Targeting HIV-1 proviral transcription
Alex Olson¹,⁴, Binita Basukala²,⁴, Wilson W Wong³ and Andrew J Henderson¹

Despite the success of antiretroviral therapies, there is no cure for HIV-1 infection due to the establishment of a long-lived latent reservoir that fuels viral rebound upon treatment interruption. ‘Shock-and-kill’ strategies to diminish the latent reservoir have had modest impact on the reservoir leading to considerations of alternative approaches to target HIV-1 proviruses. This review explores approaches to target HIV-1 transcription as a way to block the provirus expression.

Addresses
¹ Department of Medicine, Section of Infectious Diseases, Boston University School of Medicine, United States
² Cell & Molecular Biology, Biology, Boston University, United States
³ Biomedical Engineering, Boston University, United States

Corresponding author: Henderson, Andrew J (andrew.henderson@bmc.org)
⁴ These authors contributed equally to the writing of this manuscript.

Introduction
Antiretroviral therapy suppresses HIV-1 replication and decreases morbidity and mortality of HIV-associated diseases; however, viral replication rapidly rebounds once treatment is discontinued indicating the presence of a long-lived latently infected HIV reservoir. This long-lived reservoir presents a major barrier to curing HIV-1 infection [1–6].

Latently infected cells have decreased or inactive proviral transcription resulting from multiple biochemical mechanisms that include chromatin organization, lack of key transcription factors, fluctuations in Tat expression and/or inefficient proviral transcriptional elongation [7–13]. One strategy for reducing the HIV-1 reservoir is ‘shock and kill’, in which latency reversing compounds induce HIV-1 expression and the reservoir would be eliminated by immune clearance and HIV-1 triggered cell death [14–17]. Compounds that target chromatin remodeling and/or act as pan-T cell activators have been examined for their potential to reverse latency and purge HIV-1 but have failed to decrease the reservoir size in HIV-1 infected patients [14,18,19]. The lack of success of latency reversing agents may reflect the paucity of latently infected cells, which are estimated to be approximately one in 10⁷ cells. Challenges include targeting such a small population of latent cells as well as finding treatments that can effectively deliver both shock and kill signals. Furthermore, an absence of latency markers and the DNA background signal associated with abundant defective proviral DNAs that are resistant to induction, create difficulties for monitoring reservoir size following treatments [20,21,22]. The lack of efficacy of latency reversal also reflects the complexity of the reservoir which includes multiple cell types in different tissues including targeting HIV-1 provirus harbored in CD4+ T memory subsets, macrophages, dendritic cells and other myeloid cells in lymph nodes, gastrointestinal tissues, and the brain [1,23,24]. The poor outcomes of latency reversal strategies have shifted cure strategies away from broad activation of T cells to those that specifically target HIV-1 proviruses including long-term repression of infected cells [18,19,25,26]. Drugging HIV-1-specific transcription has been a long-standing interest and it is impossible to extensively review the large literature focusing on latency, Tat function, and molecular mechanisms of HIV transcription. We will highlight a subset of more recent attempts to engineer RNAs and proteins to target HIV provirus transcription as a potential treatment strategy.

Transcriptional regulation and mechanisms that contribute to HIV latency
HIV-1 transcription is a combinatorial process that includes coordination of transcription factor recruitment, epigenetic regulation and RNA polymerase II (RNAPII) activity [10,11,13]. A number of cellular transcription factors are recruited to the HIV long terminal repeat (LTR) which functions as the promoter and enhancer for the HIV-1 provirus. More than 40 cellular factors have been reported to bind and influence HIV transcription [Los Alamos Data Base; URL: http://www.hiv.lanl.gov/] [27], suggesting that the LTR has a broad range of activities in different cells and is capable of responding to cellular signals and metabolic states. Therefore, HIV-1 expression versus latency will reflect the active transcriptional networks of the infected cell [28–30]. Latent HIV-1 is primarily harbored in quiescent CD4+ T cell subsets, including T memory stem cells, and central, transitional
and effector memory cells, although macrophages and dendritic cells contribute to the latent reservoir in various tissues [24**,28,31,32]. These cell populations may lack transcription factors that mediate transcriptional activation, like NF-κB, and express a wide array of transcriptional repressors that are responsible for maintaining their quiescent state and facilitate transcriptional repression of HIV provirus. Examples of transcriptional repressors that are expressed in quiescent T cells and repress HIV-1 transcription include CBF-1 [33,34], Blimp-1 [35], AP-4 [36], YY-1 [37,38], Ets-2 [39] and Forkhead box proteins (FoxO) [39,40]. These transcriptional repressors bind the HIV-1 LTR and recruit complexes that modify chromatin, regulate RNAPII function or processiveness, and restrict the availability of positive transcription factors such as P-TEFb. For example, BLIMP-1 represses HIV transcription by recruiting HDACs to the HIV LTR which mediated post-translational modification of the positioned nucleosome Nuc-1 as well as enhance negative elongation factor to pause the polymerase [35]. CBF-1 represses HIV-1 transcription through limiting P-TEFb accessibility and mediating nucleosome modifications through HDACs and methyl-transferases [33,34].

