20.

Editor Guido Silvestri, Emory University

Address correspondence to Udaykumar Ranga, udaykumar@jncasr.ac.in.

* Present address: Sutanuka Chakraborty, Chemical Engineering Division, CSIR-National Chemical Laboratory, Pune, India; Manisha Kabi, Genome Architecture, Gene Regulation, Stem Cells and Cancer Programme, Centre for Genomic Regulation (CRG), Barcelona, Spain.

Received 24 March 2020
Accepted 7 July 2020
Accepted manuscript posted online 15 July 2020
Published 15 September 2020

attempted to develop theoretical models to explain the Tat feedback-mediated latency decision in HIV-1 (7, 12).

Master transcriptional regulatory circuits (MTRCs) have been identified in prokaryotic and eukaryotic viruses, functioning as integral molecular switches to toggle between transcriptional on and off phases. In this context, the lysis-lysogeny decision in bacteriophage λ has been widely researched (13). Fate selection in λ phage is based on the preferential expression of two different key viral proteins, CI and Cro, from a bidirectional promoter, in a mutually exclusive fashion, and on the cooperativity of the CI repressor to establish a "bistable" circuit manifesting lysis or lysogeny (14). Transcriptional circuits in several latency-establishing eukaryotic viruses function through a rate-versus-level trade-off where the rapid upregulation of a viral protein is essential for efficient viral replication, but the same molecule is cytotoxic at saturating levels. The immediate early 2 (IE2) transactivator protein of cytomegalovirus (CMV) is a typical example of this phenomenon (15–17). The Rta and ICP0 proteins of Epstein-Barr virus (EBV) and herpes simplex virus 1 (HSV-1), respectively, exploit the phenomenon of cooperativity to control alternate replication fates (18–22).

Importantly, the transcription regulatory circuit of HIV-1 appears to differ significantly from those of the λ phage or the eukaryotic viruses mentioned above. First, there is no evidence for a repressor molecule or negative feedback loop controlling the HIV-1 circuit with Tat functioning as only an integral component of a positive feedback circuit. Second, the Tat feedback circuit seems to lack bistability such that Tat transactivates the LTR as a monomer with no self-cooperativity to form multimers ($H = 1$) (23). Tat was proposed to undergo posttranslational modifications (PTMs) at specific sites to modulate latency and to account for Tat monostability and the lack of self-cooperativity. It was proposed that the enzymatic conversion of acetylated Tat (Tat_A) to a more stable deacetylated form (Tat_D) constitutes a feedback resistor module with a predominant off state (12). Importantly, the models described above are based mostly on mathematical simulations supported by simple reaction parameters with minimal experimental validation. Of note, these studies exclusively modeled the HIV-1 subtype B (HIV-1B) system, although other genetic families of HIV-1 contain subtype-specific molecular features.

The examination of transcriptional silence is expected to be technically more challenging in HIV-1C than in the other subtypes of HIV-1, including HIV-1B, for two different reasons. First, HIV-1C contains several subtype-specific variations in nearly all types of transcription factor binding sites (TFBSs) present in the viral promoter, including those of NF-κB, Sp1, RBF-2, and other elements. Among these variations, the copy number difference of the NF-κB binding elements is the most striking one and unique to HIV-1C. While other subtypes of HIV-1 contain a single (HIV-1E) or two (all others, including HIV-1B) NF-κB motifs in the viral enhancer, HIV-1C contains three or four of these motifs (Fig. 1). Second, the additional NF-κB binding elements present in HIV-1C are genetically diverse (24). Three different kinds of NF-κB motifs, H, C, and F, may be found in the long terminal repeat of HIV-1C (C-LTR). We demonstrated previously that the progressive acquisition of the additional NF-κB motifs enhances the transcriptional strength of the viral promoter in HIV-1C and confers replication superiority over the canonical viral strains in natural infection and under experimental conditions (25). Given the positive correlation between the transcriptional strength of the viral promoter and the enhanced strength of transcriptional feedback, it is intriguing how viral latency is favored in variant viral strains containing a higher number of NF-κB motifs. In this backdrop, the present study is an attempt to examine the influence of variation in the number of NF-κB binding elements in the C-LTR. Of note, the focus of the present study is only on the copy number difference of the NF-κB binding sites, and, therefore, on the overall strength of transcription, and its influence on viral latency. The present study does not aim to examine the impact of the genetic diversity of NF-κB binding motifs on viral latency.

Using subgenomic HIV-1C reporter viruses that differed in the LTR-Tat transcriptional feedback architecture and using panels of LTR variant viral strains that varied in the copy number of NF-κB motifs, we demonstrate for the first time that the enhanced

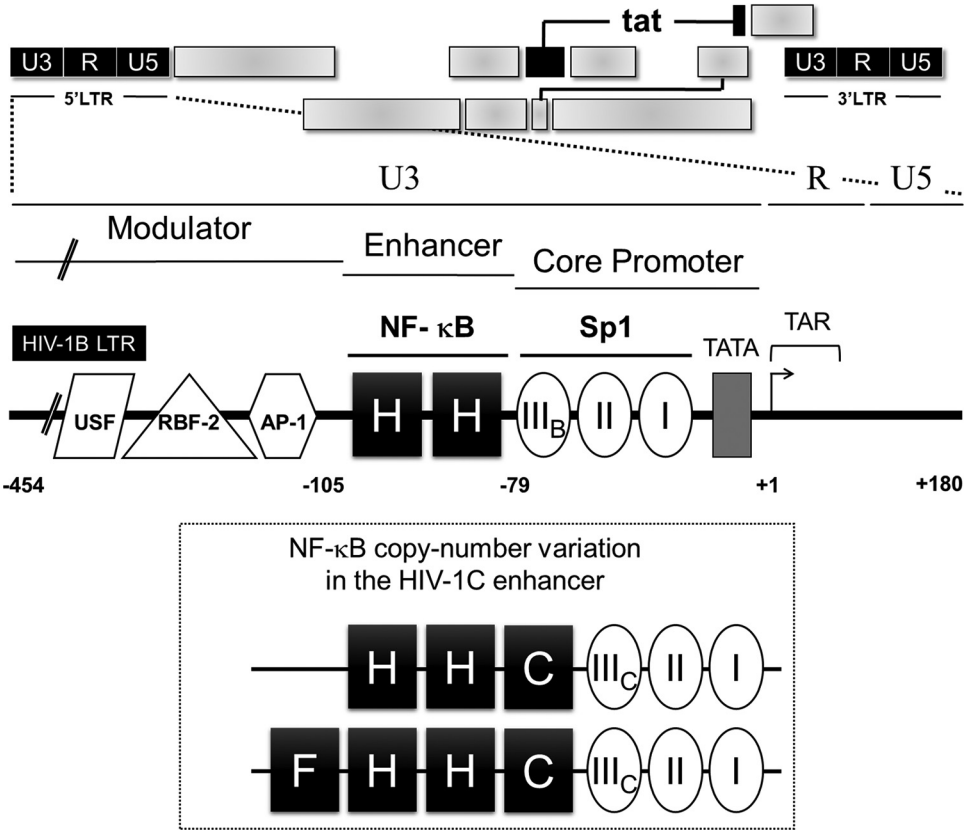

FIG 1 Schematic representation of NF-κB motif diversity in the HIV-1C LTR. The canonical HIV-1B LTR containing two identical NF-κB motifs is presented at the top. Distinct regulatory regions (U3, R, U5, the modulator, the enhancer, and the core promoter) and important transcription factor binding sites are depicted. Not only does the HIV-1C LTR contain more copies of the NF-κB motif (3 or 4 copies), but the additional motifs are also genetically variable (bottom). The three genetically distinct NF-κB motifs present in the C-LTR are denoted H (GGGACTTTCC), C (GGGGCGTTCC [differences are underlined]), and F (GGGACTTTCT). Note that the Sp1III site also contains subtype-specific variations as presented, with B and C representing the respective viral subtypes.

transcriptional strength of the LTR leads to a rapid establishment of viral latency. Furthermore, we explain the apparent paradox by demonstrating that stronger transcriptional activity of the LTR leads to a higher level of the cellular Tat protein, which, above a certain threshold, possibly establishes negative feedback on viral transcription. Importantly, using indirect immunofluorescence (IF) and a highly sensitive proximity ligation assay (PLA), for the first time, we demonstrate the presence of Tat in cells harboring an active or a latent provirus. We also show the recruitment of Tat to not only the active but also the latent proviral LTR albeit at a magnitude severalfold lower. Our data thus collectively allude to Tat playing a deterministic role in initiating transcriptional silence through negative feedback regulation.

RESULTS

The strengths of viral gene expression as well as Tat transactivation are directly proportional to the number of functional NF-κB motifs in the HIV-1C enhancer. The present study is an attempt to examine the influence of variation in the number of NF-κB binding elements in the C-LTR on viral gene expression and latency. To this end, we employed two different Jurkat T cell models, the autonomous Tat feedback (ATF) and the tunable Tat feedback (TTF) models, to examine HIV-1C latency. Subgenomic viral vectors encoding enhanced green fluorescent protein (EGFP) (or d2EGFP, a variant GFP with a shorter half-life) were pseudotyped with vesicular stomatitis virus G (VSV-G) envelope. The two experimental models differ from each other in the manner in which the LTR-Tat feedback axis is regulated. Using these two

experimental systems, we examined transcriptional activation and silencing as a function of the promoter strength (by varying the number of copies of NF-κB motifs, ranging from 0 to 4 copies, in the ATF model), Tat feedback strength (by modulating the physiological concentration of Tat in the TTF model), or both from a panel of HIV-1C LTRs.

The ATF model of HIV-1 comprises the presence of only the LTR and Tat, with all other viral factors being absent, thus retaining the natural functional association between the two major viral factors, as reported previously (7, 26). Several groups have adopted the ATF model to elucidate the mechanisms governing HIV-1 latency. In the present study, we modified the pLGIT subgenomic reporter vector (7) by replacing the Tat open reading frame (ORF) and the 3′ LTR of the parental vector, both of HIV-1B origin, with the homologs of HIV-1C to construct pcLGIT. In the pcLGIT (cLTR-EGFP-IRES [internal ribosome entry site]-cTat) vector, the expressions of EGFP and C-Tat are under the control of the C-LTR.

Given the natural propensity of HIV-1C to contain more copies of the NF-κB motif in the enhancer, 3 copies typically and up to 4 copies frequently (24), we constructed a panel of cLGIT viral strains comprising NF-κB copy number variant LTRs (p911a series [see Materials and Methods]). Using the prototype C-LTR containing four functional NF-κB binding sites (FHHC), we introduced inactivating point mutations sequentially into the enhancer to reduce the number of functional NF-κB motifs progressively from 4 copies to 0 copies (Fig. 2A). In some of the subsequent experiments involving the ATF model, we also resorted to the cLdGIT panel of viral strains, where the EGFP reporter was replaced with d2EGFP in the pcLdGIT vector backbone with the same set of NF-κB variant LTRs as for pcLGIT (p911b series [see Materials and Methods]). The viral stocks of the panel pseudotyped with VSV-G envelope were generated in HEK293T cells, and the relative infectious units (RIU) of the stocks were determined in Jurkat cells using EGFP/d2EGFP fluorescence.

First, we compared the levels of EGFP expression from the LTR variant cLGIT panel in the context of a functional, positive Tat feedback loop, where both the reporter gene and the concomitant Tat feedback strengths are expected to vary based on the autoregulatory circuit. Jurkat cells infected with each viral strain of the cLGIT panel independently at ~0.5 RIU were either activated with a combination of global T cell activators [40 ng/ml phorbol myristate acetate (PMA), 40 ng/ml tumor necrosis factor alpha (TNF-α), 200 nM trichostatin A (TSA), and 2.5 mM *N*,*N*′-hexamethylene bis(acetamide) (HMBA)] or maintained without activation, and 24 h following treatment, both EGFP fluorescence and Tat transcript levels were examined using flow cytometry and Tat reverse transcription-PCR (RT-PCR), respectively (Fig. 2B). Representative stacked histograms depicting the three conditions of treatment, uninfected Jurkat cells (black dotted histograms), infected but untreated cells (black hollow histograms), and infected and activated cells (solid gray histograms), corresponding to all five NF-κB variant strains are presented in Fig. 2C. Importantly, when the cell population in each histogram was demarcated into three categories based on the intensity of EGFP expression (EGFP negative [$EGFP^{-}$], $EGFP^{Low}$, and $EGFP^{High}$), the $EGFP^{High}$ fraction displayed the most pronounced impact of the NF-κB site copy number difference on transactivation. The percentage of the $EGFP^{High}$ fraction was directly proportional to the number of NF-κB motifs in the LTR, which was also reflected in the peak height of the $EGFP^{High}$ cluster in the stacked histogram profile.

We quantitated EGFP fluorescence in terms of the mean fluorescence intensity (MFI) as a function of the copy numbers of NF-κB motifs in the LTR and found a direct proportionality between them (Fig. 2D), although the percentages of viral infectivity were comparable (Fig. 2D, inset). The LTR containing four NF-κB motifs (FHHC) (4-κB) demonstrated the highest fluorescence intensity with (82,917.51 ± 825.7 relative fluorescence units [RFU]) and without (12,365.13 ± 179.3 RFU) activation, while the LTR in which all four NF-κB motifs had been mutated (OOOO) (0-κB) demonstrated the lowest levels of reporter expression with (22,190.38 ± 668.1 RFU) and without (6,083.36 ± 290.5 RFU) activation. The activities of the other three LTRs containing

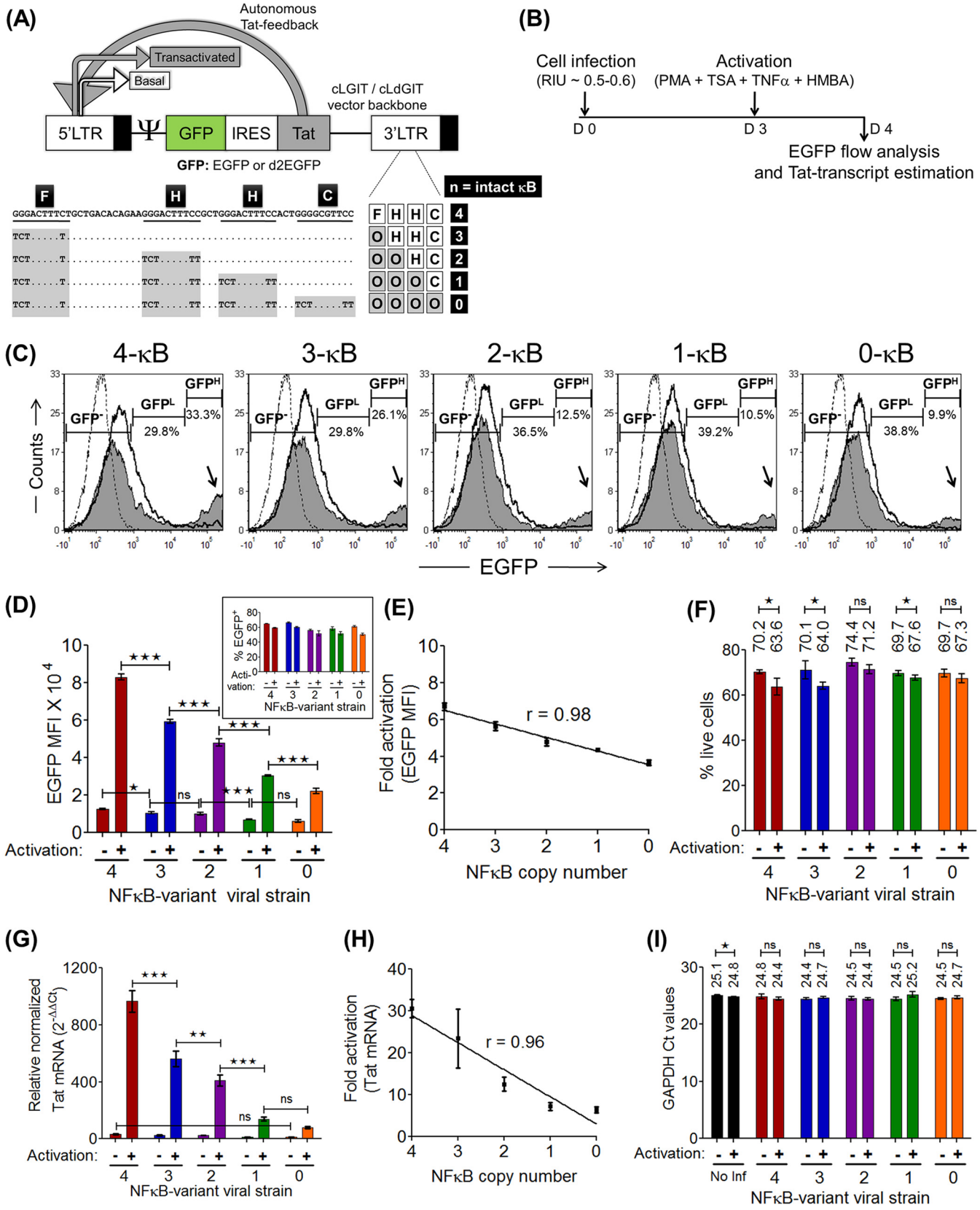

FIG 2 Viral gene expression is proportional to the number of NF-κB motifs in the C-LTR. (A) Schematic of the cLGIT/cLdGIT subgenomic viral vector panels of the autonomous Tat feedback (ATF) model. Note that both the 3′ LTR and Tat are of HIV-1C origin. The nucleotide sequences of the four NF-κB motifs and the inactivating mutations introduced into the motifs of the variant viral strains are depicted. The EGFP reporter gene (half-life, ~48 h) of the cLGIT vector backbone

(Continued on next page)

3 (OHHC) (3-κB), 2 (OOHC) (2-κB), or 1 (OOOC) (1-κB) functional NF-κB motif remained between the two extremes. The fold enhancement in EGFP expression was directly proportional to the number of functional NF-κB motifs in the LTR, with a linear correlation ($r = 0.98$) between transcriptional activity and functional NF-κB motifs in the LTR (Fig. 2E). A viability assay performed using a live/dead stain before the analysis of EGFP expression confirmed minimal cell death following activation (Fig. 2F). Similar to the EGFP MFI profile, the level of Tat transcript expression (Fig. 2G) and fold transactivation (Fig. 2H) were directly proportional to the number of NF-κB copies in the LTR ($r = 0.96$), with or without activation. The evaluation of the transcripts of glyceraldehyde-3-phosphate dehydrogenase (GAPDH), as an internal cellular reference control for real-time RT-PCR, validated the expression levels of Tat mRNA from the viral panel under diverse cell activation conditions (Fig. 2I). It is evident from the expression profile that there is a perfect correlation between the number of NF-κB motifs and the level of gene expression from the promoter. Importantly, the expression of EGFP can be used as a surrogate marker for the expression of Tat since there is a perfect correlation between the two genes coexpressed from the viral promoter. In subsequent assays, we routinely used the expression of EGFP as a measure of the transcriptional activity of the viral promoter, with frequent confirmation of Tat expression.

A stronger viral promoter establishes latency at a higher rate. A major paradox in the transcriptional regulation of HIV-1C is that a virus that must establish latency tends to acquire a stronger promoter containing more NF-κB motifs, especially when other genetic families of HIV-1 do not employ such a strategy. To understand this paradox, we used the NF-κB copy number variant strains of the ATF panel to determine the kinetics of latency establishment. Using the experimental strategy depicted in Fig. 3A, we infected Jurkat cells at a low RIU (~0.1 to 0.2) to ensure a single integration event per cell. The cells were allowed to expand before inducing them with a cocktail of global activators, and the EGFP-positive ($EGFP^+$) cells were recovered by sorting. The kinetics of the EGFP switch-off was subsequently monitored every 4 days for 16 days by flow cytometry. A representative strategy of cell gating and sorting is presented in Fig. 3B.

Using the experimental strategy described above and the cLGIT panel of NF-κB copy number variant viral strains containing 4 to 0 copies of the TFBS, we evaluated how the transcriptional strength of the HIV-1 LTR would influence the kinetics of latency establishment over 16 days. The analysis found a profound impact of the NF-κB motif copy number on the kinetics of HIV-1 latency establishment. Representative stacked histogram profiles of the LTR variant strains depicting the temporal expression pattern

FIG 2 Legend (Continued)
is replaced with its shorter-half-life variant d2EGFP (half-life, ~2 h) in the cLdGIT backbone. (B) Time schematic of gene expression analysis with the cLGIT vector panel. One million Jurkat cells were infected with ~0.5 to 0.6 RIU, independently with each LTR variant strain. After 72 h of infection, half of the infected cells were activated with a cocktail of global T cell activators (PMA, TNF-α, TSA, and HMBA), and at 24 h postactivation, EGFP and Tat transcript expression levels were estimated for both the unactivated and activated fractions using a flow cytometer and Tat-specific RT-PCR, respectively. (C) Representative EGFP histograms of the five variant LTRs. The black dotted histograms represent Jurkat cells not infected and not activated, the black hollow histograms represent cells infected but not activated, and the solid gray histograms represent cells infected and activated. The intensity ranges of $EGFP^-$, $EGFP^{Low}$, and $EGFP^{High}$ cells are indicated as GFP^-, GFP^L, and GFP^H, respectively, along with the frequency of each fraction (percent values). The black arrows point to the gradual decline in the $EGFP^{High}$ (MFI, $>10^4$ RFU) population representing the Tat-transactivated population with decreasing copies of NF-κB elements. (D) Mean EGFP MFI values from experimental quadruplicates ± standard deviations (SD). Data are representative of results from two independent experiments. Two-way analysis of variance (ANOVA) with Bonferroni posttest correction was used for the statistical evaluation (*, $P < 0.05$; ***, $P < 0.001$; ns, nonsignificant). The percent GFP^+ profile in the inset confirms the nearly equivalent infection of target cells with the LTR variant viral strains at ~0.5 to 0.6 RIU. (E) EGFP expression manifests a positive linear correlation with the NF-κB copy number as indicated by the fold enhancement in the EGFP MFI (ratios of the EGFP MFI values of the activated and uninduced fractions from each variant LTR) versus the NF-κB copy number plot. (F) Live/dead assay to compare the percentages of live cells between the activated and uninduced pairs. Shown are mean values ± SD from experimental quadruplicates; data are representative of results from two independent experiments (*, $P < 0.05$; ns, nonsignificant [two-tailed, unpaired t test]). (G) Tat expression was evaluated by RT-PCR using the $\Delta\Delta C_T$ method and GAPDH as the reference gene from total mRNA extracted from 0.2 million to 0.5 million cells of the unactivated and activated populations. Mean values of relative Tat mRNA expression from three independent experiments ± standard errors of the means (SEM) are plotted. Two-way ANOVA with Bonferroni posttest correction was used for statistical analyses. (**, $P < 0.01$; ***, $P < 0.001$; ns, nonsignificant). (H) Tat transcript expression is directly proportional to the NF-κB copy number as observed from the fold enhancement of Tat transcript levels of the activated fraction over the uninduced fraction. (I) Comparable GAPDH C_T (threshold cycle) values under the different stimulation conditions as well as across the variant viral strains (*, $P < 0.05$; ns, nonsignificant [two-tailed, unpaired t test]). No Inf, Jurkat cells not infected.

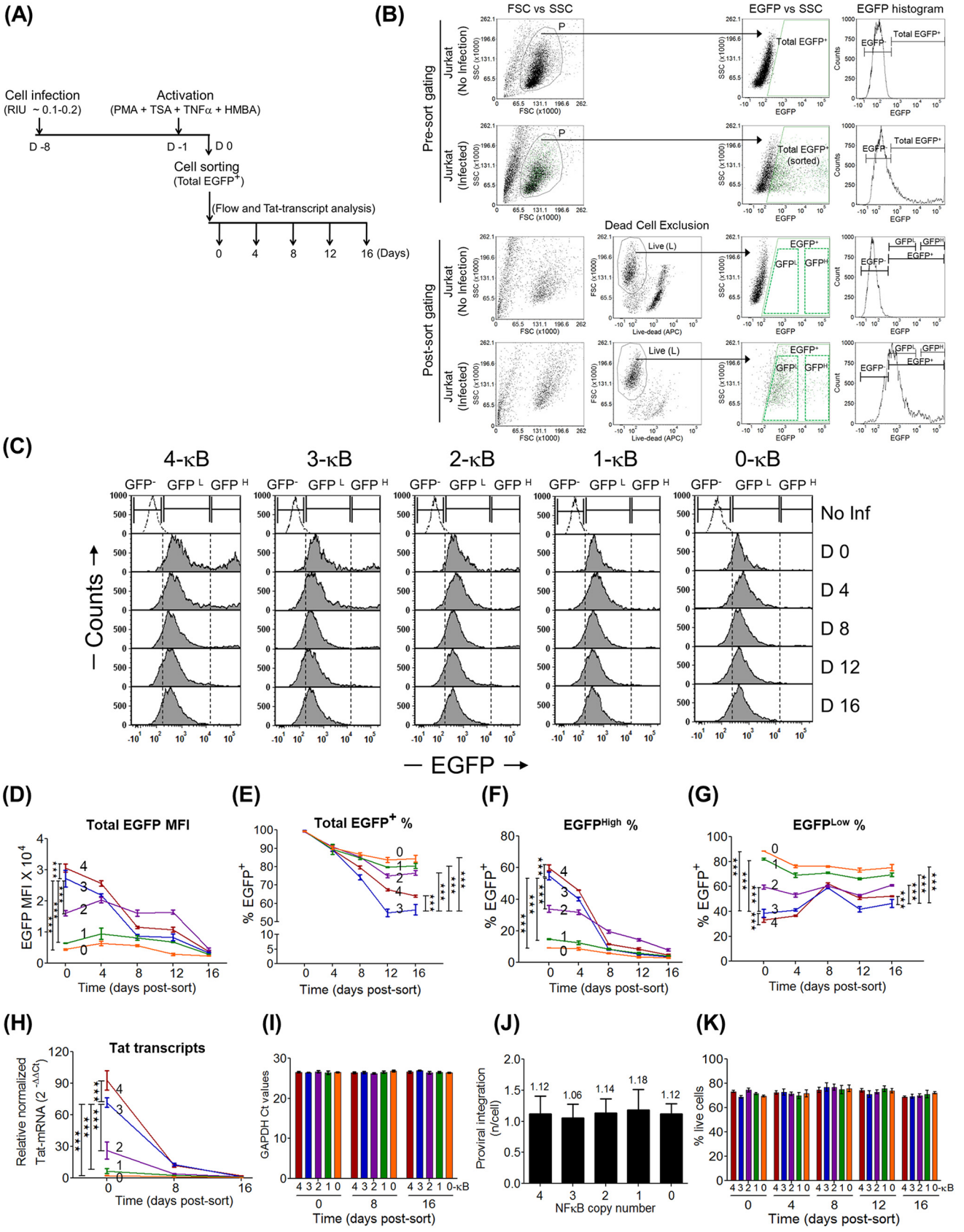

FIG 3 In the ATF model, the stronger the promoter, the faster the latency establishment. (A) Schematic representation of the experimental protocol. One million Jurkat cells were infected with individual strains of the cLGIT panel at a low infectious titer of ~0.1 to 0.2 RIU, allowed to relax for a week, and treated

(Continued on next page)

of the EGFP$^+$ sorted cells are presented in Fig. 3C. Although latency establishment was evident for all five LTRs examined, the rapidity of latency establishment unexpectedly was directly proportional to the number of NF-κB motifs in the viral enhancer (Fig. 3D and E). In other words, the stronger the transcriptional activity of the LTR, the faster the latency establishment. Based on the slope of EGFP downregulation, the LTRs could be classified into two broader groups: the two strong promoters, the 3- and 4-κB LTRs, downregulated EGFP expression at a significantly higher rate than the other three not-strong promoters, the 2-, 1-, and 0-κB LTRs. In the present study, we classify the LTRs into two groups, "strong" and "weak," based on the difference in transcriptional strength, a categorization consistent with many other properties that we analyzed subsequently, although the 2-κB LTR occasionally occupied an intermediate position (see below). For instance, the EGFP intensity values (RFU) of the 4-κB LTR were reduced approximately 8-fold from a value of 30,631.64 ± 1,278.3 on day 0 to 3,771.06 ± 245.2 on day 16, whereas the corresponding values for the weakest LTR, the 0-κB LTR, were modest and reduced by only 2-fold during the same period, from 4,455.11 ± 258.9 to 2,371.98 ± 59.3. Of note, although both the 3- and 4-κB LTRs demonstrated rapid EGFP downregulation, the 3-κB LTR established viral latency at a higher rate, and the difference between the two promoters was highly reproducible and significant. It is not clear if this difference may have implications for the relative replication fitness of the two viral strains.

The expression profile of the Tat transcripts determined using RT-PCR on days 0, 8, and 16 also correlated directly with the NF-κB copy number in the LTR, as expected (Fig. 3H), and resembled the EGFP MFI profile of the LTRs. A profound reduction in Tat expression was observed for all the viral promoters between days 0 and 8. The 4-κB LTR showed the highest level of Tat expression, 92.94 ± 5.4 at day 0, which dropped to 12.02 ± 0.8 at day 8 and subsequently to 2.3 ± 0.01 at day 16. The corresponding values for the 3-κB LTR are 71.76 ± 2.5, 12.8 ± 0.73, and 1.15 ± 0.1, respectively. Of note, comparable expression of GAPDH transcripts was observed from the viral panel at all time points of latency establishment, as quantitated using real-time PCR (Fig. 3I). Furthermore, using TaqMan quantitative PCR (qPCR), we confirmed a single integration event per cell in all five stable cell pools, thus ruling out the possibility that the difference in the integration frequency influenced the outcomes of the analyses (Fig. 3J). A live/dead exclusion assay indicated comparable percentages of live cells among the panel members, and uniform cell viability was maintained temporally throughout latency establishment (Fig. 3K). Importantly, the live/dead gating excluded dead cells before the EGFP analysis (postsort gating) (Fig. 3B), thus precluding the possibility of EGFP autofluorescence from dead cells influencing the data analysis. In summary, our data suggest that the enhanced strength of the HIV-1C LTR due to the increase in the number of NF-κB sites could play a decisive role in regulating viral latency. The positive correlation between the Tat transcript levels and the high rate of

FIG 3 Legend (Continued)
with the cocktail of global T cell activators, and 24 h later, all the EGFP$^+$ cells (harboring active provirus) (MFI, $>10^3$ RFU) were sorted. The sorted, total EGFP$^+$ cells were then maintained in culture, EGFP expression was monitored by flow cytometry every 4 days, and that of Tat transcripts was monitored on days 0, 8, and 16. (B) Gating strategies used for sorting of the total EGFP$^+$ population and for the subsequent latency establishment assay. Initially, the debris was excluded in the forward-scatter (FSC)-versus-side-scatter (SSC) scatterplot, and the total EGFP$^+$ cells from the population (P) were sorted. Importantly, a live/dead exclusion dye was used to stain the postsorted cells to include only the live, EGFP$^+$ cells for the latency kinetics analyses. The total EGFP$^+$ gate was further subgated into EGFP^{Low} and EGFP^{High} fractions for subsequent analyses (panels G and F, respectively). APC, allophycocyanin. (C) Representative, postsort, stacked histograms representing the temporal events during transcriptional silencing. The NF-κB variant viral strains demonstrate varying proportions of EGFP^{High} and EGFP^{Low} cells in the total EGFP$^+$ sort. No Inf, Jurkat cells not infected. (D to G) Kinetic curves corresponding to the total EGFP MFI and the percentages of total EGFP$^+$, EGFP^{High}, and EGFP^{Low} cells, respectively. Data are representative of results from three independent experiments. Mean values from experimental triplicates ± SD are plotted. Two-way ANOVA with Bonferroni posttest correction was used for the statistical evaluation (***, $P < 0.001$). (H) Kinetic curves of relative Tat mRNA levels of the NF-κB variant strains. (I) Absolute C_T values of the GAPDH transcripts at different time points. Data are representative of results from two independent experiments. Mean values from experimental triplicates ± SD are plotted. Two-way ANOVA with Bonferroni posttest correction was used for the statistical evaluation (***, $P < 0.001$). (J) Integration frequencies of the five LTR variant viral strains were estimated using a standard curve and regression analysis. The level of viral integration was found to be ~1.0 per cell for all five variants. Data are representative of results from two independent experiments. Mean values from experimental triplicates ± SD are plotted. (K) Live/dead analysis confirms comparable percentages of live cells between the NF-κB variants and across different time points.

EGFP switch-off by the strong viral promoters is strongly indicative of the Tat-mediated positive feedback loop playing a critical role in establishing viral latency.

The kinetics of latency establishment is predominantly a function of EGFP^{High} cells displaying a biphasic mode of transcriptional silence. At the baseline of the above-described assay, all the variant viral strains were represented by nearly 100% EGFP$^+$ cells but with a varying range of EGFP fluorescence intensities (referred to as total EGFP$^+$ cells throughout the manuscript). However, a marked difference in the mean intensities of EGFP among the LTR variants was noted at the day 0 time point postsorting (compare Fig. 3D and E). This apparent paradox could be explained by analyzing only EGFP^{High} cells but not the total EGFP$^+$ population. In the present assay, we therefore gated the cells into two additional subpopulations, EGFP^{High} (MFI, $>10^4$ RFU) and EGFP^{Low} (MFI, $\sim10^2$ to 10^4 RFU), as depicted in the postsort gating strategy in Fig. 3B and as evident in the histogram profile of each NF-κB variant strain in Fig. 3C. Kinetic curves of the percentages of EGFP^{High} and EGFP^{Low} cells were then constructed from the above-described gated subpopulations. Importantly, the reductions in the total EGFP MFI (Fig. 3D) as well as the Tat transcript levels (Fig. 3H) corresponded perfectly only with the percentage of EGFP^{High} cells (Fig. 3F) but not with the percentage of EGFP^{Low} cells (Fig. 3G). Thus, EGFP^{High} cells, and not the total EGFP$^+$ cells, are decisive in regulating viral latency. Additionally, the percent EGFP^{High} temporal curves of the strong (3- and 4-κB) versus weak (0-, 1-, and even 2-κB) LTRs were profoundly different. First, on day 0, the strong LTRs produced the highest percentage of EGFP^{High} cells compared to the weak LTRs. Second, the latency establishment of the strong LTRs appeared to have manifested in two distinct phases, a rapid reduction of EGFP expression between days 0 and 8 and a lower rate of decrease after day 8; the biphasic latency profile was either absent or not prominent for the weak LTRs. Third, the rapid fall in EGFP expression of the EGFP^{High} pool of the strong LTRs between days 0 and 8 synchronized with a significant rise in the EGFP^{Low} cell pool, peaking on day 8. These data collectively allude to the critical role that the transcriptional strength of the HIV-1 LTR plays in latency establishment. In summary, the EGFP^{High} cell pool, and not that of the EGFP^{Low} cells, plays a decisive role in the population latency kinetics of the virus.

LTR silencing in GFP^{High} cells implicates Tat feedback. Given the apparent significance of the EGFP^{High} phenotype for HIV-1 latency establishment, we investigated the phenomenon further by sorting only the EGFP^{High} cell pools for all the NF-κB variant strains that represented a population with a comparable fluorescence intensity (Fig. 4). At day 0, the EGFP MFI values were uniform among the variant viral strains of the panel, and we monitored the downregulation of green fluorescence every 4 days for 24 days (Fig. 4A). A clear distinction between the strong (4- and 3-κB) and the weak (2-, 1-, and 0-κB) LTRs was evident in the EGFP MFI profile (Fig. 4B) or when the percentage of EGFP$^+$ cells was considered (Fig. 4C), although the 2-κB LTR sometimes occupied an intermediary position; the rate of latency establishment was significantly high for strong LTRs. The biphasic mode of latency establishment, rather than a gradual and monophasic mode, was evident from the stacked histogram profiles of the sorted EGFP^{High} pool in Fig. 4F. We demonstrated above that a progressively increasing number of NF-κB sites in the LTR steadily enhances the transcriptional strength as well as the physiological concentration of Tat (Fig. 2D and G). We therefore speculate that higher cellular Tat levels, an invariable outcome of stronger positive transcriptional feedback, are necessary for the rapid silencing of the LTR, as manifested by the EGFP^{High} cells of the strong LTRs.

Of note, the process of latency establishment described above was not complete with any of the LTRs of the cLGIT panel, regardless of the transcriptional strength. The percentage of EGFP$^+$ cells reached only the halfway mark after 24 days of sorting, even for the strong LTRs that established latency at a higher rate (Fig. 4B and C). Importantly, the long half-life of the EGFP, $\sim$48 h, used in these vectors as a surrogate marker for latency did not represent the actual dynamics of LTR transcriptional activity faithfully. The cells continued to be scored as being positive for EGFP fluorescence for a

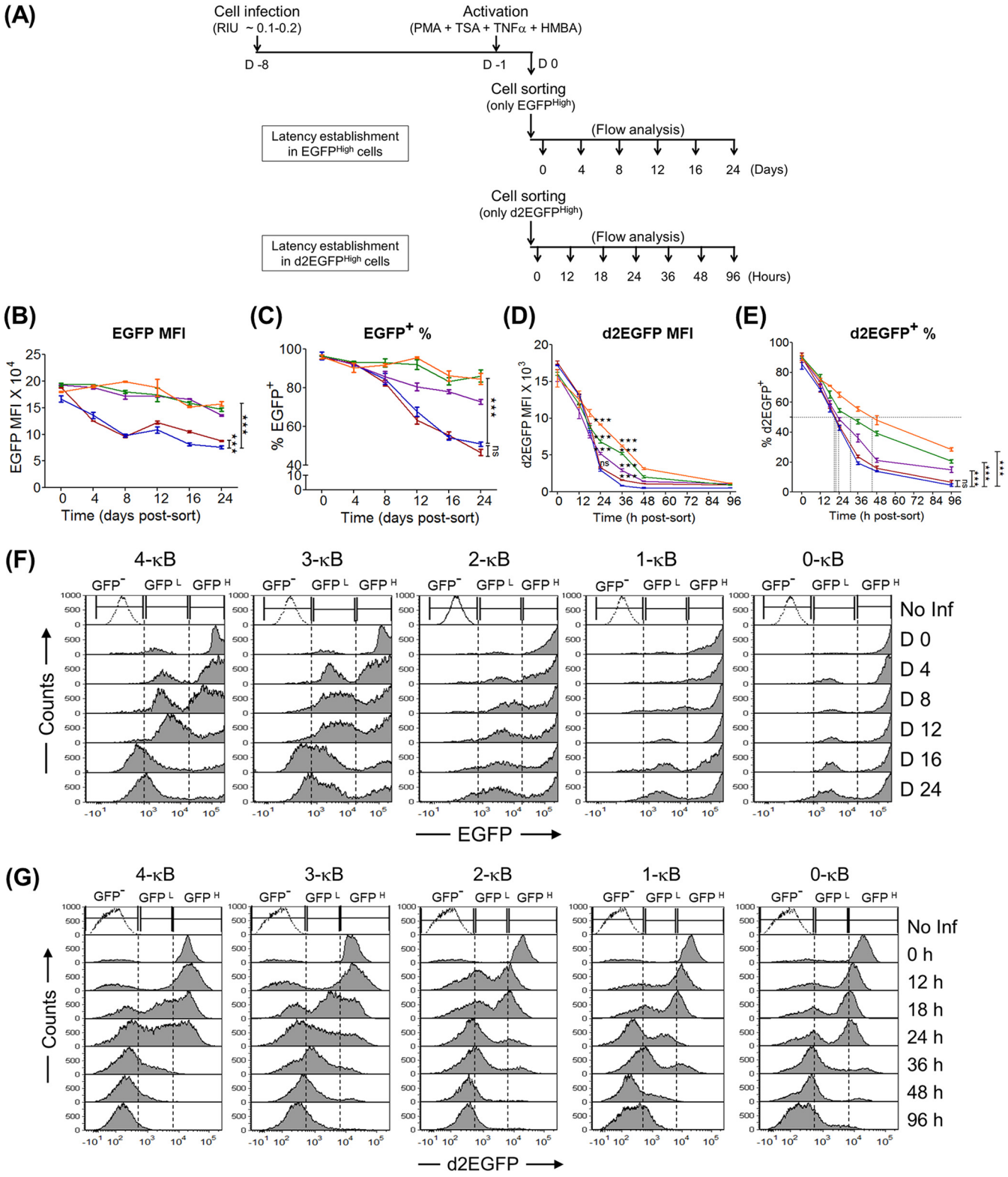

FIG 4 The binary latency trajectory of the GFP^{High} population delineates the viral promoter into strong (3- and 4-κB) and weak (2-, 1-, and 0-κB) LTRs. (A) Experimental schemes to study the kinetics of latency establishment in EGFP^{High} and d2EGFP^{High} cells. (B to E) Comparative kinetic profiles of EGFP MFIs, percentages of EGFP^{+} cells, d2EGFP MFIs, and percentages of d2EGFP^{+} cells, respectively. Mean values from experimental triplicates ± SD are shown and are representative of results from two independent experiments. Two-way ANOVA with Bonferroni posttest correction was used for the statistical evaluation (***, $P < 0.001$; ns, nonsignificant). (F) Stacked histogram profiles of EGFP^{High} cells during latency establishment. The sorted EGFP^{High} cells (GFP^{H}) (MFI, $>10^4$ RFU), comprising a homogeneous population of Tat-mediated transactivated cells, transitioned to the EGFP^{-} phenotype (GFP^{-}) through an EGFP^{Low} cluster (GFP^{L}) (MFI, ~10^2 to 10^4 RFU), representing cells with a basal level of transcription without an intermediate phenotype. No Inf, Jurkat cells not infected. (G) Stacked

(Continued on next page)

significant period even after the LTR was switched off, leading to false-positive scoring. To rectify this problem, we replaced EGFP in the reporter viral strains with d2EGFP, characterized by a significantly shorter half-life (2 versus 48 h) (27). The viral strains of the new panel (cLdGIT) are analogous to those of the previous panel.

Using the new panel, we sorted the d2EGFP^{High} cells as described above to establish the profiles of latency. Several differences in the profiles of latency were readily evident between the cLGIT and cLdGIT panels (compare Fig. 4B and D and Fig. 4C and E). Unlike the cLGIT panel, the cLdGIT variants successfully established nearly complete viral latency, and all the members of the panel demonstrated latency establishment at a higher rate; the d2EGFP MFI values were reduced to the baseline within 96 h following sorting (Fig. 4D). Although the replacement of EGFP with d2EGFP masked the differences in latency kinetics among the members of the cLdGIT panel to some extent, the overall pattern of latency establishment was consistent with that of the cLGIT counterparts. The percentage of cells downregulating d2EGFP expression was directly proportional to the number of NF-κB motifs in the viral promoter (Fig. 4E). For instance, the times required for the loss of fluorescence in half of the cells (FL_{50}s) were estimated to be 23.3, 22.1, 24.64, 32.9, and 48 h for the 4-, 3-, 2-, 1-, and 0-κB viral strains, respectively. Thus, a direct correlation between the transcriptional strength of the viral promoter and the rate of latency establishment was consistent between the cLdGIT and cLGIT panels. The biphasic mode of latency establishment was also evident in the cLdGIT model (Fig. 4G).

Collectively, our data firmly conclude that the transcriptional strength of the HIV-1 promoter is an essential regulatory parameter for viral latency. Furthermore, the latency kinetics in the Tat-transactivated population (GFP^{High} cells) of two different models, the cLGIT and cLdGIT vectors of the ATF panel, followed NF-κB site copy number-dependent transcriptional silencing.

Bimodal (on or off) latency establishment in the pools of cloned cell lines. The observation that the transcriptional strength of the LTR and the feedback loop of Tat function synergistically to silence the viral promoter was made based on cell pools. Since individual cells in a pool are heterogeneous in several biological properties, including the site of proviral integration, we examined the nature of the latency profile in multiple cloned cell lines of all five LTR variants. Jurkat cells were infected with the viral strains of the cLGIT (ATF) panel and stimulated with the global activation cocktail, single EGFP^{High} cells were sorted into individual wells of a 96-well culture plate, the sorted cells were allowed to expand for 3 to 4 weeks, and the EGFP expression profiles were assessed by flow cytometry (Fig. 5A). We recovered 16 to 25 clones from each NF-κB variant, and 16 clones from each variant were randomly selected for latency analysis. Of note, since each cell line descended from a single parental cell, all the daughter cells derived from the parental cell are expected to have a common site of integration.