One function of cellular transcription factors is to recruit RNAPII and coactivator complexes to the HIV LTR. In addition, transcription factors recruit chromatin modifiers to the promoter, which through histone post-translational modifications including ubiquitination, methylation, acetylation, SUMOylation, ADP-ribosylation, phosphorylation and, corotenylation [41,42], modify chromatin organization and interactions between transcriptional complexes; many of these post-translational modifications of histones have been reported to influence HIV-1 transcription and latency [11,43*,44,45]. The role of histone modifying complexes also underscores the importance of chromatin as a key check-point for HIV transcription and are consistent with observations in several latency models that repression of HIV transcription is associated with a positioned repressive nucleosome, Nuc-1, which presents a barrier to RNAPII processivity and transcription [11–13]. Best studied are the effects of histone acetylation/deacetylation and methylation acting as histone marks associated with transcriptional activation and repression. For example, histone acetylation, which is mediated by histone acetyltransferases (HATs), positively correlates with HIV transcription whereas histone deacetylation, which is directed by histone deacetyltransferase complexes (HDACs) is associated with HIV transcriptional repression. The activity of HATs and HDACs provided the initial rationale for exploring HDAC inhibitors as HIV latency reversing agents. Histone lysine methyltransferases (HKMTs) including SUV39H1, EZH2, G9a, SMYD2 are associated with HIV-1 transcriptional repression and if these HKMTs are inhibited reactivation of provirus transcription in cell lines and resting primary T cells isolated from HIV-1 infected patients on HAART has been observed suggesting these are potential druggable targets [45–50]. The ATP-dependent SWI/SNF chromatin remodeling complexes PBAF and BAF, influences nucleosome organization at the LTR and can be a transcriptional activators or repressors, respectively [51–55]. Despite the functionally distinct roles of SWI/SNF complexes in regulating HIV-1 transcription, BAF and PBAF complexes share many proteins including the key functional subunit, the BRG1 ATPase, making it challenging to specifically target their activities [56].

Specific histone marks can also be read by transcriptional coregulators to recruit specific complexes to promoters. One example relevant to HIV latency is the bromodomain and extra terminal domain (BET) family of protein, BRD4. BRD4 binds hyperacetylated H3 and recruits positive-transcription elongation factor b (P-TEFb) to promoters including the HIV-LTR in the absence of the HIV transactivator, Tat [57]. It has been proposed that BRD4 represses efficient HIV transcription by competing with the HIV transactivator, Tat, for P-TEFb binding which provided the rationale for using BET inhibitors such as JQ1 as latency reversing agents [58–60]. A short isoform of BRD4 binds BRG-1 to bridge BRD and SWI/SNF activities repressing HIV-1 transcription and possibly suggesting a pathway that can be usurped for a block and lock strategy [61*].

Robust HIV transcription requires the HIV encoded transactivator protein, Tat. Tat binds TAR, an RNA stem loop structure that is generated on the first 80 bases of the HIV RNA and promotes efficient transcription elongation by recruiting P-TEFb to the LTR [11,12,62]. P-TEFb enhances RNAPII activity by mediating phosphorylation of RNAPII carboxy terminal domain (CTD) and negative regulators of RNAPII activity, NELF and DSIF, as well as recruiting other coactivator complexes, in particular, the Super Elongation Complex to the active RNAPII complex [11–13,63]. Tat and TAR are required for HIV transcription and stochastic episodic changes in Tat expression have been proposed as a key determinant of HIV proviral transcription, repression, and/or latency [64,65].

P-TEFb consists of Cyclin T1, which physically interacts with Tat, as well as Cdk9, which phosphorylates RNAPII CTD, DSIF, and NELF, releasing the paused RNAPII to enhance HIV-1 transcriptional elongation. In the absence of Tat, transcription is initiated but RNAPII will pause at +45 to +50, generating aborted transcripts [11,63,66]. The propensity of RNAPII to pause on the HIV LTR and the ability of Tat to interact with histone modifiers places Tat as a key regulator of active proviral transcription versus latency.