Based on the EGFP expression pattern, the clones could be categorized into three distinct types (Fig. 5B). The persisters, all the daughter cells descending from a single parental cell, sustain the expression of high-intensity EGFP throughout the observation period of 28 days and even beyond, comparable to that of the original parental cell, indicative of a provirus transcribing actively in all the daughter cells. The relaxers, all the daughter cells of the EGFP^{High} parental cell, have EGFP expression switched off entirely during the period of observation. The bimodallers, the third clonal type, demonstrated a distinctive feature of the simultaneous existence of both phenotypes among the daughter cells, although all the cells in the cluster were derived from the same EGFP^{High} parental cell. One subset of the cells maintained high EGFP expression (EGFP^{High}),

FIG 4 Legend (Continued)
histogram profiles of d2EGFP^{High} cells during latency establishment identify regions of d2EGFP^{-} (GFP^{-}) (MFI, $<10^3$ RFU), d2EGFP^{Low} (GFP^{L}) (MFI, $\sim10^3$ to 10^4 RFU), and d2EGFP^{High} (GFP^{H}) (MFI, $>10^4$ RFU) phenotypes. Of note, given the shorter half-life of d2EGFP, the stability of the d2EGFP^{Low} phenotype was extremely transient; hence, the present system lacked a distinct d2EGFP^{Low} cluster at any time point, unlike in the EGFP system (Fig. 3C).

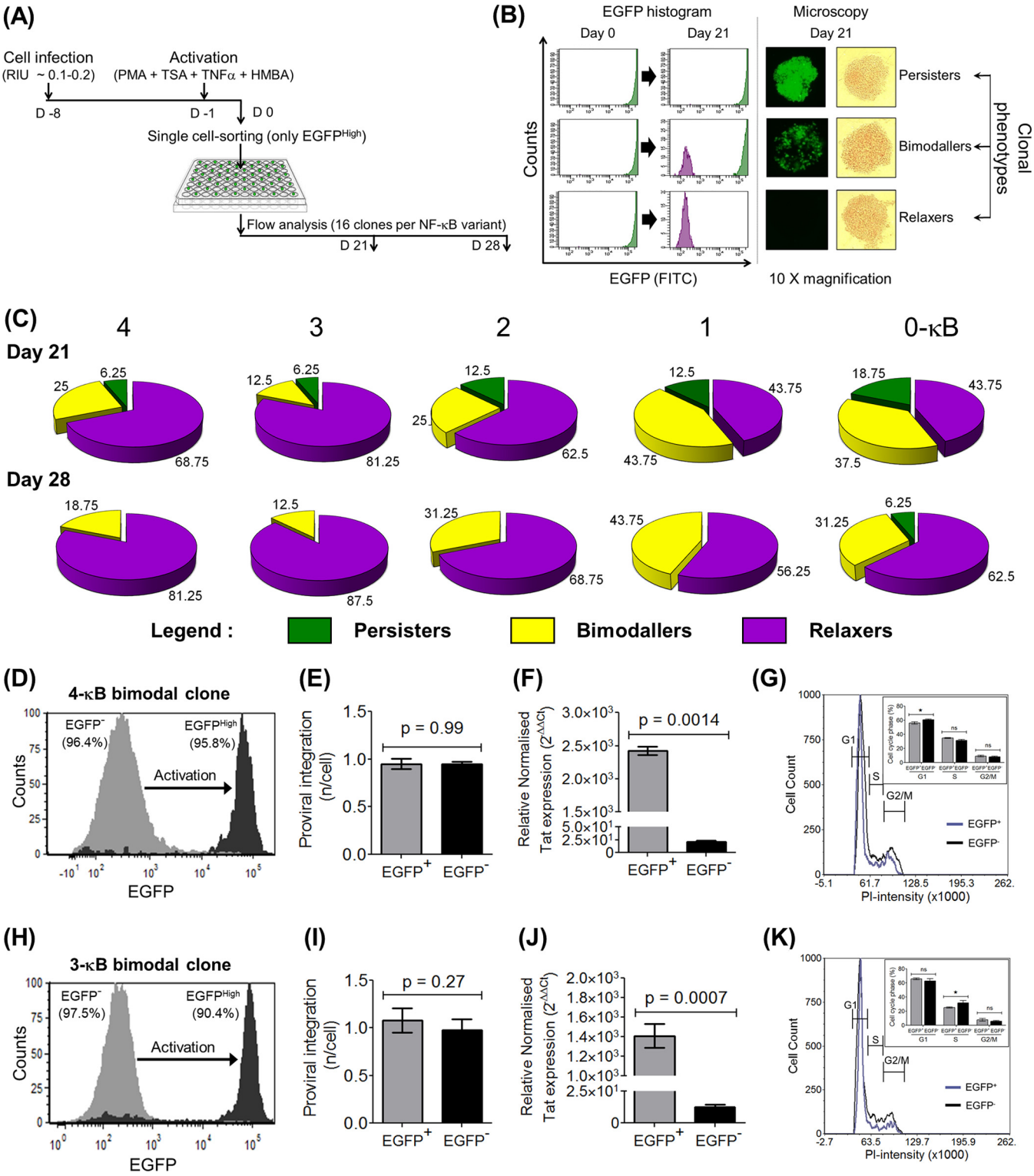

FIG 5 Manifestation of three distinct latency phenotypes of single EGFP^{High} cells. (A) The experimental layout of latency establishment in single-cell clones is essentially similar to that of nonclonal population kinetics as shown in Fig. 4A. EGFP^{High} cells (MFI, $>10^4$ RFU) were single-cell sorted and expanded for 3 to 4 weeks, and the pattern of EGFP expression was assessed on days 21 and 28 postsorting by flow cytometry and fluorescence microscopy. The EGFP expression levels of 16 randomly selected clones corresponding to each viral variant were measured. (B) Based on the fluorescence profile, three distinct categories of clones, persisters (EGFP^{High}) (MFI, $>10^4$ RFU), relaxers (EGFP$^-$) (MFI, $<10^3$), and bimodallers (binary population of persisters and relaxers), were identified. FITC, fluorescein isothiocyanate. (C) Relative proportions of the above-described three phenotypes among the LTR variants as a function of time, as indicated at days 21 and 28 postsorting. (D and H) Two bimodal cell lines, 3c and 8c, representing the 4-κB and 3-κB LTRs, respectively, demonstrate bimodal gene expression. The sorted EGFP$^-$ cells from the 4-κB and 3-κB clones generated 95.8% (D) and 90.4% (H) EGFP^{High} cells, respectively, following activation with the global activation cocktail, with a negligible proportion of cells displaying the intermediate phenotype. (E and I) TaqMan qPCR targeting a region of the LTR, performed

(Continued on next page)

whereas the other subset downregulated the reporter gene completely (EGFP$^-$), with a minimal manifestation of an intermediate phenotype.

Importantly, all five viral strains of the panel displayed the three clonal phenotypes described above, with the distinction that the proportion of the three phenotypes is directly correlated to the copy number of NF-κB sites in the LTR. Given the limitation of available cells for flow analysis, we could determine the phenotype of the clonal cells only at day 21 and day 28 but not earlier. We analyzed 16 randomly selected clones for each of the five LTRs of the panel (Fig. 5C). The profile of the three phenotypes varied significantly among the members of the panel and appeared to associate with the transcriptional strength of the LTR. On day 21, a large proportion of cell lines representing the strong viral promoters (4- and 3-κB LTRs) transited to the off state compared to those of weak promoters (1- and 0-κB LTRs); in contrast, cells of the 2-κB LTR occupied an intermediate position. On day 28, the strong LTRs contained no persistent phenotype and fewer bimodal clones. The strong viral promoters downregulated EGFP expression, both the persistent and bimodal phenotypes, at a significantly higher rate than the other three promoter variants. Despite the limitation of the small number of clonal cell lines used in the analysis, these data are broadly consistent with the results for cell pools (Fig. 4). Thus, for cell pools and clonal cell populations, both models demonstrated a direct correlation between the transcriptional strength of the LTR and the rate of latency establishment. Furthermore, both experimental models are also consistent with each other in demonstrating bimodal, and not gradual, latency establishment.

The clonal cell lines that display the bimodal EGFP phenotype offer an excellent experimental model as these clonal lines demonstrate two contrasting phenotypes (EGFP^{High} and EGFP$^-$ expression) despite identical viral genotypes, chromatin backgrounds, and host cell activation. We selected two clones, 3c and 8c, representing the strong 4-κB and 3-κB LTRs, respectively, and characterized their bimodality. The two daughter populations, EGFP^{High} and EGFP$^-$, were subsequently enriched using fluorescence-activated cell sorter (FACS) analysis and examined for the nature of transcription complexes recruited to the active and silent LTRs using chromatin immunoprecipitation (ChIP) analyses (see below). Importantly, a vast majority of the EGFP$^-$ cells of the bimodal clones representing the 4- or 3-κB LTR could be fully reactivated to the EGFP^{High} phenotype, 95.8% (Fig. 5D) and 90.4% (Fig. 5H) EGFP^{High} cells, respectively, following global activation. The levels of proviral integration between the two subpopulations of each bimodal clone were comparable and close to ~1.0 (Fig. 5E and I), ruling out the possibility of integration frequency differences underlying the bimodal phenotype. Importantly, the Tat transcript levels in the EGFP$^+$ subfractions of both clonal cell lines were significantly higher than those of their EGFP$^-$ counterparts, approximately 112-fold for the 4-κB (Fig. 5F) and 80-fold for the 3-κB (Fig. 5J) clones. We also excluded the possibility of the bimodal phenotype arising from cell cycle differences between the two phenotypes. We compared the proportions of cells in different phases of the cell cycle (G_1, S, and G_2/M) between the EGFP$^-$ and EGFP^{High} subpopulations. We found that the cell proportions were comparable for both phenotypes. The data were reproducible in both the 4-κB (Fig. 5G) and the 3-κB (Fig. 5K) clones; the manifestation of the two contrasting phenotypes in the bimodallers was therefore unlikely to be a consequence of cell cycle differences.

A tunable regulatory circuit of HIV-1 transcription alludes to the direct role of Tat in latency establishment. The stronger transcriptional activity of the LTR is

FIG 5 Legend (Continued)
as described in the legend of Fig. 3J, confirms comparable integration frequencies (~1.0 provirus per cell) between the EGFP$^-$ and EGFP^{High} fractions in both the 4-κB (E) and the 3-κB (I) clones. (F and J) Quantitative real-time PCR of Tat transcripts demonstrates significantly higher levels of Tat transcripts in the EGFP^{High} fraction than in the EGFP$^-$ fraction for both clones. Mean values from three independent experiments ± SEM are plotted. A two-tailed, unpaired *t* test was used for the statistical evaluation. (G and K) DNA cell cycle analysis was performed on the EGFP$^-$ and EGFP^{High} subfractions of the 4-κB (G) and the 3-κB (K) clones according to the standard PI staining protocol. Overlay histograms for the two subfractions showing the G_1, S, and G_2/M phases are presented. The proportions of cells in the G_1, S, and G_2/M stages were calculated for both subfractions and are depicted in the insets. Mean values from triplicates, representative of results from two independent reactions, ± SD are plotted. A two-tailed, unpaired *t* test was used for the statistical evaluation.

expected to lead to a proportionately higher expression level of Tat, which in turn should increase the transcriptional activity of the LTR further. As a consequence of the unique arrangement, the two principal regulatory elements collectively modulate viral gene expression. In this backdrop, the profile of latency kinetics observed using the ATF model described above cannot be ascribed to the different functional activity of either of the elements alone. It was therefore necessary to employ a strategy where Tat transactivation alone becomes a variable factor while the transcriptional strength of the LTR remains constant. To this end, we constructed a new HIV-1–Jurkat cell line model, the "tunable Tat feedback" (TTF) model, where the transactivation strength of Tat can be modulated independently while keeping the transcriptional strength of the LTR constant. Tat in the TTF model was engineered to possess two unique properties compared to the ATF model (Fig. 6A). First, Tat was fused with DsRed2-red fluorescent protein (RFP) (referred to as RFP throughout the manuscript) to express it as a fusion protein, enabling the direct visualization of its expression. The new HIV-1 reporter vector pcLdGITRD (cLTR-d2EGFP-IRES-Tat:RFP:DD) thus coexpressed two different fluorescent proteins, d2EGFP and Tat-RFP, under the control of the LTR. Second, the Tat-RFP fusion protein was tagged with the C-terminal destabilization domain (DD) of FK506 binding protein (FKBP) (Tat:RFP:DD). The DD can target the fusion protein for rapid proteasome-mediated degradation (28). Shield1, a small-molecule ligand, however, can rescue DD-mediated degradation by specifically interacting with the DD motif and stabilizing the target protein in a dose-responsive manner (29). The subgenomic HIV-1 reporter vector pcLdGITRD, representing the TTF model, can thus fine-tune the intracellular concentration of the Tat:RFP:DD fusion protein by changing the concentration of Shield1 in the culture medium, in the context of a fixed LTR strength.

We constructed a panel (cLdGITRD, p913 series [see Materials and Methods]) of two LTR variant viral strains consisting of 3 or 1 NF-κB motif, representing the strong and weak LTRs, respectively (Fig. 6A). A direct correlation between the Shield1 concentration in the medium, ranging from 0 to 5 μM, and the intensity of Tat-RFP expression was observed in HEK293T cells using the 3-κB viral reporter vector (Fig. 6B). Importantly, the viruses could infect target Jurkat cells. We evaluated the levels of d2EGFP expression and Tat-mediated transactivation, with an increasing concentration of Shield1 in the medium, using the experimental strategy depicted in Fig. 6C. Interestingly, the effect of the Shield1 concentration was directly manifested on the d2EGFP^{High} population in the stacked histogram profile (Fig. 6D, black arrows), indicating Shield1 dose-dependent Tat transactivation and also confirming that the d2EGFP^{High} phenotype represented Tat-transactivated cells (Fig. 6D). A direct correlation was also established in stable Jurkat cells between the Shield1 concentration and the d2EGFP MFI or Tat:RFP:DD expression (Fig. 6E and H, respectively), suggesting Shield1-dependent stabilization of the Tat:RFP:DD cassette and the subsequent Tat-mediated LTR transactivation. Of note, although we normalized viral infection, the percentages of d2EGFP^{+} cells demonstrated a dose response proportional to the Shield1 concentration, even though d2EGFP itself does not contain the DD of FKBP (Fig. 6G). The optimal fold activation of d2EGFP expression (Fig. 6F) and Tat transcript levels (Fig. 6I) were found to be 1 μM and 2.5 μM, respectively. In subsequent experiments, we therefore used Shield1 in the range of 0 to 3 μM.

Importantly, the fusion of Tat with DsRed2-RFP offered the advantage of tracking the expression of Tat in real time during latency establishment. To determine the kinetics of latency establishment in Jurkat cells, we used the experimental scheme depicted in Fig. 7A. Jurkat cells were infected with the 3- or 1-κB viral strain at ~0.1 to 0.2 RIU in the presence of 1 μM Shield1 and expanded for a week in the presence of Shield1. Subsequently, the cells were activated with the global activators for 24 h, the d2EGFP^{High} population (MFI, ~10^4 RFU) was sorted, the sorted cells were maintained separately at four different concentrations of Shield1 (0, 0.5, 1.0, and 3.0 μM), and the levels of d2EGFP and Tat-RFP expression were monitored every 24 h by flow cytometry.

The TTF model of latency offered several essential insights. Importantly, the ability to visualize two different fluorescent proteins (d2EGFP and Tat:RFP:DD) coexpressed

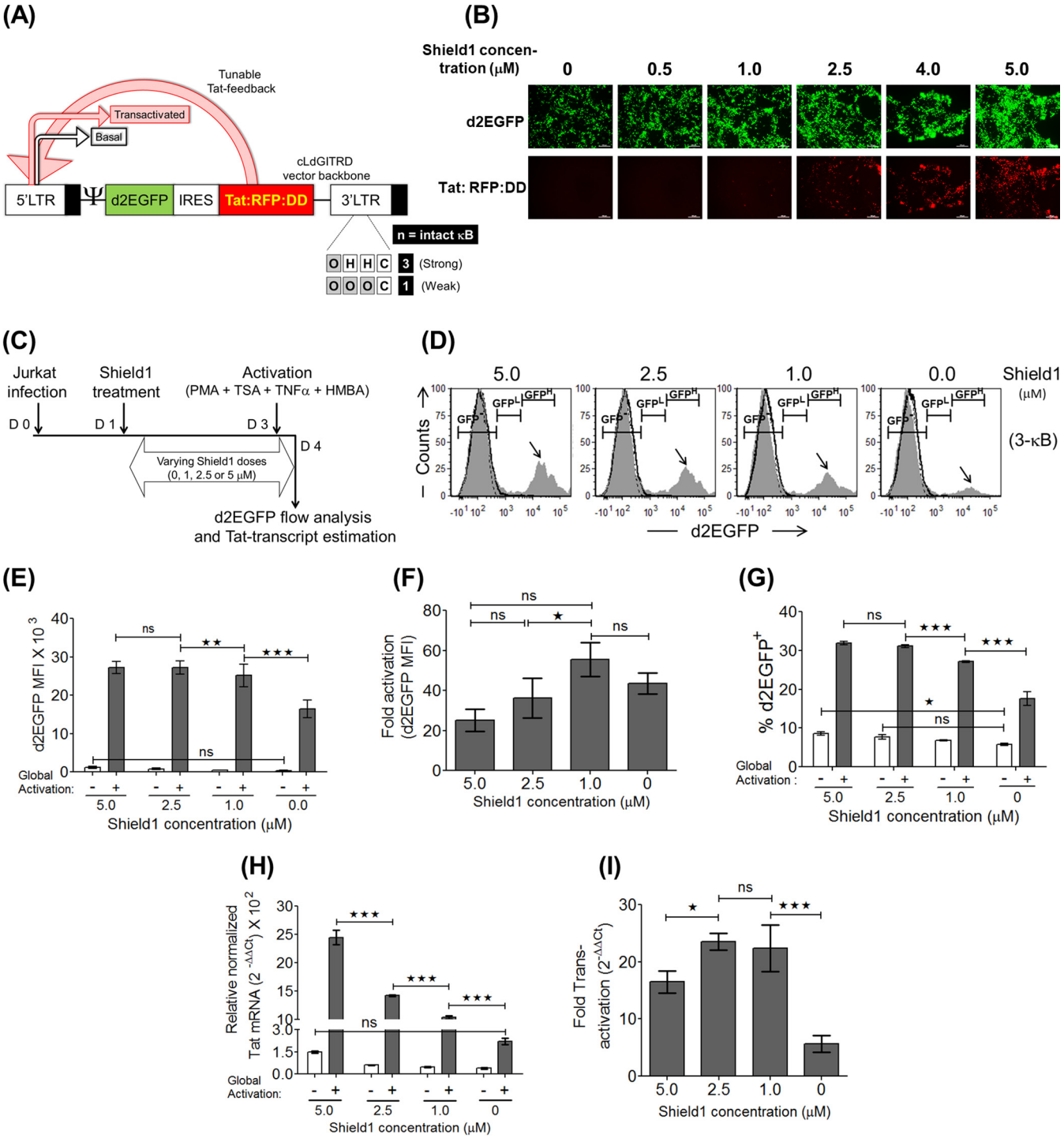

FIG 6 In the tunable Tat feedback (TTF) model, the stronger the LTR-Tat feedback, the higher the viral gene expression level. (A) Schematic of the subgenomic HIV-1 vector backbone cLdGITRD representing the TTF model. The 3- and 1-κB LTRs, representing a strong and a weak promoter, respectively, are used in the present study. The small molecule Shield1 stabilizes the Tat:RFP:DD cassette in a dose-dependent fashion, making it available for subsequent transactivation events at the LTR. (B) Validation of Shield1 dose-dependent stabilization of the Tat:RFP:DD cassette in HEK293T cells. One milligram of the cLdGITRD-3-κB vector was transfected into 0.6 million HEK293T cells in separate wells in the presence of various concentrations of Shield1, as indicated, and the images were captured at 48 h posttransfection. The experiment was repeated twice, with comparable results. (C) Experimental layout to confirm Shield1 dose-dependent gene expression and Tat transactivation in Jurkat cells. Approximately 0.3 million Jurkat cells were infected with the LdGITRD-3-κB strain (20 ng/ml p24 equivalent), and after 24 h, the infected cells were split into four fractions, each treated with a different concentration of Shield1, as indicated. After 48 h of Shield1 treatment, half of the cells from each fraction were activated for 24 h, followed by the quantitation of d2EGFP and Tat mRNA expression levels for both the induced and uninduced fractions. (D) Representative stacked histograms indicate a Shield1 dose-dependent Tat-transactivated population (black arrows) at a fixed LTR strength (fixed number of NF-κB motifs). The peak height of the d2EGFP^{High} population (MFI, $>10^4$ RFU) was proportionally reduced with the Shield1 dose; the TTF model thus confirmed that the d2EGFP^{High} cluster represents the Tat-transactivated population. (E to G) Shield1 concentration-dependent d2EGFP MFI (E), fold d2EGFP enhancement (F), and percent d2EGFP$^+$ (G) plots. Values from experimental triplicates ± SD, representing the results of two independent

(Continued on next page)

under the control of the LTR permitted the identification of the different stages of viral gene expression and latency, which we collectively refer to as the viral "latency cycle." Although both fluorescent proteins were expressed under the control of the same viral promoter, d2EGFP was expressed perceptibly earlier and at a higher intensity than the Tat:RFP:DD fusion protein. The increased molecular size of the Tat:RFP:DD fusion protein, the slow maturation of DsRed2, and the compromised translation efficiency due to the IRES element may all have contributed to the observed difference between the d2EGFP and Tat-RFP expression profiles (Fig. 7B, day 0). The profiles of gene expression throughout the different phases of the latency cycle are remarkably different between the two viral promoters. The transiting of the cells through successive phases of the latency cycle is illustrated explicitly when the 3-κB LTR profile is examined (Fig. 7B, top). At day 0 following the d2EGFP^{High} sort, the vast majority of cells (92.8%) were d2EGFP$^+$ Tat-RFP$^-$, representing a transcriptionally active viral promoter (Fig. 7B, top, day 0, and Fig. 7C). During the following 24 h, the d2EGFP$^+$ Tat-RFP$^-$ cells exited this compartment via two distinct and diametrically opposite routes. While a significant proportion of these cells (approximately 15%) switched off d2EGFP expression to directly return to the d2EGFP$^-$ Tat-RFP$^-$ compartment, approximately 6.6% of cells upregulated Tat-RFP expression from the 3-κB LTR to transit to the d2EGFP$^+$ Tat-RFP$^+$ compartment, alluding to strong Tat-dependent transcriptional activity (Fig. 7B, top, day 1, and Fig. 7D). At subsequent time points, d2EGFP$^+$ Tat-RFP$^-$ cells continued to vacate this compartment using both exit routes to reach the d2EGFP$^-$ Tat-RFP$^-$ compartment such that on day 6, 84.3% of the viral strains reestablished latency under the control of the strong viral promoter. Importantly, cells in the d2EGFP$^+$ RFP$^+$ compartment, unlike those in the d2GFP$^+$ Tat-RFP$^-$ compartment, appeared to move to latency only in one direction to the d2EGFP$^-$ Tat-RFP$^+$ compartment (Fig. 7B, top, day 3, and Fig. 7E). The relative proportion of the cells present in the d2EGFP$^-$ Tat-RFP$^+$ compartment was significantly higher than that in the d2EGFP$^+$ Tat-RFP$^+$ compartment at time points after day 1, alluding to the unidirectional movement of these cells to latency. Importantly, the d2EGFP$^-$ Tat-RFP$^+$ compartment is unique since this quadrant represents the proviruses that have "recently" switched off transcription, with significant levels of physiological Tat still being persistent in the system, as indicated by the RFP$^+$ phenotype. The proviruses of the d2EGFP$^-$ Tat-RFP$^+$ compartment also transited to latency only in one direction and entered the d2EGFP$^-$ Tat-RFP$^-$ compartment (Fig. 7B, top, day 5, and Fig. 7F).

In contrast, the 1-κB LTR predominantly displayed Tat-independent transactivation (Fig. 7B, bottom). Although approximately 4% of these cells expressed Tat-RFP at a Shield1 concentration of 3 μM, Tat-RFP expression was delayed by 24 h compared to that of the 3-κB LTR, with Tat-RFP expression reaching a peak only on day 3. Importantly, despite the presence of Tat, these dual-positive cells of the 1-κB LTR (d2EGFP$^+$ Tat-RFP$^+$) did not move forward to the d2EGFP$^-$ Tat-RFP$^+$ compartment, unlike those of the 3-κB LTR, but returned to the d2EGFP$^+$ Tat-RFP$^-$ quadrant (Fig. 7B, bottom, day 3). The proviruses activated by Tat-independent transactivation primarily manifested the d2EGFP$^+$ Tat-RFP$^-$ phenotype, and these viruses returned to latency by switching off d2EGFP expression and typically not inducing Tat-RFP expression. While a large majority of the 3-κB LTR viral strains and nearly all viral strains of the 1-κB LTR followed this route of latency, a smaller proportion of proviruses of the 3-κB LTR moved forward activated by Tat-dependent transactivation that manifested the d2EGFP$^+$ Tat-RFP$^+$ phenotype. Approximately 14% of the 3-κB LTR viral strains were activated by Tat-dependent transactivation that followed a unidirectional trajectory to latency via the d2EGFP$^+$ Tat-RFP$^+$ and d2GFP$^-$ Tat-RFP$^+$ compartments. In contrast, only approxi-

FIG 6 Legend (Continued)
experiments, are plotted. Two-way ANOVA with Bonferroni posttest correction was used for the statistical evaluation (**, $P < 0.01$; ***, $P < 0.001$; ns, nonsignificant). (H and I) Relative Tat transcript levels (H) and fold Tat-mediated transactivation (I) were evaluated as described in the legends of Fig. 2G and H, respectively. The mean values from three independent experiments ± SEM are plotted. Two-way ANOVA with Bonferroni posttest correction was used for the statistical evaluation (*, $P < 0.05$; ***, $P < 0.001$; ns, nonsignificant).

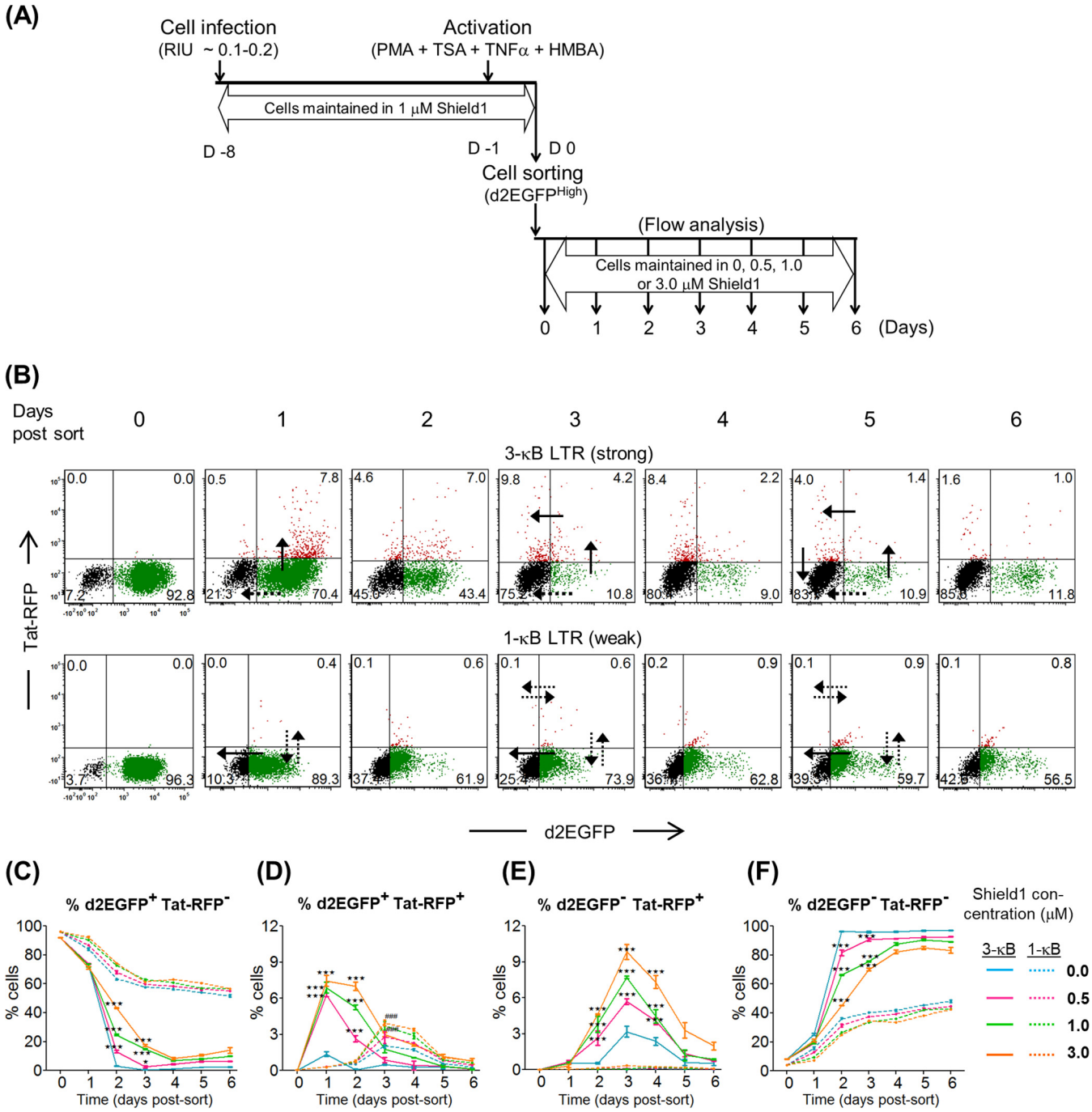

FIG 7 The TTF model identifies two distinct modes of latency establishment in the strong and weak LTRs. (A) Experimental scheme to study latency establishment as described in the legend of Fig. 4A, with slight modifications. The sorted d2EGFP^{High} cells were divided into four separate fractions and maintained at different concentrations of Shield1, as depicted. The cells were analyzed every 24 h for d2EGFP and Tat-RFP expression using flow cytometry. (B) Temporal d2EGFP and Tat-RFP trajectories for the strong (3-κB) (top) and weak (1-κB) (bottom) LTRs at a Shield1 concentration of 3 μM. The black solid and the black dotted arrows denote the dominant and less dominant routes of transit for each LTR variant, respectively. (C to F) Individual kinetic curves of the percentages of four distinct populations, d2EGFP$^+$ Tat-RFP$^-$ (C), d2EGFP$^+$ Tat-RFP$^+$ (D), d2EGFP$^-$ Tat-RFP$^+$ (E), and d2EGFP$^-$ Tat-RFP$^-$ (F). Mean values from experimental triplicates ± SD are plotted. Data represent the results of three independent experiments. Two-way ANOVA with a Bonferroni posttest was used for the statistical evaluation (*, $P < 0.05$; ***, $P < 0.001$; ns, nonsignificant). The solid and dotted colored curves represent various concentrations of Shield1 for the 3-κB and 1-κB LTRs, respectively.

mately 1% of 1-κB LTR viral strains could follow Tat-dependent transactivation, while the remainder were induced only by Tat-independent activation and returned to latency directly from the d2EGFP$^+$ Tat-RFP$^-$ compartment. Thus, the transcriptional strength of the viral promoter appears to play a critical role in regulating not only the

activation of viral gene expression but also the latency kinetics and whether or not Tat-dependent transactivation is recruited to the LTR. Only the strong 3-κB LTR, but not the weak 1-κB LTR, could successfully undergo Tat-dependent transactivation. The major routes of entry into latency (Tat dependent or Tat independent) exhibited by the two LTR variants are shown as solid black arrows, while their less dominant trajectories are indicated as dotted black arrows, in Fig. 7B. Individual trajectories of the percentages of the four distinct fluorescent populations are presented in Fig. 7C to F.

Interestingly, a clear demarcation in the profiles of the weak and strong LTRs is evident at the level of Tat-independent transactivation, in the Tat-RFP-negative cell populations (Fig. 7C and F). d2EGFP$^+$ Tat-RFP$^-$ cells of the 3-κB LTR downregulated d2EGFP at all concentrations of Shield1 by day 4. In contrast, latency establishment in the same population of the 1-κB LTR was incomplete, and nearly half of these cells remained d2EGFP$^+$ on day 6. This was primarily because a subset of the d2EGFP$^+$ Tat-RFP$^-$ cells at later time points (day 1 and beyond) followed the Tat-dependent route to latency in the case of the strong 3-κB but not the weak 1-κB promoter. Therefore, from the data from the TTF model, it appears that Tat-dependent transactivation can silence the promoter at a higher rate than that of the Tat-independent pathway. Furthermore, the kinetics of the percent d2EGFP$^+$-to-d2EGFP$^-$ transition, irrespective of Tat-RFP expression, demonstrated an identical pattern of promoter silencing in the TTF model, with the strong promoter (3-κB) facilitating a higher rate of silencing than the weak promoter (1-κB), compared with the ATF model (3- and 4-κB versus 2-, 1-, and 0-κB LTRs). At all concentrations of Shield1, the strong 3-κB LTR switched off faster than the weak 1-κB LTR. Thus, the data obtained from the TTF model strongly suggest 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 strong LTR not only is faster in establishing viral latency but also is rapid in revival kinetics from latency, whereas a weak viral promoter appears to be restricted in both functions.

Sustained presence of Tat in the nucleus following LTR switch-off. The latency kinetics of two different cellular models (ATF and TTF) alluded to the direct involvement of Tat in the transcriptional suppression of the viral promoter, in a concentration-dependent manner. Furthermore, we could detect the presence of the Tat-RFP fusion protein in cells harboring a transcriptionally silent provirus (d2EGFP$^-$ Tat-RFP$^+$) containing a strong viral promoter (3-κB LTR) (Fig. 7). Therefore, it was necessary to evaluate the physiological levels and relative distributions of Tat in cells concomitant with LTR silencing. To this end, we tracked the expression pattern of the Tat protein in Jurkat cells using indirect immunofluorescence while the cells transited from the "on" to the "off" state. Jurkat cells infected with the cLdGIT-3-κB viral strain encoding d2EGFP (ATF model) were monitored at 4-day intervals up to day 16 for d2EGFP expression using flow cytometry (Fig. 8A, left). At day 0, the d2EGFP^{High} cells were sorted and subjected to indirect immunofluorescence analysis of Tat at different points (day 0, day 4, day 8, day 12, day 14, and day 16) (Fig. 8A, right). A combination of a high-titer, polyclonal, rabbit anti-Tat primary antibody and an anti-rabbit, Alexa Fluor 568-conjugated secondary antibody was used in the assay.

d2EGFP expression analysis by flow cytometry found a progressive downregulation of fluorescence, and by day 8 and day 16, only 6.9% and 1% of the cells, respectively, remained positive. The profile of d2EGFP expression of individual cells captured by confocal microscopy was perfectly consistent with that of the flow analysis, and visible fluorescence could not be detected at day 8 and beyond. However, trace levels of Tat expression as determined by the indirect immunofluorescence of Tat (Tat-Alexa Fluor 568 signal) could be visually noted above the background level on day 12 and day 16 despite the complete downregulation of d2EGFP, indicating the sustained presence of Tat in an LTR-off context. Of note, Tat expression was found in two different compartments of the cells, nuclear and extranuclear, with the latter being localized mostly to the cell membrane. The fluorescence intensities of d2EGFP expression, as well as those of Tat-Alexa Fluor 568, were determined independently in the nuclear and extranuclear

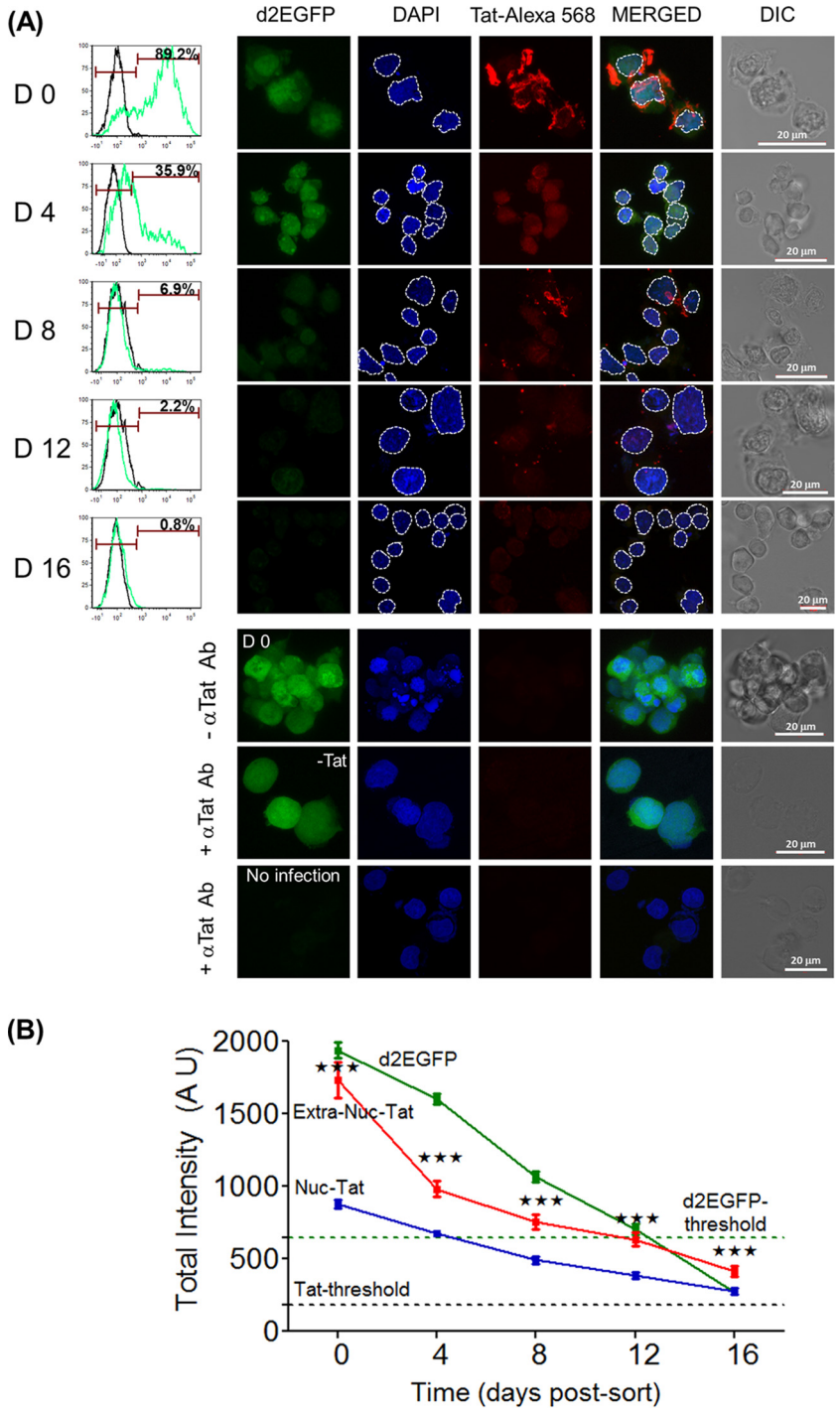

FIG 8 Persistent presence of Tat in latent cells. (A) Temporal profile of Tat expression during LTR silencing in a stable J-LdGIT-3-κB cell pool of the ATF model. The experimental strategy of latency establishment was similar to the one described in the legend of Fig. 4A. (Left) At defined time points, a small fraction of the sorted d2EGFP^{High} cells was flow analyzed for d2EGFP expression, while the remaining cells were subjected to an indirect immunofluorescence assay for Tat using a rabbit anti-Tat primary antibody (Ab) and an anti-rabbit Alexa Fluor 568-conjugated secondary antibody. DAPI was used to stain the nucleus. (Right) Representative confocal images of Tat and d2EGFP expression at the indicated time points. Appropriate negative controls for Tat IF are presented at the bottom. The white dotted lines demarcate the nucleus from the extranuclear compartment in each cell. DIC, differential interference contrast. (B) Quantitative analysis of d2EGFP and Tat-Alexa Fluor 568 intensity levels in the nuclear and extranuclear compartments (arbitrary units) at multiple time points. Data from 150 individual cells at each time point and three independent experiments are presented. The threshold values for total cellular d2EGFP and Tat-Alexa Fluor 568 intensities were obtained from uninfected Jurkat cells (A, 8th row) and infected unstained cells (A, 6th row), respectively. Mean values ± SEM are plotted. One-way ANOVA was used for statistical evaluation (***, $P < 0.001$).

compartments of 150 individual cells at all time points (Fig. 8B). The threshold levels of fluorescent protein expression were determined by using an uninfected Jurkat cell control for d2EGFP (Fig. 8A, 8th row) ($n = 10$) and a no-primary-antibody control for Tat (6th row) ($n = 10$). Importantly, while the fluorescence of d2EGFP was reduced progressively with time and fell below the threshold by day 12, representing the establishment of latency, the fluorescence of Tat in either compartment did not drop below the Tat-Alexa Fluor 568 threshold, even at day 16. The slopes of the reduction of the Tat intensities during the initial phases (day 0 to day 4) of latency establishment were estimated to be -74.54 ± 16.8 and -37.28 ± 3.2 in the extranuclear and nuclear compartments, respectively. At later time points (day 8, day 12, and day 16), there was only a moderate reduction in the Tat levels in either compartment. These data thus suggest a higher level of stability of Tat in the nucleus, with possible implications for HIV latency. Importantly, the data for Tat immunofluorescence are in perfect agreement with the results of the TTF model, where the few d2EGFP$^-$ Tat-RFP$^+$ cells at the later stages of promoter silencing indicated the sustained presence of low levels of Tat molecules in the LTR-switched-off cells (Fig. 7B, top). In summary, immunofluorescence not only detected the presence of Tat in latently infected cells as late as day 16 postsorting but also demonstrated a rapid loss of Tat from the extranuclear compartment but its relative stability in the nucleus.

The *in situ* proximity ligation assay detects the presence of Tat in latently infected cells. In the indirect immunofluorescence assay, the overall intensity of Tat at day 12, day 14, and day 16 in both cellular compartments was only marginally above the background level. To increase the sensitivity and detect limited quantities of Tat in the "LTR-off" cells, we used the highly sensitive proximity ligation assay (PLA), which conjugates immunostaining with rolling-circle replication and outperforms traditional immunoassays in sensitivity to detect trace amounts of endogenous proteins (30, 31). We optimized the Tat PLA in HEK293T cells using a pair of anti-Tat primary antibodies raised in different hosts (rabbit and mouse). Since the PLA does not work well in nonadherent cells, and our attempts to adapt the protocol to Jurkat cells were not successful, we used HEK293/HEK293T cells in this assay. Using subgenomic viral vectors encoding Tat representing HIV-1B (pLGIT) or HIV-1C (pcLGIT), we optimized the PLA (Fig. 9A and B). The B-Tat protein could be detected as distinct white dots, as opposed to sparse dots in the no-antibody and single-antibody controls (Fig. 9A). Moreover, a dose response in the intensity of PLA dots and plasmid concentrations as well as a good correlation between the number of PLA dots and GFP MFI are evident in the case of C-Tat (Fig. 9B).

Using the optimized PLA protocol for Tat, we asked if the Tat protein could be detected in d2EGFP-off cells. To this end, HEK293 cells stably and independently infected with the 4- and the 3-κB variants of the ATF cLdGIT panel were sorted for d2EGFP^{High} cells. After a week of incubation following enrichment, approximately 50% of the cells expressed d2EGFP, and the cell pool contained both active (d2EGFP$^+$) and latent (d2EGFP$^-$) cell clusters. The Tat PLA was then performed using the mixed d2EGFP pool corresponding to both strong LTR variants (3- and 4-κB) to quantitate the Tat PLA signals in the alternate phenotypes (active and latent). Cells stained with either of the antibodies alone did not show any Tat-specific signals, confirming the specificity of the assay (Fig. 9C, left, top two rows). Tat-specific staining was evident only in the presence of both antibodies, not only in d2EGFP$^+$ cells but also in d2EGFP$^-$ cells (Fig. 9C, left, bottom two rows). The average numbers of Tat PLA dots per cell was determined in a total of 164 d2GFP$^+$ cells (128 and 36 cells for the 4-κB and 3-κB variants, respectively) and 168 d2EGFP$^-$ cells (123 and 45 cells for the 4-κB and 3-κB variants, respectively), comprising data from three independent experiments (Fig. 9C, right). These values were found to be 3.35 ± 0.77 and 2.8 ± 0.59 for d2EGFP$^+$ and d2EGFP$^-$ cells, respectively, although the difference was not statistically significant. The Tat PLA data for HEK293 cells confirmed the presence of Tat in latent cells at a concentration comparable to that for active viral transcription.