The necessity of Tat for efficient HIV transcriptional elongation also demonstrates how recruitment of P-TEFb to the HIV LTR is a limiting step for provirus
transcription. P-TEFb is sequestered by an inhibitory complex, 7SK RNP, which includes HEXIM1, HEXIM2, 7SK RNA, MEPCE and LARP7 [63,66–68]. P-TEFb in complex with the 7SK RNP is sequestered in nuclear speckles [69,70] and its release is facilitated by phosphorylation of the T loop of CDK9 and exchange with Tat [71–73]. An alternative model is 7SK RNP is recruited to the LTR through interactions with BRD4 and KAP1, poising the HIV-1 LTR for the exchange of P-TEFb to Tat and transcriptional activation [74,75]. Exchange of P-TEFb from 7SK RNP to Tat is regulated by CDK7 and PPI [63,71,76,77]. Strategies to target components of P-TEFb include engineering a dominant negative mutant CycT1 protein that inhibits HIV transcription by competing with wild-type CycT1 for binding to Tat as well as blocking Cdk9 activation and subsequent phosphorylation of RNAPII CTD and proviral transcriptional elongation [78]. Although, this approach works in latent cell lines, the efficacy and toxicity of this engineered CycT1 protein have not been examined in primary cells or animal models [78].

The importance of Tat and P-TEFb interactions for HIV transcription has led to efforts to develop specific inhibitors of this complex. One particularly promising compound, Didehydro-cortistatin A (dCA), an analog of a steroidal alkaloid found in the marine sponge Corticium simplex [79], has been shown by Valente et al. to be a potent Tat inhibitor [80**]. The primary mechanism of action for dCA is disrupting Tat–TAR interactions by directly binding the Tat basic domain and blocking P-TEFb recruitment and activity [81]. dCA inhibits HIV-1 transcription and reactivation in a variety of cell lines, primary cells and samples from patients treated with ARTs. Furthermore, dCA can be removed and the repression of HIV persists for nearly a month suggesting a long term mechanism of repression [81] which may reflect direct or indirect actions of dCA on histone post-translational modifications, recruitment of repressive BAF complexes and diminished recruitment of RNAPII [82]. dCA has also been reported to inhibit the expression of inflammatory cytokines suggesting additional benefits in treating HIV comorbidities [83], although how dCA is targeting inflammation has not been examined in detail. Importantly, recent studies with a humanized mouse model has shown efficacy in vivo with inhibition of HIV replication and reactivation upon antiretroviral therapy interruption [83].

### Engineering repressors of HIV-1 transcription

In addition to efforts to develop small molecules that target the Tat–TAR–P-TEFb axis engineering approaches utilizing chimeric proteins with dominant negative activities, inhibitory RNAs or gene editing have been explored.

### Dominant negative proteins

Tat has five key functional domains that facilitate its activity as a transcriptional activator: an acidic/proline-rich domain; a zinc-finger/cystine-rich domain that influences folding and structure; a core domain; a basic-arginine-rich motif (ARM) that mediates RNA binding, nuclear localization, and Cyclin T1 binding; a glutamic acid rich activation motif [12,13,62]. Efforts over the years to disrupt Tat activity by generating dominant negative Tat proteins and chimeric inhibitors have focused on these domains. Some recent approaches to disrupt Tat activity include the generation of the dominant negative Nullbasic, in which the ARM domain of Tat was mutated by introducing a stretch of glycines; this inhibits P-TEFb recruitment to the LTR by Tat and reduces HIV transcription and reactivation of latent HIV [84]. Similarly, a Tat variant that included two domains from HEXIM1, HT-1, was engineered to inhibit HIV transcription. HEXIM1 is part of a 7SK RNP which sequesters P-TEFb to this inhibitory complex. The chimeric Tat-HEXIM1 (HT-1) competed with functional Tat for TAR binding as well as reduced available P-TEFb. In cell lines HT-1 repress HIV transcription and reactivation of latent virus [85].

### RNA strategies to blocking Tat function

In the context of HIV-1 infection, several miRNAs have been suggested to regulate HIV-1 transcription either indirectly by influencing T cell maturation and function or directly by generating antiviral RNAs. For example, miRNAs enriched in resting CD4 T cells potently inhibit HIV transcription by targeting 3’ ends of HIV-1 mRNA [86]. Similarly, several groups have reported that cellular miRNAs inhibit HIV-1 expression by interfering with the Tat–TAR–P-TEFb axis. In particular, the expression level of Cyclin T1 is regulated by several sncRNA, miRNA-198, miR-27b, and IncRNA, and NEAT1, which when overexpressed reduce Cyclin T1 and repress HIV-1 transcription [87–89]. In addition to cellular RNAs, it has been reported that HIV generates miRNAs that target the HIV LTR in primary macrophages [90]; however, the role these RNAs play in controlling HIV-1 expression and latency is still not understood.