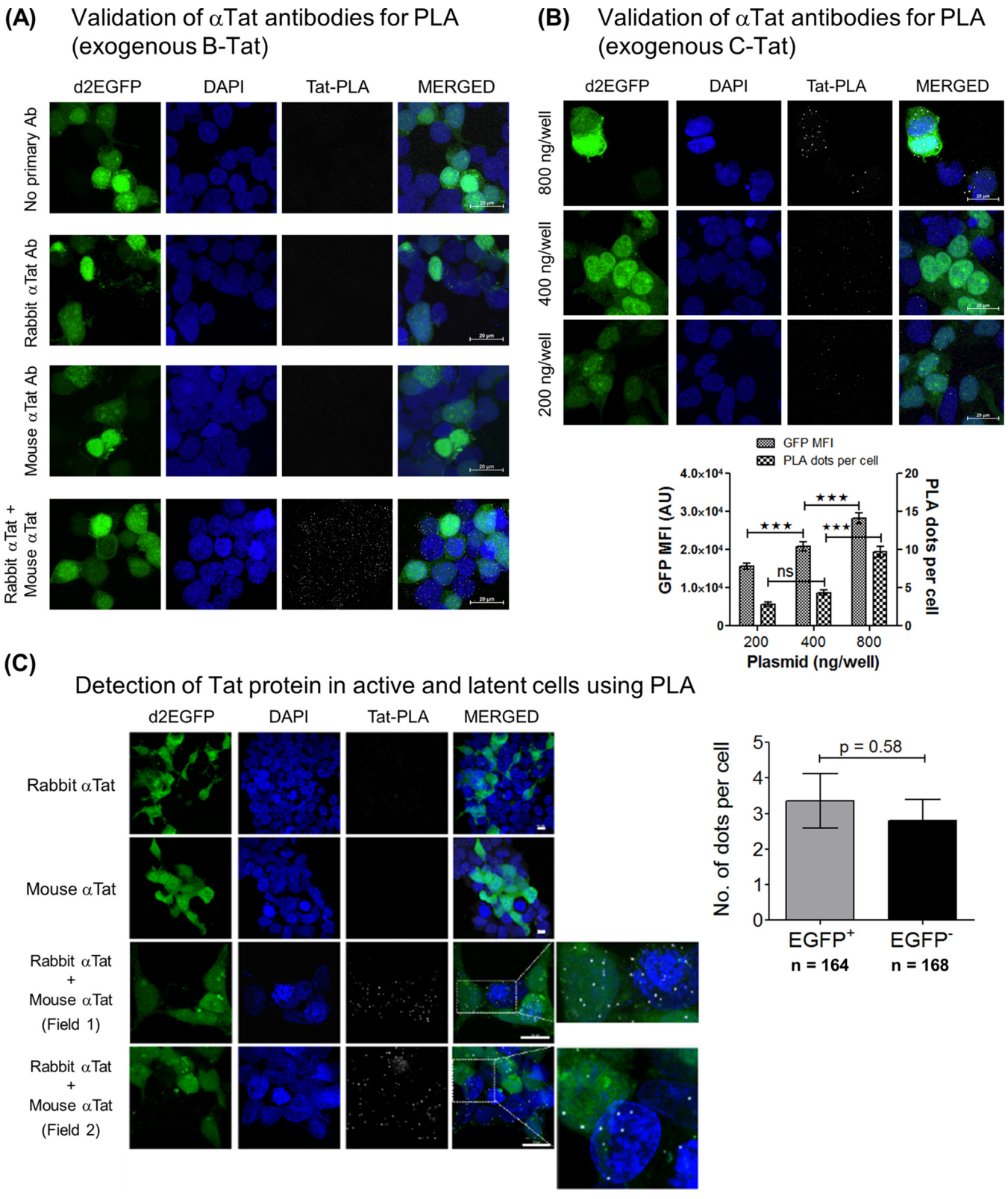

FIG 9 Presence of Tat in latent cells as confirmed by the highly sensitive proximity ligation assay (PLA). (A) Detection of exogenous B-Tat by the PLA. Approximately 0.5 million HEK293T cells seeded per well in a 12-well culture dish were transfected with 800 ng of pLGIT, an expression vector encoding Tat of HIV-1B, on poly-L-lysine-coated coverslips. The Tat PLA was performed according to the manufacturer's protocol, using a pair of anti-Tat primary antibodies (rabbit polyclonal [catalog number ab43014; Abcam] and mouse monoclonal [catalog number 7377; NIH AIDS Reagent Program]). Representative confocal images of a "no-antibody" control (1st row), single anti-Tat antibody controls (2nd and 3rd rows), and both antibodies (4th row) are presented. The mean values from three independent experiments ± SEM are plotted. One-way ANOVA with Bonferroni's multiple-comparison posttest was used for statistical analyses (***,

(Continued on next page)

Differential occupancy of cellular complexes on active and silent LTRs of bimodal clones. J. C. Burnett et al. examined the differential occupancy of NF-κB factors (p50 and p65) at each of two identical NF-κB motifs (motifs I and II) in the HIV-1B LTR by introducing inactivation mutations into each of these sites individually and the corresponding impacts on transcriptional activity (26). A similar examination at the C-LTR has not been performed. A previous report from our laboratory demonstrated that the NFAT1 and -2 proteins could be recruited to the C-κB motif, the variant NF-κB motif unique to HIV-1C, with an affinity superior to that of the canonical H-κB site (32). We attempted to compare the identities of the transcription factors and other host factors binding to the viral promoter between the active and suppressed states under identical experimental conditions.

Having demonstrated the presence of Tat in cells containing the latent provirus in both the TTF (Fig. 7) and ATF (Fig. 8 and 9) models by flow cytometry and confocal imaging, respectively, we next asked if Tat in these cells is recruited to the latent viral promoter. We sorted the d2EGFP^{High} and d2EGFP^{-} cell populations from the two clonal cell populations, and using the chromatin immunoprecipitation assay, we examined the presence of several essential host factors or epigenetic marks (Rel family members p50 and p65, NFAT1 and NFAT2, RNA polymerase Ser2 phosphorylation, and histone 3 lysine 9 trimethylation [H3K9Me3]) as well as Tat in the chromatin preparations of active and latent cells. ChIP assays were performed by amplifying a 240-bp fragment spanning the enhancer-core promoter region in the LTR using semiquantitative PCR and also using independent TaqMan probe-based real-time PCR amplifying a 127-bp region spanning the NF-κB and Sp1 sites in the LTR (Fig. 10A).

The comparative analysis of the nature of the host factors recruited between the active and latent promoters was highly reproducible and consistent between the 4- and 3-κB LTRs (Fig. 10C, D, F, and G). While the transcription-promoting host factors p65 and NFAT2 and the epigenetic mark RNA polymerase II (Pol II) S2 were found associated with the active viral promoters at significantly higher levels, the transcription-repressive factors p50 and NFAT1 and the epigenetic mark H3K9Me3 were preferentially associated with the latent viral promoters. Because the p50-p65 heterodimer is transcription promoting, the presence of a significantly higher concentration of p65 at the active promoter is expected (33–35). 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 (36). Similarly, our data are also in agreement with previous reports regarding the transcription-suppressive and -supportive functions of NFAT1 and NFAT2, respectively (37–39).

The most crucial finding of the present study is the detection of the association of the Tat protein with the latent LTR. The Tat protein was found associated with the latent 4-κB and 3-κB promoters at levels 1.7- and 3-fold lower, respectively, than those of their active counterparts. The results were reproducible and consistent in semiquantitative PCR-based ChIP analyses between the two strong viral promoters (Fig. 10C and F). The data were also consistent between conventional and quantitative real-time PCRs performed following immunoprecipitation (IP) (Fig. 10D and G). To the best of our knowledge, the present study is the first one to demonstrate the association of Tat with the latent LTR. The above-described ChIP data were generated using a commercial

FIG 9 Legend (Continued)
$P < 0.001$). (B) DNA dose response in the PLA using pcLGIT, an expression vector encoding Tat of HIV-1C (200, 400, and 800 ng). The mean GFP intensities and the average numbers of PLA dots per cell for the amounts of the pcLGIT vector transfected are presented. (C) Tat PLA dot quantitation in active versus latent cells was performed in HEK293 cells independently and stably infected with the cLdGIT-4-κB and -3-κB strains of the ATF panel (Fig. 2A). Cells were infected with one of the viral strains (~0.5 RIU), d2EGFP^{High} cells were sorted, the cells were incubated for proviral LTR relaxation to arrive at a mixed population of d2EGFP^{+} (active) and d2EGFP^{-} (latent) cells, and both cell populations were subjected to a Tat PLA. Approximately 50,000 mixed d2EGFP cells seeded into a well of an 8-well slide chamber were subjected to the PLA. Representative confocal images depicting single-antibody controls (1st and 2nd rows) and Tat PLAs with both antibodies (3rd and 4th rows) are shown in the left panel. Two subfields with distinct Tat PLA dots (white) in both d2EGFP^{+} and d2GFP^{-} cells have been enlarged for clarity. The number of Tat PLA dots per cell was determined independently for d2EGFP^{+} as well as d2EGFP^{-} cells, and the mean values from three independent experiments ± SEM are plotted. The total number of cells counted for the d2EGFP^{+} phenotype was 164 (128 for 4-κB and 36 for 3-κB), and that for the d2EGFP^{-} phenotype was 168 (123 for 4-κB and 45 for 3-κB). A two-tailed, unpaired *t* test was used for statistical analyses.

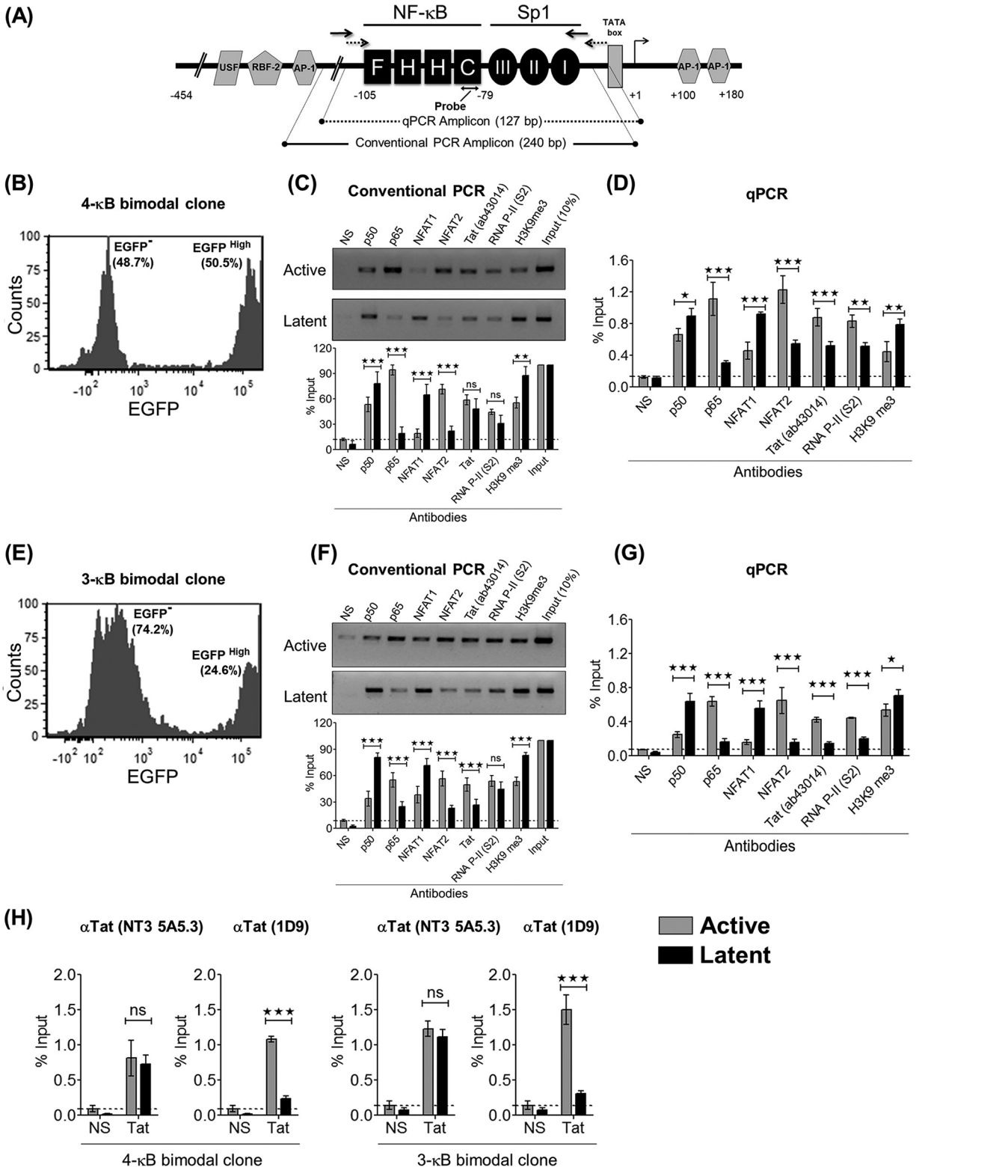

FIG 10 Active and latent LTRs recruit host factors differentially. (A) Schematic representation of the LTR sequence and the positions of the primer pairs used for conventional PCR (240 bp) (solid black) and qPCR (127 bp) (dotted black) for ChIP analysis. (B and E) Presort percentages of EGFP$^-$ and EGFP^{High} cells corresponding to the 4-κB (B) and the 3-κB (E) bimodal clones. (C and F) Data from a conventional PCR-based ChIP assay of the active (EGFP^{High}) and latent (EGFP$^-$) promoters for several host factors and the Tat protein for the 4-κB (C) and 3-κB (F) clones, respectively. The cell lysate from 2 million cells (active or latent) and 2 μg of the respective antibody were used for the individual IP reactions. A rabbit polyclonal anti-Tat antibody (catalog number ab43014; Abcam) was used for the Tat IP. One-tenth of the immunoprecipitated chromatin was used as the input control. NS refers to a matched IgG isotype control antibody. Conventional PCR was repeated three times for each IP reaction. Representative gel images and data from the corresponding densitometry analyses are presented. Data for each band are normalized to the input.

(Continued on next page)

rabbit polyclonal anti-Tat antibody (catalog number ab43014; Abcam). The data were also reproducible when two additional mouse monoclonal anti-Tat antibodies targeting different epitopes in Tat (catalog numbers 7377 and 4374; NIH AIDS Reagent Program, MD, USA) were used in the assay (Fig. 10H). All three different anti-Tat antibodies furnished positive ChIP signals for Tat at both latent viral promoters (3- and 4-κB), over and above the respective IgG isotype controls.

DISCUSSION

Significance of Tat recruitment to the latent LTR. The primary finding of the present work is the identification of a positive correlation between the transcriptional strength of the LTR and faster latency kinetics via the mediation of proportionately enhanced Tat concentrations. We found that at the time of commitment to latency and at subsequent time points, the intracellular concentration of Tat is not a limiting factor, thus ruling out the possibility that the limiting levels of Tat underlie latency establishment. To exert a positive or negative influence on the LTR, Tat must be present in the nucleus and recruited to the viral promoter. Using three different experimental strategies, flow analysis of the Tat-RFP fusion protein (Fig. 7), indirect immunofluorescence analysis (Fig. 8), and a proximity ligation assay (Fig. 9C), we successfully demonstrated the presence of Tat in the nucleus of the latent cell, through successive stages of latency establishment. The presence of Tat could be detected in both the nuclear and cytoplasmic compartments by confocal microscopy. Additionally, using ChIP, we could ascertain the recruitment of Tat to the latent LTR using three different anti-Tat antibodies, although Tat levels in the latent nuclei were typically lower than those in the active nuclei (Fig. 10). Furthermore, a weak but discernible signal of Tat transcripts was evident in the latent fractions of the bimodal clones of both strong LTRs used in the assay (Fig. 5F and J). All these data strongly suggest that Tat plays a direct role in promoting latency. Unfortunately, our attempts to adopt the PLA to suspension cells were not successful. How Tat is recruited to the chromatin complex needs to be determined. A few studies previously showed the direct binding of Tat to an extrachromosomal HIV-1 promoter and proviral DNA (40, 41); however, the tethering of Tat to the nascent transactivation response (TAR) element, as a part of the paused RNA Pol II complex proximal to the latent promoter (42), is a more likely possibility.

The underlying mechanisms regulating HIV-1 latency remain enigmatic. There have been several attempts to understand HIV-1 latency, as this question is of direct relevance for clinical management and viral purging (43). The complexity of HIV-1 latency has led to two distinct schools of thought to explain this phenomenon: the hypothesis of "epiphenomenon," where host environmental factors, including epigenetic modifications, play a deterministic role (1, 44), and that of "viral circuitry," where decision-making is hardwired in the intrinsic Tat-LTR regulatory circuit (7, 9). The two models, which need not necessarily be mutually exclusive, have been supported by considerable experimental evidence but also have specific limitations.

Experimental evidence for epigenetic modifications controlling HIV-1 latency is available from studies using clonal cell populations typically harboring subgenomic viral reporter vectors (45). The major limitation of this experimental model is the prolonged periods required for the cells to establish latency. The majority of individual clonal populations reach 50% latency on average in 30 to 80 days, which probably is not representative of the kinetics of natural latency. Given the cytotoxic properties of the viral products and the immune response and the relatively short life span of

FIG 10 Legend (Continued)
The mean values from quadruplicates ± SD are plotted. A two-tailed, unpaired *t* test was used for statistical analyses (*, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ns, nonsignificant). (D and G) A qPCR-based ChIP assay was performed independently using identical experimental conditions and the 4-κB (D) or the 3-κB (G) clone. The data for each IP were calculated using the percent input method. NS refers to a matched IgG isotype control antibody. The mean values from qPCR triplicates ± SD are plotted. A two-tailed, unpaired *t* test was used for statistical analyses (*, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ns, nonsignificant). (H) Two additional qPCR-based ChIP assays using two different mouse monoclonal anti-Tat antibodies (1D9 and NT3 5A5.3) were performed using the active and latent fractions of the two bimodal clones. The mean values from qPCR triplicates ± SD are plotted. A two-tailed, unpaired *t* test was used for statistical analyses (*, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ns, nonsignificant).

infected T cells (~2.2 days) (46), viral gene expression is expected to drive viral evolution toward rapid, and not prolonged, latency establishment in natural infection. Additionally, it is also not understood how epigenetic silencing of an active viral promoter is ever achieved, especially in the presence of abundant quantities of Tat.

The contrasting model explaining HIV-1 latency based on the intrinsic and virus-driven stochastic phenomenon is also supported by compelling experimental evidence (7, 9). The "feedback resistor" module (12) considers a single type of chemical modification, acetylation, and deacetylation, of Tat, serving as the "resistor" or dissipater of the positive transcription loop to ensure a stable latent state. This model does not take into account several other PTMs of Tat. Whereas dimethylation of the lysine residues at positions 50 and 51 (47) and the arginine residues at positions 52 and 53 (48) can suppress Tat transactivation, monomethylation of the lysine residues shows the opposite effect (49). Importantly, methylated Tat is expected to have enhanced cellular stability, with implications for latency (50). In addition to acetylation, the phosphorylation of multiple serine and threonine residues can cooperatively enhance Tat transactivation (51, 52). Polyubiquitination of Tat can also enhance the stability of Tat, thereby augmenting its transactivation function (53). Apart from chemical modifications, Tat is also known to be inactivated by the propensity of the protein to make dimer and multimer forms, although experimental evidence is scant in this regard (54). Furthermore, the differential forward and reverse reaction kinetics of Tat acetylation have been evaluated only in HeLa cells and not in cells of physiological relevance to HIV-1 infection (55). Weinberger et al. also analyzed feedback strength in terms of the noise autocorrelation function and demonstrated that stronger Tat feedback would yield transcriptional pulses of longer durations leading to cell lysis, while weaker Tat feedback and Tat-independent transcription would generate shorter transcriptional pulses leading to latency (8). This model, however, does not reconcile the fact that a small proportion of T cells can still escape cell death following Tat-mediated transcription and establish a viral reservoir (56). In summary, despite significant experimental evidence, the question regarding the critical deterministic factor(s) regulating HIV-1 latency remains unresolved.

Subtype-associated molecular features may offer vital clues to HIV-1 latency. Although the fundamental constitution of the HIV-1 promoter is highly conserved among various genetic subtypes of HIV-1, there are many subtype-specific molecular features that may modulate gene expression considerably. Such differences are evident in the copy numbers and nucleotide sequences of different TFBSs, especially those of the USF, c-Myb, LEF-1, Ets1, NF-AT, Ap-1, NF-κB, and Sp1 binding sites, and regulatory elements such as the TATA box and the TAR element (57–59). These regulatory elements play critical roles in positively regulating basal and inducible levels of viral transcription (60, 61). Most TFBSs, especially the AP-1, USF, NFAT, NF-κB, and Sp1 motifs, also play a critical role in regulating viral latency by recruiting chromatin-modifying complexes and transcription-suppressing factors such as histone deacetylases (HDACs) to the viral promoter (62–66).

Of the various TFBSs, both NF-κB and Sp1 motifs are represented by multiple and tandem binding sites in the LTR and play crucial roles in regulating gene expression and latency (36, 67–69). The most striking feature of the HIV-1C LTR is the copy number difference of NF-κB motifs, the sequence variation of the additional κB motifs (25), and the associated sequence variation of the Sp1III site (32). We demonstrated previously that NF-κB site duplication is unique to HIV-1C and not recapitulated by any other HIV-1 genetic family. Importantly, in HIV-1C, a unique NF-κB motif (the C-κB element, GGGGCGTTCC) and a genetically distinct and subtype-specific Sp1III site are located at the core of the promoter, and the two elements establish a functional association in enhancing HIV-1C transcription (32).

HIV-1C could serve as an ideal model to ask whether transcriptional strength can affect viral latency, a property not explored previously, probably due to the absence of NF-κB copy number variation in non-HIV-1C subtypes. The rapid expansion of the 4-κB viral strains in India in a short period of 10 years, from 2% to 25 to 35%, is quite

surprising (25). The ATF model that we used here also demonstrated a perfect positive correlation between the number of NF-κB binding sites in the LTR and the viral transcriptional output in the form of EGFP and Tat transcripts (Fig. 2), suggesting that all four NF-κB binding sites in the LTR are functional. In this backdrop, it remains intriguing why HIV-1C strains require enhanced transcriptional strength and, despite having a strong promoter, how these viral strains establish and maintain latency while the other HIV-1 genetic subtypes do not adopt such an evolutionary strategy.

Reciprocal binding of host factors at active and latent promoters. Gene expression is the outcome of multiple layers of regulatory events consisting of *cis*-acting TFBSs and *trans*-acting chromatin remodelers; viral factors, especially Tat; epigenetic marks; and cross talk between a wide array of proteins, ultimately leading to diverse phenotypic outcomes. Numerous studies attempted to examine how the nature of the host factor complexes recruited to the LTR regulates the dynamic switching between the active and latent states (45, 70). In an elegant analysis, Burnett et al. used PheB cells derived from GFP^{Mid} parental Jurkat cells (analogous to GFP^{Dim} in reference 7 and GFP^{Low} in the present study) and compared the natures of cellular complexes recruited between transcriptionally active and latent cells (26). This study demonstrated a nonoverlapping function of the two genetically identical NF-κB sites in regulating transcriptional activation versus suppression. We employed a similar experimental strategy, with the exception that we used $EGFP^{High}$ clonal cell populations (the bimodallers) (Fig. 5B) that manifested a bimodal phenotype of EGFP expression. The $EGFP^{High}$ and $EGFP^{-}$ subsets of bimodal cell clones offered a significant technical advantage of normalizing the inherent differences in cellular parameters.

The preferential binding of p50 and p65 (RelA) at latent and active promoters ascertained the repressive and inducing functions of the p50-p50 homodimer and the p50-RelA heterodimer, respectively, of HIV-1 transcription (Fig. 10) (26, 35, 36). 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 of NFAT1 and NFAT2 at active and latent promoters, respectively, in the context of clonal cells. Since NF-κB and NFAT factors share overlapping sites (71–73), NFAT may have a significant influence on latency in HIV-1C. 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.

The key finding of the present study, however, is the detection of the association of the Tat protein with the latent LTR. The results were highly reproducible and consistent between the two strong viral promoters (Fig. 10). The data were also consistent between conventional PCR and quantitative real-time PCR performed following immunoprecipitation. The data were reproducible when three different anti-Tat antibodies targeting different epitopes in Tat were used in the assay. The Tat protein was found associated with the active 4-κB and 3-κB promoters at levels that were 1.7- and 3-fold higher, respectively, than those for 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 than the active promoter.

Negative feedback circuits represent a powerful and common strategy that biological systems exploit to regulate gene expression. Negative feedback circuits can rapidly switch off signaling cascades; therefore, this mode of gene regulation represents the most common strategy that biological systems exploit to regulate gene expression (74). Molecules of biological significance controlling powerful signaling cascades, such as cytokines and transcription factors, often attenuate their production using negative feedback loops. The transcription factor NF-κB, which controls the expression of numerous cellular factors that regulate a wide variety of cellular processes, downregulates self-expression by activating the inhibitor protein IκBα (75).

Likewise, interleukin-2 (IL-2), the most potent cytokine that regulates T cell viability and proliferation, limits self-production by activating the expression of a FOXP3-mediated negative feedback loop (76). Given that latency establishment is central for HIV-1 survival toward evading immune surveillance and minimizing cytotoxicity, an active molecular mechanism would be necessary to suppress gene expression from the LTR rapidly. The decision-making to achieve such a critical phase of the viral life cycle must be an intrinsic characteristic of the MTRC of the virus; it could not be left to stochastic phenomena or epiphenomena regulated by cellular events. As the MTRC of HIV-1 comprises only two elements, the LTR and Tat, and the latter is the only factor encoded by the virus, Tat is the viral factor best positioned to regulate both viral transcription and transcriptional silence, perhaps at different phases of the viral life cycle following integration. The data presented here using two different latency cell models not only are consistent with this critical biological function ascribed to Tat but also provide additional information on latency. In the present work, we examined the latency profile in the context of HIV-1C only, and its validity must be examined in other genetic families of HIV-1.

Based on several facts, the master regulator of the virus is well positioned to be a potential candidate to impose negative feedback on the LTR in a temporal fashion, including the absence of a known transcription suppressor encoded by the virus, the ability of Tat to constitute the master regulatory circuit of the virus in combination with the LTR in the absence of other viral factors, the presence of Tat in the latent cell detected reproducibly and also recruited to the latent promoter, and the identification of a positive correlation between the transcriptional strength of the LTR and the rate of latency establishment.

A Tat-dependent negative feedback mechanism to establish latency? Based on the present study, we propose a novel model for the transcriptional repression of HIV-1 through a Tat-negative feedback mechanism. The attenuation of Tat-positive feedback signaling has been proposed to cause LTR silencing, triggered by extracellular cues (deterministic model) or limiting Tat levels probabilistically (stochastic model) (7, 8, 12, 26). In either case, the Tat concentration gradually falls below a threshold which is not sufficient for self-renewal or successful transcriptional elongation.

Our data allude to a concentration-dependent interconversion 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 (Fig. 11). We propose that the autonomous Tat feedback loop initially favors the steady accumulation of Tat molecules to enhance transcription. Subsequently, at a point when the Tat intracellular concentration surpasses a specific threshold level, Tat switches to the suppression mode, downregulating transcription, possibly depending on differential PTMs of Tat itself. Hence, our model proposes that the strong promoters (3- and 4-κB LTRs), characterized by stronger Tat feedback, can initiate rapid transcriptional silence compared to the weak promoters (2-, 1-, and 0-κB LTRs).

Our data raise several important questions related to HIV-1C latency, which were beyond the scope of the present study. Is the LTR of HIV-1C likely to continue to acquire additional copies of NF-κB and/or other transcription factor binding sites to augment the transcriptional strength further? Of note, unpublished data from our laboratory (D. Bhange, N. Prasad, H. Prajapati, S. Singh, K. Mehta, Y. S. Gohil, B. Anangi, S. P. Maurya, P. G. Bindu, S. Nadig, B. Das, A. Shet, M. Dias, M. Thakar, R. Gangakhedkar, D. Chaturbhuj, S. Saravanan, P. Balakrishnan, R. Paranjape, and U. Ranga, unpublished data) demonstrate a recent trend of the emergence of at least 10 different types of TFBS variant HIV strains in India. Furthermore, how do the variant NF-κB motifs unique to HIV-1C modulate viral latency? The answers to these questions will shed light on the mechanism of HIV-1 latency and will likely help in the design of novel therapeutic strategies to purge HIV infection.

We acknowledge a few technical limitations of the present work that were beyond our control. Our attempts to extend the observations to full-length HIV-1C molecular

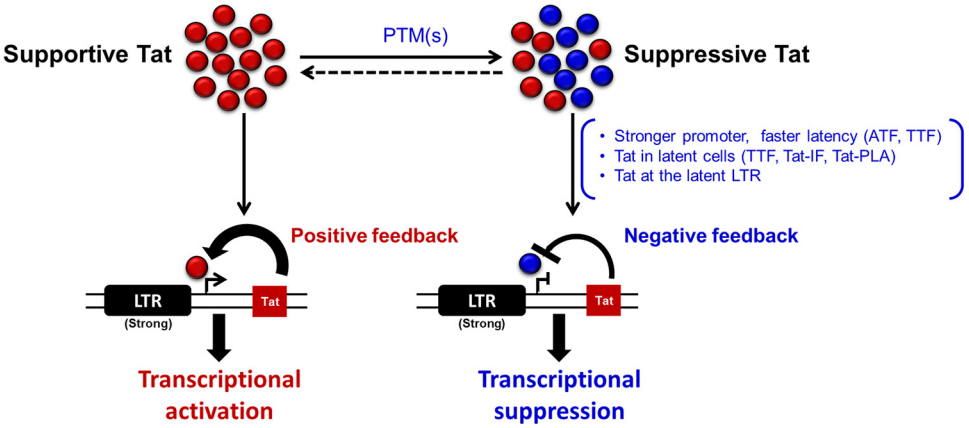

FIG 11 Hypothetical model depicting a concentration-dependent switch of the Tat protein through differential PTMs that mutually serve as an activator or suppressor at different phases of viral transcription. The model suggests that the diverse PTMs of Tat play a critical role in permitting the transactivator to toggle between being an activator and being a suppressor of the transactivation of the LTR. At the time of commitment to latency, the intracellular concentrations of Tat are not limiting. Tat may initiate a negative feedback response of viral transcription in a concentration-dependent manner by regulating the expression of host factors that control the PTM(s) of Tat. The negative feedback effect of Tat therefore follows its initial positive feedback on the viral promoter. A viral promoter characterized by stronger transcriptional activity thus mediates the establishment of latency at a higher rate by producing more Tat. The data presented in this work are consistent with this model.

clones were not successful for two main reasons: (i) there was a lack of good molecular clones representing HIV-1C (only four molecular clones of HIV-1C are available) and (ii) unlike NL4-3, an HIV-1B molecular clone that is used the most in the field, the few available HIV-1C clones do not lend themselves to significant molecular manipulation of any kind; the *in vitro* infectivity of HIV-1C strains is significantly low compared to other subtypes (see reference 77 for a recent review, which highlights this problem of HIV-1C explicitly). Furthermore, the engineering of a fluorescent protein as a reporter will make these clones nearly unviable. Our data were derived from infection of Jurkat cells, the most popular cell model in the field (26, 45, 78–81). Latently infected primary cells *in vivo* are extremely rare, in the order of 10^{-6} (1), and it would be a daunting task to isolate Tat from such a small population. Furthermore, to examine latency establishment in primary cells, a stable cell pool must be generated and reactivated. Such manipulations in primary cells would require at least 6 to 8 weeks, a time frame not conducive to sustaining primary cells in culture. Primary cell models have traditionally been used to examine latency reversal and to evaluate latency-reversing agents but have rarely been used to study latency establishment. A suitable, long-lasting, primary cell-based model to study the mechanisms involved in HIV-1 latency establishment is an absolute requirement, and we have been optimizing this model in our laboratory for future studies. Additionally, we have been raising antibodies specific to defined PTMs of Tat to investigate how Tat PTM may modulate HIV-1 latency.

MATERIALS AND METHODS

Cell culture. Jurkat cells were maintained in RPMI 1640 medium (catalog number R4130; Sigma-Aldrich, St. Louis, MO, USA) supplemented with 10% fetal bovine serum (FBS) (catalog number RM10435; HiMedia Laboratories, Mumbai, India), 2 mM glutamine (catalog number G8540; Sigma-Aldrich), 100 U/ml penicillin G (catalog number P3032; Sigma-Aldrich), and 100 g/ml streptomycin (catalog number S9137; Sigma-Aldrich).

The human embryonic kidney cell lines HEK293 and HEK293T were cultured in Dulbecco's modified Eagle's medium (DMEM) (catalog number D1152; Sigma-Aldrich) supplemented with 10% FBS. All cells were incubated at 37°C in the presence of 5% CO_2.

Design and construction of HIV-1C reporter vector panels. (i) Autonomous Tat feedback model. The pLGIT reporter vector (HIV-1 LTR-EGFP-IRES-Tat) (7), in which two different elements, the 3′ LTR and Tat, were of HIV-1B origin (NL4-3), was a kind gift from David Schaffer (University of California). We replaced these two elements with analogous counterparts of HIV-1C origin (Indie_C1; GenBank accession number AB023804) and referred to the vector as pcLGIT (cLTR-EGFP-IRES-cTat) (32). Using the pcLGIT backbone, we constructed a panel of five reporter vectors containing varying numbers of copies of

TABLE 1 Primer sets used for site-directed mutagenesis of NF-κB copy number variant LTRs in the pcLGIT/pcLdGIT vector (p911 series) (ATF model)

LTR variant	Primer pair[a]	Description of primer pair	Sequence of primer (5′–3′)
FHHC (4-κB)	N1992 FP	Inner primers	TGACACAGAAGGGACTTTCTGCTGACACAGAAGGGACTTTCCGCTGGGACTTTCCACTGGGGCGTTCC
	N1993 RP		AAGTCCCAGCGGAAAGTCCCTTCTGTGTCAGCAGAAAGTCCCTTCTGTGTCAGCAGTCTTTGTAAAACTCCG
OHHC (3-κB)	N1994 FP	Inner primers	TGACACAGAATCTACTTTTTGCTGACACAGAAGGGACTTTCCGCTGGGACTTTCCACTGGGGCGTTCC
	N1995 RP		AAGTCCCAGCGGAAAGTCCCTTCTGTGTCAGCAAAAAGTAGATTCTGTGTCAGCAGTCTTTGTAAAACTCCG
OOHC (2-κB)	N1996 FP	Inner primers	TGACACAGAATCTACTTTTTGCTGACACAGAATCTACTTTTTGCTGGGACTTTCCACTGGGGCGTTCC
	N1997 RP		AAGTCCCAGCAAAAAGTAGATTCTGTGTCAGCAAAAAGTAGATTCTGTGTCAGCAGTCTTTGTAAAACTCCG
OOOC (1-κB)	N1998 FP	Inner primers	TGACACAGAATCTACTTTTTGCTGACACAGAATCTACTTTTTGCTTCTACTTTTTACTGGGGCGTTCC
	N1999 RP		AAGTAGAAGCAAAAAGTAGATTCTGTGTCAGCAAAAAGTAGATTCTGTGTCAGCAGTCTTTGTAAAACTCCG
OOOO (0-κB)	N2000 FP	Inner primers	TGACACAGAATCTACTTTTTGCTGACACAGAATCTACTTTTTGCTTCTACTTTTTACTTCTACTTTTTAGG
	N2001 RP		AAGTAGAAGCAAAAAGTAGATTCTGTGTCAGCAAAAAGTAGATTCTGTGTCAGCAGTCTTTGTAAAACTCCG

[a]FP, forward primer; RP, reverse primer.

functional NF-κB motifs, ranging from 0 to 4 (p911a vector series). First, an LTR containing four tandem NF-κB motifs (FHHC LTR [H, GGGACTTTCC; C, GGGGCGTTCC; F, GGGACTTTCT {variations among the κB motifs are underlined}]) was generated by overlap PCR using the LTR of Indie_C1 as the template. The additional 22-bp sequence constituting the F-κB motif was adopted from the HIV-1C molecular clone BL42-02 (GenBank accession number HQ202921). The amplified FHHC-LTR was inserted into the pcLGIT vector, replacing the original 3′ LTR. Subsequently, using overlap PCR, inactivating point mutations were introduced sequentially into the FHHC (4-κB) LTR to generate the other members of the panel: OHHC (3-κB), OOHC (2-κB), OOOC (1-κB), and OOOO (0-κB) (Fig. 2A). Of note, the inactivation mutations introduced only base substitutions, and not deletions, keeping the length of the viral promoter constant among the variant viral vectors. The mutated κB motif "O" contains the sequence TCTACTTTTT (underlined bases represent inactivating mutations). The variant LTR fragments were cloned directionally between the XhoI and PmeI sites present on the outer primers N1990 FP (5′-GCGTACCTCGAGTGGAAG GGTTAATTTACTCCAAGAAAAGGC-3′) and N1991 RP (5′-TATGTCGTTTAAACCTGCTAGAGATTTTCCACACTA CCAAAAGGGTCTGAG-3′), thus replacing the original 3′ LTR of pcLGIT. Therefore, the members of the vector panel are genetically identical except for differences in the functional NF-κB motifs in the LTR. The internal primer sequences used to generate the point mutations in the LTR variants are mentioned in Table 1. The 3′-LTR sequences of all the panel members were sequence confirmed, and the expression of EGFP was ascertained in HEK293T cells.

A second panel of the five variant viral vectors, analogous to the p911a panel, was also constructed using the pcLdGIT backbone, where EGFP was replaced with d2EGFP, a variant form of the fluorescent protein characterized by a shorter half-life (p911b vector series). First, the d2EGFP ORF was amplified from the pCAG-GFPd2 plasmid (catalog number 14760; Addgene, MA, USA) using the primer pair N1142 FP (5′-CAGGAATTCGATGCTACCGGTCGCCACCATG-3′) and N1143 RP (5′-TCCTACTAGTAGGATCTGAGTCC GGACTACACATTGATCC-3′) and directionally cloned between the AgeI and BspEI sites to replace the EGFP with the d2EGFP reporter gene in the pcLGIT backbone. To generate the p911b panel, the variant LTRs of the p911a panel were transferred directionally to the pcLdGIT vector backbone between the PmeI and XhoI sites, thus replacing the original 3′ LTR. The expression of d2EGFP from all the vectors of the p911b panel was verified using HEK293T cells.

(ii) Tunable Tat feedback model. In the tunable Tat feedback (TTF) model, the HIV-1C LTR regulates the coexpression of d2EGFP and the Tat-RFP fusion protein from the vector pcLdGITRD (cLTR-d2EGFP-IRES-cTat:RFP:DD). The 5′ LTR in the pcLdGITRD vector transcribes a single transcript encoding d2EGFP and a 1,314-bp-long fusion cassette separated by an IRES element. The fusion cassette is a combination of three different ORFs: (i) the cTat expression segment (BL4-3; GenBank accession number FJ765005.1), (ii) the ORF of DsRed2-RFP, and (iii) the FKBP destabilization domain (DD) (28). The three components of the "Tat:RFP:DD" cassette were independently amplified using appropriate templates and primers, and finally, using overlap PCR, a fusion ORF was generated (primer sequences are provided in Table 2). The Tat ORF from the pcLdGIT-3-κB vector (p911b series; ATF model) was replaced with the Tat:RFP:DD ORF, thus generating the pcLdGITRD-3-κB viral vector. pcLdGITRD-3-κB was subsequently used as the parental vector to construct the other member, pcLdGITRD-1-κB, of the p913 panel (Fig. 6A) by cloning the respective 3′ LTRs between PmeI and XhoI in the pcLdGITRD backbone. d2EGFP expression from the two members of the p913 panel was confirmed in HEK293T cells.

Shield1 dose-dependent Tat:RFP:DD expression from the pcLdGITRD vector (TTF model) in HEK293T cells. The FKBP DD tag in the pcLdGITRD construct marks the Tat:RFP:DD fusion protein for rapid degradation through the proteasome pathway (28). However, Shield1, a 750-kDa cell-permeable ligand, can bind the DD motif and rescue the fusion cassette from rapid processing in a dose-responsive manner with minimum off-target effects (29). Thus, by changing the concentration of Shield1 in the culture medium, the pcLdGITRD construct permits the fine-tuning of the intracellular concentration of Tat without altering the transcriptional strength of the LTR. To validate Shield1 dose-dependent Tat:RFP:DD expression,

TABLE 2 Primers used for the construction of the pcLdGITRD vector backbone (p913 series) (TTF model)

Amplicon	Amplicon length (bp)	Primer pair	Sequence of primer (5′–3′)
d2EGFP-IRES-Tat	1,779	N2720 FP	TTTCTTCCATTGCGGCCGCGCCGCCACCATGGCCTCCTCCGAGAACGTC
		N2724 RP	GGCCATTTCGAAGTCGAAGGGGTCT
DsRed2-RFP	629	N2723 FP	ACTTCGAAATGGCCTCCTCCGAGAACG
		N2726 RP	TCCGATATCCAGGAACAGGTGGTGGC
FKBP DD	353	N2725 FP	TGTTCCTGGATATCGGAGTGCAGGTGGAAACCATC
		N2722 RP	CGTACGCGGCGCGCCTCATTCCAGTTCTAGAAGCTCC

approximately 0.6 million HEK239T cells were transfected with 1 μg of the pcLdGITRD-3-κB vector in each well of a 12-well culture dish and treated with various concentrations (0, 0.5, 1.0, 2.5, 4.0, and 5.0 μM) of Shield1 (catalog number 632189; TaKaRa Clontech). After 48 h of transfection, the expressions of both DsRed2-RFP and d2EGFP were recorded using a fluorescence microscope.

Generation of pseudotyped reporter virus and estimation of relative infectious units. Pseudotyped reporter viruses were generated in HEK293T cells. Each viral vector was transfected together with 3rd-generation lentiviral packaging vectors using the standard calcium phosphate protocol (82). Briefly, a plasmid DNA cocktail consisting of 10 μg of the individual viral vectors (NF-κB motif variants), 5 μg of psPAX2 (catalog number 11348; NIH AIDS Reagent Program, MD, USA), 3.5 μg of pHEF-VSVG (catalog number 4693; NIH AIDS Reagent Program), and 1.5 μg pCMV-rev (catalog number 1443; NIH AIDS Reagent Program) was transfected in a 100-mm dish seeded with HEK293T cells at 30% cell confluence. pCMV-RFP (0.2 μg) was used as an internal control for transfection. At 6 h posttransfection, the medium was replenished with complete DMEM. Culture supernatants were harvested at 48 h posttransfection, filtered using a 0.22-μm filter, and stored in 1-ml aliquots in a deep freezer for future use.

The relative infectious units (RIU) of the pseudotyped reporter viruses were quantified in Jurkat T cells by measuring EGFP or d2EGFP expression by flow cytometry. Precisely, 3×10^4 Jurkat cells in each well of a 12-well tissue culture plate were infected with individual viral stocks serially diluted 2-fold (from 10× to 80× dilution) in a total volume of 1 ml of 10% RPMI 1640 containing 25 μg/ml of DEAE-dextran. At 6 h postinfection, the cells were washed and replenished with 1 ml of complete RPMI 1640. After 48 h, the cells were activated with a combination of 40 ng/ml PMA (catalog number P8139; Sigma-Aldrich), 40 ng/ml TNF-α (catalog number T0157; Sigma-Aldrich), 200 nM TSA (catalog number T8552; Sigma-Aldrich), and 2.5 mM HMBA (catalog number 224235; Sigma-Aldrich) for 18 h, after which the percentages of EGFP$^+$ or d2EGFP$^+$ cells were analyzed using a flow cytometer (BD FACSAria III sorter; BD Biosciences, NJ, USA). Following this, titration curves were constructed and analyzed for 5 to 10% infectivity of the cells by regression analysis, which corresponds to ~0.05 to 0.1 RIU. For the TTF model, cells were maintained in 1 μM Shield1 throughout the procedure.