Given the importance of structured RNAs for the recruitment of Tat to the LTR and the regulation of P-TEFb availability coupled with the relative ease of engineering RNAs and their sequence specificity, several groups have engineered RNA inhibitors to target HIV transcription. Strategies have spanned siRNAs, anti-sense RNAs, aptamers, RNA TAR decoys and ribozymes to inhibit Tat and HIV replication in cell lines and primary cells [91–99]. HIV proviral transcriptional silencing and inhibition of latency reversal in cell lines has been demonstrated using shRNAs to key targets within the LTR and the HIV genome [100–102].
Engineering factors that directly target HIV

Understanding the critical biochemical interactions and functional domains necessary for transcriptional control has allowed chimeric proteins with defined targets and specific activities to be engineered. Engineering proteins that target HIV provirus and its expression has been explored. For example, recent approaches have employed engineered proteins including Zinc finger (ZnF) nucleases, transcription activator-like effector nucleases (TALENs), homing nucleases and meganucleases to mediate site-directed genome editing to directly target HIV-1 proviruses or factors associated with HIV-1 infection [103–105]. Cells resistant to HIV-1 have been generated using these engineered proteins by targeting key entry receptors such as CCR5 [106] and there are clinical trials [107] to transplant HIV-1 resistant ΔCCR5 cells into HIV-1+ individuals to recapitulate cure protocols that utilize transplanted bone marrow cells from ΔCCR5 donors [108,109]. TALENs and ZnF nucleases have also been used to directly modify or target HIV-1 provirus and shown to reduce p24 expression [110] and HIV-1 DNA [111]. While use of ZnF nucleases and TALENs to specifically target HIV-1 have shown promise in vitro and are being explored, the designing and engineering of nucleases can be iterative and, in general, are less amenable for high-throughput screening.

The discovery and development of CRISPR/Cas9 technology have reshaped gene editing partly due to ease and flexibility of targeting the Cas9 nuclease to specific DNA sequences using complementary guide RNAs [112]. In the context of HIV-1, CRISPR/Cas9 tools have been designed to target LTRs and disrupt or eliminate HIV-1 in latently infected cells [113–115]. CRISPR/Cas9 has also been used to successfully excise HIV, in vivo, with a humanized mouse model [116**]. For these experiments infected mice were treated with nanoparticles coated with antiretroviral drugs and a single injection of AAV-CRISPR-Cas9 with multiple gRNAs targeting conserved LTR and gag regions of the HIV genome [116**]. When mice were weaned from ART, two of seven mice showed no signs of viral rebound. Although this outcome seems modest, it is an important step toward showing the clinical potential of CRISPR/Cas9 as an HIV cure method. However, CRISPR/Cas9 induces double stranded breaks that are subsequently repaired through an error-prone non-homologous end joining DNA repair resulting in mutations or insertions and deletions at the targeted site [114,117]. While this can incapacitate the HIV-1 provirus, escape mutants can render CRISPR/Cas9 ineffective [118*,119,120*,121,122]. This HIV-1 escape can be mitigated by using multiple gRNAs that target the LTR making cells less susceptible to HIV infection [118*,120*]

Cas9 has also been engineered to act as a sequence-specific transcriptional activator or repressor rather than a nuclease. For example, a catalytically dead Cas9 (dCas9), which lacks endonuclease activity has been fused with transactivator or transrepressor domains and directed to specific promoters with gRNAs to modulate cellular gene regulation [123,124]. This approach has been utilized to induce HIV expression in latently infected cells by fusing dCas9 to a VP64 transcriptional activation domain and targeting the dCas9-VP64 transactivator to the HIV LTR [125,126]. dCas9 has also been repurposed to repress transcription by creating fusion proteins with repressor domains derived from Kruppel-associated box (KRAB) zinc-finger proteins [127,128**,129–131]. KRAB ZnF proteins have been proposed to maintain the integrity of the genome by silencing transcription of endogenous retroelements through the binding of their LTRs and facilitating post-translational modification of histones, recruitment of repressor complexes including KAP1 and promoting DNA methylation leading to heritable epigenetic repression [132–134]. It is interesting to speculate whether KRAB zinc finger proteins might be able to ‘block and lock’ HIV provirus. KRAB zinc fingers have been reported to repress HIV transcription [135–140] and we are exploring using dCas9-KRAB to target the HIV-LTR. Preliminary studies have shown greater than 70% reduction in HIV-1 proviral expression in cell lines, correlating with epigenetic modifications such as post-translational modification of histones (unpublished observation). Utilizing dCas9 repressor proteins may provide a novel strategy to achieve a deep repression of the latent reservoirs.