Viral gene expression analysis. Viral gene expression levels were compared among the LTR variant strains of the pcLGIT panel (ATF model). Toward this, approximately 1 million Jurkat cells were infected with each viral strain of the panel independently at ~0.5 to 0.6 RIU. Three days following infection, the cells were activated using a cocktail of global T cell activators (40 ng/ml PMA, 40 ng/ml TNF-α, 200 nM TSA, and 2.5 mM HMBA), and 24 h following activation, EGFP fluorescence (mean fluorescence intensity [MFI]) was estimated for both uninduced and activated cells using flow cytometry. Tat transcript levels in the control and activated cells were determined by Tat RT-PCR (see below). Fold enhancement in EGFP and Tat expression levels were obtained from the ratios of the EGFP MFI values of the activated and control samples and, similarly, their relative Tat mRNA values, respectively.

To compare gene expression levels from the LTR at different transcriptional strengths of Tat feedback, modulated by the Shield1 concentrations (TTF model), approximately 0.3 million Jurkat cells in a 35-mm culture dish were infected with the cLdGITRD-3-κB viral strain at a p24 equivalent of 20 ng/ml. At 24 h postinfection, the infected cells were washed and replenished with complete RPMI 1640. The cells were then equally distributed into four wells of a 6-well plate, and each well was treated with either 0, 1.0, 2.5, or 5.0 μM Shield1. After 48 h, half of the cells from each Shield1 treatment were activated using the global T cell activators, and 24 h later, both the uninduced and activated fractions from each dose of Shield1 were subjected to d2EGFP and Tat transcript expression analyses using flow cytometry and RT-PCR, respectively.

FACS analysis and generation of stable Jurkat cells and clonal lines. (i) ATF model. Individual NF-κB variant pseudotyped viral stocks of the cLGIT/cLdGIT panel were added to 1×10^6 Jurkat cells in a 35-mm culture dish at ~0.05 to 0.1 RIU in a total volume of 2 ml complete RPMI 1640 supplemented with 25 μg/ml DEAE-dextran. After 6 h of infection, cells were washed to remove DEAE-dextran, transferred to 5 ml of complete RPMI 1640 medium in a T-25 flask, and maintained under standard culture conditions. The infected cell pools were expanded over 7 days and induced with the global T cell activation cocktail as mentioned above. After 18 h of activation, total EGFP$^+$ (MFI, $>10^3$ RFU), EGFP^{High} (MFI, $>10^4$ RFU), or d2EGFP^{High} (MFI, $\sim 0.5 \times 10^3$ to 0.5×10^4) cells were FACS sorted from the respective cell pools depending on the subsequent viral latency assays. A small aliquot of the sorted cell population was reanalyzed to confirm the purity of the sorted cells. Each sorted cell pool with stable EGFP/d2EGFP expression represented a mixed population with random proviral integrations with the corresponding NF-κB variant strain.

Stable clonal cell lines of the cLGIT variant panel (expressing EGFP) were established by sorting a single cell per 3 wells in a 96-well plate. Each well of the collection plate contained 100 μl of a mix of

TABLE 3 Primers used for determination of the proviral integration frequency and chromatin immunoprecipitation-qPCR

Target region and purpose	Primer pair-probe combination	Primer or probe sequence (5′–3′)[a]
Strong-stop DNA for determination of proviral integration frequency	N2208 FP	GATCTGAGCC(T/C)GGGAGCTCTCTG
	N2209 RP	TCTGAGGGATCTCTAGTTACCAGAGTC
	N2210 probe	FAM-CTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCT-TAMRA
LTR enhancer for chromatin immunoprecipitation-qPCR	N2493 FP	CCGGAGT(A/T)TTACAAAGACTGCTG
	N2215 RP	CTGCTTATATGCAGCATCTGAGG
	N2492 probe	HEX-CACTGGGGCGTTCCAGG(G/A)GG(A/T)GT-BHQ

[a]FAM, 6-carboxyfluorescein; TAMRA, 6-carboxytetramethylrhodamine; HEX, 6-carboxy-2,4,4,5,7,7-hexachlorofluorescein; BHQ, black hole quencher.

equal proportions of complete and spent RPMI 1640 media. The cells were diluted to a cell density of 0.1×10^6 cells/ml before the sort.

(ii) TTF model. A total of 1×10^6 Jurkat cells were infected with the subgenomic cLdGITRD viral strains (3- and 1-κB variants) at ~0.1 to 0.2 RIU in 1 ml of complete RPMI 1640 medium supplemented with 25 μg/ml of DEAE-dextran and 1.0 μM Shield1. At 6 h postinfection, the infected cells (J-cLdGITRD) were washed and replenished with 1 ml of complete RPMI 1640 supplemented with 1 μM Shield1. Next, 72 h following infection, the cells were induced with the above-mentioned T cell activation cocktail for 18 h, and following activation, stable d2EGFP$^+$ J-cLdGITRD cells were sorted.

Analysis of proviral integration frequency. TaqMan qPCR was used to determine the mean number of proviral integrations per cell using genomic DNA extracted from the stable J-cLGIT cell pools and also from the EGFP^{High} and EGFP$^-$ subfractions corresponding to the representative bimodal clones of the 3-κB and the 4-κB variants. Genomic DNA was extracted from 1×10^6 stable cells using the GenElute mammalian genomic DNA kit (catalog number G1N350; Sigma-Aldrich) according to the manufacturer's instructions. The extracted DNA was dissolved in Tris-EDTA (TE), and the concentration was adjusted to 70 ng/μl. Five microliters of this solution was equivalent to approximately 10^5 copies of the human genome. The stock DNA solution was subjected to 10-fold serial dilution up to a final DNA concentration of 10^1 copies/5 μl and used as the template for PCR. A 129-bp fragment spanning the R-U5 region of the HIV-1 5′ LTR (positions +18 to +147) was amplified using the primer-probe combination N2208 FP, N2209 RP, and N2210 probe (see Table 3 for the primer sequences) for TaqMan real-time PCR. A standard curve was established simultaneously using genomic DNA extracted from J-Lat 8.4 cells that contain a single proviral copy per cell (78). The proviral copy number of the query samples was then estimated using regression analysis.

Generation of kinetic profiles of latency establishment. (i) ATF model. Kinetic curves of latency were established to compare the rates of promoter silencing among the NF-κB variant strains of the ATF panel. Toward this, the sorted cell pools (total EGFP$^+$, EGFP^{High}, or d2EGFP^{High}) were maintained under standard experimental conditions while a small aliquot was collected at regular intervals to monitor EGFP/d2EGFP expression using the FACSAria III flow cytometer. Temporal kinetic profiles for percent GFP$^+$ cells and GFP MFIs were constructed and compared among the five NF-κB variants for total EGFP$^+$, EGFP^{High}, as well as d2EGFP^{High} cells. The EGFP reporter gene, with a longer half-life (~48 h), was measured every 4 days for 16 to 24 days, while analysis of the expression of d2EGFP, with a shorter half-life (~2 h), was performed every 24 h following sorting for 7 days.

Kinetic profiles of latency were established for the clonal lines of the cLGIT panel. Sorted, single EGFP^{High} cells corresponding to each NF-κB variant strain were expanded to form clonal lines, and 16 such lines corresponding to each cLGIT viral variant were flow analyzed for their EGFP expression profiles on days 21 and 28 postsorting.

(ii) TTF model. Latency establishment profiles at varied strengths of the Tat feedback circuit were studied using the TTF model by altering the Shield1 concentrations in the culture medium. Sorted d2EGFP^{High} cells corresponding to the cLdGITRD-3-κB (strong) and the cLdGITRD-1-κB (weak) LTR variants were divided into four separate fractions and maintained at four different concentrations of Shield1 (0, 0.5, 1, and 3 μM) under standard experimental conditions, while a small aliquot was collected from all the fractions every 24 h to monitor the expression of both d2EGFP and Tat-RFP using the FACSAria III flow cytometer, for 7 days. Temporal kinetic profiles were constructed and compared between the two NF-κB variants as well as across the different concentrations of Shield1.

Live/dead analysis of cells. A live/dead viability assay was performed to exclude dead cells before every flow analysis. Cell samples were stained using the Live/Dead fixable red dead cell stain kit (catalog number L34972; Molecular Probes, Thermo Fisher Scientific, MA, USA) according to the manufacturer's protocol. Briefly, cell samples were harvested in a microcentrifuge tube, washed once with 1× phosphate-buffered saline (PBS), and resuspended in 500 μl of 1:1,000-diluted live/dead stain. The cells were then incubated for 30 min at room temperature in the dark, after which they were washed, resuspended in 500 μl of RPMI 1640 supplemented with 2% FBS, and analyzed using the flow cytometer.

DNA cell cycle analysis. Cell cycle analysis by quantitating the DNA content was performed on the EGFP^{High} and EGFP$^-$ subfractions of the bimodal clonal lines of the J-cLGIT-3-κB and -4-κB variants. The standard propidium iodide (PI) staining protocol was followed, with slight modifications to minimize the quenching of the EGFP signal (83). Briefly, 1×10^6 cells were harvested into 1.5-ml microcentrifuge tubes and washed once with 1× PBS. The pellets were resuspended thoroughly in 1 ml fix solution (2% glucose and 2% paraformaldehyde in 1× PBS) and incubated on ice for 10 min. The cells were then washed once with 1×

PBS and resuspended in 100 μl of 1× PBS, and 900 μl of ice-cold 70% ethanol was added dropwise, with gentle vortexing. The fixed cells were incubated on ice for 1 h and washed with 1× wash solution (20 mM HEPES, 0.25% NP-40, and 0.1% bovine serum albumin [BSA] in 1× PBS), and the supernatants were aspirated and resuspended in 500 μl of 1× PBS containing 10 μg/ml RNase A and 20 μg/ml PI. This was followed by incubation in the dark for 30 min at room temperature and analysis by flow cytometry. Pulse processing (pulse area versus pulse height) was used to exclude doublets, and debris and the gated singlets were applied to the PI histogram plot to determine the percentages of cells in the G_1, S, and G_2/M phases.

Quality assurance and data analysis by flow cytometry. The BD FACSAria III sorter was optimized and calibrated before every operation (flow analysis or sorting) to ensure quality performance. The instrument was calibrated for optical laser alignment, fluorescence and light scatter resolution, and fluorescence detector sensitivity using BD CS&T beads (catalog number 656504; BD Biosciences). The drop delay before every FACS sort was determined using BD FACS Accudrop beads (catalog number 345249; BD Biosciences), while fluorescence linearity and doublet discrimination before DNA cell cycle analysis were assessed using BD DNA quality control (QC) particles (catalog number 349523; BD Biosciences). Furthermore, uniform photomultiplier tube (PMT) voltage parameters were set for every fluorescence channel using appropriate negative-control samples to avoid variations in the MFI measurements at different time points of the latency kinetics experiments.

All flow cytometry data analyses were performed using FCS Express versions 4 and 6 (De Novo Software, Los Angeles, CA).

Analysis of Tat transcripts in stable Jurkat cells. We quantitated Tat transcript levels using real-time PCR as a surrogate marker of the transcriptional status of the LTR during latency establishment or latency reversal. Total mRNA was extracted from 0.5×10^6 cells using a single-step RNA isolation reagent, Tri reagent (catalog number T9424; Sigma-Aldrich), at specified time points. Using random hexamer primers, 250 to 1,000 ng of extracted RNA was converted to cDNA in a reaction mixture volume of 20 μl using the Tetro cDNA synthesis kit (catalog number BIO-65043; Bioline, London, UK). The cDNA was then amplified using a SYBR green RT-PCR kit (catalog number 06924204001; Roche Products, Mumbai, India) for a 139-bp region in Tat exon 1 using primers N1783 (5′-GGAATCATCCAGGAAGTCAG CCCGAAAC-3′) and N1784 (5′-CTTCGTCGCTGTCTCCGCTTCTTCCTG-3′). GAPDH RT-PCR was employed as an internal control (primer pair N2232 [5′-GAGCTGAACGGGAAGCTCACTG-3′] and N2233 [5′-GCTTCACC ACCTTCTTGATGTCA-3′]). Relative gene expression was calculated using the $\Delta\Delta C_T$ method.

Indirect immunofluorescence of Tat. Immunofluorescence staining of Tat was performed at multiple time points during the establishment of viral latency in stable J-cLdGIT-3-κB cells, characterized by strong d2EGFP fluorescence (MFI range, 5×10^3 to 50×10^3). The sorted d2EGFP^{High} cells were considered the day 0 sample, and Tat IF analysis was performed subsequently at an interval of every 4 days. Approximately 3×10^6 cells were collected in a 1.5-ml vial, washed once with 1× PBS, and fixed with 2% paraformaldehyde in PBS for 10 min at room temperature, with mild rocking. Fixed cells were rewashed with 1× PBS, followed by permeabilization with 0.2% Triton X-100 in PBS for 10 min, with gentle and intermittent vortexing. Fixed and permeabilized cells were then washed again with 1× PBS and blocked with 4% BSA in PBS for 30 min at room temperature, with mild rocking. The blocked cells were incubated with a rabbit polyclonal anti-Tat antibody (catalog number ab43014; Abcam, Cambridge, UK) at a 1:250 dilution for 1 h 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 (catalog number A-11010; Molecular Probes) for 20 min in the dark at room temperature, followed by a PBS wash. The nucleus was stained with 4 μg/ml of DAPI (4′,6-diamidino-2-phenylindole) for 20 min in the dark at room temperature. Cells were washed twice and mounted on coverslips with 70% glycerol for confocal imaging. Images were acquired with a Zeiss LSM 880 confocal laser scanning microscope with Airyscan using a Plan Apochromat 63×/1.4 oil immersion objective and analyzed using ZEN 2.1 software. For imaging of single cells, a 4×-higher zoom was applied. The fluorescence cutoff for d2EGFP expression was determined using uninfected Jurkat control cells, while that for Tat-Alexa Fluor 568 expression was set using the infected but no-primary Tat antibody control. At each time point, 150 cells were analyzed, fluorescence intensities (arbitrary units [AU]) for d2EGFP and Tat-Alexa Fluor 568 expressions were independently determined for the nuclear and extranuclear compartments of the cells, and temporal curves were generated.

Proximity ligation assay. We used an *in situ* proximity ligation assay (PLA) to detect Tat in HEK293 cells independently infected with the cLdGIT-3-κB and -4-κB reporter viruses. The assay was performed using a commercial kit (Duolink *in situ* red starter kit for mouse/rabbit, catalog number DUO92101; Sigma-Aldrich) according to the instructions of the manufacturer. Briefly, a heterogeneous population of HEK293 cells harboring both active and latent viruses (cLdGIT-3-κB or -4-κB), marked by the presence and absence of green fluorescence, respectively, were seeded on glass coverslips and allowed to grow to 60 to 70% confluence. The evenly distributed cells 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 three times with 1× PBS. This was followed by blocking for 1 h using the reagent supplied in the kit. The blocked cells were then treated with the rabbit polyclonal anti-Tat antibody at a 1:250 dilution (catalog number ab43014; Abcam) in combination with the mouse monoclonal anti-Tat antibody at a 1:250 dilution (catalog number 7377; NIH AIDS Reagent Program, MD, USA). The cells were incubated with a pair of probes (PLA probe anti-mouse minus [catalog number DUO92004] and PLA probe anti-rabbit plus [catalog number DUO92002]) in a 40-μl reaction mixture volume for 1 h at 37°C, followed by washing twice with 500 μl of wash buffer A for 5 min each time. The ligation and amplification reactions were performed according to the manufacturer's instructions using Duolink *in situ* red detection reagents (catalog number DUO92008). DAPI-supplemented mounting medium (catalog number DUO82040, supplied in the PLA kit) was used for mounting the cells. No-primary- and single-antibody

controls were used to assess nonspecific PLA spots. Imaging of the cells was performed using a Zeiss LSM 880 confocal laser scanning microscope with Airyscan fitted with a Plan Apochromat 63×/1.4 oil immersion objective. Signal intensities of the PLA-positive spots were quantitated manually using ImageJ software. A total of 164 cells were analyzed from the d2EGFP$^+$ category (128 from the 4-κB sample and 36 from the 3-κB sample), while 168 cells were analyzed from the d2EGFP$^-$ category (123 from the 4-κB sample and 45 from the 3-κB sample) to compare the numbers of PLA dots per cell between the two phenotypes.

The primary antibody pair, the rabbit polyclonal anti-Tat (catalog number ab43014; Abcam) and mouse monoclonal anti-Tat (catalog number 7377; AIDS Reagent Program) antibodies, was validated for Tat specificity before performing the PLA on stable HEK293 cells harboring the cLdGIT-3-κB and -4-κB proviruses. Approximately 0.5×10^6 HEK293T cells in each well of an 8-microchamber glass slide (catalog number 80826; ibidi, Grafelfing, Germany) were transfected with either 800 ng of the pcLGIT vector (B-Tat) or 200, 400, or 800 ng of the pcLGIT vector (C-Tat). After 48 h of transfection, the Tat PLA was performed as detailed above. Confocal images were captured using the same model of confocal microscope and parameters identical to the ones mentioned above. ImageJ software was used to measure d2EGFP intensity (AU), and manual quantitation was performed on Tat PLA spots from 25 cells corresponding to each dose of the pcLGIT vector.

Chromatin immunoprecipitation assay. We used a chromatin preparation equivalent of 2×10^6 cells (either EGFP^{High} or EGFP$^-$) for each immunoprecipitation assay, as described previously (32). Briefly, 2×10^6 Jurkat cells collected in a 1.5-ml vial were washed with 1× PBS, resuspended in 1 ml of RPMI 1640 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, followed by centrifugation at 3,000 rpm for 5 min at 4°C with a subsequent PBS wash (mixture containing 0.01× protease inhibitor cocktail [PIC], catalog number 11836170001; Roche Applied Science, Indianapolis, IN, USA). Following the complete removal of PBS, the cells were resuspended in 100 μl of ice-chilled lysis buffer (1% SDS, 50 mM Tris buffer [pH 8.0], 10 mM EDTA) and incubated on ice for 20 min, with occasional mixing of the lysate using a wide-bore tip. The lysate in each vial was subjected to 22 cycles of sonication in the high mode, using a 30-s-on followed by a 30-s-off pulse scheme in the Bioruptor plus sonicator (catalog number UCD-300; Diagenode, Liege, Belgium) containing prechilled water. The sonicated lysate was centrifuged at 12,000 rpm for 10 min at 4°C to remove any cellular debris, and the clear supernatant was transferred to a fresh 1.5-ml vial and stored at −80°C until use. One-tenth of the lysate (10 μl) was used to confirm the shearing of chromatin to generate 200- to 500-bp fragments. Each IP was comprised of 100 μl of the lysate and 2 μg of an antigen-specific antibody against p50 (catalog number ab7971; Abcam, Cambridge, UK), p65 (catalog number ab7970; Abcam), NFAT1 (catalog number ab2722; Abcam), NFAT2 (catalog number ab2796; Abcam), HIV-1 Tat (catalog number ab43014 [Abcam] or catalog number 7377 or 4374 [NIH AIDS Reagent Program]), RNA Pol II C-terminal domain (CTD) phospho-S2 (catalog number ab5095; Abcam), or H3K9Me3 (catalog number ab8898; Abcam). The chromatin-immunoprecipitated DNA was amplified using the primer pair N1054 FP [5′-GATCTGAGCC(T/C)GGGAGCTCTCTG-3′] and N1056 RP (5′-TCTGAGG GATCTCTAGTTACCAGAGTC-3′), spanning a 240-bp sequence within the enhancer-core promoter region in the LTR. The amplified DNA fragments were subjected to agarose gel electrophoresis, and the band intensities were normalized using the percent input method to compare the differential recruitments of each transcription factor at the active versus the latent promoter. To enhance the sensitivity of the assay, TaqMan qPCR was performed using ChIP DNA and the primer-probe combination N2493 FP, N2215 RP, and N2492 probe (Table 3). The final data were evaluated using the percent input method.

Statistics. Statistical analyses were performed using GraphPad Prism 5.0 software. The statistical tests used to calculate *P* values are indicated in the corresponding figure legends.

ACKNOWLEDGMENTS

We thank Tapas Kumar Kundu (JNCASR, India) and Ravi Manjithaya (JNCASR, India) for intellectual discussions. We thank Uttara Chakraborty, S. L. Swaroopa Yalla, and Narendra Nala of the flow cell at JNCASR; B. S. Suma of the Confocal Imaging Facility; and G. Anitha of the Sequencing Facility at JNCASR, India. We thank Neelakshi Varma and Surabhi Jirapure for their initial help in establishing the latency models. Several reagents were obtained through the AIDS Research and Reference Reagent Program.

This work was supported by grants to U.R. from the Department of Biotechnology (DBT), Government of India (DBT grant numbers BT/PR7359/29/651/2012 and BT/INF/22/SP27679/2018); the U.S. National Institutes of Health (NIH) (grant number NIDA 5RO1DA041751-02); and intramural funds from JNCASR.

S.C., conception and design, acquisition of data, analysis and interpretation of data, drafting or revising the article, and providing essential unpublished data; M.K., acquisition of data, analysis and interpretation of data, and providing essential unpublished data; U.R., conception and design, funding acquisition, validation, writing, reviewing, and editing the article.

We declare that no competing interests exist.