Conclusion

HIV persistence and latency are the major challenges to the elimination of HIV infection. Our understanding of basic mechanisms of HIV-1 transcriptional regulation has translated into exciting new treatment strategies that may enhance ‘shock and kill’ strategies or induce deep latency consistent with a block-and-lock strategy. Furthermore, advancements with gene editing approaches have made it realistic to consider directly targeting HIV-1 provirus and cofactors and coreceptors. Although these approaches have been demonstrated to work in vitro and, in a few cases, animal models, there are several practical issues that still exist in targeting HIV proviral transcription. Major challenges to implementing cure strategies include off-target effects since many of the factors are general transcriptional regulators, the ability of the virus to adapt and escape treatments and delivering expression vectors or reagents to the appropriate cells and tissues. Related to this latter issue is the paucity of latently infected cells and the lack of reliable biomarkers to monitor the reservoir. In addition, the breadth of cells that contribute to the reservoir in different tissues that harbor HIV-1 including the brain, gut, and lymph nodes remains unknown. Ideally, any cure strategy would be durable or long-lasting as well as targeting a broad range of HIV subtypes and clades. Finally, any cure would need to balance risk with
the current standard of care using ART which is a proven and effective way to control and limit HIV disease and transmissions. Despite the clear challenges for using engineered proteins and RNAs to either render cells resistant to HIV-1 infection or to target HIV-1 provirus, it is exciting to speculate, that once optimized, how these tools in combination with other cure strategies, including anti-retroviral therapies, latency reversal agents or cell-mediated transplant approaches, could lead to eradication of HIV-1.

Conflict of interest statement
Nothing declared.

Acknowledgements
This work was in part funded by N.I.H.R01 AI138960 and amfAR, The Foundation of AIDS Research Innovation Grant 109603-2-RGRL. We thank Dr Luis Agosto for critical input and suggestions.

References and recommended reading
Papers of particular interest, published within the period of review, have been highlighted as:

● of special interest
●● of outstanding interest

22. Bruner KM, Murray AJ, Pollack RA, Soliman MG, Laskey SB, Capoferri AA, Lai J, Strain MC, Lada SM, Hoh R et al.: Defective proviruses rapidly accumulate during acute HIV-1 infection. Nat Med 2016, 22:1043-1049. This study used viral outgrowth assays and limiting dilution PCR to determine the amount of intact HIV sequences capable of replication. They showed that the number of defective sequences accumulate within early stages of infection and that these are a substantial part of the reservoir.
25. This study used viral outgrowth assays and limiting dilution PCR to determine the amount of intact HIV sequences capable of replication. They showed that the number of defective sequences accumulate within early stages of infection and that these are a substantial part of the reservoir.


This study demonstrated that crotonylation is a histone lysine post-translational modification that influences the epigenetic regulation of latent HIV.


In this study, a short isoform of BRD4 is identified as a corepressor of HIV-1 latency. Using genome wide ChiP-seq and ATAC-seq, the authors show that BRD4s bind to BRG1 subunit of BAF complex, a chromatin remodeler, and together, they mediate transcriptional repression of HIV-1 by maintaining a repressive nucleosome at HIV-1 promoter.


In this study a combination treatment consisting of antiretroviral therapy (ART) and AAV delivery of CRISPR/Cas9 was employed to treat a humanized mouse model infected with HIV. A subset of mice had decreased HIV DNA and viral rebound was not detected with removal of ART providing a proof of concept that eradication of HIV in vivo is possible.


In this study, the authors excised an integrated HIV-1 provirus from infected cells by exploiting the CRISPR/Cas9 system with gRNAs targeting the HIV-1 LTRs. They also showed that cells harboring Cas9 and multiplex gRNAs targeting the HIV-1 LTRs are resistant to HIV-1 infection.


This study utilized multiple guide-RNAs to target HIV-1 and delay Cas9-mediated escape. Furthermore, the authors demonstrated that continuous Cas9 activity leads to detrimental hypermutations that inactivate the HIV-1 provirus.


The authors show that CRISPR/Cas9 system provides a platform for highly specific RNA-guided transcriptional regulation by fusing a nucleo- nase deficient Cas9 with effector domains that have distinct regulatory functions.