REFERENCES

1. Chun T-W, Carruth L, Finzi D, Shen X, DiGiuseppe JA, Taylor H, Hermankova M, Chadwick K, Margolick J, Quinn TC, Kuo Y-H, Brookmeyer R, Zeiger MA, Barditch-Crovo P, Siliciano RF. 1997. Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection. Nature 387:183–188. https://doi.org/10.1038/387183a0.
2. Finzi D, Blankson J, Siliciano JD, Margolick JB, Chadwick K, Pierson T, Smith K, Lisziewicz J, Lori F, Flexner C, Quinn TC, Chaisson RE, Rosenberg E, Walker B, Gange S, Gallant J, Siliciano RF. 1999. Latent infection of CD4+ T cells provides a mechanism for lifelong persistence of HIV-1, even in patients on effective combination therapy. Nat Med 5:512–517. https://doi.org/10.1038/8394.
3. Siliciano RF, Greene WC. 2011. HIV latency. Cold Spring Harb Perspect Med 1:a007096. https://doi.org/10.1101/cshperspect.a007096.
4. Eisele E, Siliciano RF. 2012. Redefining the viral reservoirs that prevent HIV-1 eradication. Immunity 37:377–388. https://doi.org/10.1016/j.immuni.2012.08.010.
5. Van Lint C, Bouchat S, Marcello A. 2013. HIV-1 transcription and latency: an update. Retrovirology 10:67. https://doi.org/10.1186/1742-4690-10-67.
6. Archin NM, Sung JM, Garrido C, Soriano-Sarabia N, Margolis DM. 2014. Eradicating HIV-1 infection: seeking to clear a persistent pathogen. Nat Rev Microbiol 12:750–764. https://doi.org/10.1038/nrmicro3352.
7. Weinberger LS, Burnett JC, Toettcher JE, Arkin AP, Schaffer DV. 2005. Stochastic gene expression in a lentiviral positive-feedback loop: HIV-1 Tat fluctuations drive phenotypic diversity. Cell 122:169–182. https://doi.org/10.1016/j.cell.2005.06.006.
8. Weinberger LS, Dar RD, Simpson ML. 2008. Transient-mediated fate determination in a transcriptional circuit of HIV. Nat Genet 40:466–470. https://doi.org/10.1038/ng.116.
9. Razooky BS, Pai A, Aull K, Rouzine IM, Weinberger LS. 2015. A hardwired HIV latency program. Cell 160:990–1001. https://doi.org/10.1016/j.cell.2015.02.009.
10. Ho Y-C, Shan L, Hosmane NN, Wang J, Laskey SB, Rosenbloom DIS, Lai J, Blankson JN, Siliciano JD, Siliciano RF. 2013. Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure. Cell 155:540–551. https://doi.org/10.1016/j.cell.2013.09.020.
11. Weinberger AD, Weinberger LS. 2013. Stochastic fate selection in HIV-infected patients. Cell 155:497–499. https://doi.org/10.1016/j.cell.2013.09.039.
12. Weinberger LS, Shenk T. 2007. An HIV feedback resistor: auto-regulatory circuit deactivator and noise buffer. PLoS Biol 5:e9. https://doi.org/10.1371/journal.pbio.0050009.
13. Arkin A, Ross J, McAdams HH. 1998. Stochastic kinetic analysis of developmental pathway bifurcation in phage λ-infected Escherichia coli cells. Genetics 149:1633–1648.
14. Dodd IB, Perkins AJ, Tsemitsidis D, Egan JB. 2001. Octamerization of λ CI repressor is needed for effective repression of P RM and efficient switching from lysogeny. Genes Dev 15:3013–3022. https://doi.org/10.1101/gad.937301.
15. Dwarakanath RS, Clark CL, McElroy AK, Spector DH. 2001. The use of recombinant baculoviruses for sustained expression of human cytomegalovirus immediate early proteins in fibroblasts. Virology 284:297–307. https://doi.org/10.1006/viro.2001.0924.
16. Sanders RL, Clark CL, Morello CS, Spector DH. 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. J Virol 82:7059–7077. https://doi.org/10.1128/JVI.00675-08.
17. Stinski MF, Petrik D. 2008. Functional roles of the human cytomegalovirus essential IE86 protein, p 133–152. *In* Shenk T, Stinski MF (ed), Human cytomegalovirus. Springer, New York, NY.
18. Ragoczy T, Miller G. 2001. Autostimulation of the Epstein-Barr virus BRLF1 promoter is mediated through consensus Sp1 and Sp3 binding sites. J Virol 75:5240–5251. https://doi.org/10.1128/JVI.75.11.5240-5251.2001.
19. Sarisky RT, Gao Z, Lieberman PM, Fixman ED, Hayward GS, Hayward SD. 1996. A replication function associated with the activation domain of the Epstein-Barr virus Zta transactivator. J Virol 70:8340–8347. https://doi.org/10.1128/JVI.70.12.8340-8347.1996.
20. Cai W, Astor T, Liptak L, Cho C, Coen D, Schaffer P. 1993. The herpes simplex virus type 1 regulatory protein ICP0 enhances virus replication during acute infection and reactivation from latency. J Virol 67:7501–7512. https://doi.org/10.1128/JVI.67.12.7501-7512.1993.
21. Kent JR, Kang W, Miller CG, Fraser NW. 2003. Herpes simplex virus latency-associated transcript gene function. J Neurovirol 9:285–290. https://doi.org/10.1080/13550280390200994.
22. Roizman B, Gu H, Mandel G. 2005. The first 30 minutes in the life of a virus: unREST in the nucleus. Cell Cycle 4:1019–1021. https://doi.org/10.4161/cc.4.8.1902.
23. Razooky BS, Weinberger LS. 2011. Mapping the architecture of the HIV-1 Tat circuit: a decision-making circuit that lacks bistability and exploits stochastic noise. Methods 53:68–77. https://doi.org/10.1016/j.ymeth.2010.12.006.
24. Bachu M, Mukthey AB, Murali RV, Cheedarla N, Mahadevan A, Shankar SK, Satish KS, Kundu TK, Ranga U. 2012. Sequence insertions in the HIV type 1 subtype C viral promoter predominantly generate an additional NF-κB binding site. AIDS Res Hum Retroviruses 28:1362–1368. https://doi.org/10.1089/aid.2011.0388.
25. Bachu M, Yalla S, Asokan M, Verma A, Neogi U, Sharma S, Murali RV, Mukthey AB, Bhatt R, Chatterjee S, Rajan RE, Cheedarla N, Yadavalli VS, Mahadevan A, Shankar SK, Rajagopalan N, Shet A, Saravanan S, Balakrishnan P, Solomon S, Vajpayee M, Satish KS, Kundu TK, Jeang K-T, Ranga U. 2012. Multiple NF-κB sites in HIV-1 subtype C long terminal repeat confer superior magnitude of transcription and thereby the enhanced viral predominance. J Biol Chem 287:44714–44735. https://doi.org/10.1074/jbc.M112.397158.
26. Burnett JC, Miller-Jensen K, Shah PS, Arkin AP, Schaffer DV. 2009. Control of stochastic gene expression by host factors at the HIV promoter. PLoS Pathog 5:e1000260. https://doi.org/10.1371/journal.ppat.1000260.
27. Li X, Zhao X, Fang Y, Jiang X, Duong T, Fan C, Huang C-C, Kain SR. 1998. Generation of destabilized green fluorescent protein as a transcription reporter. J Biol Chem 273:34970–34975. https://doi.org/10.1074/jbc.273.52.34970.
28. Banaszynski LA, Chen L, Maynard-Smith LA, Ooi AL, Wandless TJ. 2006. A rapid, reversible, and tunable method to regulate protein function in living cells using synthetic small molecules. Cell 126:995–1004. https://doi.org/10.1016/j.cell.2006.07.025.
29. Maynard-Smith LA, Chen L, Banaszynski LA, Ooi AL, Wandless TJ. 2007. A directed approach for engineering conditional protein stability using biologically silent small molecules. J Biol Chem 282:24866–24872. https://doi.org/10.1074/jbc.M703902200.
30. Gustafsdottir SM, Schallmeiner E, Fredriksson S, Gullberg M, Söderberg O, Jarvius M, Jarvius J, Howell M, Landegren U. 2005. Proximity ligation assays for sensitive and specific protein analyses. Anal Biochem 345:2–9. https://doi.org/10.1016/j.ab.2005.01.018.
31. Söderberg O, Gullberg M, Jarvius M, Ridderstråle K, Leuchowius K-J, Jarvius J, Wester K, Hydbring P, Bahram F, Larsson L-G, Landegren U. 2006. Direct observation of individual endogenous protein complexes in situ by proximity ligation. Nat Methods 3:995–1000. https://doi.org/10.1038/nmeth947.
32. Verma A, Rajagopalan P, Lotke R, Varghese R, Selvam D, Kundu TK, 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. J Virol 90:7046–7065. https://doi.org/10.1128/JVI.00308-16.
33. Barbeau B, Bernier R, Dumais N, Briand G, Olivier M, Faure R, Posner BI, 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. J Biol Chem 272:12968–12977. https://doi.org/10.1074/jbc.272.20.12968.
34. Chen-Park FE, Huang D-B, Noro B, Thanos D, Ghosh G. 2002. The κB DNA sequence from the HIV long terminal repeat functions as an allosteric regulator of HIV transcription. J Biol Chem 277:24701–24708. https://doi.org/10.1074/jbc.M200007200.
35. Stroud JC, Oltman A, Han A, Bates DL, 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. J Mol Biol 393:98–112. https://doi.org/10.1016/j.jmb.2009.08.023.
36. Williams SA, Chen LF, Kwon H, Ruiz-Jarabo CM, Verdin E, Greene WC. 2006. NF-κB p50 promotes HIV latency through HDAC recruitment and

repression of transcriptional initiation. EMBO J 25:139–149. https://doi.org/10.1038/sj.emboj.7600900.
37. Kinoshita S, Chen BK, Kaneshima H, Nolan GP. 1998. Host control of HIV-1 parasitism in T cells by the nuclear factor of activated T cells. Cell 95:595–604. https://doi.org/10.1016/S0092-8674(00)81630-X.
38. Kinoshita S, Su L, Amano M, Timmerman LA, Kaneshima H, Nolan GP. 1997. The T cell activation factor NF-ATc positively regulates HIV-1 replication and gene expression in T cells. Immunity 6:235–244. https://doi.org/10.1016/S1074-7613(00)80326-X.
39. Macián F, Rao A. 1999. Reciprocal modulatory interaction between human immunodeficiency virus type 1 Tat and transcription factor NFAT1. Mol Cell Biol 19:3645–3653. https://doi.org/10.1128/mcb.19.5.3645.
40. Southgate CD, Green MR. 1991. The HIV-1 Tat protein activates transcription from an upstream DNA-binding site: implications for Tat function. Genes Dev 5:2496–2507. https://doi.org/10.1101/gad.5.12b.2496.
41. Dandekar DH, Ganesh KN, Mitra D. 2004. HIV-1 Tat directly binds to NFκB enhancer sequence: role in viral and cellular gene expression. Nucleic Acids Res 32:1270–1278. https://doi.org/10.1093/nar/gkh289.
42. Barboric M, Peterlin BM. 2005. A new paradigm in eukaryotic biology: HIV Tat and the control of transcriptional elongation. PLoS Biol 3:e76. https://doi.org/10.1371/journal.pbio.0030076.
43. Mbonye U, Karn J. 2017. The molecular basis for human immunodeficiency virus latency. Annu Rev Virol 4:261–285. https://doi.org/10.1146/annurev-virology-101416-041646.
44. Pierson T, McArthur J, Siliciano RF. 2000. Reservoirs for HIV-1: mechanisms for viral persistence in the presence of antiviral immune responses and antiretroviral therapy. Annu Rev Immunol 18:665–708. https://doi.org/10.1146/annurev.immunol.18.1.665.
45. Pearson R, Kim YK, Hokello J, Lassen K, Friedman J, Tyagi M, Karn J. 2008. Epigenetic silencing of human immunodeficiency virus (HIV) transcription by formation of restrictive chromatin structures at the viral long terminal repeat drives the progressive entry of HIV into latency. J Virol 82:12291–12303. https://doi.org/10.1128/JVI.01383-08.
46. Perelson AS, Neumann AU, Markowitz M, Leonard JM, Ho DD. 1996. HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. Science 271:1582–1586. https://doi.org/10.1126/science.271.5255.1582.
47. Van Duyne R, Easley R, Wu W, Berro R, Pedati C, Klase Z, Kehn-Hall K, Flynn EK, Symer DE, Kashanchi F. 2008. Lysine methylation of HIV-1 Tat regulates transcriptional activity of the viral LTR. Retrovirology 5:40. https://doi.org/10.1186/1742-4690-5-40.
48. Xie B, Invernizzi CF, Richard S, Wainberg MA. 2007. Arginine methylation of the human immunodeficiency virus type 1 Tat protein by PRMT6 negatively affects Tat interactions with both cyclin T1 and the Tat transactivation region. J Virol 81:4226–4234. https://doi.org/10.1128/JVI.01888-06.
49. Pagans S, Kauder SE, Kaehlcke K, Sakane N, Schroeder S, Dormeyer W, Trievel RC, Verdin E, Schnolzer M, Ott M. 2010. The cellular lysine methyltransferase Set7/9-KMT7 binds HIV-1 TAR RNA, monomethylates the viral transactivator Tat, and enhances HIV transcription. Cell Host Microbe 7:234–244. https://doi.org/10.1016/j.chom.2010.02.005.
50. Sivakumaran H, van der Horst A, Fulcher AJ, Apolloni A, Lin M-H, Jans DA, Harrich D. 2009. Arginine methylation increases the stability of human immunodeficiency virus type 1 Tat. J Virol 83:11694–11703. https://doi.org/10.1128/JVI.00499-09.
51. Endo-Munoz L, Warby T, Harrich D, McMillan NA. 2005. Phosphorylation of HIV Tat by PKR increases interaction with TAR RNA and enhances transcription. Virol J 2:17. https://doi.org/10.1186/1743-422X-2-17.
52. Ammosova T, Berro R, Jerebtsova M, Jackson A, Charles S, Klase Z, Southerland W, Gordeuk VR, Kashanchi F, Nekhai S. 2006. Phosphorylation of HIV-1 Tat by CDK2 in HIV-1 transcription. Retrovirology 3:78. https://doi.org/10.1186/1742-4690-3-78.
53. Brès V, Kiernan RE, Linares LK, Chable-Bessia C, Plechakova O, Tréand C, Emiliani S, Peloponese J-M, Jeang K-T, Coux O, Scheffner M, Benkirane M. 2003. A non-proteolytic role for ubiquitin in Tat-mediated transactivation of the HIV-1 promoter. Nat Cell Biol 5:754–761. https://doi.org/10.1038/ncb1023.
54. Tosi G, Meazza R, De Lerma Barbaro A, D'Agostino A, Mazza S, Corradin G, Albini A, Noonan DM, Ferrini S, Accolla RS. 2000. Highly stable oligomerization forms of HIV-1 Tat detected by monoclonal antibodies and requirement of monomeric forms for the transactivating function on the HIV-1 LTR. Eur J Immunol 30:1120–1126. https://doi.org/10.1002/(SICI)1521-4141(200004)30:4<1120::AID-IMMU1120>3.0.CO;2-4.
55. Ott M, Dorr A, Hetzer-Egger C, Kaehlcke K, Schnolzer M, Henklein P, Cole P, Zhou M-M, Verdin E. 2004. Tat acetylation: a regulatory switch between early and late phases in HIV transcription elongation. Novartis Found Symp 259:182–193.
56. van Zyl G, Bale MJ, Kearney MF. 2018. HIV evolution and diversity in ART-treated patients. Retrovirology 15:14. https://doi.org/10.1186/s12977-018-0395-4.
57. Jeeninga RE, Hoogenkamp M, Armand-Ugon M, de Baar M, Verhoef K, Berkhout B. 2000. Functional differences between the long terminal repeat transcriptional promoters of human immunodeficiency virus type 1 subtypes A through G. J Virol 74:3740–3751. https://doi.org/10.1128/jvi.74.8.3740-3751.2000.
58. Mbondji-Wonje C, Dong M, Wang X, Zhao J, Ragupathy V, Sanchez AM, Denny TN, Hewlett I. 2018. Distinctive variation in the U3R region of the 5′ long terminal repeat from diverse HIV-1 strains. PLoS One 13:e0195661. https://doi.org/10.1371/journal.pone.0195661.
59. Montano MA, Novitsky VA, Blackard JT, Cho NL, Katzenstein DA, Essex M. 1997. Divergent transcriptional regulation among expanding human immunodeficiency virus type 1 subtypes. J Virol 71:8657–8665. https://doi.org/10.1128/JVI.71.11.8657-8665.1997.
60. Garcia JA, Harrich D, Pearson L, Mitsuyasu R, Gaynor RB. 1988. Functional domains required for tat-induced transcriptional activation of the HIV-1 long terminal repeat. EMBO J 7:3143–3147. https://doi.org/10.1002/j.1460-2075.1988.tb03181.x.
61. Pereira LA, Bentley K, Peeters A, Churchill MJ, Deacon NJ. 2000. A compilation of cellular transcription factor interactions with the HIV-1 LTR promoter. Nucleic Acids Res 28:663–668. https://doi.org/10.1093/nar/28.3.663.
62. Bosque A, Planelles V. 2009. Induction of HIV-1 latency and reactivation in primary memory CD4+ T cells. Blood 113:58–65. https://doi.org/10.1182/blood-2008-07-168393.
63. Chan JK, Bhattacharyya D, Lassen KG, Ruelas D, Greene WC. 2013. Calcium/calcineurin synergizes with prostratin to promote NF-κB dependent activation of latent HIV. PLoS One 8:e77749. https://doi.org/10.1371/journal.pone.0077749.
64. Colin L, Van Lint C. 2009. Molecular control of HIV-1 postintegration latency: implications for the development of new therapeutic strategies. Retrovirology 6:111. https://doi.org/10.1186/1742-4690-6-111.
65. Duverger A, Wolschendorf F, Zhang M, Wagner F, Hatcher B, Jones J, Cron RQ, van der Sluis RM, Jeeninga RE, Berkhout B, Kutsch O. 2013. An AP-1 binding site in the enhancer/core element of the HIV-1 promoter controls the ability of HIV-1 to establish latent infection. J Virol 87:2264–2277. https://doi.org/10.1128/JVI.01594-12.
66. Rohr O, Marban C, Aunis D, Schaeffer E. 2003. Regulation of HIV-1 gene transcription: from lymphocytes to microglial cells. J Leukoc Biol 74:736–749. https://doi.org/10.1189/jlb.0403180.
67. Baeuerle PA, Baltimore D. 1989. A 65-kappaD subunit of active NF-kappaB is required for inhibition of NF-kappaB by I kappaB. Genes Dev 3:1689–1698. https://doi.org/10.1101/gad.3.11.1689.
68. Doetzlhofer A, Rotheneder H, Lagger G, Koranda M, Kurtev V, Brosch G, Wintersberger E, Seiser C. 1999. Histone deacetylase 1 can repress transcription by binding to Sp1. Mol Cell Biol 19:5504–5511. https://doi.org/10.1128/mcb.19.8.5504.
69. Suzuki T, Yamamoto T, Kurabayashi M, Nagai R, Yazaki Y, Horikoshi M. 1998. Isolation and initial characterization of GBF, a novel DNA-binding zinc finger protein that binds to the GC-rich binding sites of the HIV-1 promoter. J Biochem 124:389–395. https://doi.org/10.1093/oxfordjournals.jbchem.a022124.
70. Mahmoudi T. 2012. The BAF complex and HIV latency. Transcription 3:171–176. https://doi.org/10.4161/trns.20541.
71. Pessler F, Cron R. 2004. Reciprocal regulation of the nuclear factor of activated T cells and HIV-1. Genes Immun 5:158–167. https://doi.org/10.1038/sj.gene.6364047.
72. Bates DL, Barthel KK, Wu Y, Kalhor R, Stroud JC, Giffin MJ, Chen L. 2008. Crystal structure of NFAT bound to the HIV-1 LTR tandem κB enhancer element. Structure 16:684–694. https://doi.org/10.1016/j.str.2008.01.020.
73. Giffin MJ, Stroud JC, Bates DL, von Koenig KD, Hardin J, Chen L. 2003. Structure of NFAT1 bound as a dimer to the HIV-1 LTR κB element. Nat Struct Biol 10:800–806. https://doi.org/10.1038/nsb981.
74. Chatterjee A, Kaznessis YN, Hu W-S. 2008. Tweaking biological switches through a better understanding of bistability behavior. Curr Opin Biotechnol 19:475–481. https://doi.org/10.1016/j.copbio.2008.08.010.
75. Hoffmann A, Levchenko A, Scott ML, Baltimore D. 2002. The IκB-NF-κB

signaling module: temporal control and selective gene activation. Science 298:1241–1245. https://doi.org/10.1126/science.1071914.
76. Smith KA, Popmihajlov Z. 2008. The quantal theory of immunity and the interleukin-2-dependent negative feedback regulation of the immune response. Immunol Rev 224:124–140. https://doi.org/10.1111/j.1600-065X.2008.00654.x.
77. Gartner MJ, Roche M, Churchill MJ, Gorry PR, Flynn JK. 2020. Understanding the mechanisms driving the spread of subtype C HIV-1. EBioMedicine 53:102682. https://doi.org/10.1016/j.ebiom.2020.102682.
78. Jordan A, Bisgrove D, Verdin E. 2003. HIV reproducibly establishes a latent infection after acute infection of T cells in vitro. EMBO J 22:1868–1877. https://doi.org/10.1093/emboj/cdg188.
79. Miller-Jensen K, Dey SS, Pham N, Foley JE, Arkin AP, Schaffer DV. 2012. Chromatin accessibility at the HIV LTR promoter sets a threshold for NF-κB mediated viral gene expression. Integr Biol (Camb) 4:661–671. https://doi.org/10.1039/c2ib20009k.
80. Dahabieh MS, Ooms M, Brumme C, Taylor J, Harrigan PR, Simon V, Sadowski I. 2014. Direct non-productive HIV-1 infection in a T-cell line is driven by cellular activation state and NFκB. Retrovirology 11:17. https://doi.org/10.1186/1742-4690-11-17.
81. Li Z, Guo J, Wu Y, Zhou Q. 2013. The BET bromodomain inhibitor JQ1 activates HIV latency through antagonizing Brd4 inhibition of Tat-transactivation. Nucleic Acids Res 41:277–287. https://doi.org/10.1093/nar/gks976.
82. Jordan M, Schallhorn A, Wurm FM. 1996. Transfecting mammalian cells: optimization of critical parameters affecting calcium-phosphate precipitate formation. Nucleic Acids Res 24:596–601. https://doi.org/10.1093/nar/24.4.596.
83. Zhu H, Coppinger JA, Jang C-Y, Yates JR, III, Fang G. 2008. FAM29A promotes microtubule amplification via recruitment of the NEDD1–γ-tubulin complex to the mitotic spindle. J Cell Biol 183:835–848. https://doi.org/10.1083/jcb.200807046.